



Magnitudes in the human brain: independent processing of time, space and number

Benoît Martin

► To cite this version:

Benoît Martin. Magnitudes in the human brain: independent processing of time, space and number. Cognitive Sciences. Université Pierre et Marie Curie - Paris VI, 2017. English. NNT : 2017PA066417 . tel-01760214

HAL Id: tel-01760214

<https://theses.hal.science/tel-01760214>

Submitted on 6 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

DOCTORAT DE L'UNIVERSITE PIERRE ET MARIE CURIE

Doctorate of Pierre et Marie Curie University

Spécialité – Specialty : Neurosciences Cognitives

Ecole doctorale – Doctoral School : Ecole Doctorale Cerveau-Cognition-Comportement (ED3C, ED n°158)

Laboratoire – Lab: Inserm / CEA / NeuroSpin / UNICOG

Magnitudes in the Human Brain: Independent processing of Time, Space and Number

Thèse présentée par **Benoît MARTIN**

Thèse dirigée par **Virginie van WASSENHOVE**

Thèse soutenue publiquement le 09 octobre 2017

Jury:

Jennifer COULL (Rapporteur – Laboratoire de Neurosciences Cognitives (LNC), Université Aix-Marseille, France)

David BURR (Rapporteur – Pisa Vision Lab, Université de Florence, Italie)

Jean DAUNIZEAU (Représentant UPMC – Institut du Cerveau et de la Moelle épinière (ICM), Paris, France)

Maria Dolores de HEVIA (Examineur – Laboratoire Psychologie de la Perception (LPP), Paris, France)

Valentin WYART (Examineur – Laboratoire de Neurosciences Cognitives (LNC), Ecole Normale Supérieure, Paris, France)

Virginie van WASSENHOVE (Directrice de Thèse – NeuroSpin, Unité de Neuroimagerie Cognitive (UNICOG), CEA Saclay, Gif-sur-Yvette, France)

Abstract

If you want to catch your train on time you have to estimate how far the station is from your current position and how long it will take you to go there. A growing body of evidence (behavior and neuroimaging studies) suggests that interferences may exist in the perception of different magnitude dimensions, such as time, distance, number, loudness, brightness and so on. These intriguing relationships fascinate researchers for decades and led Walsh (2003) to hypothesize that Time, Space and Number were part of a generalized magnitude system which may be located in the parietal cortex. In this thesis, we first investigated the possible existence of such system. Participants were presented with clouds of dynamic dots and were instructed to either judge the duration of the trial, the total number of dots or the cumulative surface filled by the dots. Manipulating non-temporal magnitude dimensions did not interfere with duration judgments. Instead, numerical and spatial estimates were biased when we manipulated the rate of accumulation of sensory evidence. In this first experiment we asked whether a generalized Bayesian magnitude estimation system would sample evidence using a common, amodal prior, as expected by a generalized magnitude system. Our results suggested that from a Bayesian perspective, computations would necessitate multiple priors, instead of one and unique amodal prior. This work provides substantial evidence against the existence of a generalized magnitude system in which Time, Space and Number share a common neural code and are processed along the same metric.

To further investigate interferences of numerical magnitude on perceived duration, we built a second experiment in which participants were required to reproduce intervals of time and also judge the numerical magnitude of a visual stimulus. Our results revealed that the number-time interaction depended on the numerical format that was used. Perceived duration always increased as a function of the numerosity when sets of items were used (Non-Symbolic numerical information). On the other hand, perceived duration did not always increase as a function of the tested numerosity when participants were presented with Arabic digits (Symbolic numerical information). This work suggests that the numerical magnitude is automatically processed at a non-symbolic level whereas it requires attentional resources when

numerical information is symbolically conveyed. These findings suggest that the number-time interferences may be governed by distinct mechanisms, depending on the numerical format that is used (either non-symbolic or symbolic).

Overview

The thesis is structured in five main chapters. In the introduction (**first chapter**), I will first review general findings in human magnitudes behavior. I will then introduce the idea of a common generalized magnitude system in the brain as well as the implications and predictions of such system, and finally, how magnitude estimations could be realized on the basis of Bayesian computations. The **second chapter** presents the first experiment we performed, in which we investigated the possible existence of a generalized magnitude system and to which extent Bayesian approaches may provide interesting perspectives on magnitudes estimations. The **third chapter** is dedicated to observations on existing literature of interference effects across magnitudes that motivate the second study of the thesis. In this chapter, I argue that there is no strong evidence in favor of a common metric across magnitudes and that the effect of numerical magnitude on perceived duration is task-dependent. In the **fourth chapter**, we aimed to clarify the interaction of numerical magnitude on perceived duration. Our results indicated that perceived duration increased as a function of the numerical magnitude of the stimulus, but depended on the task that was used and on the instructions provided to the participants. In this chapter I discuss a possible automaticity in magnitude processing at a non-symbolic level, but not at a symbolic one. The last chapter (**fifth chapter**) is a general discussion on the implications and predictions of a generalized magnitude system and to which extent they are supported (or not) by empirical data.

The work presented in Chapter 2 has been published in a peer reviewed journal under the following reference: **Martin, Benoît, Martin Wiener, and Virginie van Wassenhove. 2017. “A Bayesian Perspective on Accumulation in the Magnitude System.” *Scientific Reports* 7 (1): 630.** Results presented in Chapter 4 are actually in preparation for publication. Lastly, I contributed to another study (results not presented in this thesis) which is also actually in preparation: **Polti, I, Martin, B, van Wassenhove, V (in prep.) “Distinct effects of attention and working memory on the estimation of duration”.**

Contents

1	Introduction.....	3
1.1	Magnitudes in the Brain	3
1.1.1	The Weber-Fechner law	4
1.1.3	Regression effect	5
1.2	Same mechanisms for the processing of time, space and number?.....	7
1.2.1	The parietal lobe in fMRI studies.....	7
1.2.2	The parietal lobe in clinical studies	8
1.3	A Theory Of Magnitudes (ATOM)	9
1.3.1	A common metric for time, space and number	9
1.3.2	Investigating the directional symmetry imposed by a common metric	11
1.3.3	A common metric implies a scaling effect between magnitudes	14
1.4	A Bayesian Perspective on Magnitude Estimations	14
1.5	Aim of this thesis	17
2	A Bayesian Perspective on Accumulation in the Magnitude System	19
2.1	Summary.....	19
2.2	Reference	19
3	A Theory Of Magnitudes.....	34
3.1	A Common Metric for Time, Space and Number and the issue of scaling.....	34
3.2	Static vs. Dynamic displays and the direction of interference effects	38
3.3	Numerical magnitude affects temporal encoding, not temporal reproduction	44
3.4	Non symbolic magnitude is automatically processed, not the symbolic one	45
4	The larger the longer, but not all the time	47
4.1	Summary.....	47
4.2	Materials & Methods	48
4.3	Results	52
4.4	Discussion & Conclusion.....	62
5	Discussion and Conclusion.....	68
5.1	A common neural code for Time, Space and Number?	68

5.2 From a Bayesian perspective, time, space and number do not share the same priors	69
5.3 The rate of accumulation of sensory evidence interferes with Numerical and Spatial estimates ...	71
5.4 Perceived duration does not always increase as a function of the tested numerosity: no scaling effect	72
5.5 Non-temporal magnitudes do not interfere with duration at a perceptual level.....	73
5.6 The number-time interaction is modulated by attention and numerical format	74
5.7 Conclusions.....	75
Bibliography.....	78
Acknowledgements	85

1 | Introduction

1.1 Magnitudes in the Brain

According to Gallistel (2011), « *mental magnitudes [...] refer to continuous and discrete quantities an animal has experienced* », such as the representation of time, space and number. The representation of the spatial dimension, for example, takes place during path integration (dead reckoning) when an animal has to construct a mental representation of the different locations and objects encountered in its environment in order to find its way back. By summing some little successive displacements (3 steps forward, 2 to the left, 5 to the right) we can estimate quite precisely our current location in a dark room for example, or we can also record the position of different objects, places and landscapes to build a mental map of our environment. Animals are able to estimate different kinds of magnitudes: for example, estimating the number of lions in the opponent group will affect the fight or flight decision and swimming with the larger shoal will increase the probability that the fish survives. Interestingly, animals can also represent time and remember how long it has been since they cached food. In a study on scrub jays (Clayton, Dickinson, and Anthony, 2006), the birds were allowed to cache worms and peanuts, and the choice of which caches to visit depended first on their knowledge of how long had passed before the recovery. Birds initially searched for worms but switched to searching for peanuts if the retention period was too long, displaying the knowledge that the worms may not be edible anymore.

Like other animals, humans are able to estimate the magnitude of different dimensions, such as the duration of an event, the number of apples in a basket, the size of an object, and so on. In the initial definition of mental magnitudes previously provided, Gallistel (2011) also suggested that “*mental magnitudes [...] enter into arithmetic processing*”. In fact, the discriminability of two magnitudes (e.g., different weights), also called the just-noticeable-difference (JND) follows a mathematical relationship which corresponds to the Weber’s law, and is a function of their ratio. While it is quite easy to discriminate the difference between 10

grams and 40 grams, it is more difficult to discriminate the difference between 1.01 kg and 1.04 kg. One possible explanation of the Weber's law is that magnitudes are represented along a logarithmic scale (Cantlon *et al.*, 2009; Dehaene *et al.*, 2008). Such representation would imply that the sensitivity to notice a difference between two successive magnitudes is lower (higher) for large (small) magnitudes. Another possible explanation is the scalar variability of magnitudes according to which the noise (variability) in the representation of the magnitude increases as a function of the magnitude (Petzschner, Glasauer, and Stephan, 2015).

1.1.1 The Weber-Fechner law

The Weber–Fechner refers to a psychophysical law of human perception, which establishes the relation between the physical change of a stimulus and its perceived change. This law is looking for a mathematical relationship between physical and perceived quantities. Perceived quantities are not limited to the number of items in a set, and this law applies to a lot of different dimensions, such as the quantity of time, of space, of brightness, and so on. In a study conducted by Weber (1850), participants were presented with two different magnitude stimulus that differed by small increment, in order to determine the threshold at which they perceived a difference. The JND between two stimulus magnitudes was a function of their ratio (Figure 1.1). Weber's law followed a proportional relationship between the JND (ΔS) and the absolute value of the stimulus magnitude (S). Weber's law can be expressed by the following equation, with a constant factor (k) which depends on the type of tested stimulus:

$$\frac{\Delta S}{S} = k$$

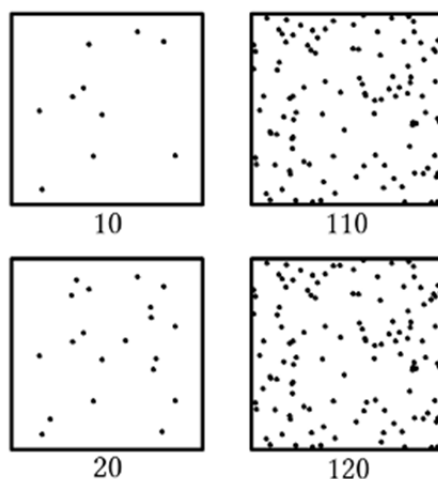


Figure 1.1: Illustration of Weber's law. Lower panels contain 10 more dots than upper panels, but the difference is easier to notice on the left than on the right.

1.1.3 Regression effect

The regression effect, also known as the central tendency effect or regression to the mean, is an effect that is usually reported in psychophysics experiments, in the case of magnitude estimations. This effect describes the tendency for participants to be biased towards the center of the tested range, resulting in the over- (under-) estimation of small (large) magnitudes and in an accurate estimations of magnitudes localized around the center of the distribution (Figure 1.2). For example, if participants have to reproduce durations ranging from 500 to 1000ms, the reproduced durations will be biased towards 750ms.

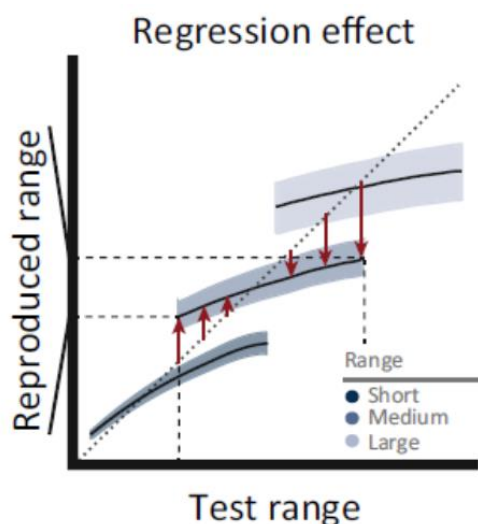


Figure 1.2: From Petzschner *et al.*, 2015. Illustration of the regression effect. The dotted diagonal line represents the ideal performance. Red arrows highlight the regression to the mean, with a bias towards the center of the tested range.

Interestingly, because these features (regression effect and Weber's law) are similar across magnitude dimensions, some authors have argued that similar mechanisms might be used for the perception of different magnitudes (Petzschner, Glasauer, and Stephan, 2015). One of the most important regions of the cortex associated with spatial, numerical and temporal processing is the parietal cortex. Brain damages (lesions) in this area have often been reported to induce deficits in the perception of these dimensions (Saj *et al.*, 2014; Zorzi, Priftis, and Umiltà, 2002). Additionally, clinical and fMRI studies (for a review, see Buetti and Walsh, 2009) gave support to the possible existence of a shared mechanism for the perception of time, space and number.

1.2 Same mechanisms for the processing of time, space and number?

1.2.1 The parietal lobe in fMRI studies

If time, space and number share similar mechanisms, one prediction is that we should observe overlap in brain activations when processing the magnitude of these different dimensions. Moreover, these dimensions should also interfere and the amount of overlap in brain activation should reflect the size and direction of the interaction. Pinel *et al.* (2004) found that the amount of overlap in brain regions activated during numerical, spatial and luminance judgements could predict the amount of interference between dimensions, with number and size engaging a common parietal spatial code, supporting the hypothesis of a common neural code. In Kaufmann *et al.* (2008) study, participants showed overlapping activations for space and number in the posterior superior parietal lobe. The same authors (Kaufmann *et al.* 2008) also compared activations for number and size using a Stroop-like paradigm in which subjects viewed pairs of digits that varied in numerical value and/or size. Participants were required to perform judgements on the value or the physical size and the results revealed that the IntraParietal Sulcus (IPS) was involved in spatial and numerical judgments. Another neuroimaging study found that the bilateral IPS and surrounding areas were activated when processing either spatial or numerical magnitudes (Hubbard *et al.* 2005). Interestingly, the IPS has also been found to be implicated in temporal perception: Coull & Nobre (1998) found an increased BOLD signal in the left IPS in a task requiring attention to temporal intervals. Moreover, transcranial magnetic stimulation (TMS) used to induce temporary disruptions of the posterior parietal cortex caused selective deficits in the processing of temporal, spatial, and numerical magnitudes (for a review, see Sandrini and Rusconi, 2009). Several clinical studies also revealed that brain lesions in and around the parietal lobes were associated with deficits (such as unilateral neglect) in the processing of temporal, numerical and spatial information (Saj *et al.* 2014; Zorzi, Priftis, and Umiltà 2002). Similar activations have also been observed in non-human primates. Regions of macaque parietal cortex (homologous to human IPS) were activated when processing temporal (Leon and Shadlen, 2003), numerical (Sawamura, Shima,

and Tanji, 2002) and spatial (Stein, 1989) magnitudes. Moreover, Tudusciuc & Nieder (2007) have also found neurons in the IPS that are tuned to numerical and spatial magnitudes.

1.2.2 The parietal lobe in clinical studies

Some neurological disorders are associated with deficits in spatial and numerical behaviors, such as dyscalculia (Butterworth, Varma, and Laurillard, 2011) and Gerstmann's syndrome (Benton, 1992). Evidence for a common neural code in the perception of space and number also comes from patients with hemineglect, which is due to right parietal damage. In such cases, patients cannot attend to their left visual field due to contralateral neglect: for example, when hemineglect patients were instructed to set the midpoint of a horizontal line, they typically shifted to the right compared to the objective line midpoint (Schenkenberg, Bradford, and Ajax, 1980).

Interestingly, such bias is also found when neglect patients are asked to perform numerical bisection tasks. Doricchi and colleagues (2005) found that some of the patients they tested presented a strong rightward bias when instructed to bisect number intervals. Another interesting finding is that patients who showed a rightward shift when bisecting number intervals were those with the most severe spatial memory impairments. In a classical clock drawing task patients had to fill an empty clock face with the appropriate numbers (Rossetti *et al.* 2011). The results revealed that patients started with the number 12 (top of the clock) and then continued with the number 13 to 23 instead of 2 to 11. The authors suggested that these findings were due to a strong numerical bias towards large numbers. Vuilleumier *et al.* (2004) instructed the patients to press a left key when a target number was smaller than a reference number, and a right key when it was larger. Analyses of the reaction times revealed that the time to represent the numbers increased when the target number was smaller than the reference; these findings were interpreted as a failure of hemineglect patients to access digits of low numerical information which would be represented along a mental number line. Such neglect has also been reported in temporal tasks (Saj *et al.* 2014) suggesting that time may also

be spatialized on a mental line: left hemispatial neglect patients exhibited deficits in representing temporal events to the past. The authors suggested that patients represented the events along a mental time line and neglected the left part of this line. Altogether, clinical and fMRI studies support the idea that temporal, spatial and numerical magnitude processing may share neural resources, mainly located in the parietal cortex. In line with this idea, Walsh (2003) suggested that time, space and number were part of a generalized magnitude system in which a population of neurons may encode quantity at large, including length, area, volume, duration, numerosity, loudness and so on.

1.3 A Theory Of Magnitudes (ATOM)

1.3.1 A common metric for time, space and number

Walsh proposed that the seemingly distinct domains of space, time and number may be processed by a single cross-domain magnitude system in the brain, a proposal that he named ATOM, for A Theory Of Magnitudes (Walsh, 2003). This domain-general magnitude system is thought to be involved in processing temporal, spatial, and numerical magnitudes, including various dimensions such as size, area, length, density for example. ATOM addresses domains that we experience in terms of “more than” or “less than”. ATOM argues that this shared neural substrate confers benefits because it supports the coordination of magnitudes that are relevant for action (Walsh, 2003; Bueti and Walsh, 2009). For example, when human and non-human animals want to grasp an object, magnitude is relevant to perceive the size of the object, how distant the object is and when we should close our hand to grasp it. Two different schemas can be drawn for processing time, space, number and other magnitude dimensions. In the first case, the different magnitudes can be independently analyzed, processed and compared, according to each individual metric (Figure 1.3 A). The second possibility is to consider a generalized magnitude system (ATOM) in which all the different magnitudes are similarly processed, according to a common metric (Figure 1.3 B).

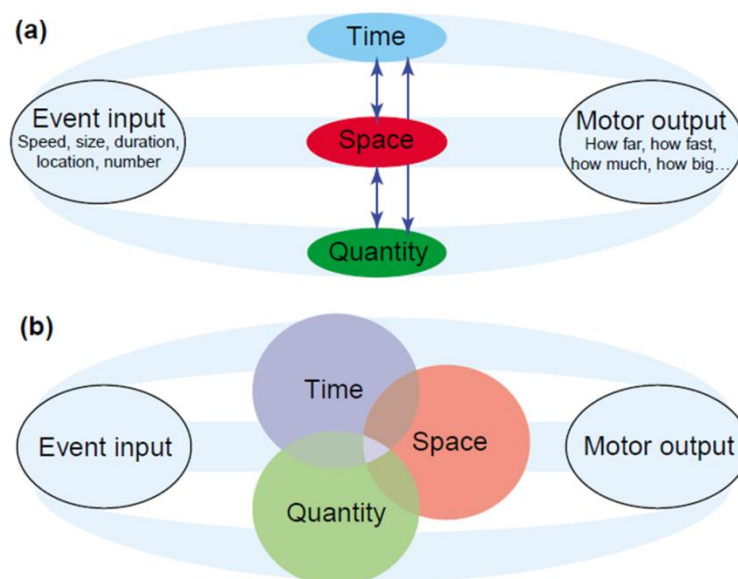


Figure 1.3: From Walsh, 2003. (a) Schema in which the processing of temporal, spatial and numerical information is independent. (b) Schema of a generalized magnitude system in which time, space and number are processed according to a common metric.

The idea of a generalized magnitude system in the brain in which different magnitudes are processed according to a common metric and share neural resources implies that some interactions should exist in the perception of time, space and number. The existence of such a system has been supported by a large number of studies (Burr *et al.*, 2010; Xuan *et al.*, 2007; Oliveri *et al.*, 2008; Dormal and Pesenti, 2007; Casasanto and Boroditsky, 2008; Javadi and Aichelburg, 2012; Mo, 1971; Pinel *et al.*, 2004; Hubbard *et al.*, 2005; Henik and Tzelgov, 1982; Hurewitz, Gelman, and Schnitzer, 2006; Yamamoto, Sasaki, and Watanabe, 2016; Bonato, Zorzi, and Umiltà, 2012). Indeed, one first prediction is that the estimation of the magnitude of a target dimension should be affected by the manipulation of the magnitude in another non-target dimension, such that the larger the magnitude of the non-target, the larger one should perceive the target magnitude dimension to be. In other words, a common metric implies that increasing (decreasing) the magnitude of one dimension should increase (decrease) the perceived magnitude of another dimension. Following this first prediction, a second implication

of a common metric is that the different magnitudes should equally interact between each other and interferences should be bi-directional across dimensions.

1.3.2 Investigating the directional symmetry imposed by a common metric

One prediction from ATOM and the possible existence of a common neural code is that we should observe bi-directional interactions between all the different magnitude dimensions (Winter, Marghetis, and Matlock, 2015). In other words, time, space and number should equally interfere between each other. One known effect in the literature is the size-congruency effect. When instructed to judge the magnitude of a digit, participants tend to respond slower if the physical size of the digit is incongruent with its magnitude (Pinel *et al.*, 2004; Henik and Tzelgov, 1982). In other words, if the size of the digit is congruent with its value (e.g. “1” and “9”), the numerical judgment will be facilitated compared to the condition in which the size of the digit is incongruent with its value (e.g. “1” and “9”). Xuan *et al.* (2007) used a congruent vs. incongruent paradigm and asked the participants to perform a temporal judgments task between two stimuli. They had to judge if the second stimulus was presented for a shorter or a longer duration than the first one. Stimulus consisted in an open square that could vary in size. Results showed that temporal estimations were influenced by variations of the physical size of the stimulus. Similarly, Dormal & Pesenti (2007) designed a Stroop task in which participants were required to compare the length or the numerosity of two linear arrays of dots. Results showed a significant main effect on response latencies: responses were provided faster in the congruent condition than in the incongruent one, when participants performed numerical judgments. However, in the spatial task, the number of dots did not interfere with the processing of spatial information. Such asymmetry has also been reported by Hurewitz *et al.* (2006). In their study, participants were presented with pairs of arrays of dots with varying circle sizes and were required to make numerosity judgments. The authors also investigated if varying the number of circles interfered with judgments of the cumulative filled area. They found an

interference effect of the size of the circles when participants had to judge the number of circles such that reaction times and error rates were larger in the incongruent condition than in the congruent one. However, the effect of numerosity on the area comparison was weaker. In these reports of interference between magnitude dimensions, behavioral effects were concluded on the basis of increased reaction times and error rates in incongruent conditions (e.g., small number presented with a long duration) which prevented the direct evaluation of participants' magnitude perception per se. As such, no clear direction of interference effects could be concluded from the studies beyond the existence of an interaction.

To further investigate interferences across dimensions and the possible existence of bi-directional interactions, Casasanto and Boroditsky (2008) conducted a series of experiment in which participants had to perform spatial and temporal reproduction task. Such design allowed the authors to quantify the size of the interaction (e.g. reproduced duration compared to the objective duration) and to test for possible asymmetry in the interference. In each task participants saw lines or dots on the screen and had to reproduce the spatial displacement or the duration of the trial. The results showed that for a given duration, participants over- (under) estimated the duration when the line traveled a long (short) distance on the screen. No effect of duration on spatial reproduction was found. In another study (Bottini and Casasanto, 2010), duration judgments have been found to be biased by the semantic of words: the estimation of the duration increased as a function of the implicit spatial length of the word. For a given duration, participants judged that the word "Highway" stayed longer on the screen than the word "Pencil", for example. However, the implicit duration of a word did not interfere with spatial judgments. Here again this finding highlights directional asymmetries between magnitude dimensions, which is not consistent with one of our prediction, according to which bidirectional interactions should be observed across magnitudes. Similar asymmetries have been reported between number and time, with numerical information interfering with duration judgments but not the reverse (Dormal, Seron, and Pesenti, 2006). Dormal *et al.* (2006) used a Stroop task in which participants were presented with visually flashing dots and had to compare either the numerosity or the duration of each trial. In the duration task, results showed that congruent condition was answered faster than the incongruent one. No effect was observed in

the numerosity comparison task. Droit-Volet *et al.* (2003) also investigated the interferences in the processing of time and number information. Results showed that an increase of the number of stimuli induced an increase of the “long” responses in the duration discrimination task. In the numerical bisection task, no time interference on the processing of number was found. In this experiment, the effect of number on duration was stronger in 5-years-old children than in 8-years-old children and in adults, suggesting that these asymmetries appear early in development. Supporting this idea, de Hevia & Spelke (2009) found that non-symbolic numerical displays affected the subjective midpoint of a horizontal line in both adults, 3- and 5-year-old children. While number is often reported to affect duration judgments manipulating the duration of events has seldom been reported to affect numerical and spatial magnitudes (Javadi and Aichelburg, 2012; Lambrechts, Walsh, and van Wassenhove, 2013; Cai and Connell, 2015; Martin, Wiener, and van Wassenhove, 2017).

A literal interpretation of ATOM predicts bidirectional interactions between time, space and number but most of the time, asymmetries are reported, with duration being the most labile dimension. To explain such asymmetries, some authors suggest that non-temporal magnitudes are processed automatically (Xuan *et al.* 2007; Dehaene and Akhavein, 1995), and that a more automatic processing interferes with a less automatic one (Hurewitz, Gelman, and Schnitzer, 2006; Dormal, Seron, and Pesenti 2006), which corresponds to a facilitation effect. In a recent paper, Bonn & Cantlon (2012) suggested that *“asymmetries in interference would arise from the different amount of weight given to each dimension in estimating a particular stimulus’ value”*. If one magnitude dimension is processed more automatically than another, we face to possible situations: (1) the difficulty of the task may not be balanced. When investigating interferences between different dimensions, if the numerical task (for example) is easier to perform than the temporal task, it is more likely that numerical magnitude will interfere with perceived duration. A possible difference in task difficulty can lead to an asymmetrical interaction, by facilitating the processing of one magnitude compared to the other. This highlights the need to use a design in which the difficulty of the task is balanced across conditions. The second possibility (2) is to consider that asymmetries are in fact due to an unequal distribution of computational resources when processing numerical and temporal

information for example. Whereas no clear answer can yet be provided about a possible unequal distribution of resources, it appears necessary when building an experimental design to ensure that the difficulty of the task is the same across the different tested dimensions.

1.3.3 A common metric implies a scaling effect between magnitudes

Another prediction directly derived from ATOM and the hypothetical existence of a common neural code is the existence of a scaling effect when estimating different magnitudes. A scaling between different magnitudes implies that the larger the non-temporal magnitude, the larger the perceived duration should be. Such prediction is not always confirmed by empirical data in studies which investigated this specific point (Chang *et al.*, 2011; Rammsayer and Verner, 2016). Indeed, several studies found that the perceived duration did not increase as a function of the absolute numerical value but results were rather better explained by the relative numerical magnitude (“small” and “large”). In other words, the perceived duration was longer for large numerosities than for small ones, but did not linearly increase as a function of the numerical value. At first glance, the lack of scaling effect in these studies is inconsistent with the idea of a common metric for time, space and number. However, this particular point will be investigated in Chapters 3 and 4.

1.4 A Bayesian Perspective on Magnitude Estimations

Recent discussions in the field suggest that the combination and evaluation of quantities in a common representational system would be realized on the basis of Bayesian computations (Petzschner, Glasauer, and Stephan, 2015; Shi, Church, and Meck, 2013). When performing temporal, spatial or numerical judgements, the information that we receive comes from a noisy environment. One concept from the signal detection theory (Green and Swets, 1989; Peterson, Birdsall, and Fox, 1954) is that when observers are instructed to detect the presence of a signal,

observers will be correct in some cases (hit) and incorrect in others (false alarm). Errors may be due to an uncertainty in the decision process, coming from the noisy sensory input but also coming from the previous situations an observer has experienced. This is where Bayes' theorem becomes useful in magnitude estimations. In a Bayesian framework, decision is made by combining a priori information (prior) with noisy sensory input (likelihood), weighing the two information sources by their relative uncertainty. This can be summarized with the following equation (Petzschner, Glasauer, and Stephan, 2015):

$$P(\pi|S) \propto P(S|\pi) \cdot P(\pi)$$

Where $P(S|\pi)$ represents the noisy likelihood, $P(\pi)$ represents the a priori knowledge and $P(\pi|S)$ correspond to the posterior. Figure 1.4 compares a classical model of magnitude estimation and a generative model based on Bayesian probabilities.

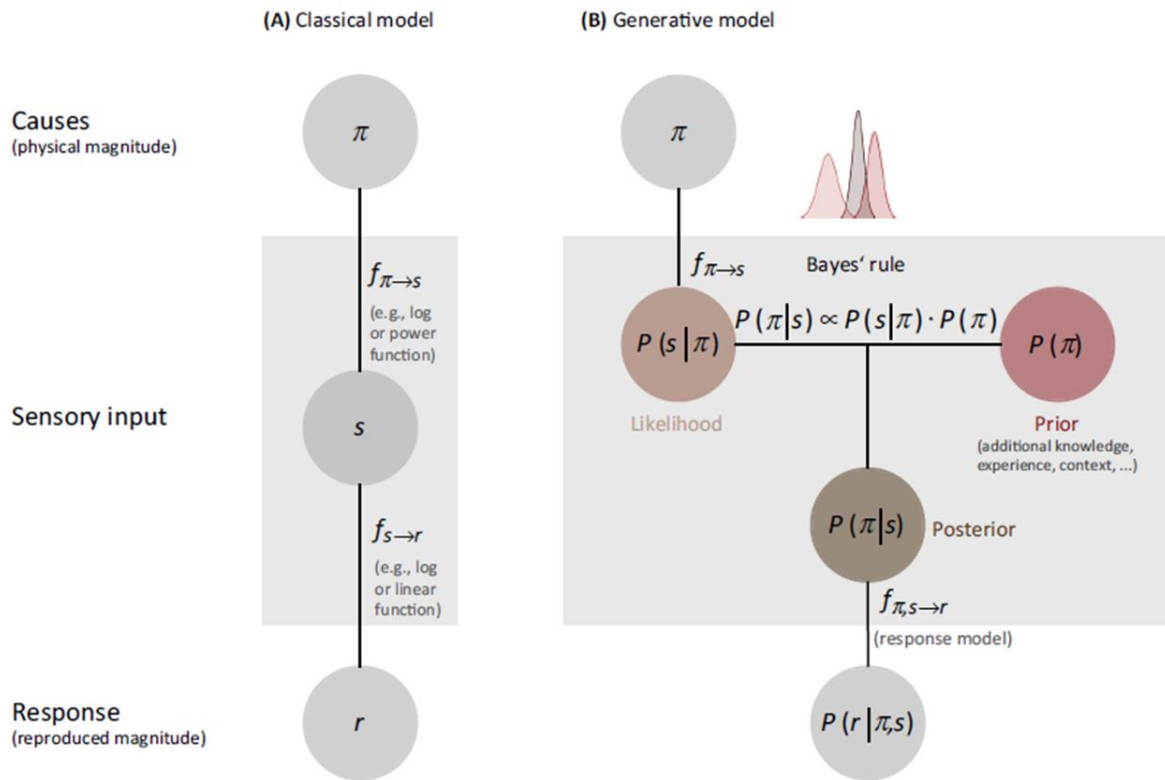


Figure 1.4: From Petzschner et al., 2015. **(A)** Classical model of magnitude estimation. The sensory input is translated into a motor response (reproduction task for example). **(B)** Generative model in which the response provided by the participants takes into account the noisy sensory input (likelihood) and his previous experience / knowledge about the task (prior).

Bayesian approaches may provide interesting perspectives on magnitudes estimations, because Bayesian models can be applied to explain classical effects that are usually reported in psychophysics experiments. For example, one of these effects is the regression effect (see section 1.1.3 of this chapter), which causes an under- (over-) estimation of large (small) magnitudes, in the tested range. In a temporal reproduction task for example, if participants have to reproduce duration intervals ranging from 500ms to 1000ms, after several trials, the prior (*a priori* information) will be located around the center of the distribution (750ms). In a generative model, if the duration of the sensory input (likelihood) is 500ms and if the prior is around 750ms, participants will be biased towards the center of the distribution and will overestimate the short temporal magnitude. On the other hand, if the duration is 1000ms and the prior is 750ms, participants will underestimate the long temporal magnitude (Figure 1.5).

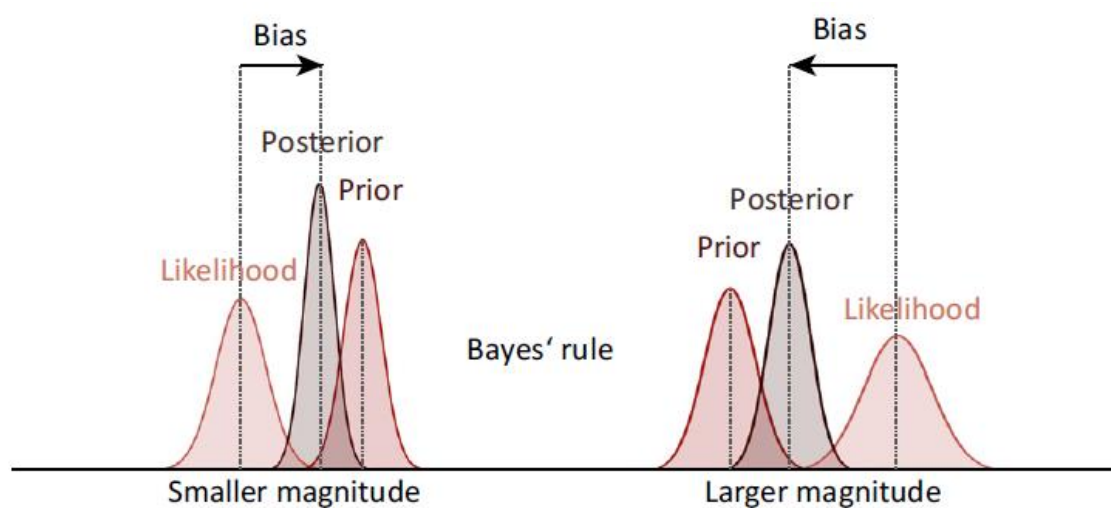


Figure 1.5: From Petzschner et al., 2015. Illustration of the regression effect in a Bayesian framework. Considering that the prior is close to the center of the distribution (tested range), the posterior will be biased towards the prior, leading to an overestimation of small magnitude (left panel) and to an underestimation of large magnitudes (right panel).

As mentioned in the previous sections, ATOM predicts that time, space and number should share a common metric. Such predictions can be formalized in Bayesian terms so that the magnitude of each dimension yields a likelihood estimate subsequently informed by an amodal prior common to all magnitude dimensions. In other words, if time, space and number share the same metric, these dimensions should share the same prior in a Bayesian framework. On the other hand, if the processing of temporal, spatial and numerical magnitudes is independent across dimension, each tested dimension should have its own prior (i.e. one prior per dimension).

1.5 Aim of this thesis

As we have seen, magnitude estimation exhibits a number of interesting features. In this thesis, the focus is on the possible existence of a generalized magnitude system (ATOM) with a common metric. To challenge this specific point, we designed a study in which participants were instructed to either perform temporal, spatial and numerical judgments while we independently manipulated the magnitude of the non-target dimensions. Such design allowed us to test for all possible interactions between magnitudes and investigate the bi-directionality of the interactions, which is one prediction from ATOM. In this experiment, we ensured that task difficulty was the same across dimensions and we also decided to use a dynamic design in which temporal, numerical and spatial information accumulated over the time course of each trial. In the first part of the thesis we also investigated to which extent Bayesian frameworks can be applied to magnitude estimations and specifically investigated the possible existence of an amodal prior for the different tested dimensions (see section 1.4). The experimental design and the results of this first experiment are presented in the second chapter of this manuscript. The second part of this thesis focused on several predictions made by ATOM and further investigated recent findings in the field of magnitude estimation, especially on the number-time interaction. In chapter 3 and 4, we investigated the possible existence of a scaling effect in time estimation. Whereas perceived duration seemed to depend on the relative magnitude of the tested digits (small or large) (Chang et al., 2011; Rammsayer and Verner, 2016), recent findings

suggested (Rammsayer and Verner, 2016) that the interferences between numerical and temporal magnitudes may be partly governed by attentional resources. Such claim is interesting and challenges the automaticity in magnitude processing initially suggested by several authors (Xuan et al., 2007; Dehaene and Akhavein, 1995). Moreover, we argue that the effect of numerical magnitude on perceived duration may be task instructions dependent and that different mechanisms may be involved in the number-time interference, depending on the numerical format that is used (symbolic or non-symbolic form). The implication and the integration of our results in the field of magnitude estimations are discussed in the last part of this manuscript (Chapter 5).

2 | A Bayesian Perspective on Accumulation in the Magnitude System

2.1 Summary

In the present article, we investigated the possible existence of a generalized magnitude system in which time, space and number were predicted to interfere. To test for all possible interactions across magnitude dimensions, we used a design in which participants were presented with clouds of dynamic dots and were either instructed to judge the duration, the total number of dots or the cumulative surface filled by the dots, while we independently manipulated the magnitude of the two non-target dimensions. We found that duration estimates were resilient to numerical and spatial manipulations whereas increasing (decreasing) the duration of the trials induced under- (over-) estimations of numerosity and surface. Results also revealed that manipulations of the rate of accumulation of sensory evidence affected numerical and spatial estimations, but did not interfere with perceived duration. From a Bayesian perspective, a generalized magnitude system predicts that time, space and number should share one common prior. Our results suggest otherwise, and a magnitude system based on Bayesian computations may necessitate multiple priors instead of one unique amodal prior.


2.2 Reference

This work was carried out under the supervision of Virginie van Wassenhove, and in collaboration with Martin Wiener. The paper was published in *Scientific Reports* under the following reference: **Martin, B., Wiener, M. & van Wassenhove, V.** A Bayesian Perspective on Accumulation in the Magnitude System. *Scientific Reports* 7, 630 (2017).

SCIENTIFIC REPORTS

OPEN

A Bayesian Perspective on Accumulation in the Magnitude System

Benoît Martin¹, Martin Wiener² & Virginie van Wassenhove¹ ¹

Received: 17 January 2017

Accepted: 8 March 2017

Published online: 04 April 2017

Several theoretical and empirical work posit the existence of a common magnitude system in the brain. Such a proposal implies that manipulating stimuli in one magnitude dimension (e.g. duration in time) should interfere with the subjective estimation of another magnitude dimension (e.g. size in space). Here, we asked whether a generalized Bayesian magnitude estimation system would sample sensory evidence using a common, amodal prior. Two psychophysical experiments separately tested participants on their perception of duration, surface, and numerosity when the non-target magnitude dimensions and the rate of sensory evidence accumulation were manipulated. First, we found that duration estimation was resilient to changes in surface and numerosity, whereas lengthening (shortening) the duration yielded under- (over-) estimations of surface and numerosity. Second, the perception of surface and numerosity were affected by changes in the rate of sensory evidence accumulation, whereas duration was not. Our results suggest that a generalized magnitude system based on Bayesian computations would minimally necessitate multiple priors.

The representation of space, time, and number is foundational to the computational brain^{1–3}, yet whether magnitudes share a common (conceptual or symbolic) format in the brain is unclear. Walsh's A Theory Of Magnitude (ATOM)³ proposes that analog quantities are mapped in a generalized magnitude system which entails that space, time, and number may share a common neural code. One additional implication for the hypothesis of a common representational system for magnitudes is that the estimation of magnitude in a target dimension (e.g., size in space) should be affected by the manipulation of the magnitude in another non-target dimension (e.g., duration in time), such that the larger the magnitude of the non-target, the larger one should perceive the target magnitude dimension to be (Fig. 1A). Such predictions can be formalized in Bayesian terms⁴ so that the magnitude of each dimension yields a likelihood estimate subsequently informed by an amodal prior common to all magnitude dimensions (Fig. 1B). In line with ATOM and the common magnitude system hypothesis, a growing body of behavioral evidence^{5–29}, for review see refs 30–32 suggests the existence of interferences across magnitude dimensions. Several neuroimaging studies also suggest the possibility of a common neural code for quantity estimations mostly implicating parietal cortices^{33–39} but see ref. 28, 40. However, while a variety of interactions between time, space and number has been reported, the directionality of these interactions is not always consistent in the literature^{13, 14} suggesting the need to moderate the claim for a common magnitude system: for instance, manipulating the duration of events has seldom been reported to affect numerical and spatial magnitudes^{12, 13, 26, 29} whereas numerosity⁶ and size^{5, 7} typically influence duration. Yet, using a literal interpretation of ATOM³, if time, number and space shared a common representational system and amodal prior, all magnitude dimensions should interact with each other in a bi-directional manner (Fig. 1A).

Recent discussions in the field suggest that the combination and evaluation of quantities in a common representational system would be realized on the basis of Bayesian computations^{13, 41}. Convergent with this proposal, recent examinations of Bayesian processing in magnitude estimation have demonstrated a number of distinct effects⁴. One primary example is the so-called central tendency effect, wherein magnitude estimates regress to the mean of the stimulus set, such that large (small) magnitudes are under (over) estimated. Crucially, central tendency effects have been demonstrated across a number of different magnitude judgments, including duration (historically known as Vierordt's law^{41, 42}, numerosity²⁷, distance and angle⁴³. Further, correlations between

¹CEA, DRF/I2BM, NeuroSpin, INSERM, U992, Cognitive Neuroimaging Unit, Université Paris-Sud, Université Paris-Saclay, F-Gif/Yvette, France. ²Department of Psychology, George Mason University, Fairfax, VA, USA. Correspondence and requests for materials should be addressed to V.v.W. (email: virginie.van.wassenhove@gmail.com)

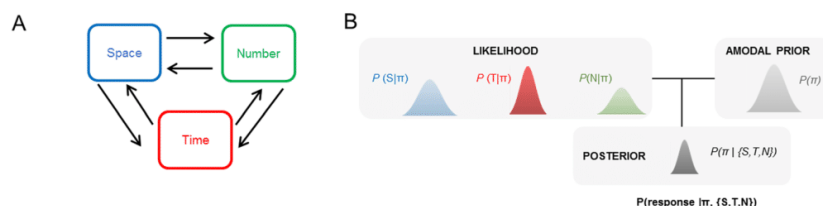


Figure 1. A Bayesian Magnitude System. **(A)** Proposals for a common magnitude system in the brain suggest that estimating one magnitude dimension (e.g. space) should be affected by the manipulation of any other magnitude dimensions (e.g. time) (Walsh, 2003). Such interactions should show bidirectional interferences so that, in our example, manipulating the spatial dimension of an event should affect the estimation of its duration in a comparable manner as manipulating its time dimension would affect the estimation of its spatial dimension. **(B)** To account for bidirectional interferences across magnitudes in a Bayesian framework for magnitude estimation, a common global or amodal prior can be posited to constrain the estimations of all magnitude dimensions⁴. Under such model, increasing (decreasing) the value of one magnitude should increase (decrease) the estimation of another magnitude. This panel is only for illustration purposes and do not reflect the actual distributions of the magnitudes used in the study.

the degree of central tendency have been found between the magnitude of different dimensions⁴³, suggesting the existence of “global” priors for common magnitude estimation⁴. The notion of global priors is compatible with a literal read of Walsh’s ATOM model by suggesting that, though differences in the initial processing of different magnitude dimensions may exist, the representation of these magnitudes is amodally stored. Further, the existence of global priors would provide an explanation for congruency effects between different magnitude dimensions. However, a single global prior for magnitude estimations would not explain why congruency effects may be inconsistent between dimensions or why the directionality of interferences may differ across magnitude dimensions.

To address this first working hypothesis, we used a paradigm in which stimuli consisted of clouds of dynamic dots characterized by the total duration of the trial (D), the total number of dots (N) and the overall surface filled by the dots (S) cumulating over the trial. The duration, the number and the cumulative surface of the dots will be thereafter used in reference to the magnitude of the time, number and space dimensions. Two experiments were conducted (Fig. 2A,B). In a first experiment (Experiment 1), while participants estimated the magnitude of a target dimension (e.g. D), we independently manipulated the magnitude of non-target dimensions (e.g. S and N). This design allowed us to test all possible combinations and investigate possible interactions between magnitudes (Table 1). If the magnitude of different dimensions interact, increasing or decreasing the N or the S should lead to an under/overestimation of D (Fig. 2D). These effects should be bidirectional so that when participants estimate N or S, increasing or decreasing the magnitude of non-target dimensions D should lead to under- or overestimation of the target magnitude dimension N or S. On the other hand, if dimensions are independent, manipulating the number of events in a given trial should not affect duration or surface estimates. Similarly, decreasing or increasing the duration should not affect numerical or spatial judgments if magnitudes are independent.

In a second working hypothesis, we manipulated the accumulation regime of sensory evidence for the estimation of N and S (Fig. 2B). The accumulation of sensory evidence in time for space and number has seldom been controlled for or manipulated during magnitude estimations. In a prior experiment⁴³, constraining the duration of sensory evidence accumulation in the S and N dimensions, the estimation of duration remained resilient to changes in the other dimensions, whereas D affected the estimation of S and N: curiously, the longer (shorter) durations decreased (increased) the estimation of S and N. These results were discussed in the context of a possible Bayesian integration of magnitudes across dimensions. Similarly, here, using a dynamic paradigm in which N and S accumulate over time raises the question of the implications of varying the speed or rate of sensory evidence delivery: for a given N or S, if D increases, the speed of presentation decreases, and vice versa. Hence, while the number of dots and the cumulative surface accumulated linearly in time in Experiment 1 (Fig. 1B), in Experiment 2, we investigated whether changes in the rate of presentation of visual information affected the estimation of D, N, and S. Two evidence accumulation regimes were tested: a fast-slow (FastSlow) and a slow-fast distribution (SlowFast), see Stimuli part in Material & Methods section.

In a third question, we wished to investigate the extent to which Bayesian models could explain the behavioral results obtained in magnitude estimation, independent of the means by which participants provided their estimates. Thus far, studies demonstrating central tendency effects^{42–44} have all relied on continuous estimation procedures, wherein participants estimated a particular magnitude value with a motor response. In particular, these studies utilized reproduction tasks, which required participants to demarcate where (when) a particular magnitude matched a previously presented standard. In contrast, the majority of studies demonstrating congruency effects in magnitude estimation have all employed two-alternative forced choice (2AFC) designs. This difference may be particularly relevant as recent studies have demonstrated that the size-time congruency effect, one of the most heavily studied and replicated, depends on the type of decision being made⁴⁵ (but see ref. 25, 46 for congruency effects with temporal reproduction). As such, for both Experiment 1 and 2, we provide systematic quantifications of the magnitude estimates as categorical estimations together with analysis of continuous reports.

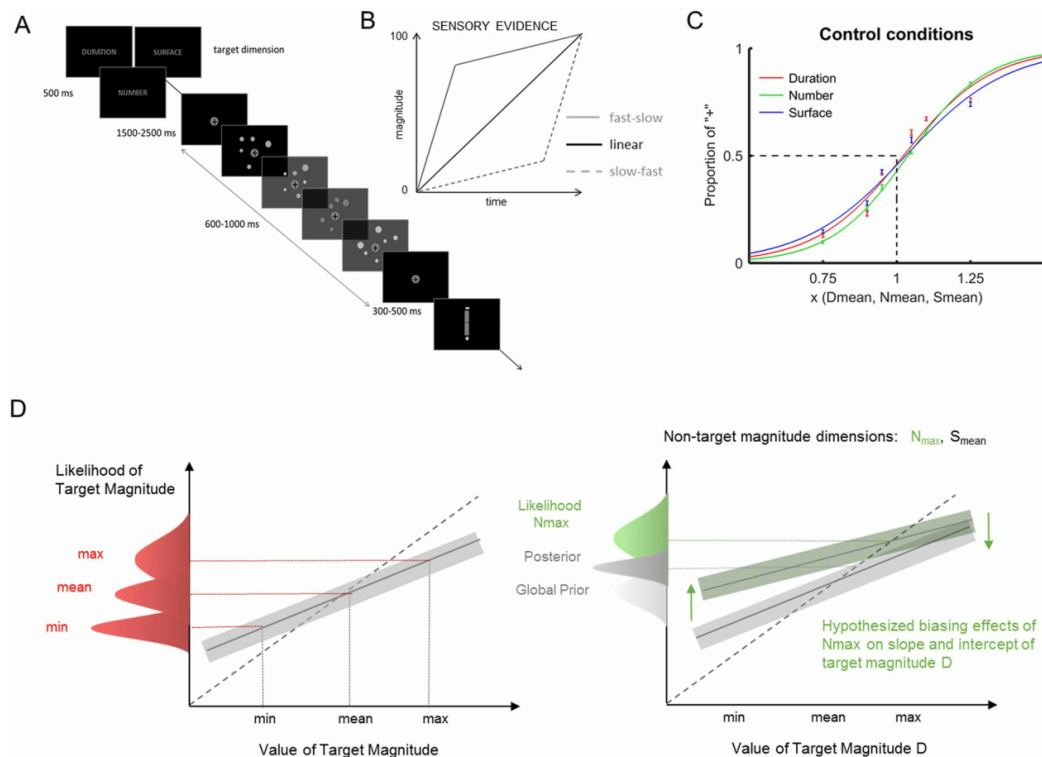


Figure 2. Experimental Design. (A) On a given trial, participants were presented with a word (“durée”, “nombre” or “surface”) indicating the dimension to estimate. In Experiment 1, one magnitude dimension could vary $\pm 25\%$, 10% , and 5% of its mean value, while the second one was set to its minimal or maximal value, and the third one to its mean value (Table 1). At the end of the stimulus presentation, participants used a vertical scale to estimate the target magnitude. (B) Three distributions were used for evidence accumulation: while D linearly accumulates over time (black trace), the rate of dot presentation could be manipulated to control N and S. Experiment 1 tested a linear distribution (filled black trace); Experiment 2 tested two distributions: a fast-slow (filled grey trace) and a slow-fast (dotted grey trace) distribution. The different stimulus distributions can be experienced with the videos Linear, FastSlow and SlowFast in Supp. Material. (C) Equated task difficulty across magnitudes. For illustration purposes, the psychometric curve captures the grand average performance obtained for the estimation of Duration, Number and Surface when all non-target dimensions were kept at their mean value. The task difficulty was equated across magnitude dimensions so that no differences in discriminability (PSE and WR) existed between the tested dimensions. Bars are 2 s.e.m. (D) Predictions for the effect of non-target manipulations on the estimation of the target magnitude dimension. Left panel: varying the target magnitude while keeping the non-target dimensions to their mean values provided the control central tendency and intercept. In a common Bayesian magnitude estimation system [4], comparable tendency and intercept should be predicted pending controlled matching between magnitudes and task difficulty (panel C). Right panel: estimation of D while N is set to its maximal value (in green, N_{max}). Maximal value in non-target magnitude may affect the central tendency and the intercept if an amodal global prior common to D and N is implicated in the estimation of duration. In this example, N_{max} would bias the lowest (highest) duration values towards larger (smaller) values and lead the intercept to move upwards.

Materials and Methods

Participants. A total of 45 participants were tested. 3 participants did not come to the second session and 10 were disregarded for poor performance: 10 participants either never reach 50% of “+” responses for the largest values, or the p-values associated with the goodness-of-fits in the control conditions were >0.05 (see Procedure). Hence, a total of 17 healthy volunteers (7 males, 10 females, mean age 24.9 ± 5.8 y.o.) took part in Experiment 1, and 15 participants (8 males, 7 females, mean age 26.5 ± 7 y.o.) in Experiment 2. All had normal or corrected-to-normal vision. Both experiments took place in two sessions one week apart. Prior to the experiment, participants gave a written informed consent. The study was conducted in agreement with the Declaration of Helsinki (2008) and was approved by the Ethics Committee on Human Research at Neurospin (Gif-sur-Yvette, France). Participants were compensated for their participation.

	DURATION [0.75; 0.9; 0.95; 1.05; 1.1; 1.25] \times D_{mean}	SURFACE [0.75; 0.9; 0.95; 1.05; 1.1; 1.25] \times S_{mean}	NUMEROSITY [0.75; 0.9; 0.95; 1.05; 1.1; 1.25] \times N_{mean}
CONTROL	$S_{\text{mean}}, N_{\text{mean}}$	$D_{\text{mean}}, N_{\text{mean}}$	$S_{\text{mean}}, D_{\text{mean}}$
MIN	$S_{\text{min}}, N_{\text{mean}}$	$D_{\text{min}}, N_{\text{mean}}$	$S_{\text{min}}, D_{\text{mean}}$
	$S_{\text{mean}}, N_{\text{min}}$	$D_{\text{mean}}, N_{\text{min}}$	$S_{\text{mean}}, D_{\text{min}}$
MAX	$S_{\text{max}}, N_{\text{mean}}$	$D_{\text{max}}, N_{\text{mean}}$	$S_{\text{max}}, D_{\text{mean}}$
	$S_{\text{mean}}, N_{\text{max}}$	$D_{\text{mean}}, N_{\text{max}}$	$S_{\text{mean}}, D_{\text{max}}$

Table 1. Full design of Experiment 1 testing the interactions across magnitudes with a linear sensory evidence accumulation regime. The tested magnitude dimension could take 6 possible values corresponding to 75, 90, 95, 105, 110 and 125% of its mean value. In CONTROL trials (first row) of the duration condition (D, second column), participants estimated the duration of the trial when D varied between the 6 possible values while the surface and number dimensions were kept to their mean values (S_{mean} and N_{mean} , respectively). In MIN trials (second row), participants estimated the magnitude of a given dimension (e.g. D) varying between the 6 possible values while one of the two non-target magnitude dimensions was kept at its mean value (e.g. N_{mean}) and the other was set to its minimal value (e.g. S_{min}). In MAX trials (third row), participants estimated the magnitude of a given dimension varying between the 6 possible values while one of the two non-target dimensions was kept at its mean value and the other was set to its maximal value. In total, 72 trials per experimental condition were collected (i.e. 12 trials per tested magnitude value in all possible combinations). D: duration; S: surface; N: number; min = minimal; max = maximal.

Stimuli. The experiment was coded using Matlab 8.4 with Psychtoolbox (v 3.0.12) and built on a published experimental design (see ref. 13). Visual stimuli were clouds of grey dots which appeared dynamically on a black computer screen (1024×768 pixels, 85 Hz refresh rate). Dots appeared within a virtual disk of diameter 12.3 to 15.2 degrees of visual angle; no dots could appear around the central fixation protected by an invisible inner disk of 3.3 degrees. Two dots could not overlap. The duration of each dot varied between 35 ms and 294 ms, and their diameter between 0.35 and 1.14 degrees. A cloud of dots was characterized by its duration (D: total duration of the trial during which dots were presented), its numerosity (N: cumulative number of dots presented on the screen in the entire trial) and its surface (S: cumulative surface covered by the dots during the entire trial). On any given trial, D, N and S could each take 6 possible values corresponding to 75, 90, 95, 105, 110 and 125% of the mean value. We fixed D to 800 ms ($D_{\text{mean}} = 800$ ms) and initially picked $N_{\text{mean}} = 30$ dots and $S_{\text{mean}} = 432 \text{ mm}^2$. The initial values of N_{mean} and S_{mean} were then individually calibrated in the calibration session of the experiment (see Procedure). To ensure that luminance could not be used as a cue to perform the task, the relative luminance of dots varied randomly across all durations among 57, 64, 73, 85, 102 and 128 in the RGB-code.

In Experiment 1, the total number of dots accumulated linearly over time (see Linear video), 2 to 7 dots at a time in steps of 9 to 13 iterations (Fig. 2A). In Experiment 2, the total number of dots accumulated in a fast-to-slow or in a slow-to-fast progression: in FastSlow, 75% \pm 10% of the total number of dots in the trial were presented in the first 25% of the duration of the trial, whereas in SlowFast, 25 \pm 10% of the total number of dots was shown during 75% of the total duration of the trial (Fig. 2B; see FastSlow and SlowFast videos).

Procedure. Participants were seated in a quiet room ~60 cm away from the computer screen with their head maintained on a chinrest. The main task consisted in estimating the magnitude of the trial along one of its three possible dimensions (D, N, or S). Each experiment consisted of two sessions: in the first or calibration session, stimuli were calibrated to elicit an identical discrimination threshold in all three dimensions on a per individual basis (see ref. 13) and the main objective of the first session was to calibrate an individual's N_{mean} and S_{mean} with the chosen D_{mean} in order to match task difficulty across dimensions. D_{mean} was kept constant for all participants. The second experimental session consisted in the experiment proper.

In the calibration session of Experiment 1 and 2, the task difficulty across magnitudes was individually calibrated by computing the participant's Point of Subjective Equality (PSE: 50% discrimination threshold) and the Weber Ratio (WR) for each dimension D, N, and S. The PSE traditionally provides an estimate of an individual's perceptual threshold: here, the PSE specifically corresponded to the magnitude value in the target dimension at which participants responded at chance level. The WR provided the estimate of the steepness of the fitted psychometric curve, and thus of an individual's perceptual sensitivity in discriminating magnitudes of the target dimension. A smaller (larger) WR indicates a steeper (flatter) curve and a better (worse) sensitivity. Participants were passively presented with exemplars of the minimum and maximum value for each dimension and were then required to classify 10 of these extremes as minimum '−' or maximum '+' by pressing 'h' or 'j' on an AZERTY keyboard, respectively. Participants then received feedback indicating the actual number of good answers they provided. Subsequently, the PSE and the WR were independently assessed for each magnitude dimension by varying the magnitude in one dimension and keeping the magnitude in the other two dimensions at their mean values (e.g. if D varied among its 6 possible values, S was S_{mean} and N was N_{mean}). 5 trials per magnitude value of the target dimension were collected yielding a total of 30 trials (5 trials \times 6 values) per dimension from which the individual's PSE and WR could be computed and compared. This process (~15 min) was iterated until the individual's PSE for the target dimension was stable and the WR similar across dimensions. The PSE were considered as matching across dimensions when all of them were between 95 and 105% of the mean magnitude value. For each participant, the mean of the three WR was also calculated, and the WR were considered as matching when: (1)

	DURATION [0.75; 0.9; 0.95; 1.05; 1.1; 1.25] \times D_{mean}	SURFACE [0.75; 0.9; 0.95; 1.05; 1.1; 1.25] \times S_{mean}	NUMEROSITY [0.75; 0.9; 0.95; 1.05; 1.1; 1.25] \times N_{mean}
CONTROL	$S_{\text{mean}}, N_{\text{mean}}$	$D_{\text{mean}}, N_{\text{mean}}$	$S_{\text{mean}}, D_{\text{mean}}$
MIN		$D_{\text{min}}, N_{\text{mean}}$	$S_{\text{mean}}, D_{\text{min}}$
MAX		$D_{\text{max}}, N_{\text{mean}}$	$S_{\text{mean}}, D_{\text{max}}$

Table 2. Full design of Experiment 2 testing the effects of duration and FastSlow and SlowFast sensory evidence accumulation regime on magnitude estimation. In CONTROL trials, as in Experiment 1, participants estimated the value of each target magnitude dimension when non-target dimensions were kept to their mean values. In MIN and MAX trials, participants estimated S or N when D was the shortest (D_{min}) or the longest (D_{max}) and the other dimension was kept to its mean value (N_{mean} and S_{mean} , respectively). Importantly in Experiment 2, the clouds of dots could accumulate over time according to two accumulation regimes (FS:FastSlow, and SF:SlowFast). A total of 144 trials per experimental condition was collected (i.e. 12 trials for a given magnitude value in a specific condition \times 2 distributions).

the three WR were in the mean ($SD = 2$), (2) when the ratio between the larger and the smaller WR was < 3 and (3) each WR value was smaller than 0.25. The final mean magnitude values for each dimension were $N_{\text{mean}} = 32$ ($SD = 3$ dots) and $S_{\text{mean}} = 476$ ($SD = 58 \text{ mm}^2$) for Experiment 1 and $N_{\text{mean}} = 32$ ($SD = 2$ dots) and $S_{\text{mean}} = 490$ ($SD = 41 \text{ mm}^2$) in Experiment 2.

In the experimental session of both Experiment 1 and 2, participants were first tested again with 30 trials calibrated to their individual N_{mean} and S_{mean} calculated in the calibration session to ensure that their PSE and WR remained identical so that task difficulty was balanced across magnitude dimension. Only two participants in Experiment 1 required the recalibration procedure to be performed again. Subsequently, participants performed the magnitude estimation task *proper* in which participants were asked to provide a continuous magnitude estimation of the target dimension by moving a cursor on a vertical axis whose extremes were the minimal and maximal magnitude values of the target dimension. In a given trial, participants were provided with the written word 'Durée' (Duration), 'Nombre' (Number) or 'Surface' (Surface) which indicated which target magnitude dimension they had to estimate (Fig. 2A). At the end of a trial, the vertical axis appeared on the screen with the relative position of '+' and '-' pseudo-randomly assigned to the extreme bottom or the top of the axis. The cursor was always initially set in the middle position on the axis. Participants used the mouse to vertically move the slider along the axis and made a click to validate their response. They were asked to emphasize accuracy over speed. Trials were pseudo-randomized across dimensions and conditions.

In Experiment 1, five experimental conditions were tested per dimension: in the control condition, the two non-target dimensions were kept to their mean magnitude values; in the four remaining conditions, one of the other non-target dimension was set to its minimal or maximal value, while the other was kept to its mean value (Table 1). A total of 1080 trials were tested in Experiment 1 (3 dimensions \times 5 conditions \times 6 magnitude values \times 12 repetitions).

In Experiment 2, two main sensory accumulation regimes were tested (FastSlow, and SlowFast) and the emphasis was on the effect of D on S and N. The main control condition consisted in assessing the estimation of duration with S_{mean} and N_{mean} , and in testing whether the rate of sensory evidence delivery affected the estimation of duration. Two control conditions investigated the effect of the rate of stimulus presentation on the estimation of S and N without varying non-target dimensions. In light of the results obtained in Experiment 1, Experiment 2 did not investigate the interactions of N or S on D, nor the interactions between N and S. Ten experimental blocks alternated between FastSlow and SlowFast presentations counterbalanced across participants. 12 repetitions of each possible combination were tested yielding a total of 144 trials for D (2 distributions \times 6 durations \times 12 repetitions), 432 trials for N (2 distributions \times 3 conditions \times 6 numerosities \times 12 repetitions) and 432 trials for S (2 distributions \times 3 conditions \times 6 surfaces \times 12 repetitions) for a grand total of 1008 trials in Experiment 2 (Table 2).

Statistical Analyses. To analyze the point-of-subjective equality (PSE) and the Weber Ratio (WR), participants' continuous estimates were first transformed into categorical values: a click between the center of the axis and the extreme demarcating '+' ('-') was considered as a '+' ('-') response. Proportions of '+' were computed on a per individual basis and separately for each target dimension and each experimental condition. Proportions of '+' responses were fitted using the logit function (Matlab 8.4) on a per individual basis. Goodness-of-fits were individually assessed and participants for whom the associated p-values in the control conditions were > 0.05 were excluded from the analysis. On a per condition basis, PSE and WR that were 2 standard deviations away from the mean were disregarded and replaced by the mean of the group. This procedure affected a maximum of 2 values per condition across all individuals. Statistics were run using R (Version 3.2.2). PSE and WR were defined as:

$$\text{PSE} = f^{-1}(0.5) \quad (1)$$

$$\text{WR} = 0.5 \times (f^{-1}(0.75) - f^{-1}(0.25)) \times \frac{1}{\text{PSE}} \quad (2)$$

where f is the logit function used to fit individuals' responses, and PSE is the magnitude value of the target dimension when the proportion of '+' responses is equal to 0.5. Using the inverse of the f function, the WR was calculated as the mean difference between the just-noticeable-differences (aka magnitude values at 25% and 75% performance) normalized by the individual's PSE. Additionally, to specifically address central tendency effects, continuous estimates were analyzed. For each magnitude dimension, continuous estimates were expressed as the relative position on the slider that participants selected on each given trial, with higher percentages indicating closer proximity to '+'. To measure the central tendency effects, continuous estimates were plotted against the corresponding magnitude dimension for each condition, also expressed as a percentage – where 0 indicated the smallest magnitude and 100 indicated the largest – and fits with a linear regression, and the slope and y-intercept of the best fitting line were extracted^{47, 48}. Slope values closer to 1 indicated veridical responding (participants responded with perfect accuracy), whereas values closer to 0 indicated a complete regression to the mean (participants provided the same estimate for every magnitude). In contrast, intercept values of these regressions could indicate an overall relative bias for over- or under-estimation so that higher (lower) intercept values would indicate that participants overestimated (underestimated) the magnitude in the target dimension⁴⁹. To compare central tendency effects between magnitude dimensions, correlation matrices between slope values for D, N, and S were constructed. Bonferroni corrections were applied to control for multiple comparisons.

Results

To examine Bayesian effects in the magnitude system, we evaluated both categorical and continuous judgments in two magnitude estimation experiments using variations of the same paradigm (Fig. 2A). To evaluate choice responses, continuous estimates were binned according to which end of the scale they were closer to. Previous work has demonstrated that bisection tasks and continuous estimations are compatible and provide similar estimates of duration^{44, 50}. Our intention was thus to first replicate the effects of Lambrechts and colleagues¹³ with a modified design, and second, to measure central tendency effects in our sample to examine whether these effects correlated between magnitude dimensions, which would suggest the existence of global priors (see ref. 4, Fig. 2D).

Control conditions: matching task difficulty across magnitude dimensions. Two independent repeated-measures ANOVAs with the PSE or WR as dependent variables using magnitude dimensions (3: D, N, S), control conditions (3: Linear (Experiment 1), SlowFast and FastSlow (Experiment 2) distributions as within-subject factors did not reveal any significant differences (all $p > 0.05$). This suggested that participants' ability to discriminate the different values presented in the tested magnitudes was well matched across magnitude dimensions (Fig. 2C).

Experiment 1: Duration affects the estimation of Number and Surface. We first analyzed the data of Experiment 1 as categorical choices. Figure 3A illustrates the grand average estimations of duration, numerosity, and surface for all experimental manipulations (colored traces) along with changes of PSE (insets).

Separate 2×2 repeated-measures ANOVAs were run on the PSE obtained during the magnitude evaluation of each target dimension using the non-target dimensions (2) and their magnitude values (2: min, max) as within-subject factors. No main effects of non-target dimension ($F[1,16] = 0.078, p = 0.780$), magnitude value of the non-target dimension ($F[1,16] = 0.025, p = 0.875$) or their interaction ($F[1,16] = 0.003, p = 0.957$) were found on duration (D) indicating that manipulating N or S did not change participants' estimation of duration (Fig. 3A, left panel). In the estimation of N, main effects of non-target dimensions ($F[1,16] = 7.931, p = 0.0124$), their magnitudes ($F[1,16] = 25.53, p = 0.000118$) and their interactions ($F[1,16] = 23.38, p = 0.000183$) were found. Specifically, when the magnitude of the non-target dimensions were at their minimal value, the PSE obtained in the estimation of N was lower than when the magnitude of the non-target dimensions were at their maximal value. Additionally, in the estimation of N, D_{\min} lowered the PSE more than S_{\min} , and D_{\max} raised the PSE more than S_{\max} . Paired t-tests were run contrasting the PSE obtained in the estimation of N during the control ($D_{\text{mean}}, S_{\text{mean}}$) and other experimental conditions: D_{\min} significantly increased [$\text{PSE}(D_{\min}) < \text{PSE}(D_{\text{mean}})$; $p = 4.1 \times 10^{-5}$] whereas D_{\max} significantly decreased [$\text{PSE}(D_{\max}) > \text{PSE}(D_{\text{mean}})$; $p = 0.0032$] the perceived number of dots (Fig. 3A, middle panel, inset). There were no significant effects of S on the estimation of N. Altogether, these results suggest that the main effect of changing the magnitude in the non-target dimension on numerosity estimation was driven by the duration of the stimuli.

In the estimation of S, we found no main effect of the non-target dimension ($F[1,16] = 1.571, p = 0.228$) but a significant main effect of the magnitude values in non-target dimensions ($F[1,16] = 22.63, p = 0.000215$). The interaction was on the edge of significance ($F[1,16] = 3.773, p = 0.0699$) suggesting that, as for N, only the magnitude of one non-target dimension may be the main driver of the significant results observed in the effect. Paired t-tests contrasting the PSE obtained in the estimation of S during the control ($D_{\text{mean}}, N_{\text{mean}}$) and other conditions showed that D_{\min} significantly increased [$\text{PSE}(D_{\min}) < \text{PSE}(D_{\text{mean}})$; $p = 8.7 \times 10^{-4}$], whereas D_{\max} significantly decreased [$\text{PSE}(D_{\max}) > \text{PSE}(D_{\text{mean}})$; $p = 0.035$] the perceived surface (Fig. 3A, right panel). No significant effects of N on S were found. As observed for the estimation of N, these results suggest that the main effect of non-target magnitude on the estimation of S was entirely driven by the time dimension.

Overall, the analysis of PSE indicated that participants significantly overestimated N and S when dots were presented over the shortest duration, and underestimated N and S when dots accumulated over the longest duration. Additionally, manipulating N or S did not significantly alter the estimation of duration. No significant interactions between N and S were found. To ensure that these results could not be accounted for by changes in participants' perceptual discriminability in the course of the experiment, repeated-measures ANOVA were conducted independently for each target dimension (3: D, N, S) with the WR as dependent variable and the experimental conditions (5) as main within-subject factors. No significant differences (all $p > 0.05$) were found

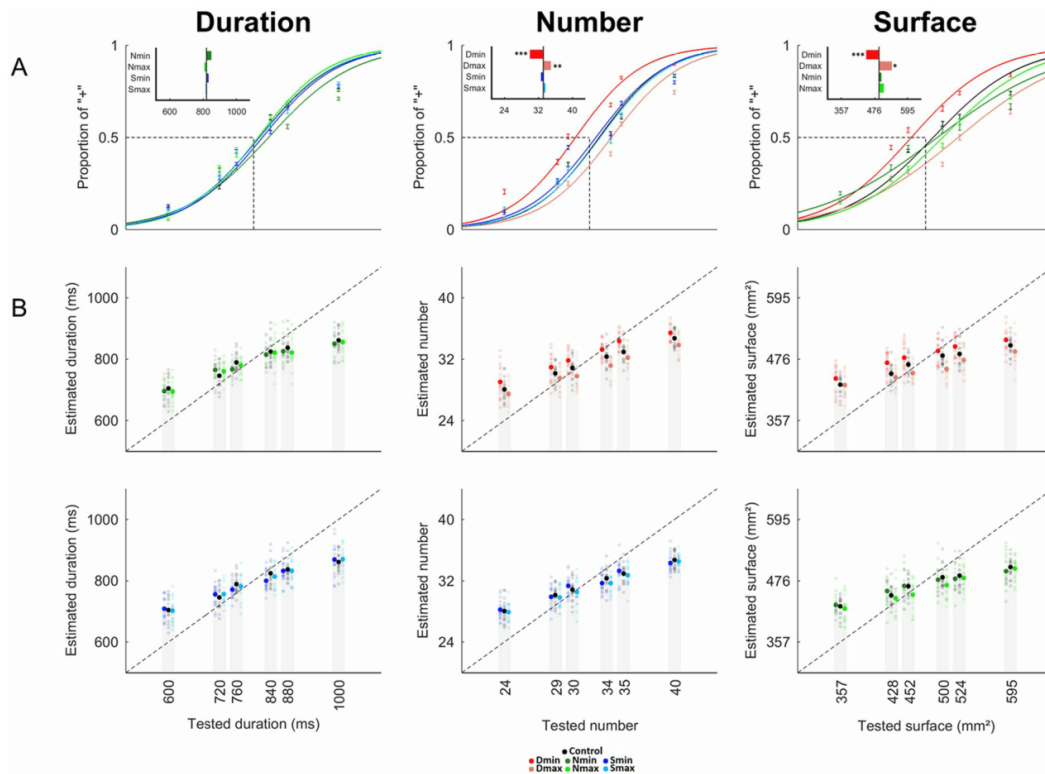


Figure 3. Duration affects the estimation of S and N (Experiment 1). **(A)** Categorical quantifications and PSE. The percentage of (+) responses as a function of the target magnitude dimension (D: left panel, N: middle panel, S: right panel) were fitted on a per individual basis. The inset in the top left of each figure depicts the shift in PSE for each experimental condition compared to the control condition represented by the black vertical line. No effects of N or S on D were found; no effects of S on N or of N on S were found; both N and S were significantly overestimated when presented during the shortest duration (Dmin, red) and underestimated when presented with the longest duration (Dmax, orange). **(B)** Continuous judgments. Individual performances (transparent dots) and mean performances (filled dots) for the estimation of D (left panels) with changing N (green) and S (blue); of N (middle panels) with changing D (red) and S (blue); and of S (left panels) with changing D (red) and N (green). The continuous scale was mapped from 0 to 100. The dotted line is the ideal observer's performance. All experimental conditions showed a central tendency. No effects (intercept or central tendency) of N or S on D were found (left top and bottom graphs, respectively); no effects of S on N (middle bottom graph) or of N on S (left bottom graph) were found. Significant main effects of D were found on the central tendency and the intercept of N (middle top) and S (left top). N_{min}: minimal numerosity value; N_{max}: maximal numerosity value; S_{min}: minimal surface value; S_{max}: maximal surface value; D_{min}: minimal duration value; D_{max}: maximal duration value. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; bars are 2 s.e.m.

suggesting that the WRs were stable over time, and that task difficulty remained well matched across dimensions in the course of the experiment.

For the analysis of continuous estimates (Fig. 3B), we first examined the effect of central tendency for each target dimension, collapsing the data across the non-target dimensions (Fig. 4A). A repeated measures ANOVA of slope values with magnitude as a within-subjects factor revealed a main effect of the magnitude of the target dimension [$F(2,32) = 13.284, p = 0.000063$]. Post-hoc paired t-tests identified this effect as driven by a lower slope, indicating a greater regression to the mean for S as compared to D [$t(16) = 3.495, p = 0.003$] and N [$t(16) = 5.773, p = 0.000029$], with no differences in slope values between D and N [$t(16) = 0.133, p = 0.896$] (Fig. 4B).

Further analyses revealed comparable findings as in the categorical analysis: separate 2×3 repeated measures ANOVAs were run for each magnitude dimension, with the non-target dimension and its magnitude value as within-subject factors. Analysis of slope values revealed no significant main effects or interactions for any of the tested magnitudes (all $p > 0.05$), indicating no change in central tendency, as a function of the non-target magnitudes. However, an analysis of intercept values demonstrated a significant main effect of the non-target dimension for S [$F(2,32) = 24.571, p < 0.00001$] and N [$F(2,32) = 39.901, p < 0.00001$], but not for D [$F(2,32) = 0.010, p = 0.99$]. Specifically, intercept values were shifted higher (lower) when D_{min} (D_{max}) was the non-target dimension

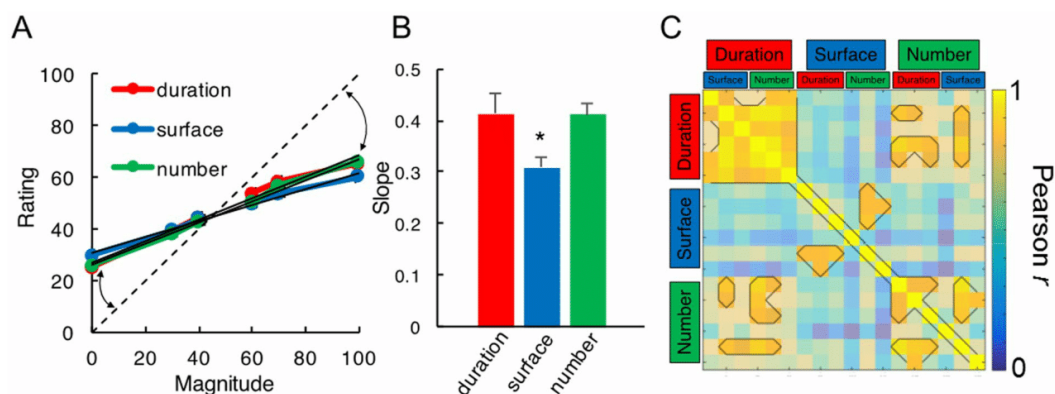


Figure 4. Central Tendency Effects in Magnitude Estimation (Experiment 1). (A) Average continuous estimates for all three magnitude dimensions as target, collapsed across all non-target magnitude dimensions. Magnitudes were normalized as a percentage of the maximum presented magnitude value, with zero representing the smallest and 100 representing the largest magnitude presented. Continuous estimates were similarly normalized as a percentage of the sliding scale, with zero representing a minimal estimate of “–” and 100 representing a maximal estimate of “+”. The dashed identity line indicates where estimates should lie for veridical performance. Deviations from this identity line (arrows) exhibited central tendency, wherein smaller and larger magnitudes were over- and under-estimated, respectively. (B) Slope values extracted from a best fitting linear regression in A quantify the degree of central tendency, with smaller values indicating greater regression to the mean. Similar slope values were observed for duration and number, but significantly lower estimates were found for surface estimates than either duration or number. (C) Correlation matrix of slope values for every target and non-target magnitude trial type. Each target magnitude dimension was tested in the presence of three possible values (min, mean, max) of a non-target dimension. Outlined pixels represent those Pearson correlation coefficients that survived a multiple comparison correction. Slope values across non-target dimensions were correlated within each target dimension (sections along the diagonal); crucially, slope values for duration and number were correlated with each other when each one was the target magnitude (lower left and upper right sections). Further, no correlation between surface and either duration or number as target dimensions were observed (middle top and bottom sections). *Indicates $p < 0.05$; bars are 2 s.e.m.

in both the S and N tasks (Fig. 3B, red hues), but not when the non-target dimension was N for S, S for N, or for either S or N when D was the target magnitude dimension.

To examine the central tendency effects across magnitude dimensions, we correlated the slope values between target magnitude dimensions (Fig. 4C). Collapsing across the non-target dimensions, we found that all three slope values significantly correlated with one another [D to S: Pearson $r = 0.594$; D to N: $r = 0.896$; S to N: $r = 0.662$]. Given that S exhibited a greater central tendency than D or N, we compared the Pearson correlation coefficients with Fisher’s z-test for the differences of correlations; this analysis revealed that the D to N correlation was significantly higher than the D to S correlation [$Z = 2.03$, $p = 0.04$], and marginally higher than the S to N correlation [$Z = 1.73$, $p = 0.083$], suggesting that D and N dimensions, which had similar slope values, were also more strongly correlated with each other than with S. To further explore this possibility, we conducted partial Pearson correlations of slope values; here, the only correlation to remain significant was D to N, when controlling for S [$r = 0.8352$], whereas D to S, controlling for N, and N to S, when controlling for D were no longer significant [$r = 0.0018$ and 0.3627 , respectively].

The results of the correlation analysis revealed that D and N tasks were highly correlated in slope, indicating that individual subjects exhibited a similar degree of central tendency for these two magnitude dimensions (Fig. 4A). From a Bayesian perspective, these results suggest that the priors for D and N may be more correlated than the priors for D and S, or for N and S; thus D, N and S do not share one single prior, but may rather rely on different priors which would be more or less correlated between each other. To explore this at a more granular level, we expanded our correlation analysis to include all non-target dimensions (Fig. 4C). The result of this analysis, with a conservative Bonferroni correction ($r > 0.8$) for multiple comparisons confirmed the above results, demonstrating that D and N dimensions were correlated across most non-target dimensions, but that D and N dimensions were weakly and not significantly correlated with S. This finding suggests that D and N estimation may rely on a shared prior, that is separate from S; however, a shared (D, N) prior would not explain why D estimates were unaffected by changes in N, nor would it explain why S estimates are affected by changes in D.

Lastly, we sought to compare the quantifications based on continuous data with those from the categorical analysis. Previous work has demonstrated that the WR on a temporal bisection task correlates with the central tendency effect from temporal reproduction⁴⁴. To confirm this, we measured the correlations between the slope values of continuous magnitude estimates with the WR from the categorical analysis. As predicted, we found a significant negative correlation between slope and WR for D ($r = -0.69$) and N ($r = -0.57$); however, the correlation for S failed to reach significance ($r = -0.41$, $p = 0.1$), indicating that greater central tendency (lower slope

values) were associated with increased variability (larger WR). This finding is notable, as the analysis of WR values did not reveal any difference between magnitude dimensions. This suggests that the slope of continuous estimate judgments may be a better measure of perceptual uncertainty than the coefficient of variation derived from categorical responses.

Experiment 2: Duration is robust to accumulation rate, not N and S. In Experiment 2, participants estimated D, N or S while the accumulation regime was manipulated as either FastSlow or SlowFast (Fig. 2B, Table 2). As previously, we systematically analyzed the categorical and the continuous reports. First, we tested the effect of the accumulation regime on the estimation of each magnitude dimension by using a 2×3 repeated-measures ANOVA with PSE measured in control conditions (Table 2, 1st row) as independent variable and distribution (2: FastSlow, SlowFast) and magnitude dimension (3: N, D, S) as within-subject factors. Marginal main effects of accumulation regime ($F[1,28] = 2.872, p = 0.0734$) and magnitude dimensions ($F[1,14] = 4.574, p = 0.0506$) were observed. Their interaction was significant ($F[2,28] = 10.54, p = 0.0004$). A post-hoc t-test revealed no significant effects of accumulation regime on the estimation of D ($p = 0.23$), but significant effects of accumulation regime in the estimation of N ($p = 0.016$) and S ($p = 0.0045$) (Fig. 5A).

Second, we tested the effect of D and accumulation regime on the estimation of N and S (Fig. 5A, top insets). We conducted a $2 \times 2 \times 2$ repeated-measures ANOVA with PSE as an independent variable and magnitude dimension (2: N, S), accumulation regime (2: FastSlow, SlowFast), and duration (2: D_{\min} , D_{\max}) as within-subject factors. Main effects of accumulation regime ($F[1,14] = 22.12, p = 0.000339$) and duration ($F[1,14] = 27.65, p = 0.000121$) were found, suggesting that both N and S were affected by the distribution of sensory evidence over time, and by the duration of the sensory evidence accumulation. No other main effects or interactions were significant although two interactions trended towards significance, namely the two-way interaction between accumulation regime and duration ($F[1,14] = 3.482, p = 0.0831$) and the three-way interaction between dimension, accumulation regime, and duration ($F[1,14] = 3.66, p = 0.0764$). These trends were likely driven by the SlowFast condition as can be seen in Fig. 5A.

For the analysis of continuous data (Fig. 5B), we first examined any overall differences in slope values for different accumulation regimes (FastSlow vs. SlowFast) across all three target magnitude dimensions (D, N, S). A (3×2) repeated measures ANOVA with the above as within-subjects factors revealed a main effect of magnitude dimension [$F(2,32) = 7.878, p = 0.002$], with S once again demonstrating the largest slope value, but no effect of accumulation regime or interaction (both $p > 0.05$), suggesting that the rate of accumulation did not influence the central tendency effect. However, on the basis of our *a priori* hypothesis, post-hoc tests revealed a significantly lower slope value for N in SlowFast compared to FastSlow [$t(17) = 3.067, p = 0.007$], suggesting that participants exhibited more central tendency for numerosity when the accumulation rate was slow in the first half of the trial (Fig. 5B). The analysis of intercept values did not reveal any effects of accumulation regime or magnitude dimension (all $p > 0.05$). However, on the basis of our *a priori* hypothesis, post-hoc tests demonstrated that S exhibited a significantly lower intercept for SlowFast compared to FastSlow [$t(17) = 3.609, p = 0.002$], with no changes for either D or N (both $p > 0.05$), indicating that participants underestimated surface when the rate of evidence accumulated slowly in the first half of the trial.

For S and N, further examination of slope values for the three possible durations using a $2 \times 2 \times 3$ repeated measures ANOVA with magnitude dimension (2: S, N), accumulation regime (2: FastSlow, SlowFast), and duration (3: D_{\min} , D_{mean} , D_{\max}) as within-subjects factors, revealed a significant main effect of magnitude dimension [$F(2,32) = 7.717, p = 0.013$] and of accumulation regime [$F(2,32) = 11.345, p = 0.004$], but not of duration [$F(2,32) = 1.403, p = 0.261$]. Using the same analysis for intercept values, we found no main effects of magnitude dimension [$F(2,32) = 1.296, p = 0.272$], but a significant effect of accumulation regime [$F(2,32) = 5.540, p = 0.032$] and of duration [$F(2,32) = 21.103, p = 0.000001$]. More specifically, we found that intercept values were lower for longer durations, indicating greater underestimation when the interval tested was longer. No other effects reached significance (all $p > 0.05$).

Overall, these findings indicate that duration estimations were immune to changes in the rate of accumulation of non-target magnitudes, similar to the findings of Experiment 1. Also similar, we found that estimates of S and N were affected by duration as non-target magnitude dimension, with longer durations associated with greater underestimation of S and N (Fig. 6). In addition, our results demonstrate a difference between accumulation regimes for S and N, with SlowFast regimes associated with greater underestimation than FastSlow, regardless of duration. Lastly, we observed that SlowFast accumulation regimes led to an increase in the central tendency effect, suggesting that slower rates of accumulation may increase reliance on the magnitude priors.

Discussion

In this study, we report that when sensory evidence steadily accumulates over time, and when task difficulty is equated across magnitude dimensions (space, time, number), duration estimates are resilient to manipulations of number and surface, whereas number and surface estimates are biased by the temporal properties of sensory evidence accumulation. Specifically, we replicated the findings of Lambrechts and colleagues¹³ by demonstrating that number and surface estimates are under- and overestimated when presented for long and short durations, respectively. These results complement the findings that duration can be resilient to numerosity interference²⁷, and that the direction of the interference between space and time may go in opposite direction when using dynamic displays^{13,28}. Although prior findings have reported asymmetrical effects in magnitude estimation, our findings differ in several ways. First, participants provided a quantified estimation of a given magnitude dimension, allowing a direct assessment of performance within a Bayesian framework in mind (specifically characterizing properties of the central tendency effects as a function of magnitude interferences). In previous reports of interference of number on duration^{5,6}, behavioral effects were concluded on the basis of increased reaction times and error rates during incongruent condition presentation (e.g., small number presented with a long duration) which prevented

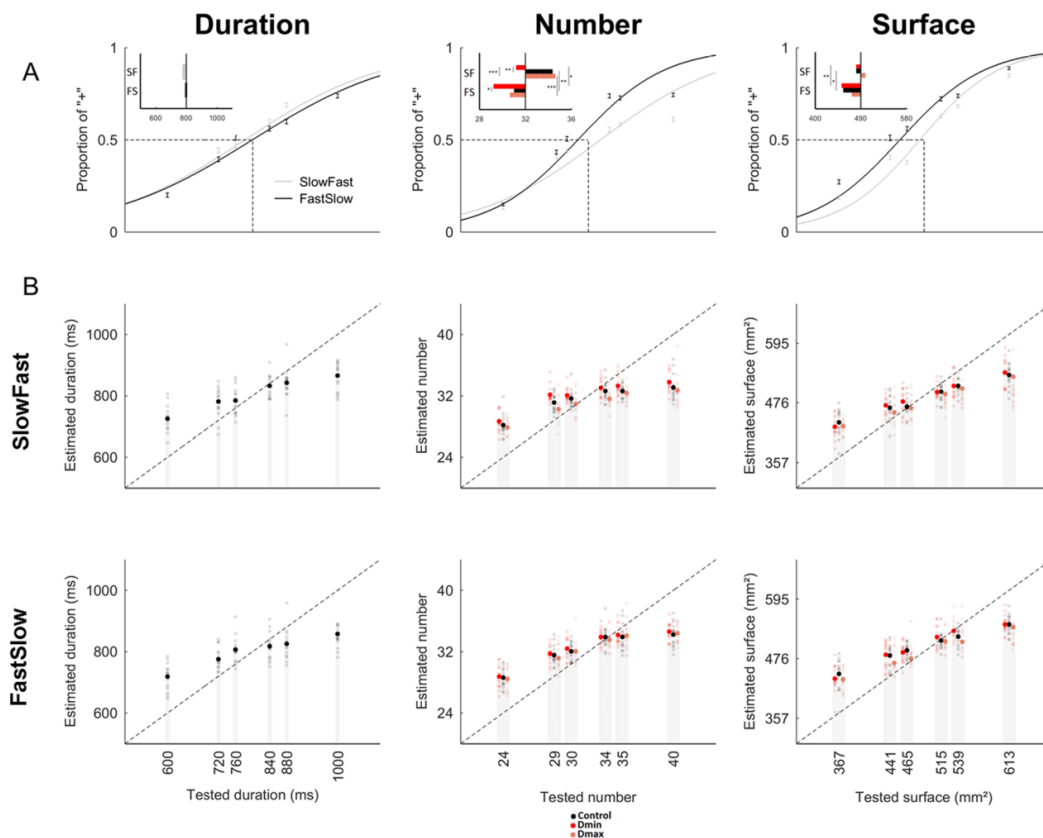


Figure 5. The Accumulation Regime affects Numerosity and Surface but not Duration (Experiment 2). **(A)** Psychometric curves illustrate the grand average proportion of “+” as a function of duration (left panel), number (middle panel) or surface (right panel) when the sensory evidence accumulation was manipulated. The SlowFast (SF) results are reported in light gray, the FastSlow (FS) results are reported in black. Left panel: top inset reports the mean PSE of duration estimation observed for each distribution as compared to the ideal observer (vertical black line). Middle panel: top inset reports the mean PSE of number estimation observed for each distribution (SF, FS) and manipulation of D (D_{mean} : black, D_{min} : red, D_{max} : pink). Right panel: top inset reports the mean PSE of surface estimation observed for each distribution (SF, FS) and manipulation of D (D_{mean} : black, D_{min} : red, D_{max} : pink). The estimation of D was not affected by the accumulation regime of sensory evidence whereas N and S were overestimated in the FastSlow as compared to the SlowFast distribution. N and S were also affected by manipulating the non-target duration. **(B)** Individual (transparent dots) and mean (filled dots) continuous judgments for the estimation of D (left panel), N (middle panel) and S (right panel) with SlowFast (top row) and FastSlow (bottom row) regimes of sensory evidence accumulation. The dotted line is the ideal observer’s performance. D_{min} : minimal duration value; D_{max} : maximal duration value. *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; bars are 2 s.e.m.

the direct evaluation of participants’ magnitude perception *per se*. As such, no clear direction of interference effects could be concluded from the studies beyond the existence of an interference. Our experimental design also prevented participants to explicitly enumerate the dots in a given trial unlike in previous experiments, in which the speed of dot presentation enabled counting which yielded an influence of N on D⁶.

In a study⁷ using dynamic display which addressed a question close to the ones we address with this experimental approach, participants judged the spatial length or the duration of a growing line. As previously discussed elsewhere (see ref. 13), the spatial task could have been performed using the coordinates of the line on the screen irrespective of the duration it took the line to grow, possibly explaining why duration was irrelevant for the spatial estimate. In other words, the duration in this task was likely the noisiest cue which in turn was affected by the least noisy cue (i.e. the spatial dimension). Consistent with our current results, a recent study²⁸ showed that a longer duration yielded an underestimation of length. In this experiment, the environment was dark (fMRI study) and participants had no access to visual cues to constrain their spatial estimate of the moving dot. In this context, the results showed that the shorter duration increased the distance of the moving dot, consistent with the present findings. Hence, and consistent with previous literature, the lack of robust bidirectional interactions

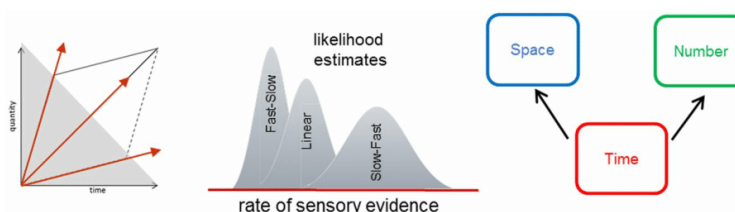


Figure 6. Accumulation regime influences Bayesian estimates. Our results suggest that the speed or rate of sensory evidence accumulation early in the trial (shaded region) and the duration of the trial affect the estimation of surface and numerosity. Additionally, slower rates of sensory evidence are associated with greater uncertainty; greater uncertainty results in increased reliance on the priors. To accommodate these findings, we suggest that the rate of sensory evidence is effectively estimated independently of the total duration although the duration may regulate the noise level in sensory evidence accumulation of quantities.

between magnitude dimensions does not support a literal interpretation of ATOM; however, we do not argue that time, number and space do not interact under certain constraints, and rather consider our results to favor a more liberal interpretation of ATOM. Specifically, by considering a Bayesian model relying on multiple priors (one for each dimension), magnitudes may interact in the context of conflicting sensory cues. Recent hypotheses suggest that a Bayesian framework can provide a general explanation for the variety of behavioral features observed in magnitude estimations independently applied to distance, loudness, numerical or temporal judgments⁴. The proposed Bayesian framework combines an estimate of the likelihood (sensory input) with a prior representation (memory). One major goal of our study was thus to determine the degree to which different magnitude dimensions might rely on an amodal global prior representation of magnitude as would be expected using a literal interpretation of a generalized magnitude system such as ATOM³. To accomplish this, participants took part in two experiments independently manipulating the congruence across magnitude dimensions (Experiment 1) and the rate of sensory evidence provided to participants (Experiment 2).

A first prediction was that if different magnitude dimensions rely on a single amodal prior, then magnitude estimates should exhibit similar levels of central tendency across magnitude dimensions (duration, surface, number; Fig. 2D). Instead, in Experiment 1, our results demonstrated that surface estimates exhibited greater central tendency than either duration or number, and surface estimates were not correlated with the degree of central tendency for either dimension. However, duration and number did exhibit correlated central tendency effects. This finding suggests that estimates of surface are distinct from estimates of duration and number, but that duration and number may be more similar to one another. Indeed, neural recording studies in the prefrontal and parietal cortex of non-human primates have revealed overlapping, yet largely separate, representations of duration and size^{51,52} and number and size^{53,54}. Further, while number, size, and time exhibit common activations of the right parietal cortex, they each engage larger networks of regions beyond this area^{28,55,56}. For size estimates, recent work suggests that comparisons of object size draw on expectations from prior experience in other brain regions⁵⁷. Yet, as no strong bidirectional effects were observed between duration and number, it is unlikely that duration and number share neuronal populations with similar tuning features.

Another possible interpretation of the results obtained in Experiment 1 is to consider multiple priors in magnitude estimations. When participants make temporal judgements, the combination of prior knowledge $P(\pi)$ and noisy sensory inputs $P(D|\pi)$ (duration of the given trial) enables participants to make an accurate posterior estimate, represented by $P(\pi|D) \propto P(D|\pi) \cdot P(\pi)$ (see ref. 4, Box 3 for more details), which explains the regression to the mean. Neither numerosity nor surface priors are present in this equation, which could explain why duration estimates are robust to numerosity or surface manipulations. Because numerosity and surface accumulate over time, one possible strategy for the participants was to estimate numerosity and surface based on both the speed of presentation of stimuli, and on the duration of the trial. Specifically, a high (low) speed over a long (short) duration of presentation corresponded to a large (small) value of numerosity or surface. In other words, these results seemed to indicate that, using dynamic displays, the speed of events influences the estimation of magnitudes and yield opposite directionality in the interference across magnitudes. If participants used the speed of event presentation and overestimated (underestimated) D, they also overestimated (underestimated) N or S. As the computation of speed also relies on the duration, both speed and duration become important cues in the estimation process but may have distinct impacts. In Experiment 1, the central tendency effect showed that the shortest duration was overestimated, which could explain why participants overestimated N and S for D_{\min} ; conversely, the longest duration was underestimated, which may explain why participants underestimated N and S. Under this hypothesis, the uncertainty related to the temporal dimension may add noise in the decision or the accumulation process, so that the perceived duration of the trial can bias numerosity and surface estimates (Fig. 6). When numerosity and surface accumulate over a given duration, if that duration is short (long) it will affect the accuracy of participants' estimations. This explanation would be compatible with the hypothesized effect of duration as introducing noise on sensory accumulation, and Experiment 2 was specifically designed to tease out the effect of speed changes in the accumulation rate on the estimation of the target magnitude.

Indeed, one noteworthy aspect unique to the time dimension is that the objective rate of presentation is fixed⁵⁸. That is, objective time by conventional measurements proceeds at a single mean rate. In contrast, we can experimentally manipulate the rate at which we present information for number and surface. In Experiment 1,

in order to keep the values of surface and number fixed when duration was manipulated, we necessarily had to change the rate of accumulation for these values. For example, between short and long durations with the same value of number, we had to change the rate of accumulation for number so that the same total value was reached at the end of the duration. This may explain the incongruent effects of duration on surface and number; shorter (longer) durations may engender larger (smaller) estimates of surface and number because the rate of accumulation is faster (slower). In this sense, surface and number are not being influenced by the duration magnitude of the time dimension *per se*, but rather the time dimension is interfering with the rate of accumulation, and so the effect of duration may be an epiphenomenon of the experimental design. To test this hypothesis, we modulated the accumulation rate of the presentation of numerosity and surface in Experiment 2.

In Experiment 2, where the rate of accumulation for number and surface experienced a rate-change a little less (more) than halfway through the presentation time from fast-to-slow (slow-to-fast), we replicated and extended our findings of Experiment 1. Specifically, we again found that shorter (longer) durations led to longer (shorter) estimates of surface and number, regardless of the rate-change in accumulation regime. However, we also found a difference in accumulation regimes: when the rate of accumulation was slower in the beginning of the session, estimates of surface and number were smaller than when the rate of accumulation was faster. It is important to remember that the ultimate value of the presented surface and number was the same, regardless of the accumulation regime. As such, participants were biased in their estimates by the rate of evidence accumulation in the first-half of the given trial, regardless of how long that trial lasted. This strongly suggests that human observers are biased by the rate of accumulation at the start of a trial, and are resistant to changes in rate throughout the trial.

This observation is important in the context of ongoing discussions on drift-diffusion processes in which the accumulation of evidence following the first end-point depends on terminated processes, guess probability (see ref. 59 case study 1 and Fig. 1; see ref. 60 Figs 2 and 4A) and on the importance of change points during the accumulation process^{61,62}. Changes in the accumulation rate performed in Experiment 2 imposed a nonlinearity in the accumulation process: the observation that an earlier rate change has a larger impact on magnitude estimation than a later rate change is reminiscent of the 'primacy effect' reported in evidence accumulation models, possibly indicative of suppression of newer information by old information⁶³. Additionally, this finding strengthens the hypothesis that the effect of duration on surface and numerosity may occur as a result of the impact on the implicit timing or accumulation rate, and not as a function of the explicitly perceived duration. This would be consistent with recent findings suggesting that noise memory - known to scale with duration - was not the primary factor of errors in decision-making but that noise in sensory evidence was instead a major contributor⁶⁴. Our results suggest that speeding up the rate of evidence and lengthening the duration of a trial may be equivalent to increasing noise in sensory accumulation of other magnitude dimensions (Fig. 6).

In Experiment 2, we also investigated the effect of accumulation regime on central tendency as participants again provided continuous magnitude estimates on a vertical sliding scale. Previous magnitude studies using continuous estimates have demonstrated a central tendency effect, with over(under)-estimations for small (large) magnitudes⁴, the degree to which depends on the uncertainty inherent in judging the magnitude in question^{44,65}. The result of this analysis revealed that, when the rate of accumulation was slow (fast) in the beginning of the trial, the degree of central tendency was greater (lesser); further, the objective duration of the trial did *not* impact central tendency. This finding suggests that slower accumulation regimes engender greater uncertainty in magnitude estimates, and that this uncertainty may be present before the ultimate decisional value is reached. Previous work in decision-making with evidence accumulation has suggested that the objective duration of a trial leads to greater reliance on prior estimates, as longer presentation times are associated with greater uncertainty⁶⁶. Specifically, Hanks and colleagues⁶⁶ found that a drift-diffusion model that incorporated a bias signal to rely on prior evidence that grows throughout the trial could explain reaction time differences in a dots-motion discrimination task. Notably, the bias signal is incorporated into the drift-diffusion process, such that longer trials push the accumulation rate towards a particular value, depending on the prior. A critical manipulation in this study was the emphasis on speed or accuracy for subjects; increased emphasis for accuracy led to longer decision times and greater reliance on the prior, as explained by the model. Our results suggest otherwise - the duration of the trial alone cannot determine reliance on the prior. If the effect of duration solely led to greater reliance on the prior, then we should have seen central tendency effects increase with longer durations, which did not occur in either Experiments 1 or 2. Instead, the rate of evidence accumulation determined reliance on the prior(s), regardless of duration, with slower rates leading to greater reliance.

Additional studies using neuroimaging techniques such as M/EEG need to investigate the neural correlates underpinning accumulation processes in the brain when estimating magnitudes (Centro-Parietal Positivity, for example, see ref. 67, 68), to fully explain the behavioral results obtained in these two experiments. Further, fMRI studies must be conducted to elucidate the neural circuits for memory representations of different magnitudes⁴⁸. Bayesian approaches may provide interesting perspectives on magnitudes estimations, and additional studies need to be performed to understand to which extent these models can be applied to explain the variety of results observed in the literature. One intriguing observation is the finding that duration estimates were not only resilient to changes in numerosity or surface, but also to the rate of sensory evidence. This finding is unexpected and runs counter-intuitive to various findings in time perception. In this task, these robust findings suggest that unlike surface and numerical estimates, duration may not rely on the accumulation of discretized sensory evidence.

References

- Gallistel, C. R. Animal cognition: the representation of space, time and number. *Annu. Rev. Psychol.* **40**, 155–189 (1989).
- Gallistel, C. R. Mental Magnitudes in *Space, Time and Number in the Brain: Searching for the Foundations of Mathematical Thought* (ed. Dehaene, S., Brannon, E.) 3–12 (Academic Press, 2011).
- Walsh, V. A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* **7**, 483–488 (2003).
- Petzschner, F. H., Glasauer, S. & Stephan, K. E. A Bayesian perspective on magnitude estimation. *Trends Cogn. Sci.* **19**, 285–293 (2015).

5. Xuan, B., Zhang, D., He, S. & Chen, X. Larger stimuli are judged to last longer. *J. Vis.* **7**(2), 1–5 (2007).
6. Dormal, V., Seron, X. & Pesenti, M. Numerosity-duration interference: A Stroop experiment. *Acta Psychol. (Amst.)* **121**, 109–124 (2006).
7. Casasanto, D. & Boroditsky, L. Time in the mind: Using space to think about time. *Cognition* **106**, 579–593 (2008).
8. Oliveri, M. *et al.* Perceiving numbers alters time perception. *Neurosci. Lett.* **438**, 308–311 (2008).
9. Bottini, R. & Casasanto, D. Implicit Spatial Length Modulates Time Estimates, But Not Vice Versa. *Spat. Cogn. VII* 152–162, doi:10.1007/978-3-642-14749-4_15 (2010).
10. Stavy, R. & Tirosh, D. *How Students (Mis-)Understand Science and Mathematics: Intuitive Rules* (Teachers' College Press, 2000).
11. Dormal, V. & Pesenti, M. Numerosity-Length Interference. *Exp. Psychol.* **54**, 289–297 (2007).
12. Javadi, A. H. & Aichelburg, C. When time and numerosity interfere: the longer the more, and the more the longer. *PLoS One* **7**, e41496 (2012).
13. Lambrechts, A., Walsh, V. & van Wassenhove, V. Evidence Accumulation in the Magnitude System. *PLoS ONE* **8**, e82122 (2013).
14. Wiener, M. & Thompson, J. C. The Effect of Background Context on the Size-Time Illusion. *Timing Amp Time Percept.* **4**, 167–186 (2016).
15. Yamamoto, K., Sasaki, K. & Watanabe, K. The number-time interaction depends on relative magnitude in the suprasecond range. *Cogn. Process.* **17**, 59–65 (2016).
16. Cai, Z. G. & Wang, R. Numerical Magnitude Affects Temporal Memories but Not Time Encoding. *PLOS ONE* **9**, e83159 (2014).
17. Xuan, B., Chen, X.-C., He, S. & Zhang, D.-R. Numerical magnitude modulates temporal comparison: An ERP study. *Brain Res.* **1269**, 135–142 (2009).
18. Roitman, J. D., Brannon, E. M., Andrews, J. R. & Platt, M. L. Nonverbal representation of time and number in adults. *Acta Psychol. (Amst.)* **124**, 296–318 (2007).
19. Calabria, M. & Rossetti, Y. Interference between number processing and line bisection: a methodology. *Neuropsychologia* **43**, 779–783 (2005).
20. Fischer, M. H., Castel, A. D., Dodd, M. D. & Pratt, J. Perceiving numbers causes spatial shifts of attention. *Nat. Neurosci.* **6**, 555 (2003).
21. Viarouge, A. & de Hevia, M. D. The role of numerical magnitude and order in the illusory perception of size and brightness. *Front. Psychol.* **4** (2013).
22. Dehaene, S., Bossini, S. & Giraux, P. The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* **122**, 371–396 (1993).
23. Ishihara, M., Keller, P. E., Rossetti, Y. & Prinz, W. Horizontal spatial representations of time: Evidence for the STEARC effect. *Cortex* **44**, 454–461 (2008).
24. Rammsayer, T. H. & Verner, M. The effect of nontemporal stimulus size on perceived duration as assessed by the method of reproduction. *J. Vis.* **14**, 17–17 (2014).
25. Rammsayer, T. H. & Verner, M. Larger visual stimuli are perceived to last longer from time to time: The internal clock is not affected by nontemporal visual stimulus size. *J. Vis.* **15**, 5–5 (2015).
26. Vallesi, A., McIntosh, A. R. & Stuss, D. T. How time modulates spatial responses. *Cortex* **47**, 148–156 (2011).
27. Agrillo, C., Ranpura, A. & Butterworth, B. Time and numerosity estimation are independent: Behavioral evidence for two different systems using a conflict paradigm. *Cogn. Neurosci.* **1**, 96–101 (2010).
28. Coull, J. T., Charras, P., Donadieu, M., Droit-Volet, S. & Vidal, F. SMA Selectively Codes the Active Accumulation of Temporal, Not Spatial, Magnitude. *J. Cogn. Neurosci.* **27**, 2281–2298 (2015).
29. Cai, Z. G. & Connell, L. Space-time interdependence: Evidence against asymmetric mapping between time and space. *Cognition* **136**, 268–281 (2015).
30. Newcombe, N. S., Levine, S. C. & Mix, K. S. Thinking about quantity: the intertwined development of spatial and numerical cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* **6**, 491–505 (2015).
31. Winter, B., Marghetis, T. & Matlock, T. Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number. *Cortex* **64**, 209–224 (2015).
32. Bonato, M., Zorzi, M. & Umiltà, C. When time is space: Evidence for a mental time line. *Neurosci. Biobehav. Rev.* **36**, 2257–2273 (2012).
33. Buetti, D. & Walsh, V. The parietal cortex and the representation of time, space, number and other magnitudes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **364**, 1831–1840 (2009).
34. Dormal, V., Dormal, G., Joassin, F. & Pesenti, M. A common right fronto-parietal network for numerosity and duration processing: An fMRI study. *Hum. Brain Mapp.* **33**, 1490–1501 (2012).
35. Pinel, P., Piazza, M., Le Bihan, D. & Dehaene, S. Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron* **41**, 983–993 (2004).
36. Hayashi, M. J. *et al.* Time Adaptation Shows Duration Selectivity in the Human Parietal Cortex. *PLoS Biol.* **13**, e1002262 (2015).
37. Hubbard, E. M., Piazza, M., Pinel, P. & Dehaene, S. Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* **6**, 435–448 (2005).
38. Riemer, M., Diersch, N., Bublatzky, F. & Wolbers, T. Space, time, and numbers in the right posterior parietal cortex: Differences between response code associations and congruency effects. *NeuroImage* **129**, 72–79 (2016).
39. Nieder, A. Coding of abstract quantity by 'number neurons' of the primate brain. *J. Comp. Physiol. A* **199**, 1–16 (2013).
40. Javadi, A. H., Brunec, I. K., Walsh, V., Penny, W. D. & Spiers, H. J. Transcranial electrical brain stimulation modulates neuronal tuning curves in perception of numerosity and duration. *NeuroImage* **102**, Part 2, 451–457 (2014).
41. Shi, Z., Church, R. M. & Meck, W. H. Bayesian optimization of time perception. *Trends Cogn. Sci.* **17**, 556–564 (2013).
42. Jazayeri, M. & Shadlen, M. N. Temporal context calibrates interval timing. *Nat. Neurosci.* **13**, 1020–1026 (2010).
43. Petzschner, F. H. & Glasauer, S. Iterative Bayesian Estimation as an Explanation for Range and Regression Effects: A Study on Human Path Integration. *J. Neurosci.* **31**, 17220–17229 (2011).
44. Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M. & Burr, D. C. Optimal encoding of interval timing in expert percussionists. *J. Neurosci. Off. J. Soc. Neurosci.* **32**, 1056–1060 (2012).
45. Yates, M. J., Loetscher, T. & Nicholls, M. E. R. A generalized magnitude system for space, time, and quantity? A cautionary note. *J. Vis.* **12**, 9–9 (2012).
46. Rammsayer, T. H. & Verner, M. Evidence for different processes involved in the effects of nontemporal stimulus size and numerical digit value on duration judgments. *J. Vis.* **16**, 13–13 (2016).
47. Vilares, I., Howard, J. D., Fernandes, H. L., Gottfried, J. A. & Kording, K. P. Differential representations of prior and likelihood uncertainty in the human brain. *Curr. Biol. CB* **22**, 1641–1648 (2012).
48. Wiener, M., Michaelis, K. & Thompson, J. C. Functional correlates of likelihood and prior representations in a virtual distance task. *Hum. Brain Mapp.* **37**, 3172–3187 (2016).
49. Ivry, R. B. & Hazeltine, R. E. Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 3–18 (1995).
50. Levy, J. M., Namboodiri, V. M. K. & Hussain Shuler, M. G. Memory bias in the temporal bisection point. *Front. Integr. Neurosci.* **9** (2015).

51. Genovesio, A., Tsujimoto, S. & Wise, S. P. Neuronal Activity Related to Elapsed Time in Prefrontal Cortex. *J. Neurophysiol.* **95**, 3281–3285 (2006).
52. Marcos, E., Tsujimoto, S. & Genovesio, A. Independent coding of absolute duration and distance magnitudes in the prefrontal cortex. *J. Neurophysiol.* jn.00245.2016, doi:10.1152/jn.00245.2016 (2016).
53. Nieder, A., Diester, I. & Tudusciuc, O. Temporal and Spatial Enumeration Processes in the Primate Parietal Cortex. *Science* **313**, 1431–1435 (2006).
54. Tudusciuc, O. & Nieder, A. Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proc. Natl. Acad. Sci.* **104**, 14513–14518 (2007).
55. Dormal, V., Andres, M. & Pesenti, M. Dissociation of numerosity and duration processing in the left intraparietal sulcus: A transcranial magnetic stimulation study. *Cortex* **44**, 462–469 (2008).
56. Dormal, V. & Pesenti, M. Common and specific contributions of the intraparietal sulci to numerosity and length processing. *Hum. Brain Mapp.* **30**, 2466–2476 (2009).
57. Konkle, T. & Oliva, A. A Real-World Size Organization of Object Responses in Occipitotemporal Cortex. *Neuron* **74**, 1114–1124 (2012).
58. Riemer, M., Trojan, J., Kleinböhl, D. & Hölzl, R. A. 'view from nowhen' on time perception experiments. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 1118–1124 (2012).
59. Ratcliff, R. A note on modeling accumulation of information when the rate of accumulation changes over time. *J. Math. Psychol.* **21**, 178–184 (1980).
60. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. Diffusion Decision Model: Current Issues and History. *Trends Cogn. Sci.* **20**, 260–281 (2016).
61. Churchland, A. K. & Kiani, R. Three challenges for connecting model to mechanism in decision-making. *Curr. Opin. Behav. Sci.* **11**, 74–80 (2016).
62. Barack, D. L. & Gold, J. I. Temporal trade-offs in psychophysics. *Curr. Opin. Neurobiol.* **37**, 121–125 (2016).
63. Holmes, W. R., Trueblood, J. S. & Heathcote, A. A new framework for modeling decisions about changing information: The Piecewise Linear Ballistic Accumulator model. *Cognit. Psychol.* **85**, 1–29 (2016).
64. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science* **340**, 95–98 (2013).
65. Murai, Y. & Yotsumoto, Y. Timescale- and Sensory Modality-Dependency of the Central Tendency of Time Perception. *PLOS ONE* **11**, e0158921 (2016).
66. Hanks, T. D., Mazurek, M. E., Kiani, R., Hopp, E. & Shadlen, M. N. Elapsed Decision Time Affects the Weighting of Prior Probability in a Perceptual Decision Task. *J. Neurosci.* **31**, 6339–6352 (2011).
67. O'Connell, R. G., Dockree, P. M. & Kelly, S. P. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat. Neurosci.* **15**, 1729–1735 (2012).
68. Kelly, S. P. & O'Connell, R. G. Internal and external influences on the rate of sensory evidence accumulation in the human brain. *J. Neurosci. Off. J. Soc. Neurosci.* **33**, 19434–19441 (2013).

Acknowledgements

We thank members of UNIACt at NeuroSpin for their help in recruiting and scheduling participants, and Baptiste Gauthier and Tadeusz W. Kononowicz for discussions. This work was supported by an ERC-YSStG-263584 to V.v.W.

Author Contributions

B.M., M.W. and V.v.W. designed the study. B.M. collected the data. B.M., M.W. and V.v.W. analyzed the data. All authors co-wrote the manuscript and gave final approval for publication.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-00680-0

Competing Interests: The authors declare that they have no competing interests.

Data Accessibility: Data will be made available through Open Science Framework

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

3 | A Theory Of Magnitudes

Here, I introduce several observations on existing literature of interference effects across magnitudes that motivate the second study of the thesis. First, I argue that there is no strong evidence for the existence of a common metric across magnitudes. Second, I argue that the direction of the interference effects is highly task-dependent. Third, I discuss the notion that numerical magnitude interferes with perceived duration when magnitude is encoded. Finally, I discuss a possible automaticity in magnitude processing at a non-symbolic level and not at a symbolic one.

3.1 A Common Metric for Time, Space and Number and the issue of scaling

While ATOM predicts scaling across magnitudes (i.e. a mapping factor across magnitude dimensions should exist), an alternative and more parsimonious interpretation of existing number-time interference effects in behavioral data may be a task-driven categorization of magnitudes. In fact, according to ATOM, perceived duration should increase as a function of the tested digit (Figure 3.1, panel A), and the difference in perceived duration between two digits should also increase as a function of the tested duration (Figure 3.1, panel B). However, results in several studies (Rammsayer and Verner, 2016; Chang *et al.*, 2011) suggested that perceived duration did not increase as a function of digit value but rather increased along a dichotomical categorization of numerosity (small or large numerical value). I argue in this section that the number-time interaction may depend on the instructions given to the participants.

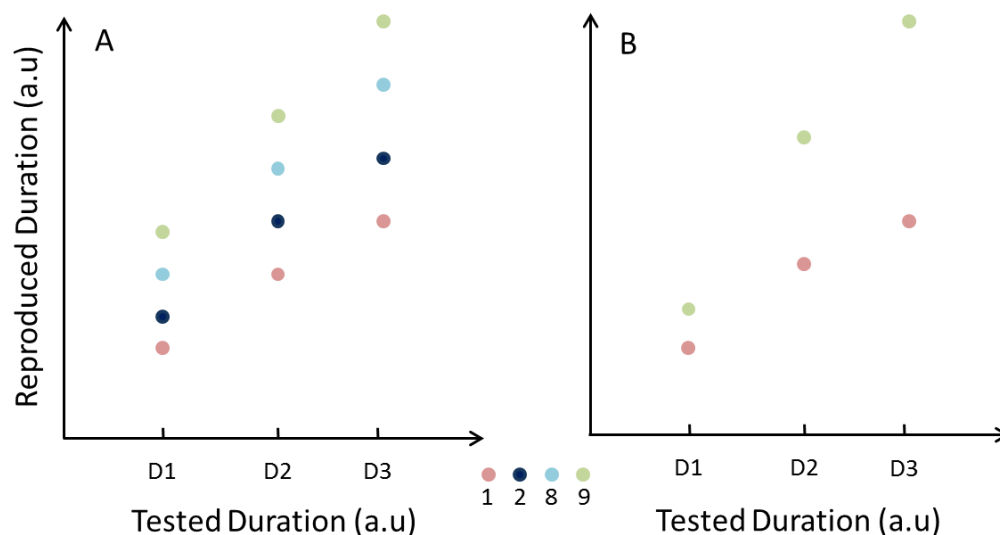


Figure 3.1: **A:** For a given tested duration (D1, D2 or D3), the reproduced duration increases as a function of the magnitude of the interfering non-temporal dimension (e.g., here, digits from 1 to 9 represented by different colors). Note that the interference effect increases with duration. **B:** The difference in reproduced duration between digits 1 (pink) and 9 (green) is predicted to increase as a function of the tested duration.

Rammsayer and Verner (2016) recently investigated the effect of the value on reproduced duration. Participants were instructed to reproduce the duration of presentation of a digit, and to judge its numerical value. The mean reproduced durations were significantly longer when large digits (8 and 9) were presented as compared to small digits (1 and 2). However, visual inspection and additional analysis of their results (see Figure 3.2) did not reveal any significant increase of the reproduced duration as a function of the numerical value. These findings go against the previous consideration (see Figure 1A) and do not support the hypothesis of a general system in which different magnitudes share a common metric.

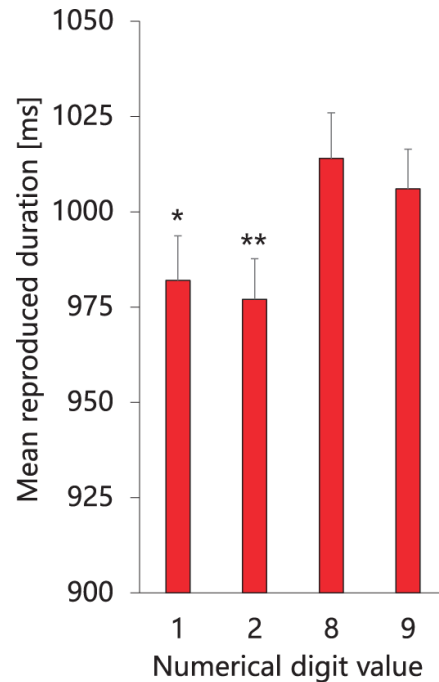


Figure 3.2: From Rammsayer and Verner (2016). Mean reproduced duration as a function of the numerical digit value. Results clearly indicate the absence of a scaling effect (i.e. the mean reproduced duration does not increase as a function of digit value). Post hoc tests only showed that digit 8 ($1014 \pm 174\text{ms}$) was reproduced longer than digit 1 ($982 \pm 156\text{ms}$, $p < .05$) and 2 ($977 \pm 148\text{ms}$, $p < .01$). * $p < .05$; ** $p < .01$

Instead, one possible hypothesis would be that the effect of numerical digit on duration reproduction could be explained by a categorical decision-making imposed by the task, when classifying the digits in relative low and high digit sets. This is particularly relevant given the task instructions provided to participants who “*were instructed that there was a low value and a high value digit set consisting of the digits 1 and 2 and the digits 8 and 9, respectively*” (see Rammsayer and Verner (2016), Experiment 2, stimuli and procedure part). One hypothesis is thus that a linear increase in reproduced duration as a function of numerical value might be observed if participants were instructed to pay attention to the exact numerical value, instead of just classifying the digit as low or high. A similar interpretation was also entertained in a line

bisection task (de Hevia and Spelke, 2009) in which lines were flanked by digits: in their study, de Hevia and Spelke (2009) found that, on average, participants bisected the lines towards the large digit. Four possible numerical distances between the small and the large digit were tested (either a difference of one, two, four or six). No effect of numerical distance was found. According to the authors, *“the null effect of the numerical distance suggests that numerical flankers exert an influence on the bisection performance mostly in terms of relative magnitude, i.e., along a dichotomical small/large distinction”*. In other words, the hypothesis of a common metric was not supported in the bisection task and rather, the relative magnitude was deemed to be important. In a similar fashion, using a time reproduction task, Yamamoto et al. (2016) found that the effect of numerical value on reproduced duration depended on the relative numerical distance, not on the absolute distance. Participants were presented with single-digit numerals (1, 3, 5, 7 and 9) and double-digit numerals (10, 30, 50, 70 and 90). For each tested pair (1-10, 3-30, 5-50, 7-70 and 9-90), the relative numerical distance was kept constant (the large number was always ten times larger than the small one) while the absolute distance varied. With the literal read of a generalized magnitude system with a common metric, one prediction would be that the larger the difference between two numbers, the larger the difference in the reproduced durations (absolute differences between pairs in this study were 9, 27, 45, 63 and 81. i.e. the absolute difference increased by step of 18 from the first pair (1-10) to the last one (9-90). Yet in the Yamamoto et al (2016) study, results showed no significant increase of the reproduced duration as a function of the absolute numerical distance (Figure 3.3) and results could again be simply explained by considering a small/large categorization of the numerical values.

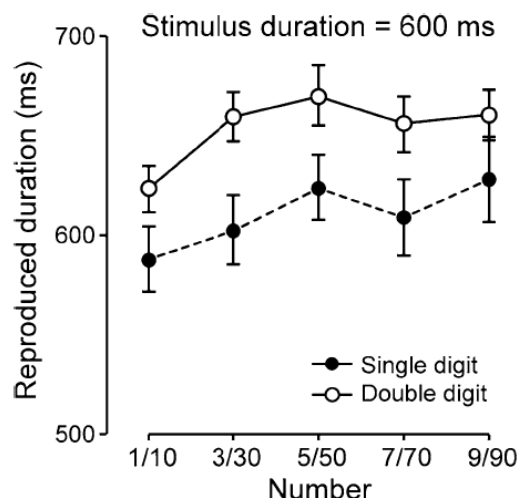


Figure 3.3: From Yamamoto *et al.* (2016). This figure shows the mean reproduced duration for single and double digits numerals. Results indicate that the reproduced duration does not readily increase as a function of the absolute numerical distance.

Altogether, the pattern of results in number-time interference is not clearly consistent with the hypothesis of a common metric across magnitudes. However, challenging this conclusion, Cai and Wang (2014), found that perceived duration linearly increased as a function of the digit when participants had to reproduce its exact numerical value (see Lu *et al.* 2009 for similar findings). The absence of a linear increase in reproduced duration as a function of the tested numerosity will thus be investigated in Chapter 4 which will specifically test two possible factors that may intervene in time-number interferences, namely: (1) the numerical format that is used (symbolic, non-symbolic numerosity) and (2) the amount of attentional resources when judging the numerical magnitude (as proposed by Rammsayer and Verner 2016).

3.2 Static vs. Dynamic displays and the direction of interference effects

The directionality of the interactions between non-temporal and temporal magnitude dimensions reported in the literature may largely depend on the experimental paradigm being

used. There is a growing body of evidence showing that the numerical value of a digit interferes with temporal judgments (Rammsayer and Verner 2016; Oliveri *et al.* 2008; Chang *et al.* 2011), and similar effects are reported when numerical information is presented in non-symbolic form: the larger the number of items in a set, the longer the perceived the duration (Dormal, Seron, and Pesenti, 2006; Mo, 1971, 1974, 1975; Xuan *et al.*, 2007; Javadi and Aichelburg 2012). Yet, while duration necessarily accumulates over time, numerical information does not. Indeed, numerical information can be statically provided to the participants (all the items of a set are presented at the same time, during the entire trial) or dynamically (the number of items increases in time, to reach its maximum value at the end of the trial). Investigating the number-time interaction with a static or a dynamic design provide opposite results: whereas it is often reported that large numerosities lengthen the perceived duration when a static design is used (Dormal, Seron, and Pesenti, 2006; Mo, 1971; Javadi and Aichelburg 2012), perceived duration is resilient to non-temporal manipulation when using a dynamic design (Lambrechts, Walsh, and van Wassenhove, 2013; Martin, Wiener, and van Wassenhove, 2017). The present section investigates this specific point.

Javadi and Aichelburg (2012) instructed participants to judge which of two successive sets of items was presented longer (duration task) or which was more numerous (number task). Their results revealed a positive correlation between time and number, with more numerous sets being judged to last longer. These results are in line with a seminal series of experiments showing that temporal estimation increased as a function of numerosity. Mo (1971) initially reported that the proportion of “long” judgments in a duration task significantly increased as a function of the number of dots presented to the participants; in a second study (Mo, 1974), participants were instructed to judge the duration of the second stimuli in a pair, and the proportion of “longer” responses was shown to decrease when the number of dots of the first stimulus increased. These observations indicate that participants perceived the first stimulus as being longer than its actual physical duration; when the numerosity of the second stimulus was manipulated, the proportion of “longer” responses increased as a function of the number magnitude. In a third study (Mo, 1975), participants were presented with sets of dots that varied in numerosity and were instructed to reproduce the duration. Once again, results

showed a general tendency for temporal reproduction to increase as numerosity increased. More recently, Dormal and colleagues (2006) used a Stroop task in which participants were presented with visually flashing dots and had to compare either the numerosity or the duration of each trial. In the duration task, results showed that neutral and congruent conditions (in which larger (smaller) numerosity were matched with larger (shorter) durations) were answered faster than the incongruent ones. Overall, these findings suggest an asymmetry in the number-time interaction, with numerical magnitude interfering with duration judgments. The same is also true in the case of space-time interactions: most studies using a static design revealed that the larger the physical size of a stimulus, the longer the perceived duration (Xuan *et al.*, 2007; Rammsayer and Verner, 2014, 2015). For instance, in a reproduction task, Casasanto and Boroditsky (2008) showed that participants (under-) over- estimated the duration when a line covered a (short) long distance. In another experiment (Xuan *et al.*, 2007), manipulating the size of stimuli interfered with temporal estimations such that reaction times and error rates increased in incongruent conditions compared to congruent ones. Altogether, the pattern of results supports the general idea of a generalized magnitude system with perceived duration increasing as a function of non-temporal (i.e. spatial and numerical) magnitude in dynamic designs. However, duration seems to be resilient to numerical and spatial manipulations when sensory evidence accumulates over time (Lambrechts, Walsh, and van Wassenhove, 2013; Martin, Wiener, and van Wassenhove, 2017). The next paragraph investigates this specific point.

Several studies also investigated the number-time and space-time interactions when numerical and spatial information accumulates in time, namely using dynamic displays. In such cases, reported results were quite different: numerical (spatial) magnitude did not interfere anymore with temporal estimates, whereas manipulating the duration of stimuli altered numerical (spatial) judgments. Furthermore, while with static designs, increasing non-temporal magnitudes increase perceived duration, in dynamic designs, increasing the duration diminished perceived numerosity or size. In Agrillo *et al.* (2010), participants were presented with successive tones and were asked to either judge the duration of the trial or the total number of tones presented in each trial. The results revealed that the number of tones which accumulated over time did not interfere with the duration task. Lambrechts *et al.* (2013) and Martin *et al.*

(2017) used an experiment in which sensory evidence accumulated over time. Experimental trials consisted of clouds of dynamic dots. Participants were asked to judge the duration of the trial, the total number of dots or the cumulative surface filled by the dots while the two non-target dimensions were manipulated. Results showed that duration estimations were resilient to spatial and numerical changes when perceptual evidence accumulated over time. Surprisingly, varying the duration of presentation interfered with spatial estimations, in a counter-intuitive way. When the duration of presentation decreased (increased) participants over- (under-) estimated the number and surface filled by the dots (see Chapter 2).

To account for differences between static and dynamic designs, Casasanto and Boroditsky (2008) presented to the participants a static line or a growing line. Both conditions revealed that the larger the distance, the longer the perceived duration. Such findings are not consistent with the previous hypothesis according to which, in dynamic designs, numerical or spatial manipulations should not interfere with duration estimates. To explain this discrepancy, it has been argued (Lambrechts, Walsh, and van Wassenhove, 2013) that the spatial task could be performed by computing the coordinates of the line on the screen, no matter the time it took for the line to grow. In other words, participants could have used visual cues, such as the edges of the screen to perform the spatial task, preventing the temporal dimension to interfere with spatial judgments. Supporting this possible explanation, in Coull and colleagues' (2015) experiment, the environment was dark (fMRI study) and participants had no access to visual cues to estimate the distance traveled by a moving dot. In this context, the results showed that the shorter duration increased the perceived distance of the moving dot, consistent with previous findings. Hence, in dynamic (and more ecological) designs, the observation that the longer the duration, the smaller the perceived numerosity (or distance) is inconsistent with a generalized magnitude system.

Indeed, in the previous section, we argued that the effect of a non-target magnitude dimension on the dimension of interest should increase as a function of the magnitude of the non-target dimension. In other words, increasing the duration should lead to an increase of the perceived number or surface, not a decrease. In a dynamic design, for a given magnitude, decreasing the duration of presentation implies increasing the rate of stimuli being displayed:

for example, the speed will be higher if 10 dots are sequentially presented in 1 second than if they are in 2 seconds. In order to perform numerical judgments, increasing the rate of accumulation implies that the number of items to be processed per unit of time will necessarily increase. It is reasonable to think that such increase will add uncertainty in the accumulation process (larger magnitudes are associated to larger standard deviation, which is called the scalar variability), which implies that perceived numerical magnitude will be noisier. On the other hand, decreasing the rate of accumulation of evidence will decrease the uncertainty in the accumulation process (we can imagine that at a given point, if the rate is slow, participants will have an accurate estimation of the numerical magnitude). In Agrillo *et al.* (2010), participants were presented with successive tones. On average, 15 tones were presented in 9 seconds (~1.67 tones/s), giving them enough time to count, which can explain why manipulating duration did not interfere with numerical estimation, supporting the idea of an accumulation process more or less noisy. In their recent paper, Ratcliff and colleagues (2016) argue that *“Sequential-sampling models like the diffusion model [...] view decision making as a process of noisy accumulation of evidence from a stimulus”*. Supporting this statement, our results (see Chapter 2) suggest that speeding up the rate of presentation of sensory evidence may be equivalent to increasing noise in sensory accumulation of other magnitude dimensions. In this sense, surface and number are not being influenced by the duration *per se*, but rather the time dimension is interfering with the rate of accumulation. This may explain the incongruent effects of duration on surface and numerosity with shorter (longer) durations that engender larger (smaller) estimates of surface and number, because the rate of accumulation is faster (slower).

To summarize, when sensory accumulation of numerical or spatial information is driven by the external rate of information (e.g. dots sequentially presented), manipulating the duration adds noise in the accumulation process. As such, shorter durations become associated with larger estimations of non-temporal magnitude dimensions. On the other hand, when the numerical (or spatial) information is statically presented, accumulation is likely driven by an internal sampling rate of sensory evidence, yielding no effect of duration on numerical estimation. All empirical predictions are synthesized in Figure 3.4.

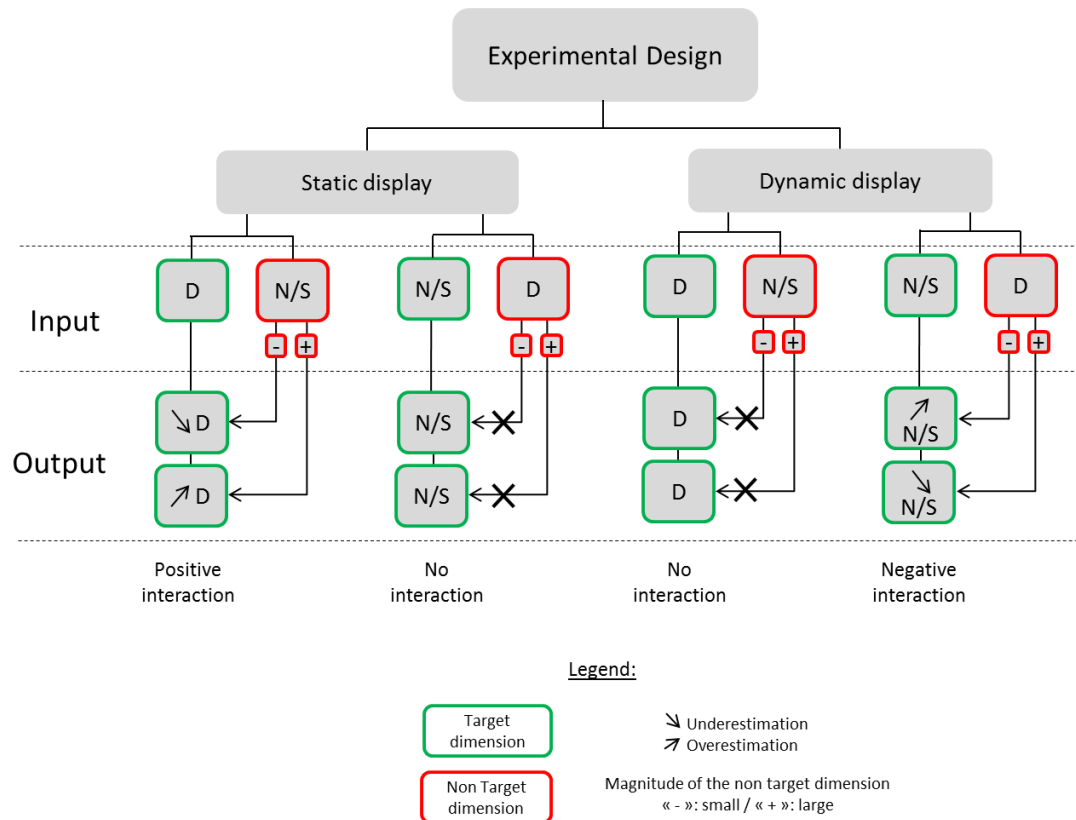


Figure 3.4: Summary of the interactions between temporal and non-temporal dimensions. Static: The larger the magnitude of the non-target dimension, the longer the perceived duration (positive interaction), no effect of duration on non-temporal dimensions. Dynamic: the longer the duration the smaller the perceived magnitude of non-temporal dimension (negative interaction), no effect of numerical or spatial information on perceived duration. Input: information presented to the participant. Output: response provided by the participant.

3.3 Numerical magnitude affects temporal encoding, not temporal reproduction

In Chang *et al.* experiment (2011), participants were instructed to reproduce the duration of a stimulus. In the first experiment, participants reproduced a standard duration during which a digit was presented; in the second experiment, the digit was only presented during the reproduction phase. In their design, if participants reproduced longer durations for large digits than for small digits in the two conditions, an implicit or automatic association of non-temporal and temporal magnitudes was predicted; if a difference in mean reproduced durations between the two experiments was found, this would rule out the existence of such association. In Experiment 1, large digits were reproduced longer than small ones and such difference was not found in the Experiment 2, when the digit was presented during the reproduction phase. In line with the literature (Xuan *et al.*, 2007; Dehaene and Akhavein, 1995; de Hevia and Spelke, 2009; Dormal and Pesenti, 2007), Chang and colleagues (2011) concluded on the existence of “*the automatic effect of numerical magnitude on temporal reproduction*”. However, this automaticity in the number-time interactions seems to only occur at the encoding or memory stages (Figure 3.5). The idea of an automatic effect of numerical magnitude on duration estimates is challenged in the next section of this chapter (section 3.4) and in the next chapter (Chapter 4). While it seems that the number-time interaction does not occur at the reproduction step, it would be interesting in the future to understand exactly when the presentation of a digit interferes with temporal estimation, either at the encoding phase or when the duration to be reproduced is stored in memory, using for example M/EEG techniques .

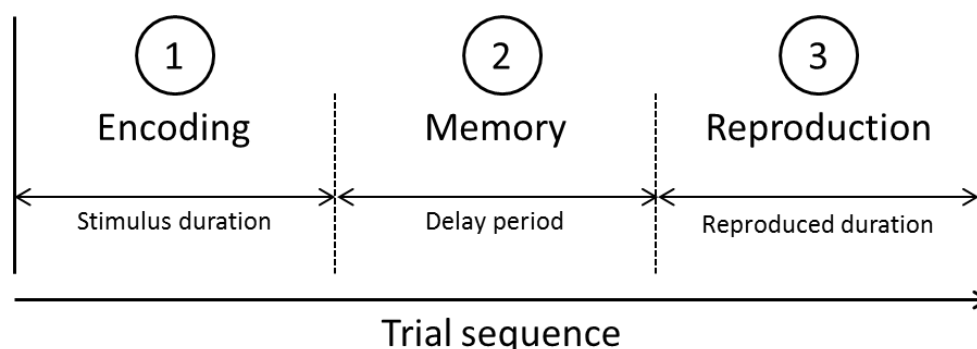


Figure 4: Illustration of a trial sequence. Numerical magnitude interferes with duration estimates either during encoding (1) or memory (2) phase. Future work on this topic should be done using neuroimaging techniques with good temporal resolution such as EEG and/or MEG, to clearly understand when this interaction takes place in temporal reproduction tasks.

3.4 Non symbolic magnitude is automatically processed, not the symbolic one

A large number of studies report that the perceived duration of a stimulus is related to various aspects of non-temporal stimulus properties such as stimulus size and numerosity (Mo, 1971; Xuan *et al.*, 2007; Rammsayer and Verner, 2015; Dormal, Seron, and Pesenti, 2006; Oliveri *et al.*, 2008). While it has often been argued that participants did not have to pay attention to non-temporal magnitude to observe an interference on perceived duration (Dormal and Pesenti, 2007; de Hevia and Spelke, 2009), a recent study suggests that it may be more complicated. Rammsayer and Verner (2016) instructed the participants to reproduce the duration of presentation of Arabic digits that varied in numerical value and physical size. After having reproduced the duration, participants had to judge the numerical value or the physical size as small or large. As expected, larger stimuli and larger digits were reproduced longer on average. However, the effect of numerical value seemed to require attentional resources to be effective. Indeed, large digits were reproduced longer only when participants' attention was directed to digit value, not when they paid attention to stimulus size. The estimation of the size of a stimulus is made at a lower level of representation than the numerical value of a digit,

suggesting that non symbolic magnitudes are automatically extracted and processed whereas symbolic ones are not.

To further elucidate the role of attention in the number-time interaction, the authors designed a second experiment in which participants had to pay attention to the numerical value of a digit or to its parity. Interestingly, the results showed that large digits were reproduced longer only when participants paid attention to the digit value. No significant differences in reproduced duration were observed when participants had to judge the parity. Overall, these results indicate that the physical size of a stimulus is automatically processed while the numerical value of a digit is not. Additional work needs to be done to better understand to which extent a magnitude is more automatically processed at a non-symbolic level than at a symbolic level. In a recent study (see Chapter 4) we designed an experiment in which participants were presented with Arabic digits or sets of items (dots or squares). When participants had to pay attention to the number of items, we found that the reproduced duration increased as a function of the numerosity. Interestingly, when participants had to judge the shape of the stimuli, a similar effect was found, with large number of items reproduced longer than small sets of items. When participants were presented with Arabic digits, they had to either pay attention to the numerical value of the digits or to the font in which the digit was written. Supporting Rammsayer and Verner's results (2016), the reproduced duration increased as a function of the digit value only when participants paid attention to the numerical magnitude.

Overall, these findings indicate that non symbolic quantities (and stimulus size) are automatically processed while attentional resources are needed to extract and process the numerical value of a digit. This suggests that different mechanisms are involved in the number-time interaction, depending on the level at which the numerical information is represented.

4 | The larger the longer, but not all the time

4.1 Summary

In this experiment we investigated to which extent the numerical magnitude of a stimulus was automatically processed, when numerosity was symbolically or non-symbolically presented to the participants. We also investigated to which extent the number-time interaction was task instructions dependent when numerical information was symbolically conveyed. In their study, Rammsayer and Verner (2016) argue that numerical magnitude is not automatically processed and that only relative (not absolute) numerical magnitude interferes with perceived duration. We hypothesize that perceived duration did not increase as a function of the digit value in their experiment because participants were instructed to categorize the digit value as “small” or “large”. Our first prediction is that perceived duration should linearly increase as a function of the digit when participants have to pay attention to its exact value. Contrary to the digit value, Rammsayer and Verner (2016) found that the size of the stimulus was automatically processed and interfered with perceived duration. Our second prediction is that, if numerical magnitude is non-symbolically conveyed, it should be automatically processed and perceived duration should increase as a function of the tested numerosity.

The main task we used was a temporal reproduction task. Participants were instructed to reproduce the duration of presentation of a visual stimulus and to make judgements regarding the numerical magnitude of the stimulus, its shape or its font. We used 6 different experimental conditions: (1) Symbolic 2-AFC: participants had to reproduce the duration of presentation of a digit and report its magnitude (small or large). (2) Non-Symbolic 2-AFC: same as (1) but numerical magnitude consisted in sets of dots or squares instead of digits. (3) Symbolic 4-AFC: after the temporal reproduction task, participants had to report the exact value of the digit (4 possible propositions). (4) Non-symbolic 4-AFC: same as (3) but participants were

presented with sets of dots or squares instead of digits. (5) Font: participants were instructed to pay attention to the font in which the digits were written (2-AFC). (6) Shape: participants were instructed to pay attention to the shape of the stimuli (dots or squares, 2-AFC).

Results showed that at a non-symbolic level (Non-symbolic 2-AFC, 4-AFC and Shape conditions), larger numerosities were always reproduced longer. We also found that the reproduced duration increased linearly as a function of the numerosity, which is consistent with the hypothesis of a common metric (ATOM), and with our second prediction. However, at a symbolic level, the number-time interaction depended on task instructions. In the Symbolic 2-AFC condition, large digits were reproduced longer than small digits, but the reproduced duration did not increase as a function of the digit value. Results revealed a categorical response profile, as previously observed by Rammsayer & Verner (2016) and Cai and Wang (2014). Interestingly, the reproduced duration increased linearly as a function of the digit value in the Symbolic 4-AFC condition, when participants paid attention to the exact numerical magnitude, which is consistent with our first prediction. Finally, when participants paid attention to the font, numerical magnitude did not interfere with perceived duration, suggesting that the numerical magnitude was not processed.

In this study we found that the numerical magnitude was automatically processed at a non-symbolic level of representation and interfered with perceived duration beyond any attentional control. However, at a symbolic level, the numerical value of a digit was not automatically processed and the number-time interaction depended on the task instructions (2-AFC, 4-AFC) and on the amount of attentional resources dedicated to the processing of the digit value.

4.2 Materials & Methods

Participants. A total of 21 participants (9 males, 12 females, mean age 24.7 ± 4.3 y.o.) were tested. Three participants were removed from the study: 1 participant decided to stop the experiment after the first half (approximately 1 hour). 1 participant was removed due to poor performance (mean reproduced duration (MRD) for this participant was above 2 standard

deviation away from the mean of the group, in 4 of the 7 experimental conditions). 1 participant was removed due to technical problems (data were not saved in Shape and Font conditions). Hence, a total of 18 participants were effectively analyzed (7 males, 11 females, 24.6 ± 4.3 y.o.). All had normal or corrected-to-normal vision. Prior to the experiment, participants signed a written informed consent. The study was conducted in agreement with the Declaration of Helsinki (2008) and was approved by the Ethics Committee on Human Research at Neurospin (Gif-sur-Yvette, France). Participants were compensated for their participation.

Stimuli. The experiment was coded using Matlab 8.4 with Psychtoolbox (v 3.0.12). In the Symbolic conditions, visual stimuli were black Arabic digits; in the Non-symbolic condition, we used static clouds of non-overlapping dots or squares. The computer screen displayed a grey background ([128 128 128] in the RGB-code; 1024×768 pixels; 60 Hz refresh rate) and visual stimuli appeared within a black circle of 15.3° visual angle centered on the display. Stimuli were characterized by their duration (450, 750 or 1500 ms), their numerosity (1, 2, 8 or 9), their font (in the Symbolic conditions) or their shape (in the Non-symbolic conditions). In the control condition the number of items was kept constant, and participants were presented with a square made of four “#” (Figure 4.1). In Symbolic conditions, Arabic digits were presented at the center of the screen in two possible fonts, which were pseudo-randomized in each trial and for each participant (Font 1 was Helvetica, e.g. “1289”; Font 2 was Lucida Handwriting “1289”; both of them in their 40 pt format). In Non-symbolic conditions, visual stimuli consisted of dots or squares, and on a given trial participants were presented with one of the two possible shapes in a pseudo-randomized manner. To avoid any possible confound between numerosity and size, the cumulative surface filled by 1, 2, 8 or 9 dots (or squares) was kept constant (1.2 cm^2).

Procedure. Participants were seated in a quiet room approximately 60 cm away from the computer screen. In all experimental conditions, the main task (Figure 4.1) consisted in (i) *duration encoding* phase, during which participants were presented with a visual stimulus specifying a given duration, (ii) a *duration reproduction* phase, during which participants reproduced the duration of the time interval they perceived during *duration encoding*, and (iii) an *identification* phase, in which participants reported the perceived numerical magnitude, the

font or the shape of the stimuli. Figure 4.1 illustrates the training followed by the seven experimental conditions and the temporal sequence in a given trial. The experimental session started with a short training (Figure 4.1) in which the time interval was provided by the duration of letters (A, M or Z). Following the presentation of the letter, participants were prompted with a green circle indicating they could initiate their temporal reproduction by pressing the spacebar and terminate it by releasing the bar when they considered that the elapsed time was identical to the letter duration presented during the encoding phase. Following their temporal reproduction, participants were asked to report which letter was presented during the time interval by pressing one of the three possible response keys (“V”, “B” or “N”; Figure 4.1). During training, participants performed a total of 18 trials (3 letters \times 3 durations \times 2 repetitions). Results were visually inspected and if the participants could clearly distinguish that different durations were used, the experiment proper took place. 5 participants required a second training block. The main experiment was a within-participant pseudo-randomized block design. Participants were provided with the instructions and experimental condition before the start of each block. For a given experimental condition, trials were equally distributed in the first and second part of the experiment to prevent any confound with fatigue or attentional lapses (\sim 50 minutes each). There were thus 2 blocks for each experimental condition, yielding a total of 14 blocks (+ training) distributed over a 2 hours experimental session. Each experimental block included 2 breaks after which participants were free to press the spacebar to continue the experiment anytime; a 2 minutes break was also provided between two consecutive blocks and a longer break half-way in the experiment.

Experimental conditions. Following the short training, there were a total of 7 experimental conditions including one control and six dual-task conditions (Figure 1). In the encoding phase of the control condition (CONTROL), participants were presented with four symbols “#” forming a square lasting for one of each possible experimental duration (450, 750 or 1050ms). 12 repetitions of each duration were tested yielding a total of 36 trials for the CONTROL. In the six dual-task conditions, following their temporal reproduction, participants had to perform an n-Alternative-Forced-Choice (AFC) regarding the stimuli presented during the encoding phase. Specifically, in the *Symbolic* and *Non-symbolic Magnitude 2-AFC* conditions, participants

reported the numerical magnitude of the symbolic or non-symbolic stimulus presented during the encoding phase by classifying their responses as “small” or “large” (“V” and “B” on the keyboard, respectively). In the *Symbolic* and *Non-symbolic Magnitude 4-AFC* conditions: participants reported the numerical magnitude of the symbolic or non-symbolic stimuli presented during the encoding phase by classifying their responses as an exact numeral 1, 2, 8 or 9 (“C”, “V”, “B” and “N” on the keyboard, respectively). In the *Symbolic Magnitude Font (2-AFC)*, participants were presented with two strings of digits in the *identification* phase, written in the two possible fonts: “1289” or “1289”. They had to select in which font the digit was written in the *duration encoding* phase, using “V” and “B” keys, corresponding to “1289” and “1289”, respectively. In the *Non-Symbolic Magnitude Shape (2-AFC)*, participants reported the shape of the non-symbolic magnitude presented during the encoding phase by classifying their responses as “point(s)” or “carré(s)” (dot(s) and square(s) in English, respectively) (“V” and “B” on the keyboard, respectively). In all symbolic and non-symbolic conditions, 4 numerosities were thus tested (1, 2, 8 and 9) combined with each of the 3 possible durations (450, 750 or 1050ms). 12 repetitions of each possible combination were tested yielding a total of 144 trials for each of the six main experimental conditions.

Statistical Analyses. The MRD was calculated for each numerical value (1, 2, 8 and 9) and each duration (450, 750 and 1050ms) on a per participant and condition basis. A total of 12 MRD/participant/condition were thus obtained in the main experimental conditions, and 3 MRD/participant were computed in the CONTROL (one per duration). For a given participant, a reproduced duration 2 standard deviations away from its MRD was disregarded and replaced by its MRD. This procedure affected $7.2 \pm 4.1\%$ of the total number of trials across all conditions and individuals. At the group level, if the MRD of a participant was 2 standard deviations away from the group MRD, it was disregarded and replaced by the group MRD. This procedure affected $3.1 \pm 0.9\%$ of the values, across all conditions. Additionally, only MRD for which the correct numerical magnitude, font or shape was reported were considered in the analysis in order to insure that participants effectively paid attention to both temporal and non-temporal magnitudes. The percentage of errors was on average $5 \pm 4.8\%$ across all conditions. Repeated-

measures ANOVA and paired t-tests (Holm correction applied to all paired t-tests) were run using R (Version 3.2.2). For each tested numerosity (see Results section) MRD were plotted for each tested duration and were fitted with a linear regression. Y-intercept and slope values of the best fitting line (using Matlab “lsline” function) were extracted for each participant. Slope values closer to 1 indicated that participants reproduced the duration with perfect accuracy, whereas values closer to 0 indicated that participants’ MRD were the same for each tested duration.

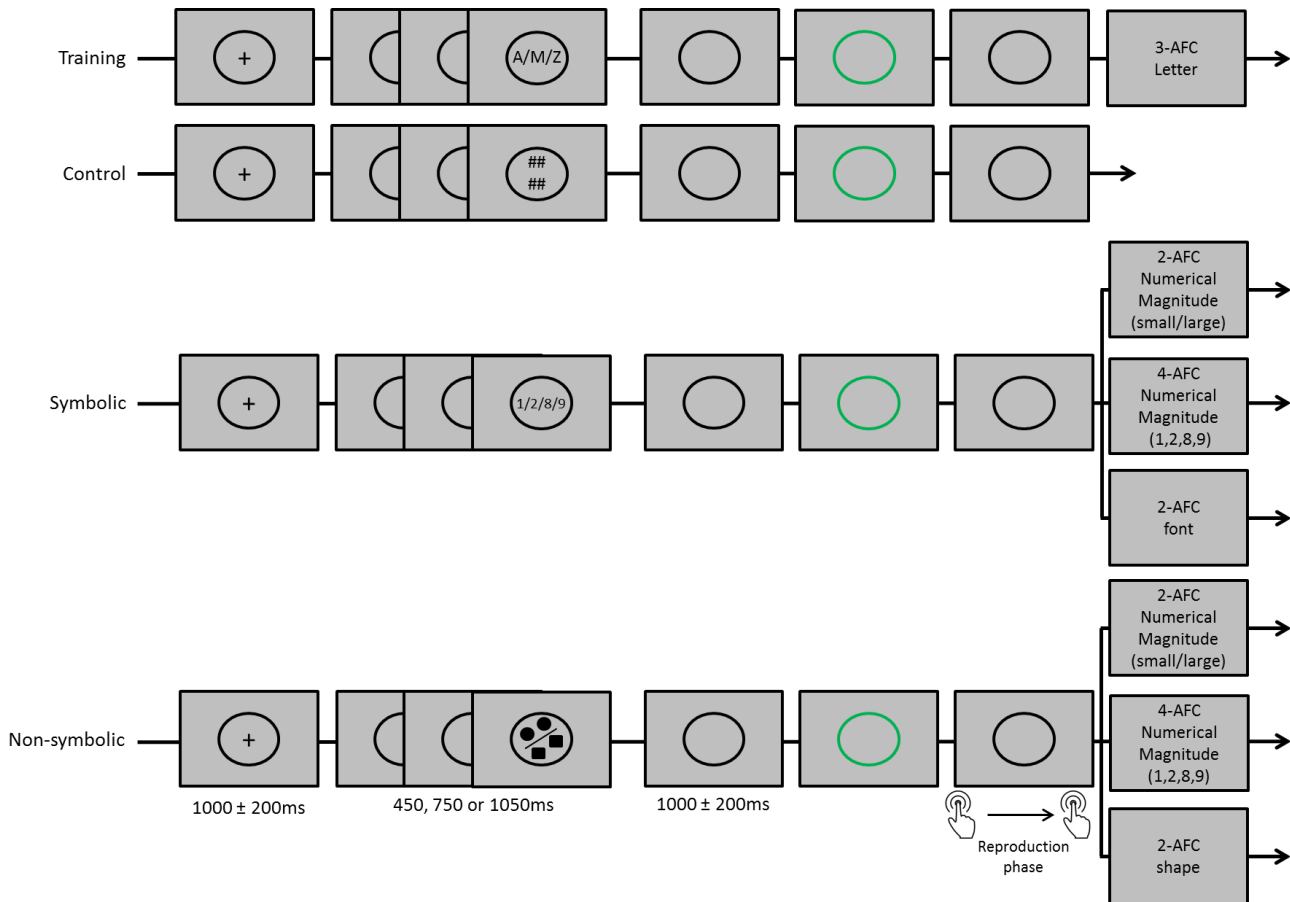


Figure 4.1: Experimental design. A trial started with the appearance of a fixation cross staying on the screen for $1000 \pm 200\text{ms}$. Following the disappearance of the fixation cross, the interval stimulus was presented for 450, 750 or 1050ms. After a delay of $1000 \pm 200\text{ms}$, the circle turned green indicated the go-signal for temporal reproduction. Participants initiated their temporal reproduction with a button press and terminated it with a release. In the CONTROL, participants only performed the temporal reproduction task. In all other dual-task conditions, participants reported the numerical magnitude, the shape, or the font (or the letter during training) following their temporal reproduction.

Control. Repeated-measures ANOVA with the MRD as dependent variable and Duration as within-subject factor (3: 450, 750, 1050ms) showed, as expected, a main effect of Duration on MRD ($F(2,38) = 132.6$, $p < 2e-16$) with MRD equal to 445 ± 100 , 696 ± 100 and 835 ± 155 ms, respectively (Figure 4.2). Results showed that participants could differentiate and reproduce the three different durations. Additionally, separate paired t-tests revealed a significant underestimation of the long duration compared to the ideal observer's performance ($p = 1.9e-6$), but no significant difference for the short duration. Typical regression to the mean profile was only observed for the long duration (Figure 4.2). Participants also significantly underestimated the 750ms duration ($p = .035$).

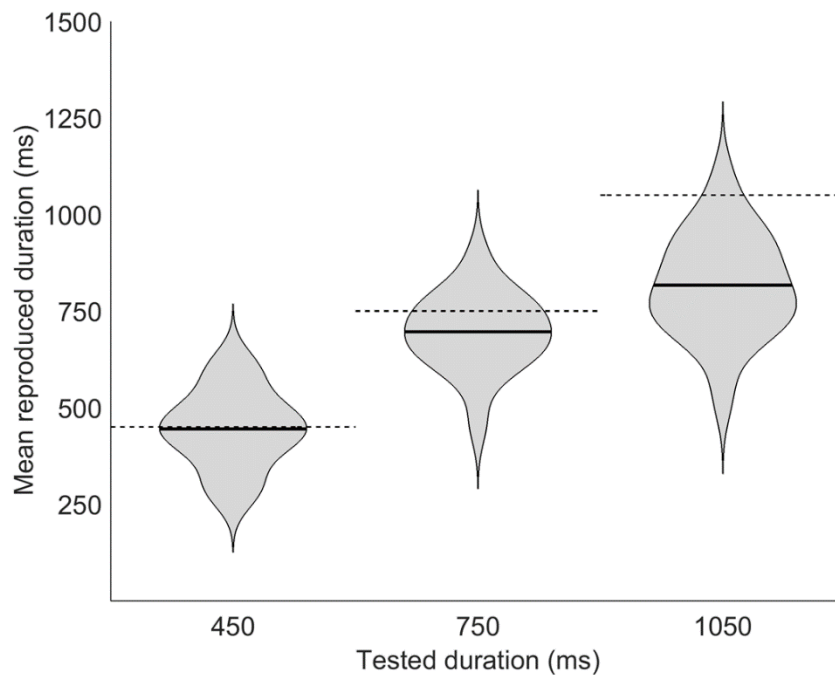


Figure 4.5: MRD in Control. Violin plots of the MRD as a function of the stimulus duration. For each tested duration, dashed lines represent the ideal observer's performance; horizontal black lines of each violin plot represent the mean; shaded gray represent the population distribution.

Large numerosities are reproduced longer than small numerosities.

To investigate the effect of numerical magnitude on perceived duration, we performed a $4 \times 3 \times 2 \times 2$ repeated-measures ANOVA with MRD as dependent variable and Numerosity (4: 1, 2, 8 and 9), Duration (3: 450, 750 and 1050ms), Format (2: Symbolic and Non-symbolic) and Decision (2: 2-AFC and 4-AFC) as within-subjects factors. Results revealed a main effect of Duration ($F[2,34] = 305.5$, $p < 2e-16$), Numerosity ($F[3,51] = 23.35$, $p = 1.19e-9$) and Format ($F[1,17] = 9.026$, $p = 0.008$). Results also revealed significant interaction between Duration and Decision ($F[2,34] = 7.52$, $p = .002$) and a significant interaction between Duration, Numerosity and Decision ($F[6,102] = 4.289$, $p = .0007$): On average, MRD were longer (shorter) in 2-AFC conditions compared to 4-AFC conditions for the 1050ms (450ms) duration. Paired t-tests revealed a significant difference only for the 1050ms duration ($p = 5.35e-6$), MRD were 842 ± 109 ms and 808 ± 111 ms for 2-AFC and 4-AFC decisions, respectively. Difference tended to be significant for the 450ms duration ($p = .06$), MRD were 463 ± 99 ms and 478 ± 117 ms for 2-AFC and 4-AFC decisions, respectively. Visual inspection of the interaction plot indicates that MRD were shorter in the 4-AFC conditions than in the 2-AFC conditions for numerosities 1, 2 and 8. On the other hand, MRD were longer for the numerosity 9 in the 4-AFC than in the 2-AFC conditions. However, separate paired t-tests performed on MRD for each numerosity as a function of the decision (2-AFC or 4-AFC) did not reveal any significant difference. MRD for numerosities 1, 2, 8 and 9 were: (1) 646 ± 188 ms and 639 ± 177 ms, (2) 650 ± 180 ms and 642 ± 175 ms, (8) 680 ± 198 ms and 673 ± 180 ms, (9) 680 ± 186 ms and 684 ± 170 ms, for 2-AFC and 4-AFC decisions, respectively. MRD for the numerosities 1, 2, 8 and 9 were 642 ± 158 ms, 647 ± 152 ms, 676 ± 150 ms and 682 ± 154 ms, respectively. Paired t-test were then performed between the four numerical values and results revealed significant different MRD for numerosities 1 vs. 8 ($p = 8.3e-10$), 1 vs. 9 ($p = 5.7e-13$), 2 vs. 8 ($p = 2.2e-7$) and 2 vs. 9 ($p = 2.2e-12$) (Figure 4.3 A). MRD for numerosities 1, 2, 8 and 9 were then compared to the CONTROL: paired t-tests revealed that the MRD for the control ($652 \text{ ms} \pm 98 \text{ ms}$) was significantly different than for experimental conditions displaying symbolic and non-symbolic numerosity 9 (682 ± 154 ms) only ($p = .026$).

across the four experimental conditions where participants paid attention to the numerical magnitude) duration. The slope and y-intercept values of the best fitting lines (four lines, one per numerosity) were extracted for each participant. A repeated-measures ANOVA with Intercept values as independent variable and Numerosity (4: 1, 2, 8 and 9) as within subject factor revealed a main effect of Numerosity on Intercept values ($F[3,51] = 17.76$, $p = 5.04e-8$). Post-hoc paired t-tests revealed that intercept value for the numerosity 1 was significantly lower than intercept value of numerosity 8 ($p = 4.2e-7$), 9 ($p = 1.6e-5$) and tended to be significantly lower than intercept value of numerosity 2 ($p = .052$). Intercept value for the numerosity 2 was significantly lower than for 8 ($p = 0.0132$) and 9 ($p = 0.0025$) (Figure 4.3B). Intercept values for numerosities 1, 2, 8 and 9 were $428 \pm 87\text{ms}$, $448 \pm 98\text{ms}$, $482 \pm 119\text{ms}$ and $480 \pm 114\text{ms}$, respectively. Slope analysis revealed no significant differences in slope between the different numerosities, indicating a similar regression to the mean. This result suggests that the differences in intercept values cannot be explained by slope variations and are only due to the magnitude of the tested numerosity. Slopes of the best fitting lines for numerosities 1, 2, 8 and 9 were 0.61 ± 0.18 , 0.59 ± 0.17 , 0.57 ± 0.19 and 0.59 ± 0.21 .

Longer MRD in Non-symbolic conditions than in Symbolic conditions

Previous repeated-measures ANOVA revealed a main effect of Format (Symbolic and Non-symbolic) on MRD. Paired t-tests revealed a significant difference in MRD between Non-Symbolic and Symbolic conditions (NS and S, respectively, in Figure 4.4) for the 450 and 750ms stimulus duration ($p = .0007$ and $p = .01$, respectively). Difference for the 1050ms duration tended to be significant ($p = .0509$). MRD were $484 \pm 115\text{ms}$ and $457 \pm 99\text{ms}$ (450ms stimulus duration), $697 \pm 110\text{ms}$ and $680 \pm 102\text{ms}$ (750ms stimulus duration), $832 \pm 108\text{ms}$ and $818 \pm 114\text{ms}$ (1050ms stimulus duration), in the Non symbolic and Symbolic conditions, respectively.

To further investigate the difference in MRD between Non-symbolic and Symbolic conditions, we performed post-hoc paired t-tests to compare MRD of the 4 tested numerosities, as a function of the Format. The numerosities 2, 8 and 9 were significantly reproduced longer (Figure 4.4 B) in the Non-symbolic conditions than in the Symbolic ones ($p = .013$, $p = .009$ and p

= .0011, respectively). MRD for numerosities 1, 2, 8 and 9 were (1) $645 \pm 184\text{ms}$ and $639 \pm 182\text{ms}$, (2) $655 \pm 175\text{ms}$ and $637 \pm 179\text{ms}$, (8) $687 \pm 181\text{ms}$ and $665 \pm 184\text{ms}$, (9) $697 \pm 183\text{ms}$ and $667 \pm 18\text{ms}$, in the Non-symbolic and Symbolic conditions, respectively.

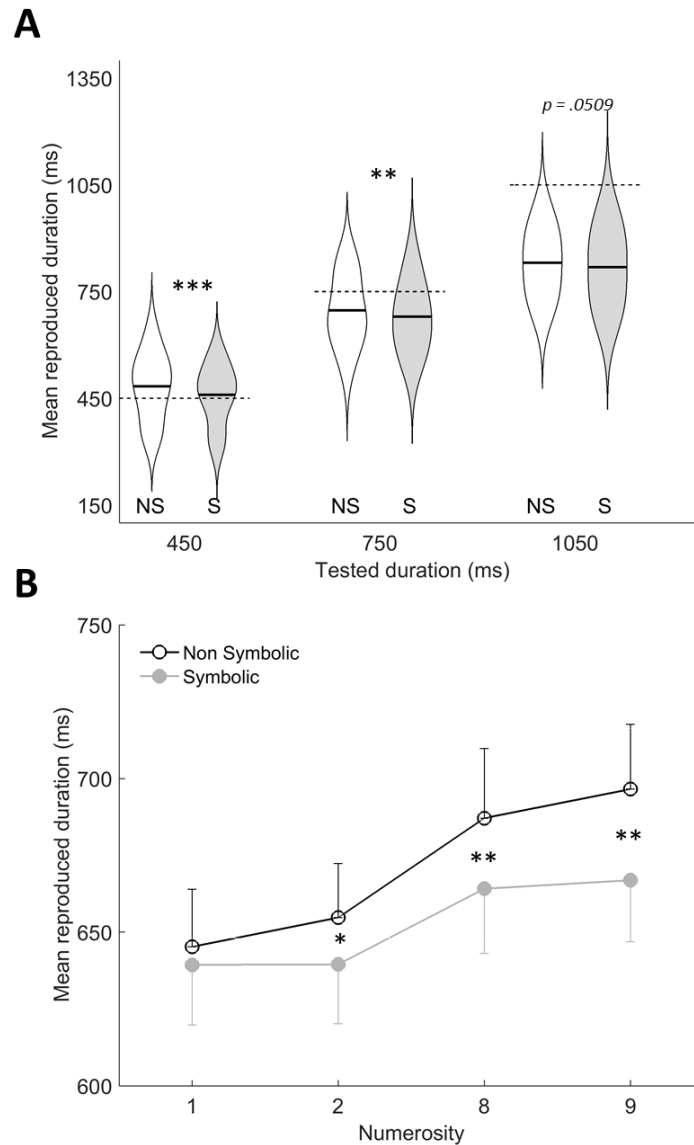


Figure 4.4: **A:** Violin plots of the MRD as a function of the stimulus duration, for Non-symbolic (NS) and Symbolic (S) conditions. For each tested duration, dashed lines represent the ideal observer's performance; horizontal black lines of each violin plot represent the mean; shaded gray represent the population distribution. **B:** MRD as a function of the tested numerosity in Non-symbolic (black) and Symbolic (grey) conditions. * $p < .05$, ** $p < .01$, *** $p < .001$, bars are 2 s.e.m.

Visual inspection of Figure 4.4 indicates that MRD increases as a function of the tested numerosity only in the Non-symbolic conditions, whereas a categorical response profile is observed in the Symbolic conditions. This latter point is inspected in the next section of the results.

MRD linearly increases as a function of the numerosity in Non-symbolic conditions only.

In Figure 4.5 A and B, MRD as a function of the 4 tested numerosities in the Non-Symbolic (left panel) and Symbolic (right) conditions are represented, for the 2 possible decisions (4-AFC, upper panels; 2-AFC, lower panels). MRD seems to increase linearly as a function of the tested numerosity in the Non-symbolic conditions (2-AFC and 4-AFC) and in the Symbolic 4-AFC condition only. In the Symbolic 2-AFC condition, a categorical response profile is observed. To further investigate possible different response profiles, we performed a linear regression in the four conditions using Matlab “fitlm” function, which creates a linear regression model. For each linear regression, the coefficient of determination R^2 is provided by the model (the better the linear regression fits the data, the closer the value R^2 is to 1). Results revealed that the linear regression did not fit to the data in the Symbolic 2-AFC condition ($R^2_{\text{adjusted}} = 0.485$), whereas it fitted to the data in the 3 other conditions (Symbolic 4-AFC: $R^2_{\text{adjusted}} = 0.907$; Non-Symbolic 2-AFC: $R^2_{\text{adjusted}} = 0.98$; Non-Symbolic 4-AFC: $R^2_{\text{adjusted}} = 0.828$). Overall, these findings indicate that the MRD in Symbolic conditions was task instructions dependent. When participants were instructed to categorize the magnitude of the digit as small or large, a categorical response profile in MRD was observed. However, when participants had to report the exact value of the digit, the MRD increased linearly as a function of the numerosity. In Non-symbolic conditions, similar response profiles were observed for both 2-AFC decision and 4-AFC decision. MRD increased as a function of the tested numerosity in the two cases, suggesting that the exact numerical magnitude was automatically processed and was not task instructions dependent. The next section investigates a possible automaticity in numerical processing in Non-symbolic conditions, compared to Symbolic conditions.

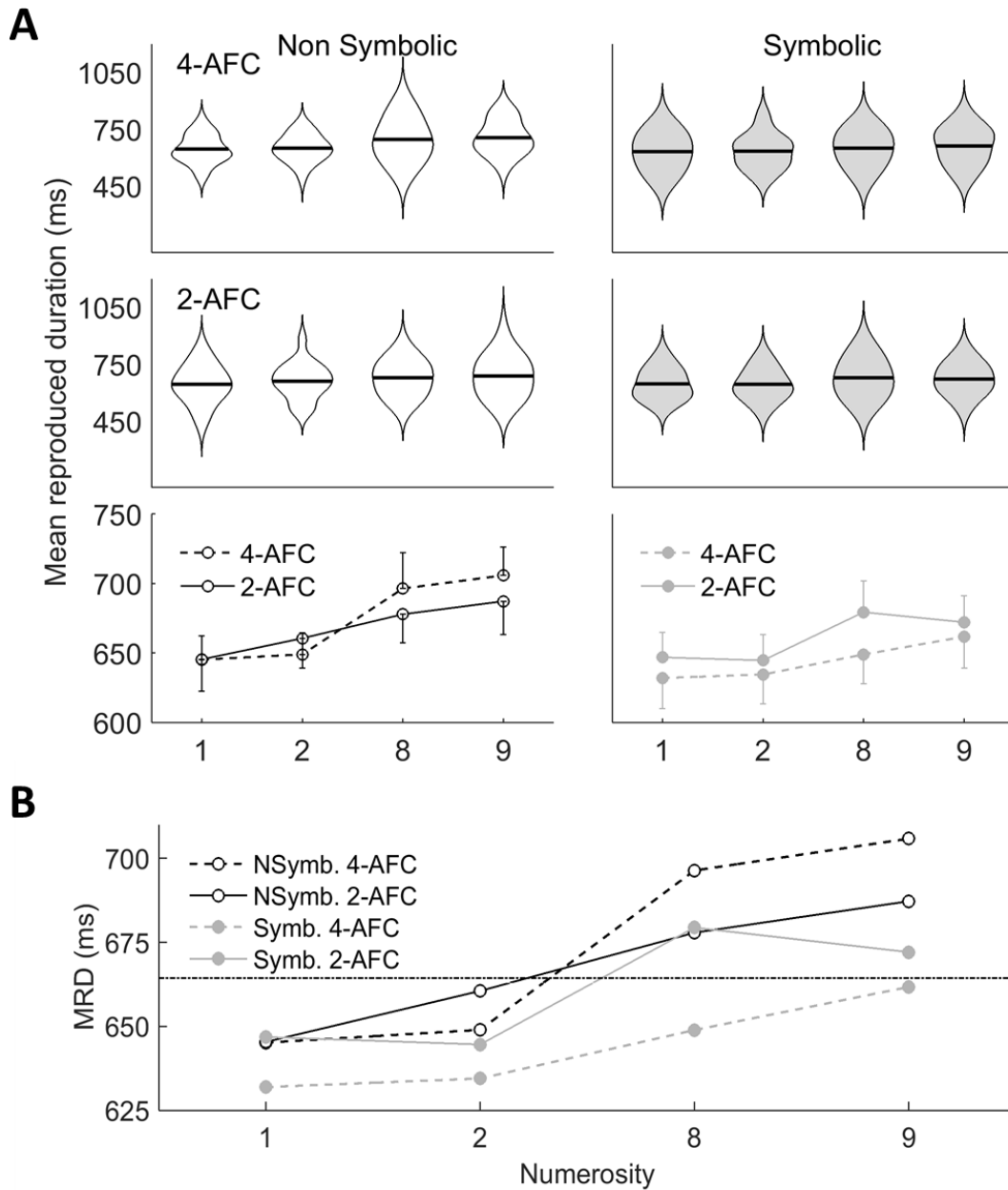


Figure 4.5: A: Violin plots of the MRD as a function of the 4 tested numerosities, for Non-symbolic (left panels) and Symbolic (right panels) conditions. Upper panels represent the MRD in the 4-AFC decision conditions, middle panels represent the MRD in the 2-AFC decision conditions. Lower panels combine the MRD for the two possible decisions (2-AFC and 4-AFC). Dashed lines represent the ideal observer's performance; horizontal black lines of each violin plot represent the mean; white and shaded gray represent the population distribution. **B:** MRD as a function of the 4 tested numerosities, in the 4 different conditions (error bars are not displayed for clarity reasons, see panel "A" if needed). Horizontal black line corresponds to the CONTROL.

Numerical magnitude is automatically processed in Non-symbolic conditions only

To further investigate to which extent the numerical magnitude was automatically processed we instructed the participants to pay attention to the shape (Non-symbolic condition) or to the font (Symbolic condition). If the numerical magnitude is automatically processed, the numerosity should interfere with MRD even if participants do not explicitly pay attention to it. We compared the effect of numerical magnitude when participants paid attention to it or not, in the Non-symbolic and Symbolic conditions, with two separated analyses.

Because 2 different shapes and 2 different fonts were used (2-AFC), we only included in the two analyses the conditions where participants had to categorize the numerical magnitude as small or large (2-AFC). In other words, we compared the results of the Font (or Shape) condition with the results of the Symbolic (Non-symbolic) Magnitude 2-AFC condition. For the Non-symbolic conditions (Attention on Shape vs. attention on Numerical magnitude): a $3 \times 4 \times 2$ repeated-measures ANOVA with MRD as independent variable and Duration (3: 450, 750 and 1050ms), Numerosity (4: 1, 2, 8 and 9) and Attention (2: Shape, Numerical magnitude) as within-subject factors revealed a main effect of Duration ($F[2,34] = 358.4, p < 2e-16$) and Numerosity ($F[3,51] = 9.39, p = 4.8e-5$) on MRD. No significant effect of Attention was found. A significant interaction between Duration and Numerosity was found ($F[6,102] = 3.679, p = .0024$). MRD increased for large numerosities only for 750 and 1050ms stimulus duration. 3 separate paired t-tests (one per stimulus duration) were performed to compare MRD as a function of the 4 tested numerosities. 450ms duration: paired t-tests revealed no significant difference in MRD between the 4 tested numerosities. 750ms duration: paired t-tests revealed significant differences in MRD between numerosities 1 vs. 8 ($666 \pm 120\text{ms}$ and $711 \pm 99\text{ms}$, respectively, $p = .00017$), 1 vs. 9 ($666 \pm 120\text{ms}$ and $711 \pm 120\text{ms}$, $p = .00576$), 2 vs. 8 ($670 \pm 106\text{ms}$ and $711 \pm 99\text{ms}$, respectively, $p = .016$) and 2 vs. 9 ($670 \pm 106\text{ms}$ and $711 \pm 120\text{ms}$, respectively, $p = .029$). 1050ms duration: paired t-tests revealed significant differences in MRD between numerosities 1 vs. 8 (MRD = $808 \pm 124\text{ms}$ and $838 \pm 111\text{ms}$, respectively, $p = .030$), 1 vs. 9 (MRD = $808 \pm 124\text{ms}$ and $876 \pm 118\text{ms}$, respectively, $p = .00035$), 2 vs. 9 (MRD = $821 \pm 101\text{ms}$ and $838 \pm 111\text{ms}$, respectively, $p = 1.3e-5$) and 8 vs. 9 (MRD = $838 \pm 111\text{ms}$ and $876 \pm 118\text{ms}$, respectively, $p =$

.014). MRD for the 4 different numerosities when participants paid attention to the numerical magnitude (Non-symbolic condition) or to the shape were respectively: (1) $645 \pm 192\text{ms}$ and $639 \pm 182\text{ms}$, (2) $661 \pm 180\text{ms}$ and $657 \pm 169\text{ms}$, (8) $678 \pm 188\text{ms}$ and $673 \pm 185\text{ms}$, (9) $687 \pm 207\text{ms}$ and $688 \pm 199\text{ms}$ (Figure 4.6, left panels).

Similar analysis has been performed in the Symbolic conditions (Font vs. Numerical magnitude) and the $3 \times 4 \times 2$ repeated-measures ANOVA revealed a significant effect of Duration ($F[2,34] = 327.5$, $p < 2e-16$), Numerosity ($F[3,51] = 11.1$, $p = 1.02e-5$) and Attention ($F[1,17] = 12.73$, $p = .0024$). Significant interactions were found between Duration and Attention ($F[2,34] = 4.137$, $p = .025$) and between Numerosity and Attention ($F[3,51] = 2.997$, $p = .039$). MRD were longer when participants paid attention to the numerical value compared to the font and were longer for large digits than for small digits. Paired t-tests revealed that MRD for numerosities 1 and 8 were significantly higher when participants paid attention to the numerical value (1) $647 \pm 187\text{ms}$ and $621 \pm 170\text{ms}$, $p = .03$; (8) $681 \pm 186\text{ms}$ and $627 \pm 182\text{ms}$, $p = 3.01e-6$). Numerosity 9 also tended to be significantly reproduced longer when participant paid attention to the numerical value (671 ± 190 and $646 \pm 181\text{ms}$, $p = .053$). No significant differences between MRD were found for numerosity 2 ($639 \pm 180\text{ms}$ and $626 \pm 177\text{ms}$, when participants paid attention to the numerical magnitude and the font, respectively (Figure 4.6, right panels).

Finally, to ensure that the difficulty to discriminate the two fonts and the two shapes was the same, we calculated the percentage of errors. Paired t-test revealed no significant difference in percentage of errors between the two conditions ($p = .68$). On average, participants made 4.82% and 4.09% of errors in the Font and Shape conditions, respectively. We previously observed that manipulating attention only interfered with MRD in the Symbolic conditions (Font vs. Numerical Magnitude), not in the Non-symbolic ones (Shape vs. Numerical magnitude). Because percentages of errors were similar, this asymmetry in the results when manipulating attention cannot be explained by different difficulties to differentiate the two shapes and the two fonts.

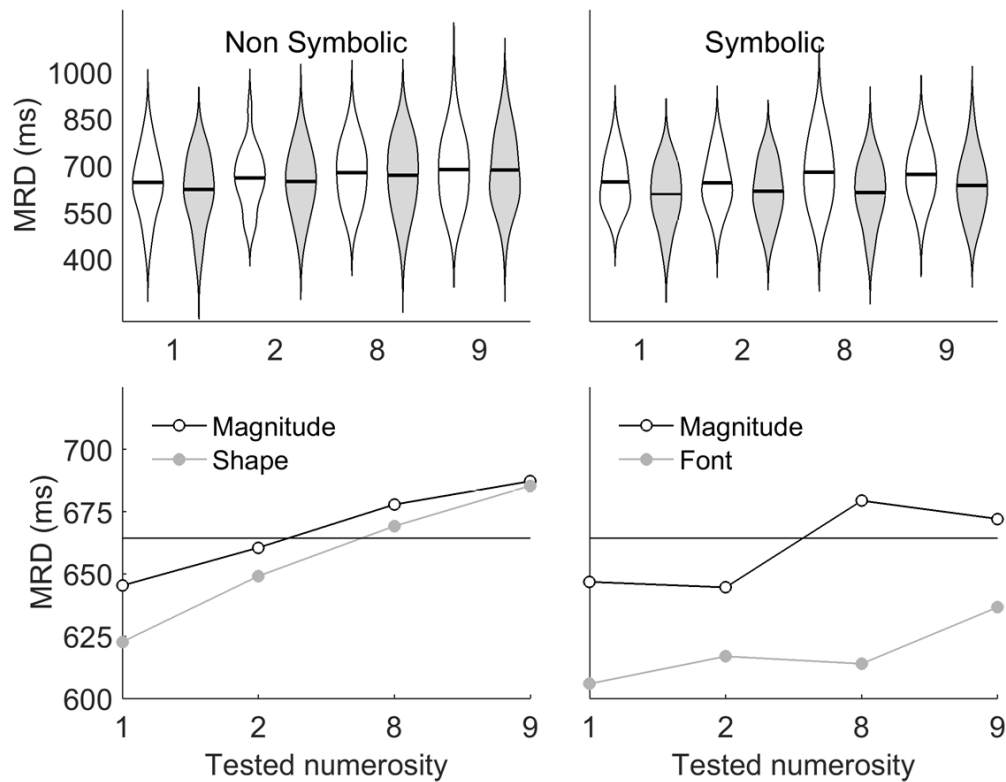


Figure 4.6: Upper panels: Violin plots of the MRD as a function of the 4 tested numerosities, when participants paid attention to the numerical magnitude or to the font/shape in Non-symbolic (left panels) and Symbolic (right panels) conditions. Horizontal black lines of each violin plot represent the mean; white and shaded gray represent the population distribution. **Lower panels:** MRD as a function of the 4 numerosities in Non-symbolic and Symbolic conditions, when participants paid attention to the numerical magnitude or to the font/shape. Error bars are not displayed for better clarity on the figure; see the population distribution in the violin plots if needed.

4.4 Discussion & Conclusion

The main goal of this experiment was to investigate to which extent numerical information was automatically processed when it was symbolically or non-symbolically presented, to better understand the number-time interaction. To account for behavioral and psychophysical interactions between different dimensions such as time and number, Walsh

(2003) suggested that temporal and non-temporal stimulus magnitude were part of a generalized magnitude system and share a common metric (Buetti and Walsh, 2009).

A growing body of studies (Dormal, Seron, and Pesenti, 2006; Dormal and Pesenti, 2007; Mo, 1971, 1974, 1975; Oliveri *et al.* 2008; Javadi and Aichelburg, 2012; Bottini and Casasanto, 2010; Casasanto and Boroditsky, 2008; Cai and Wang, 2014; Cai and Connell, 2015) that report behavioral interactions in the perception of temporal, numerical and spatial magnitudes support the hypothesis of a common magnitude system. However, the existence of such system has been challenged in recent studies (Lambrechts *et al.*, Martin *et al.*, Coull *et al.*) which revealed that perceived duration was resilient to numerical and spatial manipulations when sensory evidence accumulated over time. Moreover, it has been hypothesized that such magnitude system would be controlled by parietal cortex (Walsh, 2003; Buetti and Walsh, 2009; Winter, Marghetis, and Matlock, 2015). However, several studies showed that time, space and number engaged larger networks of regions beyond the parietal cortex, such as prefrontal cortex and SMA (Coull *et al.*, 2015; Dormal, Andres, and Pesenti, 2008; Dormal *et al.*, 2012). Recent studies (Marcos, Tsujimoto, and Genovesio, 2016; Genovesio, Tsujimoto, and Wise, 2012) suggest that different magnitudes are independently encoded before the decision process, and that interactions between magnitudes do not occur at the perceptual level but rather at the level of goal coding (decision). Such findings, and specifically the fact that magnitudes are independently encoded is not consistent with a theory of magnitude which argues that time, space and number should be encoded with the same common metric.

Several studies (Javadi and Aichelburg, 2012; Dormal, Seron, and Pesenti, 2006; Agrillo, Ranpura, and Butterworth, 2010; Lambrechts, Walsh, and van Wassenhove, 2013; Martin, Wiener, and van Wassenhove, 2017; Coull *et al.*, 2015; see also Chapter 3) revealed that perceived duration was biased by numerical and spatial magnitude, only when stimuli were statically (and not dynamically) presented. Recently, Rammsayer and Verner (2016) revealed that the number-time interaction seemed to be under attentional control, when participants had to pay attention to the numerical value of digits. However, the size of items interfered with perceived duration beyond attentional control, suggesting the existence of two different

mechanisms involved in the processing of numerical and spatial magnitudes. These authors also found that the perceived duration increased as a function of the relative numerical magnitude, i.e., along a dichotomical “small/large” distinction. However, we argued that the categorical response profile observed in this experiment was driven by task instructions (participants had to categorize the numerical value as “small” or “large”). Thereby, we hypothesized that perceived duration should linearly increase as a function of the digit value if participants have to report the exact digit value. We used Arabic digits and sets of items (dots or squares) to investigate possible differences in the automaticity of numerical processing, when numerical magnitude is symbolically or non-symbolically conveyed.

In our task, participants were instructed to pay attention to the numerical magnitude of the stimulus. Larger numerosities were found to be reproduced longer than numerosities of smaller magnitude, as expected. However, the pattern of results obtained in the different conditions tested here provided evidence against one only magnitude system underlying the effect of numerical magnitude on perceived duration. Whereas the processing of Non-symbolic numerical magnitude seemed to be automatic, the processing of Symbolic numerical magnitude seemed to require some attentional resources. Rammsayer and Verner (2016) showed that the size of a visual stimulus was automatically processed and interfered with perceived duration. To ensure that participants could not use the size as a cue to perform the numerosity task, we kept the cumulative surface constant. It is also known that perception of numerosity and density involves different mechanisms (Anobile, Cicchini, and Burr, 2014). However, we used a small number of dots and the density was too low to consider that participants could have used the density as a cue to perform the numerical task. Thereby, the effects observed in Non-symbolic conditions were only due to the processing of numerical magnitude.

In the Non-symbolic conditions, larger numerosities were always reproduced longer than smaller numerosities indicating that the number-time interaction takes place automatically and does not depend on the amount of attentional resources dedicated to the non-temporal attributes of the stimulus. This finding supports the assumption of Xuan *et al.* (2007) suggesting that the magnitude of a non-temporal stimulus does not need to be intentionally processed to interfere with duration perception. When sets of items (dots/squares) were presented to the

participants, the numerical information was automatically processed, which is consistent with a recent study from Rammsayer and Verner (2016) who suggested that stimulus size is processed *“beyond attentional control but still influences perceived duration”*. The same authors also indicated that the effect of numerical digit value on reproduced duration became effective only when participants had to explicitly pay attention to the numerical value. Our present findings also support this latter point. In the Symbolic conditions, when participants were required to make a 2-AFC judgement to classify the numerical magnitude, we observed a categorical response profile in the reproduced duration. Large digits (8 and 9) were reproduced longer than small digits (1 and 2). However, MRD for digits 1 and 2 were similar, as well as for MRD for digits 8 and 9. No linear increase of MRD as a function of digit value was observed. Other previous studies also reported the same profile of response. For example Chang *et al.*, (2011) reported that large digits (8 and 9) were reproduced longer than small digits (1 and 2). In these studies, low and high digits value were merged to small and large magnitude sets to perform the statistical analyses. To investigate a possible increase of MRD as a function of the digit value, Rammsayer and Verner (2016) also looked at the MRD for each single digit value. The results only revealed significant difference in MRD between digits 1-8 and 2-8. Visual inspection of Figure 2 (see Chapter 3) there results indicate that MRD did not increase as a function of the digit value. The authors concluded that the effect of numerical value on reproduced duration did not increase continuously as a function of digit value.

However, we previously hypothesized that the lack of linear increase of MRD as a function of digit value was due to the instructions given to the participants, in these two experiments (Chang *et al.* 2011; Rammsayer and Verner, 2016). The participants were instructed to report the magnitude of the digit as small or large, in other words they had to categorize the value of the digit as “small” or “large”, which can explain why the MRD did not increase continuously as a function of the digit value. To test this hypothesis, we designed a condition in which participants were instructed to make a 4-AFC judgment and report the exact numerical value of the digit. Firstly, we predicted that MRD should increase as a function of the digit when participants pay attention to its exact value. Secondly, we predicted that MRD will not increase linearly as a function of the digit when participants have to categorize the

numerical magnitude as “small” or “large”. Results confirmed our hypothesis. Whereas a categorical response profile was observed in the Symbolic 2-AFC condition, a linear regression better explained the data in the Symbolic 4-AFC condition. We argue that reproduced duration increases as a function of digit value, only when participants pay attention to the exact numerical value, in the Symbolic conditions only.

Interestingly, results revealed similar MRD in Non-symbolic conditions, when participants had to make 2-AFC or 4-AFC judgments, suggesting that numerical magnitude was automatically processed. On the other hand, the processing of Symbolic numerical information seems to be governed, at least partly, by an attentional control. To further investigate this possibility, we instructed the participants to pay attention to the font instead of paying attention to the numerical value, in Symbolic conditions. Whereas no differences were observed in the Non-symbolic conditions when manipulating attention (participants paid attention to the shape or the numerical magnitude), significant differences in MRD were observed in the Symbolic conditions. Interestingly, the effect of digit value on MRD was weak and almost non-existent when participants paid attention to the font, suggesting that the numerical magnitude was not processed, or not enough efficiently processed to interfere with perceived duration. Overall these results indicate that numerical information is automatically processed at a non-symbolic level, but is governed by attentional control at a symbolic level. The fact that manipulating attention only modulated MRD in the Symbolic conditions also suggest the existence of separated, at least partly, mechanisms underlying the interaction of numerical magnitude on perceived duration. The existence of different mechanisms is not compatible with a strong interpretation of ATOM. In their study, Rammsayer and Verner (2016) also concluded that two different mechanisms may underlie the influence of spatial and numerical magnitude on perceived duration. The existence of partly independent and different mechanisms for the processing of non-temporal magnitude is consistent with the results of a study from Agrillo and colleagues (2010), in which participants had to perform numerical or temporal judgments. Results showed that varying the number of tones (or the duration) did not influence duration (number) judgments, indicating that temporal and numerical estimations were independent. Additional evidence against the existence of a common magnitude system comes from studies

showing the absence of interference of non-temporal magnitude on temporal estimations (Lambrechts, Walsh, and van Wassenhove, 2013; Martin, Wiener, and van Wassenhove, 2017).

General conclusion

In this study, we investigated to which extent the numerical magnitude of a stimulus was automatically processed when performing a temporal reproduction task. At a non-symbolic level, results revealed that the reproduced duration increased as a function of the numerosity when: (1) participants had to categorize the numerical magnitude as small or large; (2) participants had to pay attention to the exact numerosity; and (3) when they paid attention to the shape of the stimulus. Overall, this indicates that the numerical magnitude of the stimulus was automatically processed when participants performed a temporal reproduction task. At a symbolic level, the number-time interaction highly depends on the amount of attentional resources and on the instructions provided to the participants. (1) When participants were instructed to categorize the value of the digit as small or large, a categorical response profile was observed in the reproduced duration; (2) when participants paid attention to the exact value of the digit, the reproduced duration increased continuously as a function of the digit value; and (3) when participants paid attention to the font, numerical value did not (or very weakly) interfere with the temporal reproduction task.

Future work is needed to better understand how numerical magnitude interferes with temporal estimations, and should specially focus on the different mechanisms underlying such interactions. Specifically, we found that non-symbolic numerical magnitude interfered more efficiently with the temporal reproduction task than the symbolic numerical magnitude did. Whereas the processing of non-symbolic numerical magnitude and duration seems to be more closely related than the processing of digit value and duration, future work should be done to clarify to which extent temporal and numerical magnitudes share common resources, depending on the numerical format (symbolic or non-symbolic) that is used.

5 | Discussion and Conclusion

5.1 A common neural code for Time, Space and Number?

One major prediction made by Walsh in his theory of magnitude was that time, space, number and, more generally, magnitudes were commonly processed in a general magnitude system which should be located in the parietal cortex. This claim was supported by several neuroimaging studies that reported that the IPS and surrounding areas were commonly activated when participants processed numerical or spatial magnitudes (Pinel *et al.*, 2004; Hubbard *et al.*, 2005). Coull and Nobre (1998) also found an increase of BOLD signal in the left IPS when performing a task which required paying attention to temporal cues. However, while number, size, and time exhibit common activations of the right parietal cortex, some studies have also shown that they engaged larger networks of regions, such as prefrontal cortex and supplementary motor areas (Coull *et al.*, 2015; Dormal, Andres, and Pesenti, 2008; Dormal and Pesenti, 2009). Several neuroimaging studies have suggested the existence of a domain-general representation of magnitude in a parieto-frontal network (Fias *et al.*, 2003; Pinel *et al.*, 2004; Walsh, 2003). It is possible to hypothesize that, because of the shared engagement of parietal and prefrontal cortices in temporal judgments and in magnitude processing, these regions may play an important role in magnitude interference effects (larger digits are reproduced longer than small ones). However, Dormal *et al.* (2008) found that numerosity and duration were processed independently. Supporting these findings, recent studies (Genovesio, Tsujimoto, and Wise, 2012; Marcos, Tsujimoto, and Genovesio, 2016) suggested that duration and relative distance were independently processed by different populations of neurons and that the common activations were due to the coding of the goal rather than a common representation of magnitude. These findings suggest that the interactions between magnitudes occur at the decision stage (i.e. decision about the magnitude of the stimulus) and not at a perceptual level

(i.e. encoding of the magnitude). This hypothesis does not support the existence of a generalized magnitude system in which different magnitudes will be processed and coded with the same neural code. To conclude, even if some neural populations are commonly activated for the processing of different magnitude dimensions, the network of activations is too broad to consider that only one brain area and only one mechanism underlie the representation of all magnitude dimensions. As there is no clear evidence from neuroimaging and clinical studies that Time, Space and Number may share neuronal populations with similar tuning features, it does not support a strong interpretation of ATOM.

5.2 From a Bayesian perspective, time, space and number do not share the same priors

As previously mentioned in the Introduction (Chapter 1) and in Chapter 2, recent discussions in the field suggest that in a common representational system, magnitude estimations would be realized on the basis of Bayesian computations. Regardless of whether participants are instructed to perform temporal, numerical or spatial judgements, some behavioral phenomena are systematically observed. One of these phenomena is the regression effect, also called regression to the mean, in which participants' estimates are biased towards the center of the distribution (Petzschner, Glasauer, and Stephan, 2015). From a Bayesian perspective, participants' estimates are made by taking into account at least 2 different variables: the prior and the likelihood. The prior corresponds to *a priori* information and is acquired over the course of the experiment; the likelihood corresponds to the noisy sensory input. A prior which encodes the mean of previously encountered stimuli (mean duration of a trial for example) will bias participants' estimates towards the center of the distribution and cause a regression to the mean. The strength of the bias is highly correlated to the noise of the prior and the likelihood. In Chapter 2, we found an increase in central tendency effect when participants performed spatial judgments, compared to temporal and numerical judgments, and suggested that a larger regression effect was linked to an increased reliance on the prior.

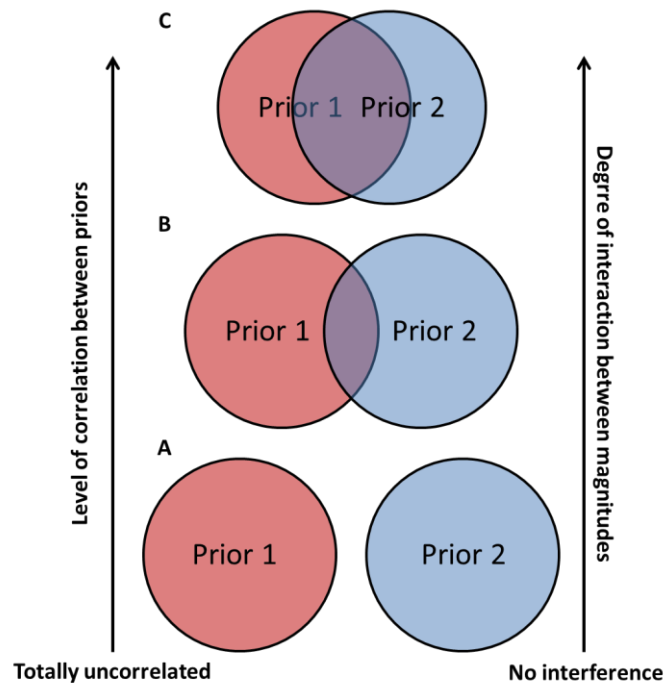


Figure 5.1: From a Bayesian perspective: **A:** if the priors of two magnitude dimensions are totally uncorrelated, the processing of one magnitude will be made independently from the processing of the second magnitude. **B, C:** increased correlation between priors predicts stronger interaction between magnitude dimensions

We concluded that a common magnitude system based on Bayesian computations would necessitate multiple priors that are more or less correlated. Pinel *et al.* (2004) found that the amount of overlap in brain regions could predict the amount of interference between dimensions. To extend this conclusion to a Bayesian perspective, a larger correlation between two priors may predict the amount of interaction between these two dimensions. If the priors of two distinct dimensions are totally uncorrelated, it is reasonable to think that these two dimensions are totally independent and will not interfere. However, two dimensions with highly

correlated priors will interact together, and the strength of the interaction will be linked to the degree of similarities between the different priors. In other words, each magnitude dimension has its own prior which is more or less correlated to other priors, with higher correlations associated to stronger interactions (Figure 5.1).

5.3 The rate of accumulation of sensory evidence interferes with Numerical and Spatial estimates

Whereas temporal information necessarily accumulates over time, numerical and spatial information do not. In Chapter 2 we used an experimental design in which task difficulty was equated across dimensions and we manipulated the rate of dot presentation to control numerical and spatial magnitude dimensions. It is important to remember that the ultimate value of the presented surface and number was the same, regardless of the accumulation regime. Participants were biased in their estimates by the rate of evidence accumulation in the first-half of the given trial, regardless of how long that trial lasted. Our main conclusion was that surface and number estimates were not influenced by the duration of the trial, but rather the duration interfered with the rate of accumulation, with shorter durations associated to a faster rate of accumulation which lead to an overestimation of numerical and spatial magnitudes. We also found that human observers are biased by the rate of accumulation at the start of a trial, and are resistant to changes in rate throughout the trial. This last observation is important in the context of drift-diffusion models in which the rate of evidence accumulation can change during the time course of the experiment. In such models, a second diffusion process with a new value of drift is considered when the rate of accumulation changes at a given time point (Ratcliff *et al.*, 2016; Ratcliff, 1980).

5.4 Perceived duration does not always increase as a function of the tested numerosity: no scaling effect

The hypothesis of a common metric implies that perceived duration should increase as a function of the tested numerosity. In other words, the larger the numerical magnitude, the larger the perceived duration should be, and the longer the duration, the larger the effect of numerosity (Figure 3.1, Chapter 3). However, such predictions are not always supported in the literature. For example, Yamamoto *et al.* (2016) found that the effect of numerical value on reproduced duration depended on the relative numerical distance, not on the absolute distance. Large digits were reproduced longer than small digits, but the difference in perceived duration did not increase as a function of the numerical distance between digits (Figure 3.3, Chapter 3). Similarly, Rammsayer and Verner (2016) reported that large digits were reproduced longer than digits of a smaller magnitude but did not observe a linear increase of reproduced duration as a function of the digit value. However, in the previous chapter, we saw that the effect of digit value on perceived duration was highly correlated to the task instructions. When participants were instructed to report the relative magnitude of a digit (small or large), we observed a categorical profile in the reproduced duration: larger digits were reproduced longer, but the effect of numerosity did not increase as a function of the digit value. Interestingly, we found a linear increase in the reproduced duration when participants had to pay attention to the exact numerical value of the digit. Supporting our results, Cai and Wang (2014) instructed the participants to reproduce the duration of presentation of a digit. Results showed that when the numerical magnitude was manipulated as a continuous variable using five digits: 1, 3, 5, 7 and 9, the reproduced duration increased linearly as a function of the digit value (Figure 5.2).

In Chapter 4, we also found that the numerical magnitude was automatically processed at a non-symbolic level of representation. To summarize, the perceived duration seems to increase as a function of the numerical value (scaling effect) when numerical magnitude is non-symbolically conveyed, or when participants pay attention to the exact value of the digit, at a symbolic level. Finally, Rammsayer and Verner (2016) found that the effect of stimulus size on reproduced duration was modulated by target duration, with larger differences in perceived

duration for long durations than for short ones, supporting the hypothesis of a scaling effect on perceived duration when the non-temporal magnitude is processed at a low level of representation.

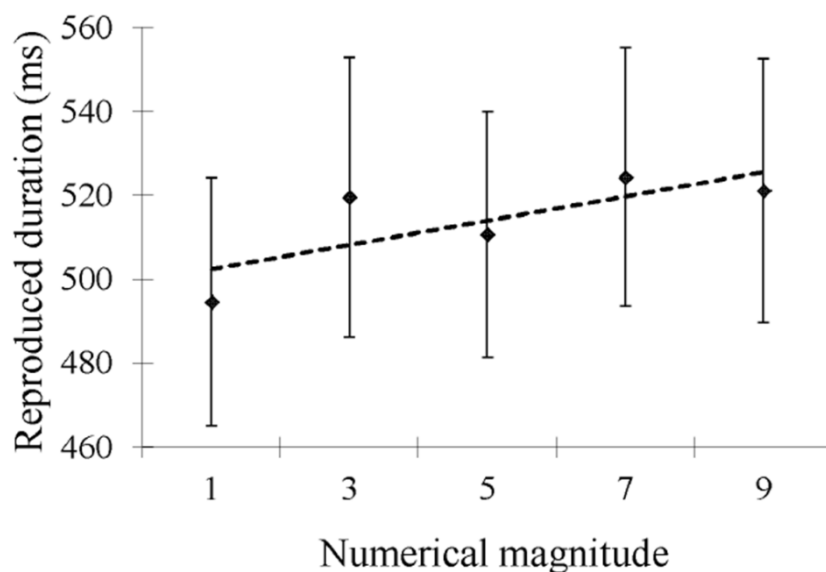


Figure 5.2: From Cai and Wang, 2014: the reproduced duration increases as a function of the numerical magnitude

5.5 Non-temporal magnitudes do not interfere with duration at a perceptual level

Several studies (Cai and Wang, 2014; Chang *et al.* 2011) revealed that the value of a digit interfered with reproduced duration when presented at the encoding stage, not during the reproduction stage. This finding suggests that the number-time interaction does not occur at a perceptual level, but at a higher level of representation. Some authors (Xuan *et al.*, 2007; Dehaene and Akhavein, 1995; de Hevia and Spelke, 2009; Dormal and Pesenti, 2007) suggested that the effect of numerical magnitude on temporal reproduction was automatic. However, we challenged the idea of an automatic processing of numerical magnitude in Chapter 4 (see also

Rammsayer and Verner, 2016). This indicates that the effect of numerosity on duration may not be as automatic as previously suggested, and may be mediated by different factors, such as numerical format and attention (see Chapter 4; Rammsayer and Verner, 2016).

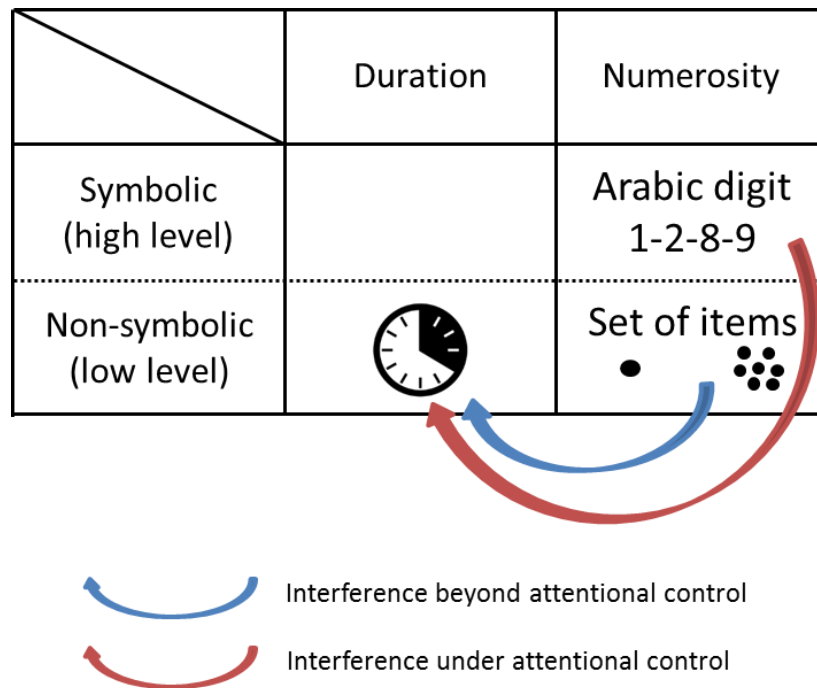


Figure 5.3: Illustration of the number-time interaction. At a symbolic level (red arrow) the number-time interaction is under attentional control, whereas it occurs automatically at a non-symbolic level (blue arrow)

5.6 The number-time interaction is modulated by attention and numerical format

As previously mentioned in Chapter 4, the effect of numerical magnitude on perceived duration depends (at least) on two distinct factors: (1) the numerical format that is used and (2) attentional resources. We saw that the number of items in a set (number of dots) was automatically processed, beyond attentional control, and interfered with reproduced duration.

However, the numerical value of a digit did not always interfere with perceived duration. The number-time interaction in this case depended on the amount of attentional resources allocated to the processing of the magnitude of the digit (Figure 5.3).

5.7 Conclusions

The main aim of this thesis was to investigate the possible existence of a generalized magnitude system, in which time, space and number may share common neural resources. One additional implication for the hypothesis of a common representational system for magnitudes is that the estimation of magnitude in a target dimension (e.g. the duration of the event) should be affected by the manipulation of the magnitude in another non-target dimension (e.g. the number of items or their size), such that the larger the magnitude of the non-target dimension, the larger the perceived target magnitude dimension. We first designed an experiment in which participants had to perform temporal, numerical or spatial judgments while we independently manipulated the magnitude of the non-target dimensions. This design allowed us to test all possible combinations and investigate possible interactions between magnitudes. The results revealed a lack of bi-directionality in the interactions, which do not support the theory of a common magnitude system. Furthermore, we found that numerical and spatial judgments were biased by the rate of accumulation of sensory evidence in dynamic designs.

In a second experiment, we found that numerical magnitude was automatically processed at a non-symbolic level, but not at a symbolic level. These findings suggest that at least two different mechanisms may underlie the influence of numerical magnitude on perceived duration, which is not consistent with the predictions made by ATOM. Additionally, the parietal cortex and specifically the IPS (Buetti and Walsh, 2009) has been suggested to be the main locus where time, space and number may share common populations of neurons. However, neural recording studies in the prefrontal and parietal cortex of non-human primates have revealed overlapping, yet largely separate, representations of duration, number and size (Coull *et al.*, 2015; Dormal *et al.*, 2012). The processing of these different magnitude dimensions may

share some neural resources but it is unlikely that only one specific and well localized brain area is responsible of the processing of time, space and number.

Whereas several authors previously suggested that numerical or spatial magnitude interfered automatically with perceived duration (Xuan *et al.*, 2007; Dormal and Pesenti, 2007), our recent findings (Chapter 4) suggest that the interaction between numerical and temporal magnitudes depends on the numerical format that is used (either non-symbolic or symbolic). On the one hand, numerical magnitude seems to be processed automatically at a non-symbolic level; on the other hand the processing of numerical digit value (symbolic level) seems to be under attentional control.

We know that humans rely on two different systems of numerical quantification. The first one is the non-symbolic system; it allows us to have approximate representations of number and is called the approximate number system (ANS) (Piazza *et al.*, 2013). The second one is the symbolic system that allows us to precisely manipulate quantities and perform exact calculations. Whereas the ANS is evolutionarily ancient and observed in adults, children and other animal species, the symbolic representation of quantities such as digits is specific to humans and has to be learnt (Feigenson, Dehaene, and Spelke, 2004; Piazza *et al.*, 2013). When children learn to count, they acquire a symbolic system to represent and manipulate digits. This system involves precise representation of quantities and allows precise comparisons. However, this new learnt system does not seem to replace the non-symbolic system, but they rather seem to be jointly mapped onto one another (Wong, Ho, and Tang, 2016).

It is possible to think that because this symbolic system is learnt, some attentional resources are needed for the numerical magnitude to interfere with perceived duration at a symbolic level. On the contrary, the non-symbolic system is deeply rooted and the number-time interaction can occur automatically in that case. Such hypothesis implies that time and number may share more similarities and neural resources when numerical information is non-symbolically conveyed, than when it is symbolically conveyed. Future lines of research, using neuroimaging techniques such as fMRI or M/EEG should focus on the neural correlates underpinning the number-time interaction, depending on the numerical format that is used, to further understand the nature and the locus / loci of this interference, as well as the temporal

dynamics of these number-time interactions. Furthermore, the use of M/EEG techniques, with a good temporal resolution could allow us to precisely identify at which moment in the process of magnitude estimation the interaction takes place (either at the encoding or at the decision stage). According to previous studies (Genovesio, Tsujimoto, and Wise, 2012; Marcos, Tsujimoto, and Genovesio, 2016; Chang *et al.*, 2011; Rammsayer and Verner, 2015), my guess would be that such interaction will not occur at the perceptual / encoding stage but rather at the decision stage.

Bibliography

- Agrillo, Christian, Ashish Ranpura, and Brian Butterworth. 2010. "Time and Numerosity Estimation Are Independent: Behavioral Evidence for Two Different Systems Using a Conflict Paradigm." *Cognitive Neuroscience* 1 (41): 96–101. doi:10.1080/17588921003632537.
- Anobile, Giovanni, Guido Marco Cicchini, and David C. Burr. 2014. "Separate Mechanisms for Perception of Numerosity and Density." *Psychological Science* 25 (1): 265–70. doi:10.1177/0956797613501520.
- Benton, A.L. 1992. "Gerstmann's Syndrome." *Archives of Neurology* 49 (5): 445–47. doi:10.1001/archneur.1992.00530290027007.
- Bonato, Mario, Marco Zorzi, and Carlo Umiltà. 2012. "When Time Is Space: Evidence for a Mental Time Line." *Neuroscience & Biobehavioral Reviews* 36 (10): 2257–73. doi:10.1016/j.neubiorev.2012.08.007.
- Bonn, Cory D., and Jessica F. Cantlon. 2012. "The Origins and Structure of Quantitative Concepts." *Cognitive Neuropsychology* 29 (1–2): 149–73. doi:10.1080/02643294.2012.707122.
- Bottini, Roberto, and Daniel Casasanto. 2010. "Implicit Spatial Length Modulates Time Estimates, But Not Vice Versa." Edited by Christoph Hölscher, Thomas F. Shipley, Marta Olivetti Belardinelli, John A. Bateman, and Nora S. Newcombe. *Spatial Cognition VII, Lecture Notes in Computer Science*, no. 09 (August): 152–62. doi:10.1007/978-3-642-14749-4_15.
- Bueti, Domenica, and Vincent Walsh. 2009. "The Parietal Cortex and the Representation of Time, Space, Number and Other Magnitudes." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 364 (30): 1831–40. doi:10.1098/rstb.2009.0028.
- Burr, David C., John Ross, Paola Binda, and M. Concetta Morrone. 2010. "Saccades Compress Space, Time and Number." *Trends in Cognitive Sciences*, Special Issue: Space, Time and Number, 14 (12): 528–33. doi:10.1016/j.tics.2010.09.005.
- Butterworth, Brian, Sashank Varma, and Diana Laurillard. 2011. "Dyscalculia: From Brain to Education." *Science* 332 (6033): 1049–53. doi:10.1126/science.1201536.

- Cai, Zhenguang G., and Louise Connell. 2015. "Space–time Interdependence: Evidence against Asymmetric Mapping between Time and Space." *Cognition* 136 (March): 268–81. doi:10.1016/j.cognition.2014.11.039.
- Cai, Zhenguang G., and Ruiming Wang. 2014. "Numerical Magnitude Affects Temporal Memories but Not Time Encoding." *PLOS ONE* 9 (16): e83159. doi:10.1371/journal.pone.0083159.
- Cantlon, Jessica F., Sara Cordes, Melissa E. Libertus, and Elizabeth M. Brannon. 2009. "Comment on 'Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures.'" *Science (New York, N.Y.)* 323 (5910): 38; author reply 38. doi:10.1126/science.1164773.
- Casasanto, Daniel, and Lera Boroditsky. 2008. "Time in the Mind: Using Space to Think about Time." *Cognition* 106 (07): 579–93. doi:10.1016/j.cognition.2007.03.004.
- Chang, Acer Y.-C., Ovid J. L. Tzeng, Daisy L. Hung, and Denise H. Wu. 2011. "Big Time Is Not Always Long: Numerical Magnitude Automatically Affects Time Reproduction." *Psychological Science* 22 (12): 1567–73. doi:10.1177/0956797611418837.
- Clayton, Nicky, Nathan Emery & Dickinson, and Anthony. 2006. "The Rationality of Animal Memory: Complex Caching Strategies of Western Scrub Jays." In *Rational Animals?*, edited by Susan Hurley and Matthew Nudds. Oxford University Press.
- Coull, Jennifer T., Pom Charras, Maxime Donadieu, Sylvie Droit-Volet, and Franck Vidal. 2015. "SMA Selectively Codes the Active Accumulation of Temporal, Not Spatial, Magnitude." *Journal of Cognitive Neuroscience* 27 (37): 2281–98. doi:10.1162/jocn_a_00854.
- Coull, Jennifer T., and Anna C. Nobre. 1998. "Where and When to Pay Attention: The Neural Systems for Directing Attention to Spatial Locations and to Time Intervals as Revealed by Both PET and fMRI." *Journal of Neuroscience* 18 (18): 7426–35.
- Dehaene, Stanislas, and Rokny Akhavein. 1995. "Attention, Automaticity, and Levels of Representation in Number Processing." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21 (2): 314–26. doi:10.1037/0278-7393.21.2.314.
- Dehaene, Stanislas, Véronique Izard, Elizabeth Spelke, and Pierre Pica. 2008. "Log or Linear? Distinct Intuitions of the Number Scale in Western and Amazonian Indigene Cultures." *Science (New York, N.y.)* 320 (5880): 1217–20. doi:10.1126/science.1156540.
- Doricchi, Fabrizio, Paola Guariglia, Marina Gasparini, and Francesco Tomaiuolo. 2005. "Dissociation between Physical and Mental Number Line Bisection in Right Hemisphere Brain Damage." *Nature Neuroscience* 8 (12): 1663–65. doi:http://dx.doi.org/10.1038/nn1563.

- Dormal, Valérie, Michael Andres, and Mauro Pesenti. 2008. "Dissociation of Numerosity and Duration Processing in the Left Intraparietal Sulcus: A Transcranial Magnetic Stimulation Study." *Cortex*, Special Issue on Numbers, Space, and Action, 44 (53): 462–69. doi:10.1016/j.cortex.2007.08.011.
- Dormal, Valérie, Giulia Dormal, Frédéric Joassin, and Mauro Pesenti. 2012. "A Common Right Fronto-Parietal Network for Numerosity and Duration Processing: An fMRI Study." *Human Brain Mapping* 33 (31): 1490–1501. doi:10.1002/hbm.21300.
- Dormal, Valérie, and Mauro Pesenti. 2007. "Numerosity-Length Interference." *Experimental Psychology* 54 (11): 289–97. doi:10.1027/1618-3169.54.4.289.
- . 2009. "Common and Specific Contributions of the Intraparietal Sulci to Numerosity and Length Processing." *Human Brain Mapping* 30 (54): 2466–76. doi:10.1002/hbm.20677.
- Dormal, Valérie, Xavier Seron, and Mauro Pesenti. 2006. "Numerosity-Duration Interference: A Stroop Experiment." *Acta Psychologica* 121 (06): 109–24. doi:10.1016/j.actpsy.2005.06.003.
- Droit-Volet, Sylvie, Angélique Clément, and Michel Fayol. 2003. "Time and Number Discrimination in a Bisection Task with a Sequence of Stimuli: A Developmental Approach." *Journal of Experimental Child Psychology* 84 (1): 63–76. doi:10.1016/S0022-0965(02)00180-7.
- Feigenson, Lisa, Stanislas Dehaene, and Elizabeth Spelke. 2004. "Core Systems of Number." *Trends in Cognitive Sciences* 8 (7): 307–14. doi:10.1016/j.tics.2004.05.002.
- Fias, Wim, Jan Lammertyn, Bert Reynvoet, Patrick Dupont, and Guy A. Orban. 2003. "Parietal Representation of Symbolic and Nonsymbolic Magnitude." *Journal of Cognitive Neuroscience* 15 (1): 47–56. doi:10.1162/089892903321107819.
- Gallistel, C. R. 2011. "Mental Magnitudes." Edited by Stanislas Dehaene and Elizabeth Brannon. *Space, Time and Number in the Brain*, no. 03: 3–12.
- Genovesio, Aldo, Satoshi Tsujimoto, and Steven P. Wise. 2012. "Encoding Goals but Not Abstract Magnitude in the Primate Prefrontal Cortex." *Neuron* 74 (4): 656–62. doi:10.1016/j.neuron.2012.02.023.
- Green, David Marvin, and John A. Swets. 1989. *Signal Detection Theory and Psychophysics*. Los Altos Hills, Calif: Peninsula Pub.

- Henik, Avishai, and Joseph Tzelgov. 1982. "Is Three Greater than Five: The Relation between Physical and Semantic Size in Comparison Tasks." *Memory & Cognition* 10 (4): 389–95. doi:10.3758/BF03202431.
- Hevia, Maria-Dolores de, and Elizabeth S. Spelke. 2009. "Spontaneous Mapping of Number and Space in Adults and Young Children." *Cognition* 110 (2): 198–207. doi:10.1016/j.cognition.2008.11.003.
- Hubbard, Edward M., Manuela Piazza, Philippe Pinel, and Stanislas Dehaene. 2005. "Interactions between Number and Space in Parietal Cortex." *Nature Reviews Neuroscience* 6 (34): 435–48. doi:10.1038/nrn1684.
- Hurewitz, Felicia, Rochel Gelman, and Brian Schnitzer. 2006. "Sometimes Area Counts More than Number." *Proceedings of the National Academy of Sciences* 103 (51): 19599–604. doi:10.1073/pnas.0609485103.
- Javadi, Amir Homayoun, and Clarisse Aichelburg. 2012. "When Time and Numerosity Interfere: The Longer the More, and the More the Longer." *PloS One* 7 (12): e41496. doi:10.1371/journal.pone.0041496.
- Kaufmann, Liane, Stephan E. Vogel, Guilherme Wood, Christian Kremser, Michael Schocke, Lothar-Bernd Zimmerhackl, and Jan W. Koten. 2008. "A Developmental fMRI Study of Nonsymbolic Numerical and Spatial Processing." *Cortex*, Special Issue on Numbers, Space, and Action, 44 (4): 376–85. doi:10.1016/j.cortex.2007.08.003.
- Lambrechts, Anna, Vincent Walsh, and Virginie van Wassenhove. 2013. "Evidence Accumulation in the Magnitude System." *PLoS ONE* 8 (13): e82122. doi:10.1371/journal.pone.0082122.
- Leon, Matthew I, and Michael N Shadlen. 2003. "Representation of Time by Neurons in the Posterior Parietal Cortex of the Macaque." *Neuron* 38 (2): 317–27. doi:10.1016/S0896-6273(03)00185-5.
- Lu, Aitao, Bert Hodges, Jijia Zhang, and John X. Zhang. 2009. "Contextual Effects on Number–time Interaction." *Cognition* 113 (1): 117–22. doi:10.1016/j.cognition.2009.07.001.
- Marcos, Encarni, Satoshi Tsujimoto, and Aldo Genovesio. 2016. "Independent Coding of Absolute Duration and Distance Magnitudes in the Prefrontal Cortex." *Journal of Neurophysiology*, no. 50 (October): jn.00245.2016. doi:10.1152/jn.00245.2016.
- Martin, Benoît, Martin Wiener, and Virginie van Wassenhove. 2017. "A Bayesian Perspective on Accumulation in the Magnitude System." *Scientific Reports* 7 (1): 630. doi:10.1038/s41598-017-00680-0.

- Mo, Suchoon S. 1971. "Judgment of Temporal Duration as a Function of Numerosity." *Psychonomic Science* 24 (2): 71–72. doi:10.3758/BF03337898.
- . 1974. "Comparative Judgment of Temporal Duration as a Function of Numerosity." *Bulletin of the Psychonomic Society* 3 (5): 377–79. doi:10.3758/BF03333504.
- . 1975. "Temporal Reproduction of Duration as a Function of Numerosity." *Bulletin of the Psychonomic Society* 5 (2): 165–67. doi:10.3758/BF03333236.
- Oliveri, Massimiliano, Carmelo Mario Vicario, Silvia Salerno, Giacomo Koch, Patrizia Turriziani, Renata Mangano, Gaetana Chillemi, and Carlo Caltagirone. 2008. "Perceiving Numbers Alters Time Perception." *Neuroscience Letters* 438 (08): 308–11. doi:10.1016/j.neulet.2008.04.051.
- Peterson, W., T. Birdsall, and W. Fox. 1954. "The Theory of Signal Detectability." *Transactions of the IRE Professional Group on Information Theory* 4 (4): 171–212. doi:10.1109/TIT.1954.1057460.
- Petzschner, Frederike H., Stefan Glasauer, and Klaas E. Stephan. 2015. "A Bayesian Perspective on Magnitude Estimation." *Trends in Cognitive Sciences* 19 (04): 285–93. doi:10.1016/j.tics.2015.03.002.
- Piazza, Manuela, Pierre Pica, Véronique Izard, Elizabeth S. Spelke, and Stanislas Dehaene. 2013. "Education Enhances the Acuity of the Nonverbal Approximate Number System." *Psychological Science* 24 (6): 1037–43. doi:10.1177/0956797612464057.
- Pinel, Philippe, Manuela Piazza, Denis Le Bihan, and Stanislas Dehaene. 2004. "Distributed and Overlapping Cerebral Representations of Number, Size, and Luminance during Comparative Judgments." *Neuron* 41 (32): 983–93.
- Rammsayer, Thomas H., and Martin Verner. 2014. "The Effect of Nontemporal Stimulus Size on Perceived Duration as Assessed by the Method of Reproduction." *Journal of Vision* 14 (24): 17–17. doi:10.1167/14.5.17.
- . 2015. "Larger Visual Stimuli Are Perceived to Last Longer from Time to Time: The Internal Clock Is Not Affected by Nontemporal Visual Stimulus Size." *Journal of Vision* 15 (25): 5–5. doi:10.1167/15.3.5.
- . 2016. "Evidence for Different Processes Involved in the Effects of Nontemporal Stimulus Size and Numerical Digit Value on Duration Judgments." *Journal of Vision* 16 (45): 13–13. doi:10.1167/16.7.13.

- Ratcliff, Roger. 1980. "A Note on Modeling Accumulation of Information When the Rate of Accumulation Changes over Time." *Journal of Mathematical Psychology* 21 (57): 178–84. doi:10.1016/0022-2496(80)90006-1.
- Ratcliff, Roger, Philip L. Smith, Scott D. Brown, and Gail McKoon. 2016. "Diffusion Decision Model: Current Issues and History." *Trends in Cognitive Sciences* 20 (58): 260–81. doi:10.1016/j.tics.2016.01.007.
- Rossetti, Y., S. Jacquin-Courtois, Aiello Marilena, M. Ishihara, C. Brozzoli, and Doricchi Fabrizio. 2011. *Neglect "Around the Clock": Dissociating Number and Spatial Neglect in Right Brain Damage*. Elsevier Inc. doi:10.1016/b978-0-12-385948-8.00011-6.
- Saj, Arnaud, Orly Fuhrman, Patrik Vuilleumier, and Lera Boroditsky. 2014. "Patients With Left Spatial Neglect Also Neglect the 'Left Side' of Time." *Psychological Science* 25 (1): 207–14. doi:10.1177/0956797612475222.
- Sandrini, Marco, and Elena Rusconi. 2009. "A Brain for Numbers." *Cortex* 45 (7): 796–803. doi:10.1016/j.cortex.2008.09.002.
- Sawamura, Hiromasa, Keisetsu Shima, and Jun Tanji. 2002. "Numerical Representation for Action in the Parietal Cortex of the Monkey." *Nature* 415 (6874): 918–22. doi:10.1038/415918a.
- Schenkenberg, T., D. C. Bradford, and E. T. Ajax. 1980. "Line Bisection and Unilateral Visual Neglect in Patients with Neurologic Impairment." *Neurology* 30 (5): 509–509. doi:10.1212/WNL.30.5.509.
- Shi, Zhuanghua, Russell M. Church, and Warren H. Meck. 2013. "Bayesian Optimization of Time Perception." *Trends in Cognitive Sciences* 17 (40): 556–64. doi:10.1016/j.tics.2013.09.009.
- Stein, jf. 1989. "Representation of Egocentric Space in the Posterior Parietal Cortex." *Quarterly Journal of Experimental Physiology (Cambridge, England)* 74 (5): 583–606. doi:10.1113/expphysiol.1989.sp003314.
- Tudusciuc, Oana, and Andreas Nieder. 2007. "Neuronal Population Coding of Continuous and Discrete Quantity in the Primate Posterior Parietal Cortex." *Proceedings of the National Academy of Sciences* 104 (52): 14513–18. doi:10.1073/pnas.0705495104.
- Vuilleumier, Patrik, Stéphanie Ortigue, and Peter Brugger. 2004. "The Number Space and Neglect." *Cortex* 40 (2): 399–410. doi:10.1016/S0010-9452(08)70134-5.

- Walsh, Vincent. 2003. "A Theory of Magnitude: Common Cortical Metrics of Time, Space and Quantity." *Trends in Cognitive Sciences* 7 (02): 483–88.
- Weber E.H (1850): Der Tastsinn Und Das Gemeingefühl - Inhalt." 2017. Accessed August 1. <http://psychologie.biphaps.uni-leipzig.de/wundt/opera/ehweber/tastsinn/Tinhalt.htm>.
- Winter, Bodo, Tyler Marghetis, and Teenie Matlock. 2015. "Of Magnitudes and Metaphors: Explaining Cognitive Interactions between Space, Time, and Number." *Cortex* 64 (28): 209–24. doi:10.1016/j.cortex.2014.10.015.
- Wong, Terry Tin-Yau, Connie Suk-Han Ho, and Joey Tang. 2016. "The Relation between ANS and Symbolic Arithmetic Skills: The Mediating Role of Number-Numerosity Mappings." *Contemporary Educational Psychology* 46 (July): 208–17. doi:10.1016/j.cedpsych.2016.06.003.
- Xuan, Bin, Daren Zhang, Sheng He, and Xiangchuan Chen. 2007. "Larger Stimuli Are Judged to Last Longer." *Journal of Vision* 7 (05): 2.1-5. doi:10.1167/7.10.2.
- Yamamoto, Kentaro, Kyoshiro Sasaki, and Katsumi Watanabe. 2016. "The Number–time Interaction Depends on Relative Magnitude in the Suprasecond Range." *Cognitive Processing* 17 (15): 59–65. doi:10.1007/s10339-015-0744-3.
- Zorzi, Marco, Konstantinos Priftis, and Carlo Umiltà. 2002. "Brain Damage: Neglect Disrupts the Mental Number Line." *Nature* 417 (6885): 138–39. doi:10.1038/417138a.

Acknowledgements