The perception of number: towards a topological approach

MSc Thesis *(Afstudeerscriptie)*

written by

Marco Bacchin
(born June 29th, 1984 in Feltre, Italy)

under the supervision of **Prof. Dr. Michiel van Lambalgen**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam*.

Date of the public defense:	**Members of the Thesis Committee:**
September 27th, 2017	Prof. Dr. Benedikt Löwe (Chair)
	Dr. Jakub Szymanik
	Dr. Luca Incurvati
	Prof. Dr. Michiel van Lambalgen

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

It has been suggested that our understanding of numbers is rooted in the perception of numerosities. A capacity, that of assessing the approximate number of objects in a scene, which is believed to be available also to other species.

The present work fits within the current debate on whether a 'true sense of number' is perceptually available. We will provide a comprehensive review of the behavioral, neurophysiological, and computational findings that seems to support the claim, and the limitations of the approaches taken.

Importantly, we will argue that without a clear stated definition of numerosity, it's impossible to answer the question. We will therefore provide a formal definition of numerosity, and show how the framework of tolerance homology might be used to cast light on the debate.

Acknowledgements

I should like to thank my supervisor, Michiel van Lambalgen. I was afraid, and by now and then I still am, that I could not live up to your expectations, and I've always felt a little out of my depth in our meetings. I've scratched pages of notes, and studied seemingly unrelated topics, only to realize the deeper connection already impressed in your mind. Most of those notes, thoughts and reflections didn't fit in this thesis. I was daunted at first, before realizing that what I've been left with is the greatest intellectual gift one can receive. A lifelong path of research.

I should moreover like to thank Tom Verguts, for his support in understanding Self Organizing Maps, and his application of unsupervised learning to numerosity perception in particular.

I wish to thank Lorenzo Galeotti, for the denumerable chats about logic and math. In these MOL years you have been my closest friend, jokingly referred to by others as my boyfriend.

I am grateful to Lisa Benossi, often the last person to leave the MOL room at night, sometimes morning. Our conversations started from the common interest in logic programming, and Kant's philosophy, and easily spread out to any topic. 'Stealing' coffees, and smoking too many cigarettes, have never been healthier.

I am also grateful to Almudena Colacito. You always pushed me in the right direction, when often I was taking the weirdest route to solve a problem.

I wish to thank Valentina Maggi, for her patience in reading, and correcting the draft of this thesis. And how I could not be grateful for 'il lattino', ' the morning milky', that became a daily ritual.

A thank to my family, for their support, and to my sister, Maria Elena, in particular.

Table of Contents

Page

CHAPTER 1

INTRODUCTION

> As man possesses the same senses as the lower animals, his fundamental intuitions must be the same.
>
> Charles Darwin, *The descent of man*

Numbers. Are they fundamental intuitions? At first, we might be inclined to say no. Numbers are such abstract, ineffable entities that they must require a great deal of intellectual machinery to be grasped. We might be inclined to consider numbers as an offspring of language. 'No language, no numbers' seems a pretty innocent expression. Indeed, the Pirahã, a small and isolated population living on the Maici river bank, don't have numerals. Admittedly they have a fancy language, one that can be whistled, and a culture without history, no god nor religion. If that of number is a cultural concept shaped in thousand of years, then it's quite natural Pirahã don't possess the concept. And yet their intellectual abilities are our own. Their perceptual capacities are the same and yet they don't have words for colors. It's surely harder to claim that colors are not "fundamental intuitions." Pirahãs are a strong case for the advocates of the Whorfian Hypothesis, but perhaps surprisingly, Frank et al. [48] have convincingly shown that, when numerical tasks don't involve a memory component, they are no worse than us. Language, they suggest, act like a compressor and the underlying perceptual faculties are unaltered. This suggested that numbers are not indivisible entities. It seems like there is something in the concept of number that might be considered a fundamental intuition. Something we share with Pirahã. In a certain sense, Darwin's idea is appealing. It suggests that we, speakers of a language with numerals, the Pirahã with no numerals and Monkeys, with no language whatsoever, share a fundamental intuition with respect to numbers.

Fundamental intuitions are fleeing and Darwin's remark is vague. The suggestion that the fundamental intuitions must be the same can be read in a Kantian or in an anti Kantian way. If we take Darwin to be an associationist then a priori intuitions have no place in his remark. However, reading the passage in a Kantian way suggests that the way we construct the world out of the manifold of sensory data "goes beyond the information given [1]." Interpreted in this fashion, having the same senses is regarded as sharing the way the sensory data are organized, up to a certain degree. In this spirit, cognitive neuroscience is advocating a Kantian project, a quest in the search of the intuitions our mind[2] contributes to shape the sensory impressions.

Replacing the term 'understanding' with the term 'mind', however, is only morphological sugar unless we take a stance on what we mean by mind. We share Minsky [105]'s view: "the mind is what the brain does." Therefore, the search for fundamental intuitions in mathematical cognition can be put simply as asking how do population of neurons encode distance, size, location, duration and number, and how numerical cognition might arise from the interaction of these neural codes. Although at first it might seem we are taking Kant's ideas too freely, it may help to recognize that we are pursuing the Transcendental Idealism, in brief the stance that we don't know anything about objects in themselves. We will return to this significant matter in chapter 2, in the interim it is sufficient to recognize that this philosophical excursus is not an exotic rambling into the wild mind of philosophers, but it's a sketchy portrait of the tacit assumption made by the working scientists.

Seeking an explanation that goes all the way down to the neural level is the long term stand of psychology in general and of mathematical cognition in particular. It's by no mean the objective of this thesis to try to lay down a theory so broad, yet, I maintain that a cognitive theory that is totally blind and uninterested in the neural level is quite ill posed [3].

I've always read with a sort of incredulity and admiration Locke's Epistle to the

[1] cf. Bruner [17] and Stenning and Van Lambalgen [159] for a discussion on how this connects to Kant.

[2] "Understanding" or "spontaneity" in Kant's terms.

[3] In particular, it seems to me there is a gap between the quest in trying to understand our concept of number and the hypothesis that this is based on a more foundational sense of numerosity, the former described in abstract terms, the latter mostly in neurophysiological terms.

reader, and I feel I am in a similar circumstance, where people much smarter and knowledgeable than I am, are tackling this problem. It's therefore even more ambitious to me *"to be employed as an under-labourer in clearing the ground a little, and removing some of the rubbish that lies in the way"*[4].

Housekeeping

Housekeeping may be felt as a mundane task. A mere reorganizing of thoughts, concepts, techniques and methods in a jumble. However, sometimes reshuffling things around it's sufficient to highlight aspects previously not considered and therefore to reorganize the material.

Before starting with any cleaning let's have a look at the building condition. There are three critical points we have to look into,

1. foundation

2. infrastructures

3. neighborhood

The edifice has no foundation. The central concept, the one of numerosity has not been formally defined. Numerosity is taken to be the number of items in a scene, where what counts as an item varies as much as what counts as a scene. The most clear definition has been given in Nieder and Dehaene [111]: *numerical quantity refers to the empirical property of cardinality of sets of objects or events (also called numerosity).* The cardinality of a set is a technical and well defined term that doesn't apply to sounds or images. The use of the term 'numerosity' instead of 'cardinality' to be technically uncommitted points in the direction of a lack of definition.

The district is under developed. This is made apparent in two situations. On one side, the main findings may be appreciated only from report of data and not directly from the data. As a matter of fact, only a handful of papers are connected to an open access repository. Metadata are totally useless in the text, but it is the common practice to give ceremonious descriptions of technical details. Some notable exceptions exist where experiments are described exceptionally well, however it is hard to make a point out of this exercise in technical writing. On the other side, for only a few models in the

[4] Locke [93], Epistle to the reader.

field the code is available[5]. By papers' inspection and by mail communications, I got the strong impression that nobody else but the authors actually "shook down" the models[6]. The fact that in the numerical cognition literature, so many models are abandoned and dismissed is weird when compared to what happens in other fields. ACT-R, Emergent, Nengo Models, to name a few, all are shared on-line. A notable exception is Testolin et al. [164]'s effort to promote the use of a simple deep network tool for neuroscientists. Coding the models in an unitary framework is therefore a work that will keep me busy well after this thesis and it might be seen as a continuation of what they have started.

Neighbors don't get along. Although interdisciplinarity is advocated and always praised, it's rarely practiced. There is an almost total lack of communication among fields pursuing closely related objectives. Philosophy of mathematics is practically absent and the brief incursions have usually more the flavor of historical curiosities rather than insightful proposals. No different is the exchange between the AI community and the neuroscience one [7].

Contributions of this thesis

Save point one, the situation is not serious, and it's quite subjective. Briefly having the code it's helpful, but not necessary, and walking into the philosophical minefield might help to discern what is possible from what is ill posed. Notwithstanding, rewriting the simulations' code from the few specifications given in the literature it's not a trivial task. This is especially the case in computational cognitive modeling. In machine learning, we wish the algorithms to be as efficient as possible, and if a different implementation of an algorithm leads to better performances, it is considered a better one. In cognitive modeling, although the architectures are similar, we have more restrictions, both with respect to learning rules and with respect to the interpretation of the network's behavior. Moreover, although it's often assumed that a network might scale flawlessly, the generalization is almost never achieved by simply adding more units. Particularly the code for Dehaene and Changeux [35] and Verguts and Fias [179] models, despite they are

[5]This makes the peer reviewed community quite small, and the code, when available, obsolete.

[6]Even in a recent review (Anobile, Cicchini, and Burr [7]) the authors claim the model of Dehaene and Changeux [35] is able to account for Weber fraction and invariance by describing summarily how the model supposedly achieves these results. When it comes to be precise about the model, alas, they admit that it's not certain that the same behavior emerges in a more powerful network.

[7]For the first we note the often quoted Kronecker's phrase "God made the integers, all the rest is the work of men", and for the latter we notice the close similarity of crowd estimation algorithms to the one studied in the cognitive literature.

considered the leading models, is not publicly available. The reader might have a look at Dehaene-Verguts model in the online support material of this thesis[8], to appreciate how easier it is to understand the models' assumptions, and how deeper the understanding of the models goes, once the code is provided [9]. For what concerns behavioral experiments, we are working on a JavaScript library, Stimulus, to run psychophysical experiments in a web browser with minimal performance loss compared to standalone softwares. This will give us, and hopefully the mathematical cognition community, an easier and faster tool to assess the models' hypotheses. Especially, we are devising this tool to assess the plausibility of the definition of numerosity we will provide in chapter 6.

The lack of a definition of 'numerosity' is, indeed, the most serious issue. Being stimuli modal by their nature, a modality independent definition that is blind to these differences is ill posed. Visual numerosity is mostly spatial and auditory numerosity is mostly temporal, as is tactile numerosity. This implies we need at least a definition of visual numerosity and a definition of auditory numerosity, and only once we have these two in place, we should seek for a way to encompass both into one definition. We will focus on visual numerosity, and the main task of this thesis is to argue that the correct way of modeling it is by considering visual numerosity as a topological invariant. The idea of considering a topological framework arose in connection with recent findings that topological properties affect numerosity judgments. Interestingly, when I was trying to lay down a definition of numerosity, I come across a, strangely neglected, paper by Kluth and Zetzsche [80] tackling the same problem[10]. Although the formalism we will propose comes from algebraic topology, whilst the one in Kluth and Zetzsche [80] is inspired by results in differential geometry, our definition aligns with theirs. The two approaches might be regarded as complementary, whilst they use infinitary methods to describe human behavior, we resort to finitary tolerance homology. Importantly, the two choices lead to different insights with respect to the proposed definition.

[8] Available at https://github.com/bramacchino/numberSense/blob/master/Competitive_VergutsFias.ipynb

[9] We invite the reader to have a look at the weights definition, and try to change their values. Interestingly an uncommitted network is unable to generate the desired behavior.

[10] The fact that this interesting paper is not considered in the mathematical cognition literature seems to us not to be associated only with the fact that it's a recent publication, but especially for that lack of communication among fields we referred to.

Overview

The definition we seek involves the concept of representation. Asking what is visual numerosity is asking how the brain represents numbers given a visual scene. Whilst the need of representations is widely accepted, what counts as mental representation is highly dependent on the perspective. In chapter 2 we will therefore offer a theory of representation that will help us in shaping a definition of numerosity. With a theory of representation at hand, we will specify the referent. Therefore in chapter 3 we will look at the behavioral results that constrain and inform the psychological theories, and we will review various effects that have been observed. In chapter 4 we will analyze the details of the numerosity representations (encoding, decoding procedure), and the corresponding computational models. In chapter 5 we will provide an overview of the current research in the neurophysiology of numerical cognition, its limits, and how it can be linked to the computational models. We will then move on chapter 6 suggesting a more modality dependent approach, more directly linked to the theory of representation proposed, namely visual numerosity as a topological invariant, and we will analyze in chapter 7 the computational models of visual numerosity proposed in the literature.

CHAPTER 2

Representing numbers

A mathematician is a device for turning coffee into theorems.

Alfréd Rényi

2.1 Abstraction, Representation and Information

If we take at face value Rényi definition of mathematics, mathematical cognition is the field of research that aims to understand how coffee can be turned into theorems. Less metaphorically, how the brain (its neurons, neurotransmitters, structures, and so on) acts during a complex mathematical task. In the present thesis we are interested in a subfield of mathematical cognition, numerical cognition, which aims to understand how numbers are cognized and represented. In Newell's [109] terms, we are interested in seeing how the "cognitive wheels turn" and "the cognitive gears grind" during the numerical cognitive behavior.

As mentioned in chapter 8, explanation in computational neuroscience invokes the notion of representation. Representations, broadly speaking, serve to relate the internal state of the agent to its environment. We might say that representations "stand-in for" some external state of affairs.

Although there is an almost universal agreement to the usefulness of the concept, the nature of what counts as representation depends on the approach. Representations are symbols in the classicist approach, whilst they are real valued vectors in a high dimensional feature space encoded via "subsymbols" in a connectionist approach. These differences are mirrored in the respective computational models. We wish to avoid falling into the sterile discussion of past decades between classicists and connectionists culminated in Fodor and Pylyshyn's [45] provocative paper [1]. In particular, a sprout of that argument is the still often claimed belief that connectionist architectures can be seen at most as implementation of classical architectures. Regrettably, the term 'implementation' is quite misguiding. In fact, it might suggest that somehow a connectionist network is at a different level of abstraction from a classic model. It can be, but it doesn't need to be [2]. The terms 'translation' and 'interpretation' seem to be more to the point. Specifically, even assuming that two "theories" can be bi-interpretable doesn't imply that we can pick one or the other indifferently. Which paradigm is better suitable can be assessed on a case per case basis. Indeed, such is the current state of practice, where depending on the situation a symbolic or a connectionist architecture is preferred, and it is quite common to pursue an hybrid approach. Moreover, to what extent the two theories are bi-interpretable is ongoing research under the label of "neural symbolic integration". When singling out a meaning of representation, the classicist/connectionist distinction can be furthermore problematic. The often spurious separation into symbolic and subsymbolic may indeed give the impression that logic might be the right tool for cognitive processes, but not for perceptual ones. On the contrary, it is difficult to find a better tool to bridge higher level cognitive representations, such as language meaning, with low level perceptual features than logic [3].

Nonetheless, the connectionist approach, if carefully designed, might provide a slight advantage, when it comes to define representations, with respect to the classical approach. A neural network is a theoretical tool that forces the experimenter to state the representational assumptions in a testable form. That is, if carefully specified, a connectionist model can be indeed taken as abstracting neuronal networks, and the representations can be seen as neuronal representations. In the theory of representation we are about to delineate, this is favorable and the majority of models analyzed in this

[1] The reader may refer to Garson [52] for a clear and up to date overview of the issue.

[2] Indeed, 'connectionist models' is just an umbrella term comprising a wide variety of approaches, some of which are better seen as statistical inference engines. For a thorough discussion on the matter, we refer the reader to Kohonen [82, chapter 2]. In particular we share the view that *especially many supervised learning models, although they look like (neuronal) networks, may not describe low level neuronal anatomy or physiology at all: they should rather be regarded as behavioral model or general models of learning, where the nodes represent abstract processor and communication channels, respectively*.

[3] An example of this approach is the logic of vision in Van Der Does and Van Lambalgen [175].

thesis are artificial neural networks. This doesn't mean that the neural grounding cannot be achieved by means of a symbolic approach. Nor, particularly, that the route down the Marr [101]'s path has to go through a connectionist approach.

No matter how representations are structured, they relate the internal state to the environment. But what grounds these representations? Prima facie it seems desirable to have

> some description of this processing that yields the right predictions without descending all the way to the neuron-by-neuron level (Lycan [95, pag. 259]).

As Van Der Does and Van Lambalgen [175] show, this can be done to a certain extent by investigating the model theoretic core underlying a mathematical construct used in psychophysics such as Gaussians, Laplacian and other operators. But the full extent of this grounding is achieved via models that are informed at the neuronal level. We agree with Eliasmith [39] that a neuron-by-neuron grounding is not a bad idea after all, and that a fruitful information theoretic view on representations as neural codes (cf. Eliasmith and Anderson [41]) better characterizes representations in a neural system. This characterization of the cognitive inquiry as an information processing task allows us to characterize the above mentioned Marr's three levels of inquiry[4]. Defining representations in information theoretical terms, and grounding them at the neuronal level, stresses the fact that the three levels cooperate to give an explanation of the task at issue. This simple move allows us to avoid a common pitfall. It's not rare to see authors claiming that their models are just 'computational' to underlie the fact that no algorithmic level, nor implementation is addressed. Grounding representations on the neuronal level implies that those models might not be computational level models after all, if it turns out such an implementation is not possible. It seems that the term 'computational' referred to that practice is just chosen to replace the disgraced 'phenomenological' term. Phenomenological models, however, have no place in Marr's framework, and the assumption that a phenomenological model might be used as a computational one is either originated by a strong abuse of terminology, or it arises from a misinterpretation of the framework.

The theory of representation we have in mind is borrowed from Eliasmith and Anderson [41] and Eliasmith [39, 40], and the textbook Dayan and Abbott [28] to which the reader is referred for a more in dept technical analysis.

[4]For a description of Marr's level as information processing stages we invite the reader to look at Stenning and Van Lambalgen [159, chapter 11].

In the next section, we will give an overview of the theory that is sufficient for our purposes. This is needed because the term representation is used in a variety of ways, sometimes to indicate an encoding, sometimes to indicate a decoding, sometimes as a vague term. With a clear stated notion in the back of one's mind, these different uses are immediately discernible. This will allow us to disentangle some seemingly incompatible positions and clearly indicates why and how a theory that goes all the way down to the neural level, as stated in the introduction, can be addressed.

2.2 Neurosemantics

The simplest communication system one can imagine is made up by a transmitter or sender, a channel, and a receiver (Figure 2.1).

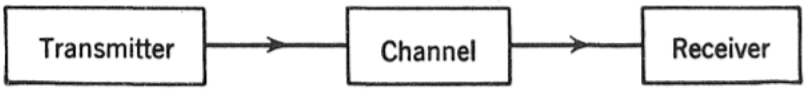


Figure 2.1: Schema of a communication system

Codes are then defined by the complementary encoding and decoding procedures. A sender sends the encoded information through a channel, possibly noisy, that is then decoded before reaching the receiver. The minimal information relation is therefore a three place relation schematizable as

$$carries(channel, information, receiver)^5.$$

By mirroring this schema the representation relation may be stated by using the standard terminology in 'vehicles represent content w.r.t. a system':

$$represent(vehicles, content, system)$$

Therefore, defining representations as codes requires defining encoding and decoding procedures, and (possibly different) input and output alphabets. Describing representations in these terms is broad enough to allow us to extend the concept of representation to that of transformation. In this way, a transformation of a representation is still a representation[6]. This gives us a powerful tool to talk directly about representations at a higher level on the hierarchy and thus of all mental representations:

[5] These three objects are necessary and sufficient to define Information in Shannon's terms.

[6] This is achieved via a transformational decoder such that the transformed representation can be extracted directly from the stimulus.

> That is, since all mental representations can be described as some combination of scalars, vectors, and functions, and those mathematical objects can be neurally represented, these methods can be used to describe all mental representations (Eliasmith [39, pag. 1043])

Although powerful, defining the representation relation in such terms misses the contribution of the referent. If we assume that referents are contents, as suggested by causal theories, then accounting for misrepresentation becomes quite cumbersome. On the other side, if we take contents to be referents, there is no place for truth conditions in determining of meaning. For these reasons, Eliasmith proposes the fourth place relation 'vehicles represent content regarding a referent w.r.t. system':

$$represent(vehicles, content, referent, system)$$

Incidentally, it could be helpful to see the introduction of referents in the relationship as somehow mirroring the Fregean sense (*Sinn*).

So far, we have been uncommitted about the four arguments of the representation relation, and we have been only moved by the close relationship between information processing and biological systems.

By *system* (the receiver), we mean the whole nervous system.

Vehicles, that we might call representations, are physical objects that carry representational content (namely, neurons and population of neurons described by the pair encoder and decoder[7]).

Referents are measurable external objects that representations assign properties to. But how are referents and vehicles related? Being they measurable quantities, we can assign random variables to them. The dependence of two random variables X, Y is summarized by the mutual information $I(X;Y)$, that captures the relation between change in one and change in the other. That is, if X, Y are independent random variables such that $P(X,Y) = P(X) * P(Y)$ then $I(X;Y) = 0$. The set of relevant events (referents) are thus the ones that maximize the mutual information $I(X;Y)$. Notice that all there is to know about the stimulus and the response relation is contained in the joint probability $P(X,Y)$, that is to compute the mutual information $I(X;Y)$ we need the joint probability or a conditional probability, and the marginal probability it is conditioned over.

This corresponds to what Eliasmith [40] dubbed Statistical Dependence Hypothesis (SDH):

[7] Clearly, this is an oversimplification: glia cells and neurodynamics might carry representational content as well. Allowing these representations into the theory simply requires the availability of an encoding and decoding procedure. The oversimplification, therefore, shouldn't be harmful.

Definition 2.1. The set of causes relevant to determining the content of neural responses (referent of a vehicle) is that set that has the highest statistical dependence with the neural responses under all stimulus conditions and does not fall into the computational description.

Where the computational description refers to the neural functioning provided by the theory of representation and computation (that is, we want the referent to be outside the system).

Contents may be taken to be the properties ascribed to a referent by a vehicle, therefore *content* is determined by decoders[8]. If no information about the stimulus can be extracted from the spiking neurons, then it makes no sense to say that it represents the stimulus.

Working with a four place relation instead of the standard three place relation prompts us to define the relation between content and referent. This requires choosing a perspective among first person perspective and third person perspective. We don't refer to the experimenter perspective as opposed to a first person perspective intended as a phenomenological favorite access (Dennett [37]). In this respect, we advocate a third person perspective. What we are interested in is a third person perspective filtered thorough the subject perspective.

Briefly the representational content problem can be addressed by means of two conditional probabilities: $p(response|stimuli)$ vs $p(stimuli|response)$. The former characterizes the observer perspective. Notice that this is what we obtain in a standard experimental settings. However, from the point of view of the subject that probability doesn't make any sense[9]. The latter characterizes the subject perspective, namely the problem of inferring the stimuli in the world from the "neural response".

A look at Fig 2.2 clarifies the point. All there is to know about the probabilistic relation between a stimulus and a response, the referent-vehicle relation, is given by the joint probability $p(r,s)$. The left part of the graph corresponds to the animal perspective. That is, from the joint distribution $P(n,v)$ (spiking rate and velocity) the conditional probability $p(v|n)$ is inferred and so it is the graph of the best estimate of the velocity given some spike rate. The right part of the graph, instead, corresponds to the observer perspective. The graphs on top show how the representational content can be highly different given a perspective.

[8]That is, a decoder tells us what properties of the encoded signal are "saved" by the neural signal.

[9]In fact, without $p(s)$ this is insufficient for a full characterization of the representation relation.

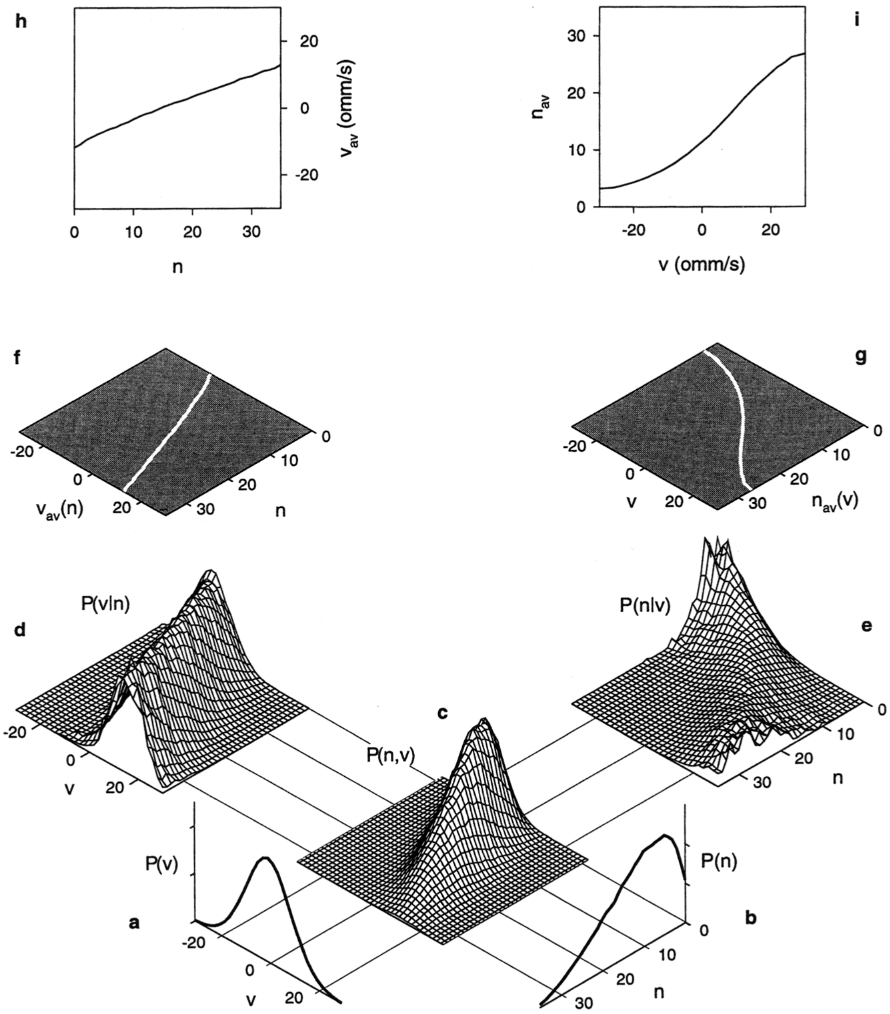


Figure 2.2: Figure from Rieke et al 1997

That representations need both perspectives is often overlooked, and seemingly contradictory point of views are in several cases just different, and importantly not incompatible perspectives[10]. This happens frequently in numerical cognition. On one side, we have those that advocate

> there is no reason to think that number is a complex parameter of the external world, one that is more abstract than other so-called objective or physical parameters such as color, position in space, or temporal duration. In fact, provided that an animal is equipped with the appropriate cerebral modules,

[10]The conditional probabilities are linked via the Bayes's law.

> computing the approximate number of objects in a set is probably no more difficult than perceiving their colors or their positions (Dehaene [32]).

On the other side, this illusory simplicity is challenged by those claiming

> it is easy to see that there is no such single visual attribute unambiguously related to the number of dots. Strictly speaking, it is impossible to see numerosity at all. The only possibility is to rely on an intermediate impression of numerosity which is formed on the basis of a certain stimulus attribute, more or less closely correlated with the number of objects. The visual number can be communicated to the observer only through a certain set of visual attributes, none of them being the visual number as such (Allik, Tuulmets, and Vos [2]).

The scenario envisaged by Dehane focuses mostly on the (external) observer side. Numerosity is out there, we know it, and we can capture how to compute the approximate number of objects by looking at $p(response|numerosity)$. Allik, on the other side, is thinking about $p(numerosity|response)$: none of the visual attributes and therefore the responses are the numerosity itself, but somehow the subject has to infer the numerosity from these responses. The representational content is therefore different and both positions are partial and need to be complemented in order to achieve a computational theory of numerical cognition.

In conclusion of this discussion on representations, we wish to highlight how the problem of misrepresentation is easily addressed within this theory. What the SDH picks is the "conceptual content" (determination of decoders over all stimulus conditions), which however may be different from the "occurent content":

Definition 2.2. The referent of an occurent representation is the cause that has the highest statistical dependency with the representation under the particular stimulus conditions in which it is occurent.

To sum up, in order to understand our capacity to deal with numbers, it is crucial to identify the type of number representations that our brain uses. We have to identify the referents, the vehicles (that is the encoding and decoding procedures), and pick a perspective[11].

[11] The most complicated part of this plan is, unsurprisingly, picking up the referents. This is especially true considering the fact that addressing this issue is almost impossible without a clear definition of what

Given that the standard psychophysical experiments in numerical cognition involve dot arrays in a screen, starting out with visual numerosity seems a promising approach.

Although numerosity is at this stage still a vague term, and it will be our effort through this thesis to give the necessary concepts in order to propose a formal definition, we remind the reader that we were prompted to investigate numerosity from the suggestion that number might not be a primitive concept, and that some kind of fundamental intuitions, which we share with other species, can be its primitives. With this we wish to point out that, although we cannot claim to have a computational theory of numbers, unless we have all the components of the representation relations, the way we fill in the details is not constrained by the representational theory proposed. A back and forth between levels reshapes, step after step, the concepts, constraining the space of possibilities on the higher levels, and narrowing down the guesses needed for reverse engineering the neural code.

In fact, although the definition that we will propose for numerosity is mainly intended for visual numerosity, there is another line of research, that starts from the higher level in the hierarchy seeking those representations that should be foundational. The starting point is therefore the representation of natural numbers.

2.3 Representing numerosity and fearing natural numbers

Natural numbers are abstract entities. This makes finding an adequate representation even harder. At first sight, how might we know a vehicle represents a referent if we don't know what this referent is? This might suggest to give an account of what things our number words and numerals name or stand for. However, this is a dangerous and possibly fallacious path, in Mayberry's [102] words,

> the beginning of wisdom is to realise that there simply are no such things as "natural numbers", that natural numbers as "mathematical objects" are illusions, non-entities, mere artifacts of our notation, reified and alienated products of our counting and calculating procedures, and that, consequently, to devise a theory of what "they" are as particular objects is utterly otiose

"we wish" to represent. Anyway, assuming what it requires to be shown, it is still possible to give the encoding and decoding procedures. The interested reader may found a clear example, in line with the present discussion, in Le Mouel and Pouget [87] (unpublished, but freely available on line).

and, indeed, productive of quite unnecessary confusion (Mayberry [102, pag. 258]).

Naturally, the above sentence provokes strong reaction, but such a "philosophical position" highlights the impasse in which the cognitive psychologist finds herself. When working at this level of abstraction, the researcher has to rest on an intuitive definition of numbers that avoids any pitfall into philosophy of mathematics. That natural numbers don't necessitate any definition seems to be, probably surprisingly, a widespread opinion even among mathematicians. That this doesn't create a problem is captured by Jouko Väänänen [174], according to whom "mathematicians argue exactly but informally", which "has worked well for centuries". An intuitive pretheoretic concept of number is the sequence of which we don't know nothing else other than that it is generated from zero by successive iterations of the operation of passing from a number to its immediate successor[12]. As far as it goes, such a characterization should suffice. Indeed, as Rips, Asmuth, and Bloomfield [139, pag. 9] pointed out, this seems to be the one implicitly or explicitly assumed by many cognitive scientists.

By this, we don't want to claim that a dialogue between mathematicians, philosophers of mathematics and cognitive scientists wouldn't be profitable. As stressed in the introduction, the opposite is advised. We just want to point out that, once one is aware of the informality of the definition, and importantly of the limitations of it, its adoption is legitimate. Failing to account for the limitations, however, might result in empirically misguided inferences. For example, from the fact that the structure of finite ordinals, with the ordinal operations $(+_o, x_o)$ is isomorphic to the structure of the numerosities with the cardinal operation $(+_c, x_c)$, one might be inclined to infer that an ordinal definition suffice[13]. However this assumption might mask comprehension. In Gelman and Gallistel's [55] proposal the infant comes to understand that the last word in the counting sequence denotes the cardinality of the enumerated set (cardinality principle). If the child associates with the number words only the ordinal position, for example, then it might appear she doesn't yet know the cardinality principle, whilst in fact, she might already have inferred it. Moreover it might not be the case that the operations an infant uses are the standard binary operations we usually associated with natural numbers,

[12] Although intuitive, the definition is in fact circular. It implies that a collection is finite if it can be put in correspondence with an initial segment of the natural numbers, that is if it can be counted out. But it can be counted out if the iteration of the successor function is finite. The appealing of the definition comes from the belief that a definition by recursion doesn't need justification. A fallacy reminiscent of the sorites paradox (cf. Mayberry [102]).

[13] This seems to be the hypothesis behind Gelman and Gallistel's [55] counting principles.

and might be the plus-one, minus-one, unary operations, applied only to 'Spelke-objects', roughly a persistent object, or a pair of persistent objects (Spelke [157]).

This brings up a fundamental distinction of intents. On the one hand, if mathematical cognition's research is trying to single out the 'intended model', that is to seek the representations underlying the intuition that there is a paradigmatic structure of the natural numbers, then a formal theory of natural numbers, such as Peano arithmetic, is of little use. In fact, if those intuitions were again formalized according to a given theory, the explanation will be circular and not informative. This is what the majority of scholars address. For them, an informal definition should suffice: paraphrasing Jouko Vaananen's statement, mathematical cognitive scientists should argue exactly but informally. On the other hand, if what is claimed is our possession of the concept of natural numbers, then mirroring it by means of a formal theory seems appropriate. Less straightforward is which theory is the one to use as benchmark. Are cognitively experienced natural numbers the "standard" natural numbers? To answer this question, one can embark in a different project and try to see which among various different mathematical theories is the closest to psychological reality. From this point of view, it could be the case that cognized natural numbers are not the PA, ZFC numbers defined in standard mathematics, but a different, and admittedly more exotic, kind of natural numbers. In this respect we observe that there are only few studies in mathematical cognition in which "big" natural numbers are investigated(e.g. Rips [137]). 9223372036854775807 is a number as much as 2, but surely doesn't behave cognitively in the same way[14]. The *exoticism* is therefore necessary. More broadly, the history of the concept of infinity, from the Greek *horror infiniti* to the embracing of its paradoxes, and the confused statements that children give when prompted, suggests us not to blindly assume that the concept is naturally available.

Therefore, what we can do with just an intuitive definition is asking ourselves what are the cognitive foundations, the representational primitives out of which the natural number representations are built. Which constraints do we have in seeking representations? Are these representations innate or are they learned from more basic representations? In the context of core cognition (Carey [20]), and in the neural equivalent (Dehaene and Cohen [36], Dehaene [29]), for example, is hypothesized that, when we learn and practice science and mathematics, we take capacities of the mind and the brain that evolved to serve other functions, and we harness them for new purposes[15]. In

[14]For the curious reader, the number presented is $2^{63}-1$, that is the largest 64 bit number, usually, the largest number representable by a computer. In Python for example type sys.maxint or sys.maxsize.

[15]This concept recalls exaptation in evolutionary theory.

the representational framework proposed, this means that the representations we seek are transformations of earlier representations. As Carey [20] (by extending Dehaene's [32] proposal) lists, those encompass number line, representation of space and continuous quantities, time, length, distance, iterative capacities, logical capacities, relational and order capacity, the syntactic/semantic representation of numbers in natural language, and the system of parallel indexing of small sets in mid level attentional systems.

This list is daunting. Way too broad and too abstract to actually have one's hands dirty. What one can do is to hypothesize how these capacities must be organized in order for the concept of natural number to emerge. From all the items in the list, particular relevance is given to the approximate number sense (ANS), a specification of the representation of continuous quantities, and to the parallel individuation system (PIS)[16].

2.3.1 Numbers in the scrum

How those representations are to be organized in order to support the concept of natural number is subject of strong debate. We can identify three major contenders:

- Natural numbers are innate, while abstract number sense and possibly other capacities in the list are just ancillary
- Core systems are foundations. Throughout life, representing and reasoning about natural numbers depends on them.
- Core systems are scaffolding. Once the natural number system is constructed, it has a life of its own.

Nativists

Nativists are of the opinion "rather be on a subs bench than in the scrum". Without denying the representations in the list might play a role, they contend their foundational nature. Being an opposition party gathers different approaches. Gallistel and Gelman [51] and Gelman [53] propose that the presence of 'preverbal numbers' is innate and that the child learns a bidirectional mapping along with them. However, this account doesn't explain the developmental trajectory, suggesting a much briefer learning pattern than the one actually observed. Leslie, Gelman, and Gallistel [90, 89], and Izard et al.

[16] In chapter 4 we will explain why the hypothesis of an approximate number sense (ANS) has emerged and what is its supposed role. For the remaining of the chapter it is sufficient to assume that numerosity are encoded approximately via overlapping Gaussians.

[74] maintain that exact equality and successor functions are sufficient and necessary to build up the natural numbers, and are not learned for other primitives. However, whether and when the child understands the successor function, and whether ordinal numbers are learned before cardinal numbers is still debated (cf. Kaminski [77]). Bloom [15] and Hauser, Chomsky, and Fitch [62] maintains that numbers are built in natural language, in particular in the recursive capacity that is a hallmark of human language. However, Gelman and Butterworth [54] strongly criticize this approach by claiming that dissociation between language and numbers is possible. Moreover the claim that natural language, as a cognitive capacity, is recursive, it's not an innocent and self evident position[17]. It's probably in Rips, Asmuth, and Bloomfield [140, 139] and Rips and Hespos [141] that the nativist approach has its peak. Here, it is claimed that there is no difference between cognitive and mathematical natural numbers. These are taken as any list that obeys Peano-Dedekind axioms, or a cognitive plausible version of Peano-Dedekind axioms, although to what this cognitive plausible version amounts to remains unanswered[18].

Foundationalists

This thesis is championed by Dehaene [32], in his words *these abilities (ANS) not only enable us to quickly work out the numerosity of sets, but also underlie our comprehension of symbolic numerals such as Arabic digits. In essence, the number sense that we inherit from our evolutionary history plays the role of a germ favoring the emergence of more advanced mathematical abilities*. (For a recent update of this view, called neuronal recycling , and its possible extension to reading and language skills see Dehaene [30] and Dehaene and Cohen [36]). In particular, as noted in Graziano [59], there is a will to distance from the nativist, in the sense that the language-less features of numerical competence are the basis of numerical cognition, and yet to accept the idea that language is necessary. The main claim is that the ANS encode both symbolic and non symbolic numerosities. In particular, it is suggested that number symbols simply operates with narrower tuning curves (Nieder and Dehaene [111], Piazza et al. [120], Verguts and Fias [179]).

[17]The classical 'competence/performance' dichotomy is blurred out at the cognitive level (see Stenning and Van Lambalgen [159, pag. 371], and the Turing Equivalence of recursion and iteration is often confounded for a claim that, cognitively and neurally speaking, one can be replaced by the other (Luuk [94]).

[18]En passant, in Rips, Asmuth, and Bloomfield [139, pg. 58, commentary)] this seems to refer to the use of the least number principle and not to the induction schema. A more thorough approach has been given by Krysztofiak [85].

Recent behavioral and neurophysiological studies, however, suggest that the non symbolic and symbolic numbers are more distinct, and that the latter form a system of discrete, categorical, representations, rather than being coded simply by narrower tuning curves (Cohen Kadosh et al. [25], Lyons, Ansari, and Beilock [97, 96], Holloway et al. [70]). The outcome of these studies suggests a greater importance on how a symbol is related to other symbols, than how its related to the quantity it represents (in terms of ANS). The fact that, at the neural level, network overlapping has been observed, prompted the hypothesis that the detachment might be a learned one, fostering the developmentalists view.

Developmentalists

Given the difficulties of the Nativists and Foundationalists, respectively violating Occam's razor and the aforementioned studies, a third position gained momentum. As a middle way between the two positions, Spelke [158] maintains that natural numbers concept emerges thorough the combination of core knowledge and natural language. And that the use of natural language to combine core representation rapidly and productively is fundamental. A stronger position is advocated in Carey [20, 21], within the core cognition proposal. Natural numbers are Bootrstrapped from the earlier representations of the ANS and the parallel individuation system (PIS). In particular it is suggested that the ANS grants the concept of *progression*, and the PIS the one of *discreteness*. Tangential to our concerns, but fundamental in Carey's system, is the "discontinuity hypothesis" underling the "Quinian bootstrapping"[19]. Carey proposes that by combining the two systems we should be able to "bootstrap" our knowledge developing the concept of exactness and successor. Importantly the new conceptual system developed is not translatable into its foundational system.

The discontinuity claim is the more problematic, and taken Carey's approach as a whole, there are no mathematical or computational models to support the theory. Indeed, as noted in Rips, Asmuth, and Bloomfield [138], the only model for Carey's Bootrstrapping Theory, Piantadosi, Tenenbaum, and Goodman [119], is not a model of Bootstrapping but of Fodorian hypothesis and testing. Indeed it's more in line with Spelke's proposal than with Carey's one (for example the ANS has no use in the model) and recursion play the same role that nativist advocates.

We therefore take some liberty from Carey's actual proposal, and we will briefly present her proposal omitting the discontinuity hypothesis. Carey's suggestion, although

[19] Natural numbers learnability is taken as a paradigmatic example of discontinuity.

interesting, is quite vague about the details, therefore our reconstruction of her argument might differ from her actual assumptions. Briefly the proposal might be summarized in four (possibly five) steps.

1. The child starts memorizing a short list of ordered numerals S (e.g. *one, two, ..., ten*) as an uninterpreted place holder structure.

2. She later links S to a mental representation of 'set numbers' induced by the Parallel Individuation System. That is the name *one* is mapped into the object file $\{O_1\}$, the numeral *two* into the object file tracking any two objects $\{O_1, O_2\}$, and *three* into $\{O_1, O_2, O_3\}$ [20].

2b. The child is then able to link S to a mental representation of "magnitude ordering" given by the ANS [21].

3. From 2 and 2b the child realizes a parallel exists between 'syntactical order' and 'representation' [22].

4. This helps the child realizing that *the meaning of the next element on the numeral list is the set size given by adding one to the set size named by the preceding numeral*.

5. The concept of natural numbers might arise by taking the limit of the sequences generated by successive applications of step 4.

Step 5 embodies a passage from a potential infinite, implicit in the numerals grammar, to an actual infinity through a limit operation. This last step is the one Rips is interested in, whilst Carey seems to have doubt that this mature stage is ever reached in numerical development. Only through a lengthy historical process the concept of natural numbers arises as a cultural invention.

[20]The limit of the infant PIS is taken to be around three objects. This mapping is acquired via what Carey dubs a "modeling process" (see Carey [20] pag. 307, 418): induction, abduction, analogy, limited case analyses and thought experimentation).

[21] Carey doesn't explain how this is achieved. We might, for example, represent the ANS ordering as the partial order naturally induced by set inclusion over the tolerance classes associated to the tolerance relation given by the Weber fraction. Assuming the child has access to the ANS implicit ordering, the link might be established via "modeling processes".

[22] This require a refinement of the partial order into a linear order. For example the child has to infer a rule of the form $\frac{x < y \quad z \in P \wedge y \in P}{x < z}$.

2.4 Summary

In the introduction we claimed that definitions in neuroscience involve the concept of representation. Here we built upon Eliasmith's [40] 'neurosemantics' proposal of grounding representations at the neural level. An information theoretic approach suggests us to see representations as neural codes. In information theory codes are seen through the complementary encoding and decoding procedures between two alphabets. Interpreting representations as codes therefore requires finding these procedures. We stressed that, to define the decoding procedure, taking the subject's perspective saves us from common mistakes. This simple shift is of paramount importance to assess whether the concept of numerosity might be linked to a representation of numerosity. Importantly, from the decoding procedures, a hierarchy of representations arises naturally, and supports our proposal of seeking lower level representations first. Although there is a great deal of research on visual numerosity, much attention comes from the goal of numerical cognition of defining our understanding of natural numbers. In the proposed framework this corresponds to a representation at the highest level in the hierarchy. Working at this level prompted a lot of speculations and we briefly exposed the main approaches taken. With the exception of hard core nativists like Rips, Asmuth, and Bloomfield [139], lower level representations are considered important to define the concept. We didn't take any stance on the debate but noted that a firmer ground is indeed needed.

CHAPTER 3

Behavioral observations, or what (almost) everyone finds

> There is a great difference between the Idols of the human mind and the Ideas of the divine. That is to say, between certain empty dogmas, and the true signatures and marks set upon the works of creation as they are found in nature.
>
> *Francis Bacon*, Novum Organon

When it comes to single out the capacities our mind possesses, an intuitive appeal to appearances may lead astray[1]. We have seen in chapter 2 that most researchers (except for 'nativists') share the view that, among the representations our mind needs to build up the concept of number, the approximate number sense and the parallel individuation system play a major role. Thus, in this chapter, and in chapter 4, we will delve into these topics deeper. In particular, here we will provide an excursus about the behavioral experiments devised in order to assess our numerical competence, while in the following chapter, we will investigate various theories that have been proposed to account for these results.

[1]The philosophically inclined reader may read this as an application of Kant's transcendental argument to cognitive science. Here we are not simply interested in what are the necessary intuitions that *a mind* must possess for a given capacity to arise, but we want both to ascertain what those capacities are and what actual conditions are necessary for the given capacity to arise in *our mind*

3.1 Behavioral observations

Behavioral observations and hypotheses explaining them are usually presented together. We have decided to disentangle the findings both from the theories that generates them and from the hypotheses that can explain such facts. Moreover the behavioral observations cluster various findings. Outward similar behavior can be elicited by different underlying mechanisms. This remains the case also when behavioral findings are correlated to neurophysiological one [2].

Special attention must be held to the fact that the behavioral observations may not be about 'numbers' (or numerosity) after all. Although this situation seems paradoxical, the reader may convince herself of the difficulty of devising a behavioral experiment that remove all confounding factors correlated with numerosity. An ideal numerosity mechanism should be insensible about the shape and spatial distribution of objects in the scene (see section 3.7). There is indeed an ongoing debate on whether there is a dedicated (visual) mechanism for the 'sense of number' (Burr and Ross [19], Ross and Burr [148], Arrighi, Togoli, and Burr [9]), or whether the representation of numerosity is linked to other visual attributes such as density, or non visual attributes like coding of duration (Tokita and Ishiguchi [168], Walsh [183], Tibber, Greenwood, and Dakin [166], Dakin et al. [26], Durgin [38]). A limitation of most behavioral, and physiological, studies is the small range of numerosities tested, usually in the range one to six[3], and almost never higher than thirty two. The assumption that on larger numerosities the behavioral findings align is thus not granted. A recent review (Raphael and Morgan [133]) in fact concluded that at the present stage, we cannot claim that numerosity is a perceptual feature. However, Cicchini, Anobile, and Burr [24] suggest that the same data speaks for two systems, one numerical, another based on density.

As a matter of fact this debate is a driving force behind this thesis, and we deem that speaking about general magnitude effects is at this stage less committing.

3.2 Distance effects

The distance effect has first been recognized in the seminal work of Moyer and Landauer [108] and it has by then occupied a prominent place in numerical cognition. Usually the various kinds of distance effects are grouped into the broader term of *distance effect*,

[2]The simple fact that a given region represents two types of stimuli does not means that the underlying neural codes are the same (see Lyons, Ansari, and Beilock [96] for an example).

[3]We will see that this range has a special characterization.

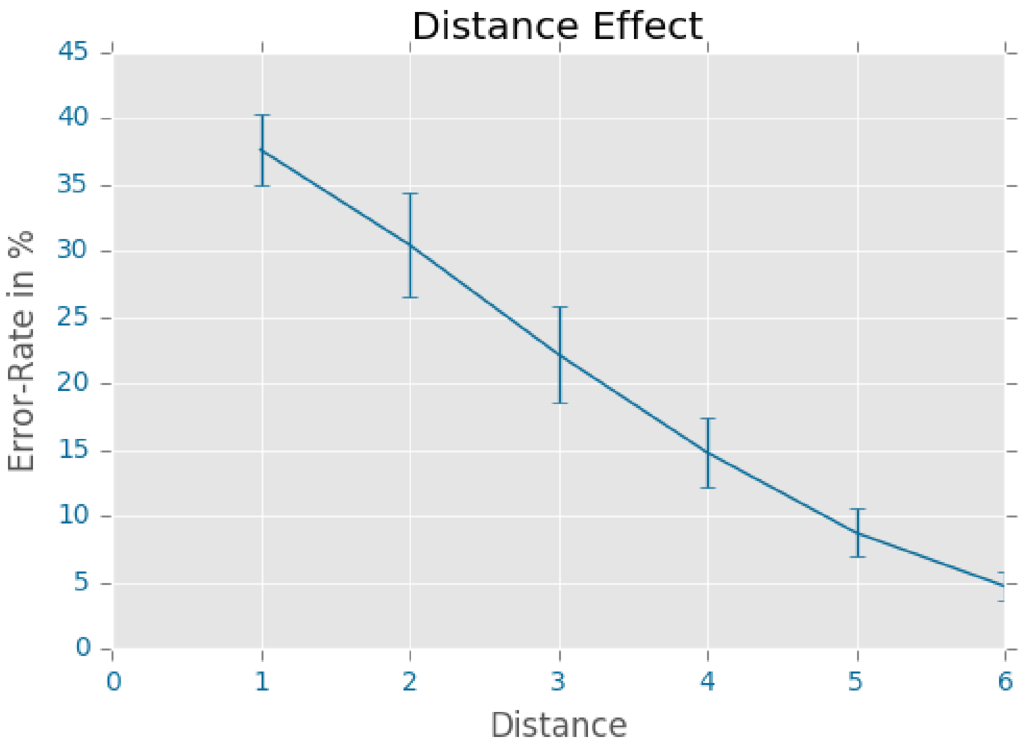
Van Opstal and Verguts [177] pointed out that this obfuscate the origin of the behavioral effects.

3.2.1 Comparison distance effect

This effect is apparent in the number comparison task and in the same-different task. In the former, participants need to select the largest (or smallest) of two numbers. In the latter, the participants have to indicate whether a item is equal or different than another (therefore removing any ordinal decision). What is observed is a systematic dependency of error rate and response time on the numerical separation between the items, where reaction time (RT) smoothly decreases with the numerical distance between them. The effect is modality independent, and has been observed both in stimuli containing dot arrays, and in numerical stimuli presented in a symbolic form.

3.2.2 Distance priming effect

By priming we refer to a temporary change in the ability to identify a stimulus as a result of a specific prior experience. A priming effects for numbers has been reported in several studies (e.g., Heyer and Briand [68]; Reynvoet, Brysbaert, and Fias [136]). The priming effect is inversely proportional to the numerical distance between the prime and the target. That is, it's easier to respond to a numerical stimulus when it is preceded by a prime (number) that is numerically close, compared to when it is preceded by a prime that is numerically far. Moreover, the effect is symmetric with respect to the priming direction, that is the size of the priming effect for a given target, n, is the same for both $n+1$ and $n-1$.



3.3 Size effect (a.k.a. magnitude effect)

For a given numerical distance, pairs of small numbers are compared faster, and more accurately, than pairs of large numbers. For example chimpanzees have no difficulty in determining that an array containing two dots is more numerous than one containing a single dot, even though these two quantities differ only by one unit. However, they fail increasingly more often as one moves to larger numbers such as three versus four, and so on. The effect is observed also when the subject is required to assess the relative magnitude of symbolic numerals, and it appears to be modality aspecific.

3.4 Subitizing effect

As early as 1871, Jevons [75] reported a subject ability to identify with considerable speed and accuracy the number of visual stimuli simultaneously presented. The term 'subitizing' has been proposed in 1949 by Kaufman et al. [78], and defined by Von Glasersfeld [181] as "the immediate correct assignation of number words to small collections of perceptual items". The upper limit of this process is debated, but there is a certain agreement to set it at three or four elements in the visual display, based on increased

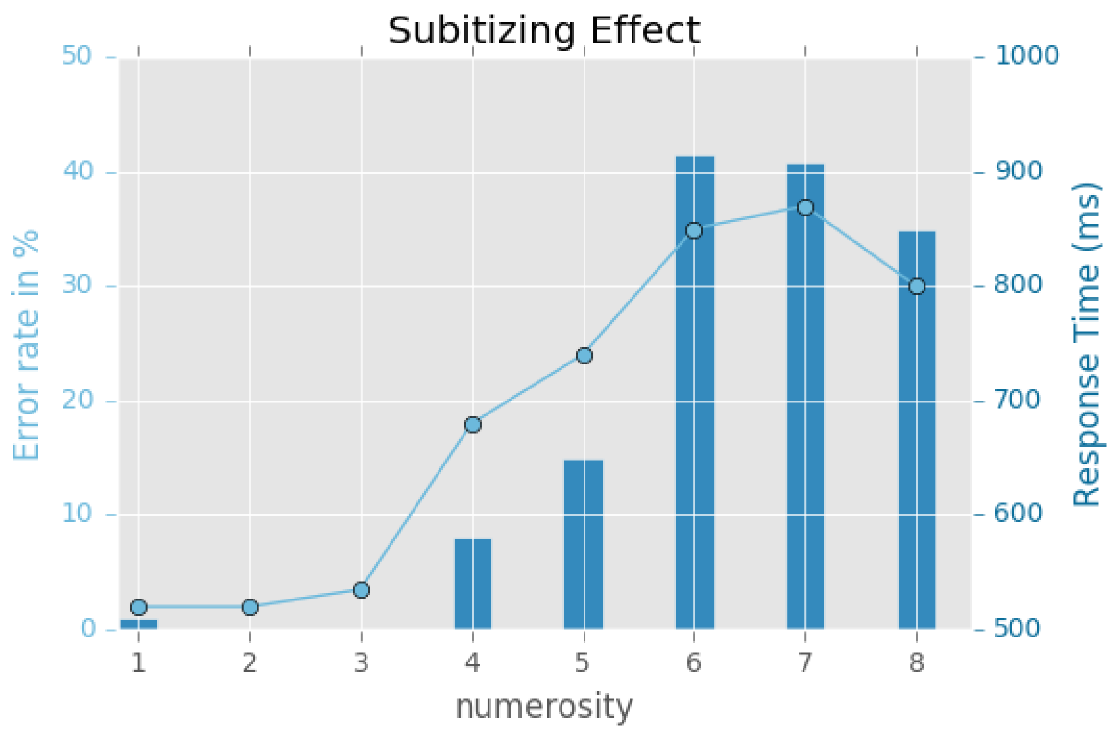


Figure 3.1: Subitizing effect. Error rate and response time are reported for numerosity 1 to 8. A discontinuity at 3 is clearly visible. Data from Piazza et al. [124].

latency and errors above this threshold. (Akin and Chase [1], Mandler and Shebo [100], Sathian et al. [151], Piazza et al. [123], Ross [147]). Although Von Glasersfeld [181] goes beyond this range. Interestingly in several animals the subitizing limit appears higher than in human subjects (Davis and Pérusse [27]).

The typical experimental design is asking people to enumerate a patch of scattered items as rapidly and accurately as possible. In the "subitizing" range error rate and reaction time increase only slightly, above that range they grow faster, as can be appreciated by inspecting Figure 3.1.

Although initial reports didn't assess any interference with attention, more recent studies highlight a connection between the subitizing ability and attentional resources. Railo et al. [132] report that manipulating attentional resources can drastically reduce the subitizing maximum to two items. Vice versa, enhancement in attention can improve the subitizing range (Gliksman, Weinbach, and Henik [56]). Interference and correlation with working memory capacity has also been observed (Piazza et al. [124]).

3.5 SNARC effect

SNARC is an acronym for "Spatial Numerical Association of Response Code", a term minted by Dehaene as a tribute to Lewis Carrol's

> wonderfully nonsensical poem. "The Hunting of the Snark", tells of the relentless quest for a mythical creature, the Snark, that no one has ever seen but whose behavior is known in exquisite detail

What is observed is that larger numbers elicit a faster response on the right side whilst smaller numbers elicit a faster response on the left side. The direction effects is associated with reading habits (Dehaene, Bossini, and Giraux [34], Shaki and Fischer [154]) and is flexible and instable in bilinguals (Fias [44]).

3.6 Transfer Effect

We take here some liberty with respect to our commitment to present the data disentangled by the hypothesis that inspired the experiments generating them. In the case of the transfer effect, it's not only the explanations that clash, but the very effect that is in question. Indeed, this is the most controversial of the behavioral findings here presented. With transfer effect we term the findings that the behavioral signature for number symbols transfer to those of symbolic manipulation. We mean this in a strong sense. As we have seen, distance and size effects are observed in both numeral and dot arrays. However the fact that the same, or better a closely similar signature, is observed, if it can suggest a relation, of course it doesn't imply one [4]. If the developmental trajectories for the symbolic acquisition, and successful use of numerals, are predicted by what we termed 'general magnitude effects', however, we would be in a better position to claim a relation. The way this is usually expressed is by saying that the Abstract Number Sense (ANS)[5] correlates and predicts mathematical abilities [6]. Piazza et al. [122] cautiously report a correlation between ANS and number knowledge, but no correlation where found between ANS and calculation abilities. On the contrary, Libertus, Feigenson, and

[4] For example (Van Opstal and Verguts [177]) show how distance effects may arise as a byproduct of a decision mechanism.

[5] That is, the subjective signature of the general magnitude effects. ANS is one widely accepted hypothesis on the explanation of those signatures, and we will dedicate much of the next chapter to explain it.

[6] The term 'math abilities' has to be taken with a grain of salt, what it's meant is simply elementary arithmetical and numerical skills.

Halberda [91] and Mazzocco, Feigenson, and Halberda [104] conclude that the ANS' precision predicts performance on school arithmetic, but does not predict non numerical abilities. A recent study (Lindskog, Winman, and Poom [92]) suggests a correlation with math anxiety. However, Sasanguie et al. [150] found no correlation between ANS acuity and symbolic number processing, and in a large scale (170 children), one year long study (Göbel et al. [57]) the authors found no correlation with ANS[7].

That the effect is dubious is not surprising. As we pointed out in section 3.1 the very nature of the behavioral observations here presented is puzzling. Recall that for both developmentalists and foundationalists (cf. chapter 2) the numerical content in the approximate number system is required to build the natural number concept, therefore, if no transfer were observed, it would be very much less direct the way to adopt the ANS as a building block. We are dubious, however, that this is the right strategy to assess that. Checking the transfer effect requires proving a negative: an empirically impossible task. That this has long been the strategy of choice is correlated to the lack of a formal definition of numerosity we highlighted in chapter 8. The definition we will put forward in 6 will gives us another strategy correlated with the interference effect we are now going to look at.

3.7 Interference effect

As we pointed out in chapter 8 invariance is assumed by "definition". As such the degree of invariance is mostly not reported, but implicitly, in the literature. However, as briefly mentioned, interference effects are quite widespread.

Allik, Tuulmets, and Vos [2] report size invariance in numerosity comparision, but dependence on spatial arrangement, in particular to proximity (two proximal dots have less impact than two far apart on numerosity estimation). Ross [147] found that mixing sizes disrupt numerosity estimation (more specifically the Weber ratio increases). Hurewitz, Gelman, and Schnitzer [72] report that cumulative area interferes with, and predicts numerosity estimation, interestingly even in the subitizing range.

Numerosity is influenced by topological invariants such as connectivity and inside/outside relationship. In He et al. [64, 65], connecting and enclosing items led to robust numerosity underestimation, with the extent of the underestimation increasing monotonically with the number of connected/enclosed items. The same effect has been reported even when participants were explicitly told the lines were irrelevant and after extensive

[7] Although predictive power with Arabic number identification.

training (Franconeri, Bemis, and Alvarez [46]). The underestimation is present no matter the contours are real or illusory (Kirjakovski and Matsumoto [79]). Similarly Im, Zhong, and Halberda [73] report that grouping dots by proximity induces underestimation.

3.8 Weber's law

Collectively, Comparison Distance Effect and Size effect are usually reported compressed into a single law: "the time to judge the numerosity of two stimuli is a function of the ratio of the numerical magnitude they represent". Weber's law states that the just-noticeable difference between two stimuli is proportional to the magnitude of the stimuli. This is an old stand finding in psychophysic that date back to 1834 [8], and has hold true as an approximation for various senses within a specific bandwidth of the external stimuli.

Let's call the initial intensity I and, let ΔI be the minimal amount needed to detect a difference (JND). In discrimination experiment we are interested to found ΔI as a function of I (s.t. I+DI is just discriminable from I). Therefore Weber's law is usually stated by the linear equation

$$\Delta I = w * I$$

or more generally $\Delta I = w * (I + I_0)$ (to account for very small values, that is $w * I_0$ is the absolute threshold, the smallest intensity we can reliably detect) where w is a constant called Weber fraction. In Fechnerian's terms the differential change in perception is a geometric progression of the differential change in stimulus

$$d\psi = w * \frac{dI}{I}$$

By integrating, $\int d\psi = \int w * \frac{dI}{I}$ we obtain Fechner'law

$$\psi = w * ln(I) + C$$

[8]For an historical overview and limits of the law see Hecht [66].

3.9 Summary

Experiments similar to the one in our demo[9] have been widely used to assess numerosity perception in both humans and animals. Whether these experiments target numerical competence, or other non numerical mechanisms are at play, is currently a matter of a strong debate. We reviewed various effects, and noticed that the *transfer effect*, to wit the observation that performances in the above tasks correlate with arithmetical achievement, faintly suggests that numerical competence is indeed addressed. The results of the experiments are however inconclusive and discordant results are reported in the literature. The *interference effect*, the observation that numerosity perception is not totally invariant to various transformations of the visual stimulus, is frequently disregarded. The interference effect will play a major role in chapter 6 where we indicate how the proposed definition of numerosity can be lifted to cast some light into the aforementioned debate. For the numerosities in the range one to three (possibly four) a *subitizing* effect is observed, reaction time and error rate in this range increase only slightly. The perception of numerosities obeys the Weber-Fechner law according to which relative differential sensitivity remains the same regardless of size, as observed in the *size effect* and in the *comparison distance effect*. An interference of numerosity with space has also been observed (*SNARC effect*) which direction is correlated to the reading direction of the tested subjects.

[9] Available at https://bramacchino.github.io/stimulus/demo.html and based on the library Stimulus current in development.

CHAPTER 4

Functional and Computational models

> There are reasons to believe that the goal of understanding the human mind strictly from observations of human behavior is ultimately untenable, except for small and limited task domains.
>
> Ron Sun, *Introduction to computational cognitive modeling*

To account for the behavioral findings reported in chapter 3 various functional and computational models have been proposed. With the notable exception of the interference effect, distance, size, SNARC, and subitizing effects might be explained by the proposed models, or by a combination of them. The transfer effect is usually not addressed by the models, and it shouldn't be, considering the fact that it's seen as the major clue to assess whether the proposed models account for numerical competence or a generic magnitude. As clearly emerges from the preceding chapters, we think that this line alone is insufficient for claiming that higher arithmetical abilities are founded on mechanisms which signature is compatible with the observed effects. The fact that the interference effect is not accounted, because contrary to the behavioral findings, a total invariance is assumed, seems to us the major drawback of these models, if one wish to apply them as building block of higher mathematical capacities. In this chapter we will avoid this issue, that will be resumed in chapter 7. Taken at face value the behavioral observations, it seems that two separate systems are at play. One responsible to the exactness of the subitizing range. Another, approximate, able to give an estimation of numerosity [1].

4.1 Functional Models

ANS

4.1.1 Barcode Magnitude representation, (BMR)

The first representation of numerical information has been presented in a series of arithmetical architectures (Anderson et al. [4], Anderson, Spoehr, and Bennett [6], and Anderson [5]). In the BMR number magnitude is coded as a moving bar of activation on a topographic scale. Numerosity is encoded as a "discrete number line" where each number is encoded by activating the corresponding node and its immediate neighbors. That is assuming a fixed length of the moving bar, this representation might be seen as a discrete version of the number line coding with logarithmic scaling described below. This representation is however too crude. In the most basic form (the one used in the above mentioned models) the BMR is represented as a finite subset of the natural numbers equipped with a tolerance relation τ (i.e. a symmetric, reflexive, not transitive relation) defined by the rule

$$n, m \in \tau \leftrightarrow |n - m| \leq 1$$

The problem for this account lies in choosing one as a threshold ad hoc. More generally for any threshold $n \in \mathbb{N}$ the maximum elements in the tolerance classes will form a linear order[2]. To avoid this conclusion, in the models of Anderson and colleagues, it's assumed a skewed frequency with which numbers are encountered during learning. With this frequency based accounts, the size effects emerges by a weakening of the connection for larger numerosities[3].

[1] Moving from a behavioral description to a neuronal implementation, there is an ongoing debate whether the two systems are still separated, or if they might be instantiated in the same network (Sengupta, Surampudi, and Melcher [153]).

[2] A winner takes all network to find the maximum is indeed quite trivial to construct in this case. For a given numerical stimulus, the elements of an ordered tolerance class are activated. If we let the connection weights to be excitatory in the right direction and inhibitory on the left direction, the network activation will settle on the maximum.

[3] For example, by imposing a weakening of the inhibitory links in the WTA network for classes representing larger numerosities so that the stable state requires longer to be reached.

4.1.2 Numerosity code a.k.a. thermometer representation, a.k.a. summation coding, a.k.a. monotonic coding

The variety of names epitomizes the popularity of this representation schema: a linear, discrete implementation of cardinal semantics. Originally proposed in Zorzi and Butterworth [187, 186], a thermometer representation encodes numerosity magnitude in a standard set theoretical way s.t. numbers are represented by sets and bigger numbers contains smaller numbers. In particular, a given number is defined as the set of active nodes, and for $N > M$ sets of active nodes, there is a $K \subset N$ s.t. M, K are in one-to-one correspondence. The size and distance effects are explained by referring to the similarity of the corresponding representation vectors. In particular, in the case of binary vectors, we have the following relations between the cosine similarity, the Tanimoto coefficient and the Jaccard coefficient.
Let N and M be two numerosity vectors, then

$$cos(\theta) = \frac{N \cdot M}{||N|| \cdot ||M||} \propto \frac{N \cdot M}{||N||^2 \cdot ||M||^2 - M \cdot N} = \frac{N \cap M}{N \cup M} = sim(N, M)$$

This gives us an extremely concise way to see the similarity between two numbers n and m, s.t. $n > m$ as m/n. The distance effect is readily explained. Let's defined the Jaccard distance as the complement of the Jaccard coefficient $dist(N, M) = 1 - sim(N, M)$. For an arbitrary number m and a set of successors of m, $N = \{m + 1 = n_1 < n_2 < ... < n_k\}$ we have $dist(m, n_1) > dist(m, n_2) > ... > dist(m, n_k)$, for any $k \in \mathbb{N}$. If we define the subject's Weber's coefficient as $w = WR - 1$, with $WR = \frac{q}{p}, q > p$ being the smallest ratio that results in a probability of discrimination greater than 75% (Just Noticeable Difference), then for every pair m, n the error rate is proportional to $dist(m, n) * WR - w$. Importantly the Weber's parameter w is contributed by the decision process. Similarly in Zorzi and Butterworth [186] the size effect is a byproduct of the decreasing weight connections learned via Hebbian Rule, s.t. for two pairs of numbers with the same ratio, the larger are connected to the output node (in the response system) via a weaker connection. Consequently the time the net needs to cycle to win the competition increases for larger numbers.

4.1.3 Number line models (a.k.a. place code, a.k.a. analog magnitude)

Analog magnitude representations predate the literature, and it's often considered the standard model for numerosity representation. A tutorial overview of the model is available In the online support material, NumberLineModel.ipynb.

In [32] Dehaene narrates that the SNARC effect, the association of number and space, promoted the metaphor of an oriented "number line". The reports of subjects' capacity of automatically visualize numbers as disposed in a line, and the positioning of number in a given line, are interpreted as decoders of an inner "number line", a conscious and enriched version of it. Less metaphorically, within the field of signal detection theory (SDT, see Macmillan and Creelman [98]), it is assumed that numerosity judgment follows Thurstone's law of comparative judgment (Thurstone [165]). Briefly, the law states that a stimulus triggers a process in which the external stimulus value n is transformed into some value on an internal *psychological continuum*.

Appreciating the model requires a little digression in SDT. The basic components of STD are a sensory process (which transforms physical stimulation into internal sensations/representations) and a decision process (which decides on responses based on this internal representation). In the simplest scenario, the response is a simple yes or no ("yes, the stimulus was larger" or "no, the stimulus was not larger"). From the hitting rate (HR) and the false alarm rate (FAR) detection sensitivity and response criterion are computed. The specific way in which they are computed from the HR and FAR depends on the specific model one adopts for the sensory process, and for the decision process. The leading model is the Gaussian model, according to which the sensory process is assumed to have a continuous output based on random Gaussian noise and that when a signal is present it combines with that noise (i.e. it is assumed that noise and stimuli are i.i.d.). By assumption, the noise distribution has a mean, μ_n , of 0 and a standard deviation, σ_n, of 1. The mean of the signal-plus-noise distribution, μ_s , and its standard deviation, σ_s, depend on the sensitivity of the sensory process and the strength of the signal. The observer is credited with computing the log posterior ratio of two alternatives, and the decision arises by comparing its value to a criterion. The noise in the observation used to compute the log posterior ratio determines the observer's errors.

Coding numerosity representation as Gaussians was originally proposed in Van Oeffelen and Vos [176], and it has been extended during a 20 years period and put together in Dehaene [30].

The 'internal representation' of numerosity is represented as a random variable X with distribution

$$p(X \in [x, x+dx]) = \frac{1}{\sigma_n \sqrt{2\pi}} exp - \frac{(x-q_n)^2}{2\sigma^2}$$

The choices of the standard deviations σ_n and of the mean q_n are restricted by

Weber's law. The simplest scenario, assumed in Van Oeffelen and Vos [176] and taken up in Dehaene [33, 31], assumes the internal variability to be the same for any n, ($\sigma_n = \sigma$) that take the value of the Weber's ratio, therefore $q_n = log(n)$. Briefly this assumption is referred to as logarithmic scale and fixed variance or as compressed number line model (Figure 4.1[b.]). Alternatively, as proposed in Gallistel and Gelman [50], assuming the internal variable scales linearly ($q_n = n$) requires that also the internal variability scales linearly ($\sigma_n = \sigma * n$), (Figure 4.1[a.]).

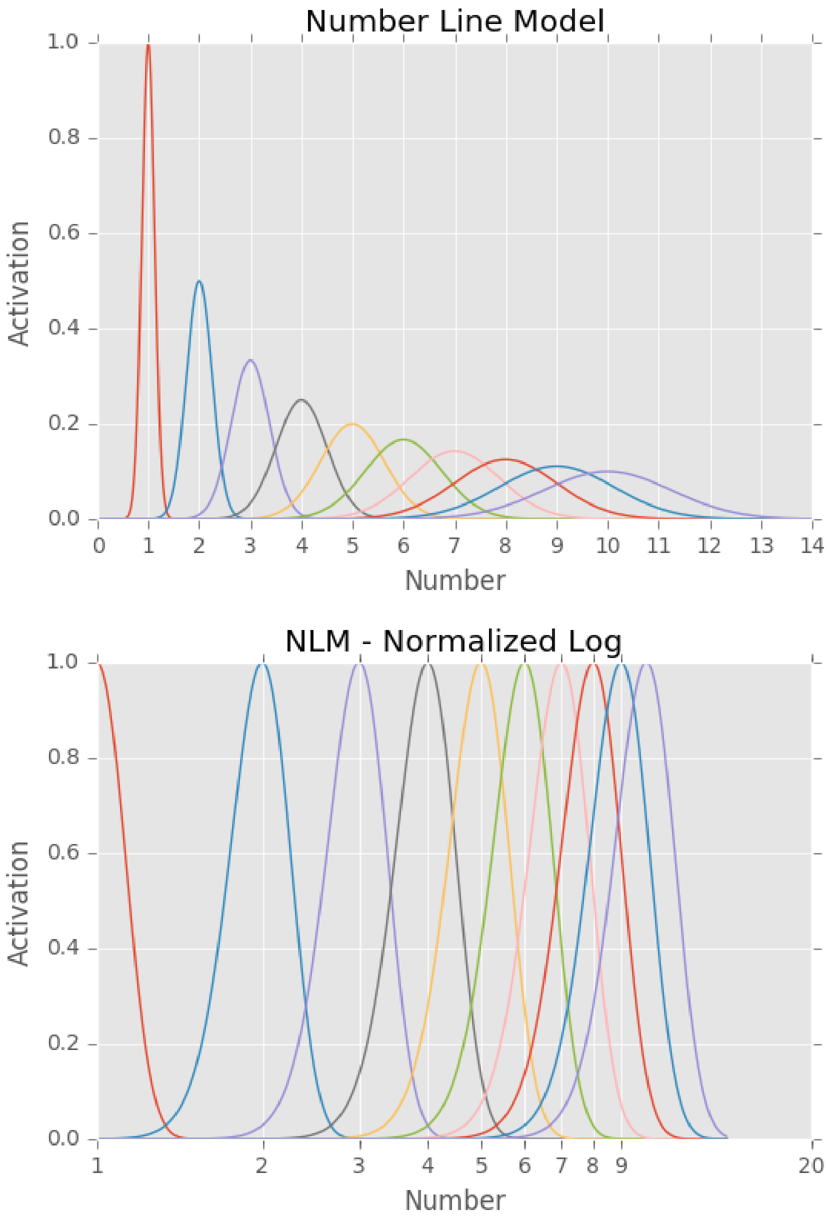
When taken as phenomenological models, the two choices make practically identical prediction w.r.t. discrimination and comparison behavior. The Gaussian hypothesis is moreover transferred to neuronal representations. In Dehaene and Changeux [35] for example the neuron tuning curves (the encoding part of the representation) mirror exactly the internal representation. For a neuron that responds preferentially to a numerosity p, the tuning curve is given by

$$f(n,p) = \frac{1}{w\sqrt{2\pi}} exp - \frac{(\log(n) - \log(p))^2}{2w^2}$$

The only difference w.r.t. the behavioral model lies in the exact value of the Weber ratio w (the coarseness of the representation).

Subitizing

There is some debate over whether subitizing and estimation might be explained by a common mechanism. In Vetter, Butterworth, and Bahrami [180] is proposed that the ANS mechanism might explain the subitizing phenomena, given that neighboring numerosities are more discriminable in the small number range. This implies that the subitizing range is dependent on the Gaussian standard deviation (Weber coefficient), therefore subjects with more acute number sense should have a larger subitizing range, and infants a smaller one. In a study by Revkin et al. [135]), aimed at testing this hypothesis. However, the authors found particularly low variability in the 1-4 range, suggesting a separate mechanism is indeed in place. As implied by Sengupta, Surampudi, and Melcher [153] study, the inhibition parameter alone in the decision system, starting from a thermometer representation, can explain the subitizing phenomena.



[a.]

[b.]

Figure 4.1: Number Line Model linear scale and normalized logarithmic scale. Figure b) makes apparent that the width of the Gaussians is constant when plotted on a log scale

4.1.4 Pattern recognition model

Mandler and Shebo [100] based the subitizing effect of the recognition of a canonical configuration. In a visual display, a small set of items, will form canonical configurations, that can be recognized in parallel. One is a dot, two items form a line, three items a triangle, and four items will look like a square (or a triangle with a dot inside). For the mechanism to operate, it's required that the items to be counted pop-out from the background, and must occupy distinct and easily identifiable positions in space. In the literature on subitizing, the pattern recognition model is mostly rejected. However the evidence against the model is weak[4]. The under appreciation of the model seems to be a lack of appeal when compared with the parallel individuation system representation we will see briefly. As we've pointed out in section 2.3.1 the PIS underlying the subitizing effect is assumed to be foundational to bootstrap the natural numbers, a subitizing mechanism that explains away the discreteness of the succession for a spatial geometric configuration therefore compromises the assumption. This is clearly not a sufficient condition to discard a model, instead it would seem to us a necessary condition to shakedown it properly, especially considering that it has recently regained attention and plausibility (see for example Krajcsi, Szabó, and Mórocz [83]).

4.1.5 Object-file/FINST representation

The subitizing range stroked for its resemblance with the multi objects tracking (MOT) ability[5] (Trick and Pylyshyn [171], Pylyshyn [131]). This prompted the hypothesis that the same mechanism that account for it might explain the subitizing range[6]. Trick and Pylyshyn [171] advocate for a parallel individuation system that identifies an object with a spatio-temporal address (analogically dubbed FINST, finger of instantiation). Once indexed in this way, the object can be represented in working memory by a mental token such as an object-file (Kahneman, Treisman, and Gibbs [76]). Working memory limitation, the number of available FINST, or both are then assumed as the cause of the tracking limitation and of the subitizing range. Notice that quantity is not represented, but implicitly in this account. There are no symbols for numbers (or magnitude), but a symbol (file) for each individual in the attended set. The pre-attentional ability to single

[4]These comprise, misinterpretation of the model, subitizing effect in other modalities, simultagnosic subjects studies, rebuked in Krajcsi, Szabó, and Mórocz [83].

[5] The simultaneous visual tracking of several moving items.

[6]The reader is informed that both theories are highly debated. In the mathematical cognition community the MOT range is taken as a basis for the subitizing range. In the attentional community the subitizing range is taken as evidence for the MOT signature

out up to four objects, therefore, accounts partially for the faster reaction time by cutting off visual search time. The parallel individuation system (PIS) is assumed to enable the subject to form mental models of small collections of items (3-4). The hypothesis is that when the first three or four numerals are learned they are connected to the pre existing PIS mental models. However how this is achieved is unclear, and often at odds with the infants' abilities. Trick and Pylyshyn [171] propose that the infant counts the FINSTs, this requires the toddler to be able to take set theoretic operations as Union and Powerset, and the ability to count the element on any subset. In Feigenson, Carey, and Hauser [43] and Feigenson and Carey [42], the ability to subitize is accounted for infants as young as 12 months, an age in which they might not have yet acquired these abilities, in particular the counting routine. This seemingly contradictory assumption is avoided by separating an implicit counting procedure, already available innately, from an explicit one learned (a strategy reminiscent of Gallistel and Gelman [51] proposal, we mentioned in **??**, from which inherits the same drawbacks). Alternatively, the counting procedure might be replaced by the ANS output which tuning curves are enough separated to be easily distinguishable. However in this latter case, it seems that the PIS doesn't give any gain. An alternative account was proposed in Le Corre and Carey [86], where it is suggested that the "fullness" of one's object files can be directly associated with a specific cardinality, without invoking the ANS or any other mechanism to tally the occupied object files.

The assumption that the subitizing and MOT range is derived from an architectural constraint has yet not received a neurophysiological counterpart[7], and might be at odds with recent behaviorally findings.

Franconeri, Jonathan, and Scimeca [47], building upon previous work, suggests that the limit on object tracking is not achieved by a fixed number of pointers (Pylyshyn [131]) or by a variable number of pointers (Alvarez and Franconeri [3]), but that there is no limit per se on the number of objects that can be tracked in parallel, and that it's spacing that influence the number of objects that can be tracked at once. The hypothesis is still under investigation (Bello, Bridewell, and Wasylyshyn [12]), and to our knowledge it hasn't yet received attention in the mathematical-cognition community.

[7]With only a small support from a neuropsychological study of neglect patients, ([182]), counterbalanced by one brain-imaging study that found instead that human parietal cortex activation increased linearly with the number of items (Piazza et al. [121]

4.2 Computational Models

4.2.1 Deahene & Changeaux, (1993)

In [35] numerosity detection system visual objects are presented as simplified one-dimensional input normalized to a size-independent code. The activation are then summed up and sent to numerosity detectors which activity mirrors the number line code. Importantly, the model rests on the postulate that numerosity detection is present at birth and therefore, was hardwired by a pattern of on-center off-surround units.

The model

Figure 4.2 The input is a vector of length 50 simulating a “retina”. Each object is coded as a local Gaussian distribution so that the standard deviation of the distribution corresponds to the object’s size. Up to five objects are randomly clumped to the input vector avoiding overlap and touching (therefore, the maximum size of object is 9). Inputs are projected into a (50×9) location map. The connections to the location map are set such that the receptive fields approximately mimic one-dimensional difference of Gaussian (DOG) filters ($\sigma \in \{1,\ldots,9\}$). Each blob in the visual input, therefore, is detected by the clusters whose receptive field approximately matches its size. Lateral inhibition ensures that only few clusters will remain active in any given location. The location map, therefore, achieves a certain size invariance. Whilst in the input retina objects’ size was coded by the number of active clusters, in the location map different sizes are coded by approximately the same number of active clusters in different positions.

Each cluster in the location map projects, with equal strength, to a layer of ‘summation clusters’ with increasing threshold (monotonic coding). Summation clusters then project to ‘numerosity clusters’ whose connections are set with central excitation and lateral inhibition. In this way numerosity clusters respond only to a preferred numerosity.

Discussion

Although the model is hardwired, and it assumes that numerosity detectors are innate components of the system, it shows the feasibility of extracting, in parallel, approximate numerosity from a visual stimulus (although a one-dimensional simplification of it). The variance in the representation derives from the approximate normalization in the location map. The filters provide only an approximate match to the actual size of the input object, this in turn implies that more than one location cluster might be activated

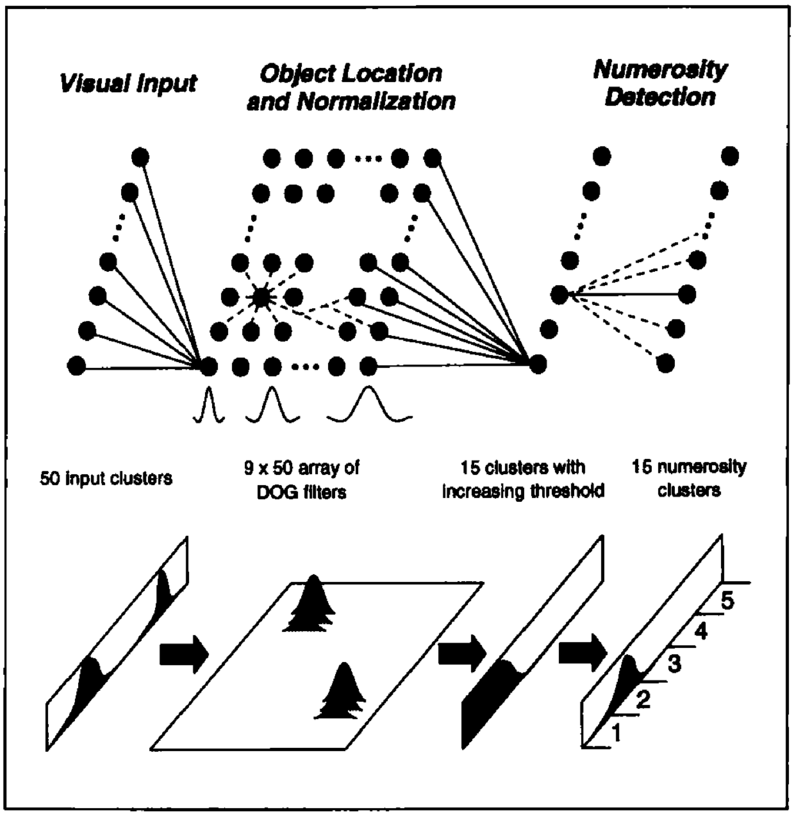


Figure 4.2: Numerosity detection

for each input object. As we will briefly discuss in chapter 7 the preprocessing stage is too coarse, and not easily generalizable. The last two parts of the network, the layers of summation and numerosity clusters, directly assumes both the monotonic and the number line code we reviewed in the previous section. Although learnability in a network and neuronal development are only loosely associable, it would be interesting to assess whether, and which kind of architecture, might learn the necessary connections. The model of Verguts and Fias [179] we are about to review will answer these questions.

4.2.2 Verguts & Fias, (2004)

Dehaene and Changeux's [35] model achieved the selection of the correct numerosities from the summation layer hard coding the weights. Given a summation field it's indeed not hard to see how to obtain this, it is sufficient that the corresponding weights are positive, and each link to a unit lower in the scale has a negative weight. In this way the winner unit is reached only by positive activations, whilst the other units are reached by positive as well as negative activations. With their unsupervised architecture Verguts and Fias [179] wanted to answer the question of whether number-neurons can develop from a thermometer representation.

Testing the thermometer representation coding

Dehaene and Changeux's [35] model assumes an intermediate layer of monotonic coded neurons between the numerosity detection system and the location map. This step was envisaged considering the fact that the object location map is number "sensitive", but not "selective". When more objects are presented, more units will activate and, as a consequence, there is more activity in the map as a whole. This information need to be converted into a number-selective coding, this entails a non linear transformation. More specifically given a spatial location coding, the task of the network is to output the correct numerosity of the inputs by activating the corresponding units (and inhibiting the other). The problem is a generalization of the XOR problem, and therefore, it cannot be processed by a single layer network. Verguts and Fias [179] trained a multi layer perceptron (MLP), with five input units, five hidden layer units and five output units, representing numerosity one hot encoded. The activation function used was a standard logistic function, and learning were achieved by back-propagating a least-square error loss function. Analyzing the hidden layer weight connections revealed a monotonic response. All connection weights to a given hidden unit were approximately equally strong. The emergence of summation units from an uncommitted general network shows that summation coding is a computationally natural solution to numerosity detection[8].

Unsupervised learning network

As for Dehaene and Changeux [35] model, it's assumed that a preprocessing stage extracts the objects in the visual scene exactly and normalizes them with respect to

[8] Although, how this is achieved in a biological network is still unaccounted.

a strong invariant. Given the results of the MLP simulation the inputs are encoded with a thermometer representation, where each input is encoded as two units[9] and normalized. Normalization is biologically plausible and allows to use the dot product as a measure for similarity. The activation is the classical linear function $\sigma_i w_i x_i$. The output layer, number field in Verguts and Fias terms (numerosity clusters in Dehaenes's terms), consists of 500 units. The learning rule is a slight modification of Kohonen learning rule where no topographic ordering is assumed

$$\Delta w_{ij} = \alpha \exp[-\beta(y_{max} - y_i)](x_i - w_{ij}).$$

Here $\exp[-\beta(y_{max} - y_i)]$ is the neighborhood function that defines how strongly the connection should be changed, where β define the size of the neighborhood.

Results

Detailed operations of the network are provided in the Jupyter notebook Competitive_VergutsFias.ipynb[10]. The network learns to represent the numerosity, and interestingly it does so embedding a number line like "representation".

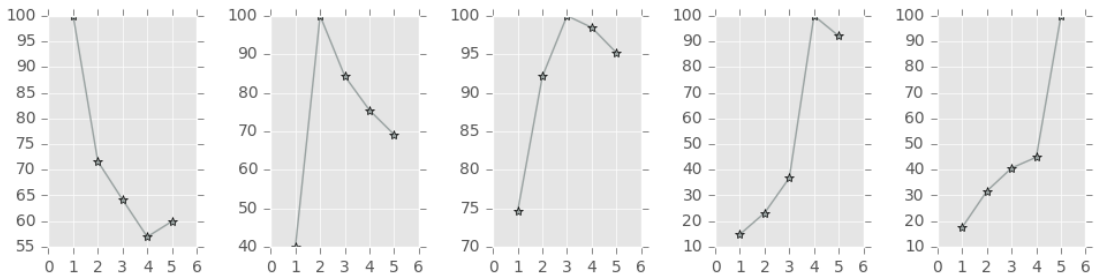


Figure 4.3: Average responses of the number selective neurons

Calling these neurons, "numerosity neurons" is tricky. The encoding of the input is indeed general enough to stand for any other feature that can be encoded via a thermometer representation. However this uncommitment on the initial input implies that a compressed coding can emerge in a quite natural way.

[9] This is quite confusing in the paper as an example given suggests otherwise. Personal communication to the author confirmed this natural assumption.

[10] Available at https://github.com/bramacchino/numberSense/blob/master/Competitive_VergutsFias.ipynb.

4.3 Summary

Diverse functional and computational models have been proposed to account for the behavioral observations we displayed in chapter 3. There is some debate over whether the subitizing effect and the ability to estimate might be explained within a common framework. With respect to the ability to estimate the number of objects within a scene, two major representations have been proposed: monotonic coding, and number line coding. The former encodes numbers as increasing pattern of activations, such that larger numbers activate more units. Number line coding emerges from the signal detection theory assumption that neuronal noise converts an external stimulus into some value on an internal psychological continuum. It has been suggested that subitizing depends upon our ability to recognize geometric shapes in stimuli consisting of few items. Although this hypothesis has been criticized, it has recently regained support, and more investigations are needed. The leading theory assumes that the ability to subitize comes from an architectural limit to the ability to track objects pre-attentively. Computational models of subitizing are scarce, and not really necessary in case the subitizing limit is achieved by an architectural constraint. After attentive consideration, we have decided to omit those models all together. However, the interested reader may have a look at [118] model, for an attempt to unify subitizing and approximation in a single architecture. The computational models for the ANS assume three processing stages: a location map, an intermediate layer of monotonic tuned units and an output layer of numerosity detectors embedding a number line code. As a whole, the architecture might be claimed to compute numerosity. However, the extraction process is in these models assumed and the last two layers are general enough to encode magnitude aspecifically. Although the models cannot show that numerosity units are biologically relevant, they show that compressed coding and summation units are computationally feasible solutions that might emerge via a learning process from an initial uncommitted network.

CHAPTER 5

NEUROPHYSIOLOGY

I am a brain, Watson. The rest of me is a mere appendix.

Arthur Conan Doyle, *The Adventure of the Mazarin Stone*

The computational models reviewed predict the existence of a hierarchy of several types of units: summation units and numerosity detector units. In this chapter, we will therefore, briefly seek for the neural representations that might be the biological counterparts. It's customary in the field to term these representations *number neurons* and we are going to do the same. Without fear of being repetitive, we stress that a less committed term would be 'magnitude neurons'. The direct terminology and the sometimes overexcited conclusions that flourish in the literature might, in fact, prompt a stronger confidence than what is justified. Nieder and Dehaene [111] provided an antidote in the section *Abstractness and Specificity of Number Coding in the IPS*. The take home message is that the term 'magnitude neurons' is more akin to be correct considering the fact that it is more plausible that 'number neurons' don't exclusively code for numerical information. From a philosophical standpoint, the fact that number, time, space are probably interconnected in our representations is extremely interesting and with a recognized Kantian flavor.

For a thorough appreciation of the chapter, a 3D functional brain atlas and connectome of the monkey brain might be helpful[1].

5.1 Number neurons

In neurophysiological terms, that is in signal processing terms, the two representations proposed correspond to two different families of filters. Summation coding is equivalent to a bundle of low-pass and high-pass filters. The number line coding is equivalent to neurons acting like bandpass filters.

5.1.1 Number line neurons: bandpass filters

In the seek of neurons, whose tuning curves resemble the number line model, Nieder and Miller [115] sampled 352 neurons from the dorso lateral prefrontal cortex (LPFC) of two rhesus monkeys. The monkeys were trained on a delayed match to quantity task (2IFC) on a standard set of stimuli. The monkey task was to release a lever if the sample and test stimuli matched in numerosity. Area, density, shape, border, geometrical arrangement covaried to exclude the results indebted to interfering visual features[2]. Approximately one third showed selectivity to numerosity, both during sample presentation and memory delay. Neural activity formed neural filter functions consistent with a "number line" representation of numerosity. That is, they fired maximally to a given number of dots, and showed decreasing firing rates when the numerosity was smaller or larger than their preferred value. Interestingly, asymmetries in the tuning curves were compatible with the log-Gaussian hypothesis of a fixed-width Gaussian tuning curve once plotted as a function of $log(n)$. As stressed in chapter 2, the encoding part of the representation is not enough to suggest that the animal is using this response to cognize about the external stimulus. Nieder and Miller therefore checked for the bandwidth ratio, obtained by dividing the neural sigma by the behavioral sigma, to infer the existence of a decoding procedure. The bandwidth ratio was constantly 1.5 showing less sensitivity on the neural representation than on the behavioral one. That the ratio is practically constant

[1]For example, see the freely available software Suma Anfi and the *macaque 3d atlas* provided by Reveley et al. [134] at https//afni.nimh.nih.gov, or, alternatively, the online services reviewed in Majka et al. [99].

[2]As we noted in chapter 3 this might be not enough. This remains the case at the neural level. Single neurons may be simultaneously tuned to several modalities and may be reached by different nets. What is extracted is the tuning function (not the response function) w.r.t. to external stimuli. However, unless the input hierarchy is given, this is not enough to associate the neuron exclusively with a given external signal. In particular, it's evident by inspecting table 1 in Nieder, Freedman, and Miller [112] that density and cumulative area cannot be mutually excluded.

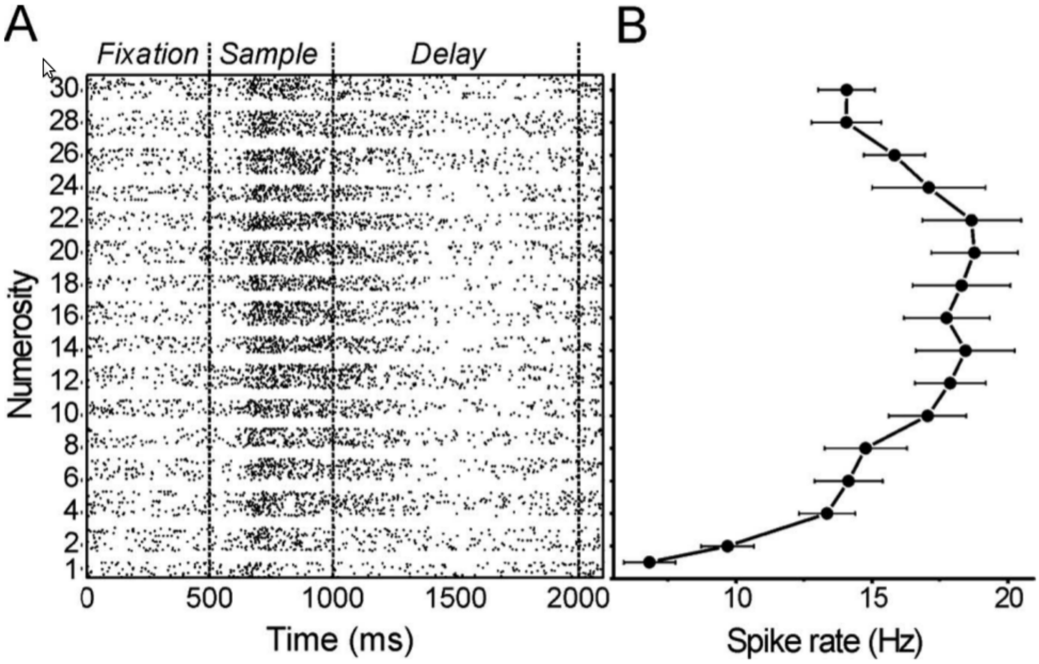


Figure 5.1: Raster dot plot and tuning function for neuron encoding 20

suggests a direct relationship between behavior and (neural) representation. The higher behavioral sensitivity is common in sensory physiology (e.g. Purushothaman and Bradley [130]) and might be explained by the *lower envelope principle* (Parker and Newsome [117]), according to which only the most sensitive neurons contribute to a decision. A limitation of this study was the small range of numerosities tested. To overcome this difficulty, Nieder and Merten [113] repeated the experiment collecting data for numerosities up to 30. This time the bandwidth ratio was almost three fold (2.8), which might suggest a limit of the neural representation capacity and an emergence of a different mechanism.

Although we are mainly focusing on visual numerosity, "numbers" are abstract entities. Thus, it's interesting to see whether the discovered representations are modality aspecific or specific. In Sawamura, Shima, and Tanji [152], the anterior part of the parietal association area was active in numerical representation for action task. Two snow monkeys were trained to select a movement (either *push* or *turn*) and repeat it five times (or four in a second experiment) before shifting to the other movement. To control for confounder, the time spent performing the block of five consecutive trials varied between twenty and forty six seconds for a particular movement. They found a focal region (of number selective cells) on a caudal portion of the superior parietal lobule

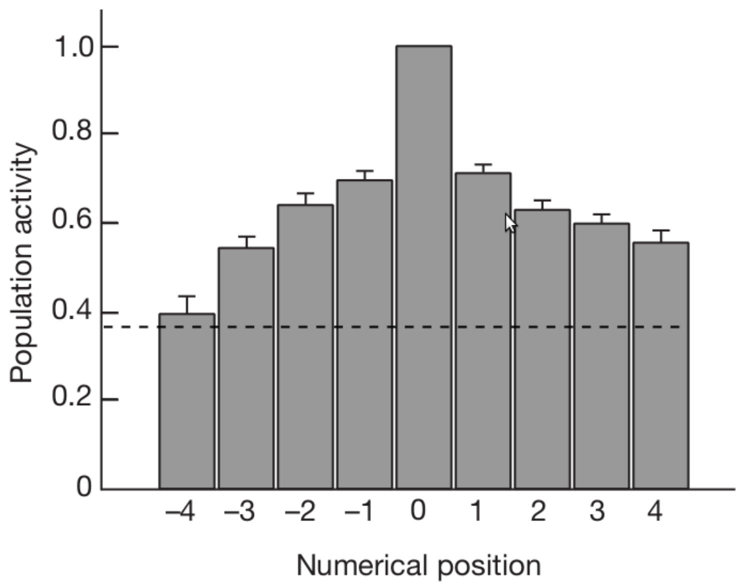


Figure 5.2: Population responses of superior parietal neurons in numerical positions relative to a position of peak response.

(SPL) in which neural activity was correlated with task performance[3].

Cellular selectivity was found too coarse to act as a bandpass filter, so they resorted to population coding. To see the activity profile of population responses relative to the peak activity for each neuron, they calculated the mean activity during performance of trials in each numerical position in a block, realigned relative to the most active position. By normalizing the activity relative to the peak values and constructing a population histogram, they found that the population responses showed unimodal distribution with a distinct peak (see Figure 5.2). The vast majority of neurons (85%) were not "abstract" number-selective neurons, but activity depended on whether the monkey's movement was *push* or *turn*.

The two studies, therefore, point two possible candidates for numerical representation in PFC and PPC. However, on this basis, and given the stimuli differences, it is not possible to infer their functional organization and respective contributions. In particular, it is not clear where in the cortical hierarchy numerosity is first extracted[4].

Thus, in a successive study, Nieder and Miller [114] analyzed the response of both

[3]Moreover, there was only a minor implication of inferior parietal lobule, somatosensory cortex, and frontal cortex.

[4] For example, it is known that the PPC, and the anterior inferior temporal cortex (aITC) provide the PFC with a major source of visual input, that in turn sends feedback projections to both (Stuss and Knight [163], see Figure 5.3). Therefore, "numerosity information" might be extracted in PPC, or aITC first and sent to the PFC for memory purposes or extracted in the PFC and then sent back to the posterior areas.

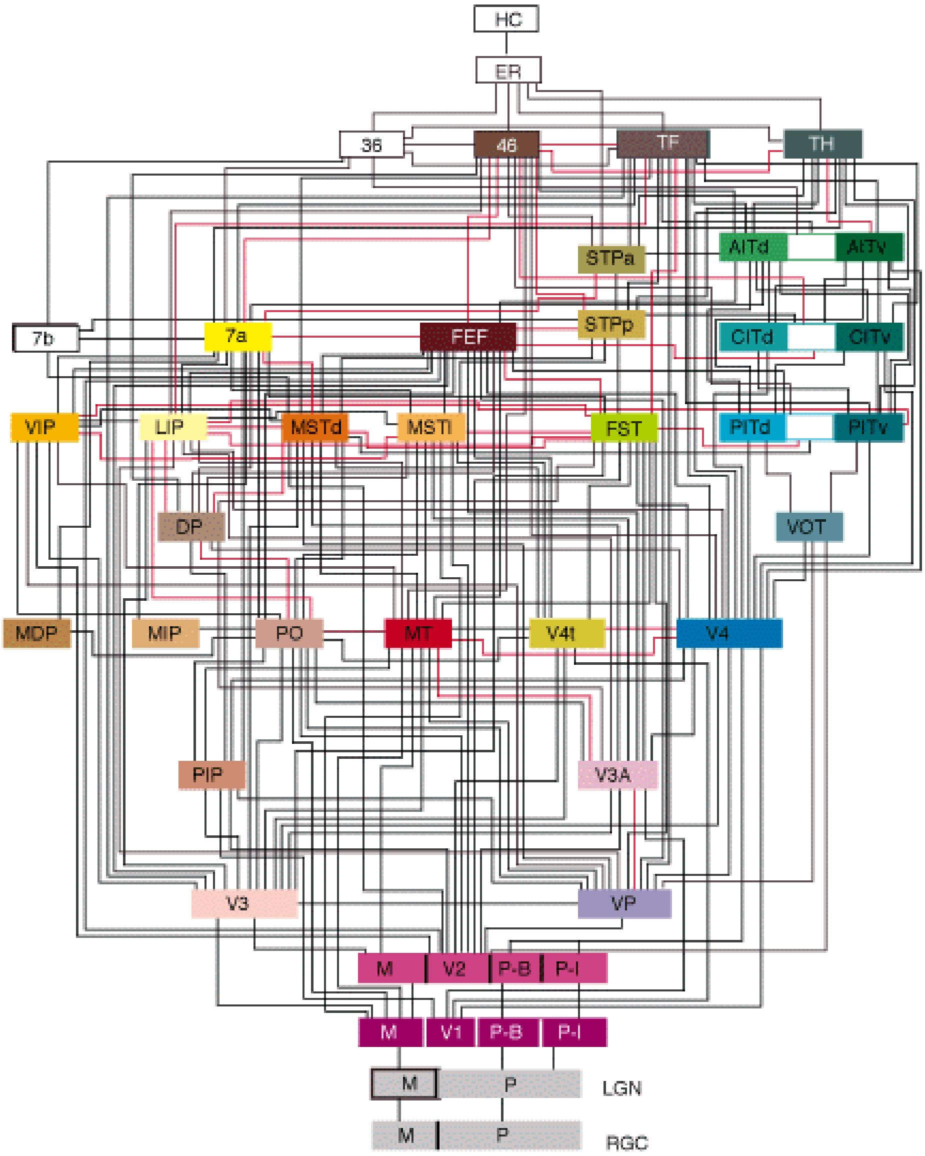


Figure 5.3: van Essen wiring diagram of the primate visual pathways

populations. They discovered that there were proportionally a higher number of neurons sensitive to numerosity in the F-IPS. Interestingly, the parietal cells differed from the prefrontal neurons in two ways: they had a significantly faster latency, and they fired less strongly during the delay. These findings suggest the F-IPS as the prime source of numerosity processing, at least with respect to the LPFC. The fact that neurons in the F-IPS were less affected by changes in stimuli appearance, whilst the tuning functions reported by Sawamura, Shima, and Tanji [152] were more modality specific, suggests moreover a more abstract representation in the F-IPS with respect to SPL.

A further analysis of the activity within the IPS has been carried on in Tudusciuc and Nieder [173]. They analyzed the response properties of individual neurons in the fundus of the intraparietal sulcus of monkeys simultaneously engaged in numerosity and length discrimination tasks. They found that single neurons might encode spatial quantity, numerical quantity, or both. The fact that those populations are intermingled in a restricted area of the IPS suggests a cross talk. Restricting the attention to the small population of neurons whose response was selective to numerosity, they found that, when spike rate was integrated by spike train coding, numerical behavior could be predicted. Although only a small range was tested (1-4), this result is suggestive. It might provide a neural basis for the SNARC effect (chapter 3), and indicates the plausibility of a decoding procedure cognitively available.

Although our divide and conquer approach suggests to capture a modality dependent numerosity first and then try to encompass other modalities (and this is what we are going to do in the next chapter), it's helpful to know where one can aim. Nieder [110] trained two rhesus monkeys to discriminate the number of sequentially presented auditory and visual items using the usual delayed match-to-sample protocol. Sequentially visual and auditory stimuli were to be matched against a multiple dot display. Given the difficulty of the task, only numerosities from 1 to 4 were tested. They discovered groups of neurons in the VIP and PFC that encoded either the number of auditory pulses, visual items, or both. In the PFC, 25 (11%) of the 42 (18%) auditory and 67 (29%) visual numerosity-selective neurons responded to both auditory and visual numerosity. In VIP, six neurons (3%) of the 20 (10%) auditory and 22 (11%) visual numerosity-selective neurons, responded to quantitative information bimodally. The proportion of neurons responding to numerosity irrespective of modality supports the idea of a most abstract, supramodal neuronal code of numerical quantity.

Somebody might say we are not just smarter monkeys. Checking if there is a human homologous of the putative number neurons was therefore tested in a series of experi-

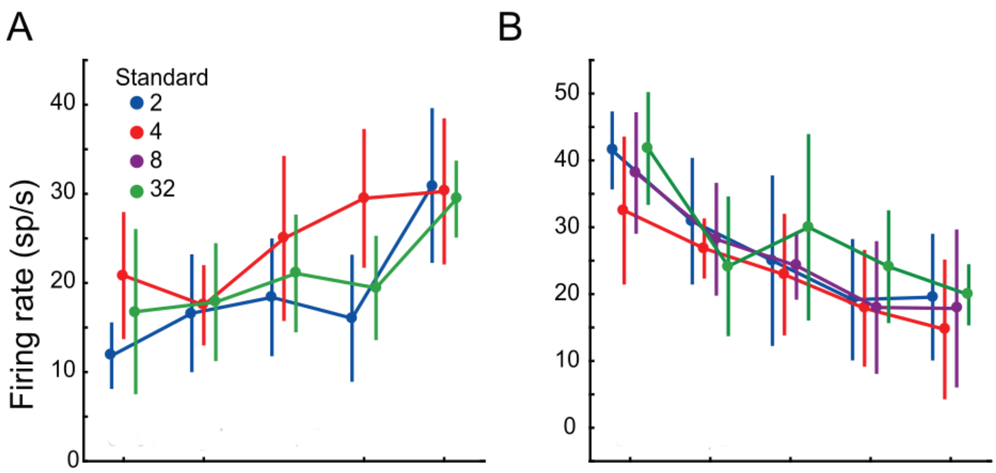


Figure 5.4: x axis numerosity from 2 to 32. y axis average spiking rate. A) represents a neuron positively tuned, whilst B) represents a neuron negatively tuned. The inset on the top left indicates the standard reward stimulus in the trial.

ments by Piazza and colleagues (Piazza et al. [125, 120]) following a behavioral paradigm in line with the one used by Nieder and colleagues in the mentioned studies. Since they could not record from single human neurons, they took advantage of the functional magnetic resonance imaging (fMRI) adaptation method. The paradigm consists of the repeated presentation of a habituated specific numerosity (either 16 or 32) to habituate neurons tuned to that number, so to elicit suppression, then displaying an occasional "deviant" number (half or double of the habituation number), which elicits an activation increasing with the numerical distance between adaptor and deviant. That is, the the Gaussian evoked by the deviant stimulus allowed to read out the state of adaptation, providing in turn tuning curves for numerosity in the hIPS.

5.1.2 Summation neurons: high/low pass filters

Most models in the previous chapter are dependent on this kind of coding. Indeed, the last step in both Dehaene and Verguts' models is reduced to a simple classification problem. Something ANNs are extremely capable of managing. Moreover, the graded, bell shaped, response might be dependent on the response of this summation neurons. It has been proposed that LIP neurons function as neural integrators (Mazurek et al. [103]). Roitman, Brannon, and Platt [145] hypothesized that the activity of LIP neurons would encode the quantity of visual elements placed within their classical receptive fields (RF) in a graded manner, independent of low level visual features. This extends the idea of neural integrators such that these neurons accumulate numerical quantity as well.

The monkeys' sole task was to shift gaze to a visible target following the disappearance of the fixation point, testing an implicit sense of number. The subtle reaction time difference observed, however, suggests the monkeys were attending numerosity despite no training nor requirement to do so. LIP neurons are particularly spatial selective, therefore the target array of numerosity 2, 4, 8, 16, 32 was totally within the RF of the LIP neurons, hence in the visual periphery, and shown for 400ms. The saccade target was located in the opposite hemifield (outside the neuron RF). Out of 57 neurons, 14 had a positive relation and 17 had a negative relation between numerosity and response, showing graded modulation. Moreover, for all neurons the initial positive modulation was followed by numerosity related activity, analogously to temporal integration, suggesting that LIP neurons might integrate the number of objects within their RF.

Investigation of summation coding in human homologous of LIP has been presented by Santens et al. [149]. The ROI analysis showed numerosity sensitivity in bilateral posterior superior parietal areas (functionally correspondent to monkey LIP) therefore suggesting the possibility that a summation coding like the one found by Roitman, Brannon, and Platt [145] might be housed in those regions. The small range of numerosities tested (1 to 5) gives few details on the limit of such system. Tentatively, from the fact that it has been suggested (Gottlieb [58]) that LIP provides a topographical salience map representation of the stimulus, it might be hypothesized that the summation coding neurons act on the basis of this map.

5.2 Summary

We have reviewed suggesting evidence pointing at possible neural substrates of the units which the computational models predict. The process that converts the visual input into a number selective neural code remains elusive, and therefore, we left the term 'numerosity' undefined. Neurons that are sensitive or "selective" for numerosity have been found in various areas of the monkey's brain, with particular prominence of the LPFC, IPS, VIP and LIP. fMRI adaptation studies found homologous areas in the human brain. The computational models predicted a hierarchy of units: object mapping, summation units, and numerosity selective units. Object mapping might be tentatively associated with saliency mapping that might be extracted in the LIP, the same area where summation units have been reported. Numerosity neurons that respond abstractly have been found in the fundus of the IPS. Considered together, these findings suggest that the posterior areas process a general sense of magnitude with both selective units and aspecific units intermingled. Whether numerosity from natural images might be extracted in the same way remains an open issue. Whilst the tuning curves provide an encoding procedure, the decoding part of the representation has only been hypothesized assuming the behavioral comparison. That is, the studies have mainly focused on stimulus evoked representations (encoding) and has not yet been addressed the issue of whether the same representations are recruited when internally generating and manipulating numerical magnitudes.

Chapter 6

Numerosity as topological invariant

> We chose to investigate connectedness because of a belief that this predicate is nonlocal in some very deep sense; therefore it should present a serious challenge to any basically local, parallel type of computation.
>
> Marvin Minsky and Seymour Papert, *Perceptrons*

Recall the definition given in the introduction:

"numerosity refers to the empirical property of cardinality of sets of objects or events".

At first, the meaning appears to be clear. However, this definition, as it stands, doesn't make sense unless one takes somehow a platonic stance on what sets are. For sure, that's not the sense the definition wants to convey. It's difficult to see how it can be useful to capture a "general property of a perceptual process". From one side, the abstract notion of a set and the arbitrariness of the construction of numbers are too weak to speak about the numerosity of objects. From the other side, the cardinality of a set requires a stronger prerequisite, i.e. the distinction between a set and its elements, that cannot be assumed given an image.

We are thus left with a worrisome question: what are the perceptual primitives of visual perception? We can identify two main positions.

1. The *computational approach*[1] holds that perceptual processing is from local to global. In Marr [101], for example, primitives are simple components like zero crossings, edges, boundaries and curves.

2. The *early holistic registration approach*[2] can be summarized with the slogan "the whole is more than the simple sum of its parts", that is holistic registration is prior to local analysis.

Depending on the stance, therefore, what might be the visual cues from which the number of objects in a scene is inferred may differ significantly.

The difficulty lies in the fact that, whilst the approach taken in the field is closer to the first view, what is meant by *object* is more in line with the holistic registration approach. Given the informality of the definition, this is not a surprise: independently from the order of application of global properties in the processing of visual inputs, the end point might be assumed to be the same. In accordance with the Gestalt's view (and common sense), an object can be identified as a connected region of the space. The Gestalt's view, moreover, maintains that connectedness is one of the most significant criteria by which the visual system decides whether an element belongs or not to a single coherent region. This should imply that the visual system is well adapted for its detection and, therefore, that "objecthood" (when identified as connectedness) might be a feature sent to the number system.

According to the feature integration theory, however, connectedness seems not to be pre attentive available[3], that is computing connectedness from low level features requires an intrinsic serial, attentional, component. That topological properties and, in particular, connectedness have high computational complexity was a foundational part of Minsky and Papert's [106] influential book. Therefore, it is natural for the computational approach to expect that discrimination based on topological properties would occur at a higher level of perception. At the end of the last chapter, we have suggested that a saliency map might be encoded in the lateral intraparietal sulcus, and that numerical information might be extracted from it. A saliency map, however, needs to be normalized for size and it's blind to topological transformations that we have seen in chapter 3 affect numerosity estimation, such as connecting or enclosing dots. A direct descendant of

[1]This approach is exemplified in Marr [101] and by *feature integration theory* (FIT) in Treisman and Gelade [169] and Treisman and Souther [170].

[2]This approach is exemplified by the Gestalt's view (Rock and Palmer [142] and Palmer and Rock [116]) and recently resumed in the perceptual topological view (Chen [23]).

[3]Cf. *experiment 11* ("Topological properties: connectedness and containment") in Treisman and Souther [170].

FIT with attention to Gestalt's properties is the Incremental Grouping Theory (IGT) proposed in Roelfsema [143] and Roelfsema and Houtkamp [144], where pre attentive base grouping is separated by attentive incremental grouping achieved via a neuronal labeling strategy. Ultimately, whether parallel algorithms for solving the connectedness problem may be found remains an open theoretical question. Moreover, it's an empirical issue whether reaction times in numerosity estimation are independent from the number of objects in the stimulus, when the objects are not simply dots within the receptive field. If this turns out to be the case, then, to compute the number of connected components in a discrete space, we can borrow the analysis from the field of digital topology, the research area dealing with the computation of topological invariants in image processing. The class of algorithms we are interested are dubbed 'connected component labeling' algorithms (CCL). The strategy adopted by CCL algorithms can be summarized in three steps:

1. *first labeling*, that assigns a provisional label to each pixel,
2. *label equivalences solving*, that is finding all equivalent labels and building an equivalence table,
3. *final labeling*, i.e. to replace temporary labels by the final label (usually, the smallest one in the equivalence table).

Thus, the number of connected components results as a by-product from the number of assigned labels[4].

If numerosities estimation time is independent from the number of objects in the stimulus, then either this is an indication towards the "holistic view' or, within FIT or IGT (and assuming no bottom up parallel algorithm for connectedness), this might suggest to associate the numerosity extracted from the size invariance saliency map (location map) with the Euler's characteristic, i.e. the number of simply connected components (without holes)[5].

For these reasons, we maintain that a topological definition is in accordance with both theories and it's closer to the intuitions of the researchers.

[4] Most CCL algorithms show a direct runtime dependence on the number of objects. In He, Chao, and Suzuki [63], for instance, the execution time for calculating the number of connected components is proportional to the number of provisional labels assigned to connected components, whose order is $O(N \times M)$.

[5] Although neurobiologically plausible, this proposal is at odds with explaining the size interference effect reported in section 3.7. In particular, the results in Ross [147] suggest that grouping by size precedes the operation of a number sensitive mechanism.

The definition provided at the beginning of the chapter therefore can be reformulated as follows:

> "visual numerosity refers to the empirical property of the number of connected components in an image (when seen as a topological space)."

6.1 Discussion

Whether this suggestion can be extended to the tactile and auditive capacities is an empirical question[6], but connectedness, and similarity, should be properties broad enough to be at play.

Seeing numerosity as a topological invariant instead of as a synonym for cardinality is not to be considered simply as a a more coherent mathematical formulation.

Defining numerosity as a topological invariant has an immediate empirical advantage. In chapter 3 we left the reader with a promise: a new definition of numerosity and the interference effects might help in resolving the debate concerning the nature of the ANS representation. Understanding numerosity become an investigation of the specific invariants the perceptual processes are sensible to. In chapter 4 we noticed that the two final layers of Dehaene and Verguts' architectures were general enough to code a general magnitude. For example a network that compares the length of two shapes might be endowed with the same systems: monotonic coding encodes directly the length of the shape, and 'numerosity' detectors are triggered by a preferred amount of activation on the preceding layer of thermometer organized units (cf. Tudusciuc and Nieder [172]). Whether and how objects are extracted and preprocessed is therefore the most important part concerning the debate. If deviations with respect to the numerosity invariant (here assumed provisionally to be connectedness) are predicted by a mechanism as the one assumed in Dehaene and Changeux [35], then speaking of a 'true sense of number' wouldn't be too stretched out. If, on the contrary, the extraction phase proceeds by means of a different mechanism, for example from texture base mechanisms, and the deviations from invariance are predicted by such a system, then speaking about a sense of number might be nothing more than a vivid metaphor.

[6] However, there is no principled objection. See, for example, Plaisier, Van Polanen, and Kappers [126] about haptic object perception.

6.2 A formal approach

The standard experimental settings are 2D still images on a screen. The subject is required to assess the numerosity of the objects, that is the number of connected components.

Both the screen images and the retinal sampling are discrete. For the moment we are assuming they are approximation of a continuous function[7]. We represent 2D pictures on a screen via their binarized luminance function, so that they can be represented as a two-dimensional vector space $\mathbb{R}^2$. Let's define a configuration C of objects O_i such that we have a function (objecthood) that assigns to each point $(x, y) \in I \subset \mathbb{R}^2$ the value 1 if the point is within an object boundary, and 0 otherwise. Let's moreover assume that the background is the complement of the union of all object regions (the points where the objecthood function assigns 0). Moreover, any two objects are "perceptibly separable", that is no two objects overlap or come in contact within a given tolerance (greater than the visual acuity). We therefore have a configuration C defined to be $C = \bigcup_i O_i$. We are interested in the property 'numerosity' of a configuration $N(C) = |\{O_i | O_i \in C\}|$. Given the image $I \subset \mathbb{R}^2$, we don't have access to O_i, but only to the union (all the points the objecthood function assigns 1 to), therefore the standard set theoretical concept of cardinality cannot be used directly in this context. Nonetheless, we want N to be invariant to any affine geometric transformation, position, orientation, size and change in shape (let T be the class of such transformations, we want $N(C) = N(T(C))$). That is, N is sought to be a topological invariant. The number of connected components is such an invariant, therefore what we are seeking is a precise way to define this notion in a computably feasible way.

6.2.1 Homological approach

We are therefore interested in the number of connected components in a stimulus. In particular, we wish to characterize N such that we can compute the number of connected components in an image. Seeing an image as generated from a topological space, and therefore computing directly from the topology the number of connected components is unwieldy, way too abstract to perform computations. Resorting to homotopy we can count the number of holes in a space (and therefore, the number of connected components, as the number of holes plus one) by computing the fundamental group (but again, although this gives us a nice characterization, it doesn't provide us with a nice computational

[7] In 6.2.3 we will take into consideration the discrete nature of the visual stimulus.

tool). For this reason it's helpful to resort to a commutative, combinatorial, alternative to homotopy: simplicial homology and give the characterization of N within this theory. This will require us a whirlwind of simplicial homology, that we will briefly introduce in the following[8]. Moreover, with a suitable choice of definitions, the algorithm to compute the homology groups uses standard linear algebra techniques, making computations easier[9].

6.2.2 Simplicial homology primer

The idea behind homology, not different from the one underlying homotopy, is to distinguish topological spaces by counting holes in a space (those are loops not contractible to a single point) by associating to each topological space a family of groups such that whenever two spaces are homeomorphic then their associated groups are isomorphic. Computing the associated groups might be unwieldy in the general case, but it becomes easier if the associated groups take the form of finitely generated abelian groups. To increase tractability, and to give a combinatorial structure to a topological space, so that it becomes easier to manipulate, the topological spaces of interest will be simplicial complexes[10].

We therefore start with simplexes, the building block of simplicial complexes, namely a n-dimensional generalization of triangles.

Definition 6.1. Given any linearly independent set $V = \{v_0, v_1, ..., v_n\}$ of $n+1$ points in $\mathbb{R}^{n+1}$, the *n-simplex* with vertices in V is the convex hull of V (i.e. the set of all points of the form $\lambda_0 v_0 + \lambda_1 v_1 + ... + \lambda_n v_n$ where $\sum_{i=0}^{n} \lambda_i = 1$ and $\lambda_i \geq 0$ for all i).

Low dimensional examples, and the ones we are mainly interested, are easy to visualize: for $n = 0$ we obtain the one point space $\{v_0\}$, for $n = 1$ we find the line segment joining v_0 and v_1, for $n = 2$ we find the triangle with vertices v_0, v_1, v_2 and for $n = 3$ we get the tetrahedron with vertices V.

Simplexes of any dimension can be 'glued' together to construct a simplicial complexes, that is a simplicial complex is a topological space realized as a union of any collection of simplices S such that any face of a simplex σ is also in S and the intersection of any two

[8] We will provide only a bare overview, for the main results and proofs of homology theory we refer the reader to Hatcher [61].

[9] The linear algebra needed is not a lot and we hint the reader to Strang [161], and Strang [162] where applications are discussed.

[10] This latter restriction is of no particular concern to us, given that all closed surfaces are triangulable spaces.

simplices is also a simplex. It's convenient to give a pure combinatorial description of a simplicial complex without worrying on how to put it into Euclidean space.

Definition 6.2. A (abstract) simplicial complex K consist of a pair (V,S) such that:

1. V is a set of vertices
2. Σ is a set of finite non-empty subsets of V.

The elements of Σ, called simplices, satisfy the following conditions:

1. if σ is a simplex $\sigma \subset V$ and $\tau \subset \sigma, \tau \neq \emptyset$, then τ is also a simplex $\tau \in \Sigma$
2. for each $v, v \in V$, the singleton $\{v\} \in \Sigma$

We say τ is a face of σ. If σ has $m+1$ elements it is said to be a $m-simplex$. The set of $m-simplices$ of K is denoted by K_m. The dimension of K is the largest m, such that K_m is not empty.

The orientation of the edges is given naturally by the implicit ordering of the vertices of a simplex (e.g. $\{i,j\}$ has orientation (i,j) if $i<j$ and (j,i) otherwise).

The process of realizing a topological space into a simplicial complex is called triangulation. That is given a topological space X and a simplicial complex K[11], a triangulation is a pair (K,h) s.t. $h : X \to K$ is a homeomorphism. Therefore the topology of a triangulable space doesn't change on how we realize the space as a simplicial complex (the composition of homeomorphism is an homeomorphism). This gives us quite a lot of freedom in how we triangulate a space, once we disregard complexity concerns[12].

Recall that we wish to characterize loops. Resorting to simplicial complexes gives as a natural way to define boundaries. Intuitively a boundary of a simplicial complex is just the set of simplexes that surrounds a face. And a boundary itself has no boundary. This property of being boundariless coincides with our intuitive idea of what it means to be a loop. We therefore need just to recast this intuition in an algebraic form. As a first step we represent simplicial complexes as algebraic objects as follows.

Definition 6.3. Let X_k be the set of $k-simplices$ in the simplicial complex X. We define the *chain group* $C_k(X)$ to be the $\mathbb{Q}-vector$ space with X_k for a basis. The elements of the $k-th$ chain group are called $k-chains$ on X.

[11]Notice that K is a topological space as well.

[12]Whether the triangulated space might be isomorphic to a 'functional space' is an empirical, or at least computational, question, that need to be carried out by neural informed models.

For example $C_0(X)$ is the linear span of the set of vertices V, with coefficient in $\mathbb{Q}$, and $C_1(X)$ is the $\mathbb{Q}-$ linear span of the set of edges (that might be geometrically visualized as a 'path') .

Let's introduce some shortcut notation: given a k-simplex $(v_0, v_1, \ldots, v_k)$, we write $(v_0, v_1, \ldots, \hat{v_i}, \ldots, v_k)$ to indicate the removal of the i-th vertex, as an abbreviation for $(v_0, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k)$.

Definition 6.4. We define the boundary operator on chain groups the linear map $\partial_k : C_k(X) \to C_{k-1}(X)$ on the standard n-simplex with orientation $(v_0, v_1, \ldots, v_n)$ via the alternating sum

$$\partial_k((v_0, v_1, \ldots, v_k)) = \sum_{i=0}^{k} (-1)^i (v_0, v_1, \ldots, \hat{v_i}, \ldots, v_k),\rangle$$

Despite the appearance this definition is quite natural, and by inspection we realize that $\partial\sigma$ changes sign whenever two adjacent vertices are interchanged. For example if we take the simplex $S = (v_0, v_1, v_2)$, $\partial(S) = (v_0, v_1) + (v_1, v_2) + (v_2, v_0) = (v_0, v_1) - (v_0, v_2) + (v_1, v_2) = \sum_{i=0}^{2} (-1)^i (v_0, \hat{v_i}, \ldots, v_2)$.

By extending ∂ linearly on chains we get the operator on the entire chain group. Notice that this characterization agrees with the informal argument, every chain that is a boundary of an higher dimensional chain is boundariless. For example if we take a boundary of a 2-simplex we get a cycle of three 1-simplices, which boundary is 0.

We can therefore state the fundamental theorem that drives the definition of homology for simplicial complexes.

Theorem 6.1. $\partial^2 = \partial_{k+1} \circ \partial_k = 0$

Proof. The proof is easy, it just involves some algebraic manipulation, and it's intuitively obvious. In fact it is sufficient to realize that the second time we apply the boundary operator, we are shifting the power of negative of one, by one index, and therefore the alternating sum cancels out.

Clearly for $k < 2$ the result is trivial, suppose then $k \geq 2$ then

$$\begin{aligned}
\partial_{k-1}\partial_k(\langle 0, 1, \ldots, k\rangle) &= \sum_{i=0}^{k} (-1)^i \partial_{k-1}(\langle 0, \ldots, \hat{i} \ldots, k\rangle) \\
&= \sum_{i=0}^{k} \sum_{j=0}^{i-1} (-1)^{i+j}(\langle 0, \ldots \hat{j}, \ldots, \hat{i}, \ldots, k\rangle) + \sum_{i=0}^{k} \sum_{j=i+1}^{i-1} (-1)^{i+j-1}(\langle 0, \ldots \hat{i}, \ldots, \hat{j}, \ldots, k\rangle \\
&= 0
\end{aligned}$$

■

Definition 6.5. Let K be a simplicial complex, a $k-chain$ C_k is said to be a $k-cycle$ if $\partial_k C_k = 0$, and we call them boundariless. A $k-chain$ C'_k is said to be a $k-boundary$ if $C'_k = \partial_{k+1} C_k$ for some $k+1$-chain C_k. We denote the group of $k-cycles$ of K, $Z_k(K)$, and the group of $k-boundaries$ of K, $B_k(K)$ [13]

$Z_k(K)$ is thus the kernel of the boundary map $\partial_k : C_k(K) \to C_{k-1}$, and therefore a subgroup (actually a subspace) of $C_k(K)$ (since Kernels are always linear subspace). $B_k(K)$ is the image of the boundary map $\partial_k : C_{k+1}(K) \to C_k$. Given 6.1 (every boundary itself is boundariless) $B_k(K) \subset Z_k(K)$ and since the image of a linear map is always a linear subspace of the range, we get that it is a subspace of $C_k(K)$ too.

We might represent these relationships by the following diagram.

$$\cdots \xrightarrow{\partial_4} \underset{\substack{\text{3-d}\\\text{chains}}}{C_3} \xrightarrow{\partial_3} \underset{\substack{\text{2-d}\\\text{chains}}}{C_2} \xrightarrow{\partial_2} \underset{\substack{\text{1-d}\\\text{chains}}}{C_1} \xrightarrow{\partial_1} \underset{\substack{\text{0-d}\\\text{chains}}}{C_0} \xrightarrow{\partial_0=0} 0$$

We can therefore form the quotient group and define the homology group.

Definition 6.6. The k-th homology group of a simplicial complex X, denoted $H_k(X)$, is the quotient abelian group (vector space)

$$Z_k(X)/B_k(X) = \ker(\partial_k)/\mathrm{im}(\partial_{k+1}).$$

Two elements of a homology group which are equivalent (their difference is a boundary) are called homologous.

The number of k-dimensional holes in X is thus realized as the dimension of $H_k(X)$ as a vector space.

Intuitively the quotient is doing the heavy lifting. Suppose we have two paths, and we wish to know if they represent two different holes or just a more or less convoluted loop of edges. We distinguish them by taking their differences and see if they bound an higher dimensional chain. If they do, then the two chains are the same, alternatively the two chains carry intrinsically different topological informations.

We can therefore give the following definition,

Definition 6.7. Given a Space X, and its associated triangulation K, the rank of the n-th homology group H_n, that is the number of $\mathbb{Q}$ summands, is called the n-th Betti number, denoted b_n.

[13] We will omit the simplicial complex K when it is clear from the context

Recall the semi-formal definition provided at the beginning of the chapter

"visual numerosity refers to the empirical property of the number of connected components in an image (when seen as a topological space)."

We can now be more specific, and recast the intuition in homological terms, and in line with what suggested in Kluth and Zetzsche [80], this might be expressed as

$$N(K) = b_0(K)$$

where b_0 is the zeroth-Betti number: the rank of the first (simplicial) homology group, and K is the simplicial complex associated to the configuration C.

The formal solution might be considered concluded at this point, what remains is the consideration of computational aspects.

1. How can we compute the homology groups, and in particular b_0 starting from a visual stimulus?

2. Moreover given an algorithm that compute b_0 how can we say whether it is biologically plausible, that is neural implementable?

3. How the topological perspective proposed might help explaining the interference effects reviewed in section 3.7 that are still unaccounted in the literature?

Answering these issues will require an excursus in digital topology that will allow us a natural way to triangulate a space. We will only briefly suggest how the theory might be linked to a neural implementation, but we will free ourself from any deeper claim. However, with respect to the second point, we notice how relaxing the connectedness condition and allowing the assumption that objects are simply connected, that is without holes, Kluth and Zetzsche [80] following Chen and Rong [22], proposed to use the Gauss-Bonnet theorem to compute the Euler's characteristic in a parallel, and biologically plausible way. In the case of a simply connected object the Euler's characteristic and b_0 coincide, as it is clear from the following theorem

Theorem 6.2. *The Euler characteristic of a topological space X, is given by the alternating sum of the associated Betti numbers:* $\chi(X) = b_0 - b_1 + b_2 - b_3 \dots$

Instead of taking a discrete image representation they start from the curvature of the luminance function and apply the Gauss-Bonnet theorem

$$\int_S K dS = 2\pi\chi(S)$$

where k is the Gaussian curvature and χ the Euler's characteristic. In this way the numerosity is expressed in term of its Euler's characteristic. The linearity of the integral allows the additivity of the invariant, so that the invariant of a configuration of n objects amounts to just n times the basic invariant. Interestingly, by injecting an additive normally distributed noise $\eta \sim N(0,\sigma)$ at the input and output filters their simulations accounted for the Weber fraction observed in human and animal psychophysical and neural experiments. This result is encouraging and might suggest that although we have been mainly driven by mathematical considerations an empirical support is not out of reach [14].

The third question will be addressed in subsection 6.2.4. In particular we will propose a possible interpretation to the "perceptual grouping problem", that is the fact that when items are close enough the subject strongly underestimates the numerosity of the visual stimulus. We will follow up a proposal by Zeeman [184] and see how tolerance spaces might be used to provide a cognitively plausible solution to this problem.

Before giving the necessary definitions and results, that will allow us to compute the Betti numbers associated with a given stimulus pattern, we pause to dispel a possible quibble. In chapter 2 we stressed the importance of disentangling the system perspective from the observer perspective. Computing the homology of a visual stimulus has an intrinsic meaning for the observer, but can be only loosely associable to the system perspective. Resorting to simplicial homology, that is a highly abstract and combinatorial view of a topological space, might moreover be seen as quite unrelated to the actual neural computations. To dispel any possible misunderstanding, we don't claim that the neural system triangulate a space (given by the retinal input) for computing its homology. We instead take the simplicial complex to stand for an abstraction of the functional disposition induced by the neural activity. In particular given an array of retina receptors, the fibers of the optic nerve reach the visual cortex in an orderly fashion (preserving topography), so to a large extent the electrochemical activity in the cortex is isomorphic to the retinal image. The functional order given by the neural activity is available to the system itself, but the spatial distribution of this activity has meaning only for the observer. In order to use homology (directly) we need a way to associate the functional order to the spatial distribution isomorphic to the retinal array. A possible approach has been given in Toet [167] via a series of biologically inspired algorithms.

[14] In particular we are here considering the problem of extracting exact numerosity, whilst the perceive numerosity is just an estimation. The fact that via decision process, as already pointed out in Verguts and Fias [178], and by adding a plausible Gaussian noise a underlying overlapping representation may emerge, bodes well for the approach.

Suppose a set $E = \{e_1, ..., e_n\}$ of neuronal elements without knowledge of the spatial layout nor any inherent ordering. Via the coincidence activity matrix $Q \subset E \times E$ the simulations showed that a partial order of functional inclusion can be extracted, and, via neural recruiting, a lattice order (I) from the simultaneous activity of the net might be learned. Moreover, if a model has enough internal coherence, that is assuming high detector densities, then the functional order may allow isomorphism with simplicial complex, at least locally. If a neural net has enough internal coherence, then the algebraic structure of the complex with which it can be identified will reflect the topological and possibly the geometrical structure of the underlying detector array. We will not extract the simplicial complex from the functional order, but we will triangulate the space directly from a binarized retina input and compute the homology from there. However, the preceding discussion points at the direction of a possible neural algorithm which operation might be mapped to operation on a simplicial complex. The underlying theory, therefore, remains almost unchanged and what variates is the triangulation procedure and the implemented algorithm.

6.2.2.1 An algorithm for computing Betti numbers

For small spaces, and small dimensions, computing homology can be done by hand, and it boils down to filling in the details. That is writing down the chain complexes, computing the kernels and images of the boundary operators and take the quotient. The choice to enrich the group structure by defining simplicial complexes in a vector space $\mathbb{Q}$ allows us to see chain groups as vector spaces and boundaries as linear maps allowing a matrix representation. Therefore granting us a combinatorial way to compute homology. The interested reader may found linear algebra algorithms and some 'numerosities' examples by looking at the Jupyter notebook BettiNumbers.ipynb on the on-line support material[15].

In the following we give a classical example that should suffice to clarify the computation of the homology groups.

We are going to compute the homology of the sphere with an extra handle, that is homeomorphic to a triangle glued to a tetrahedron (figure 6.1). There is one connected component, that is zero 0-dimensional holes, one 1-dimensional hole, and one 2-dimensional hole.

[15] Available at https://github.com/bramacchino/numberSense/blob/master/BettiNumbers.ipynb.

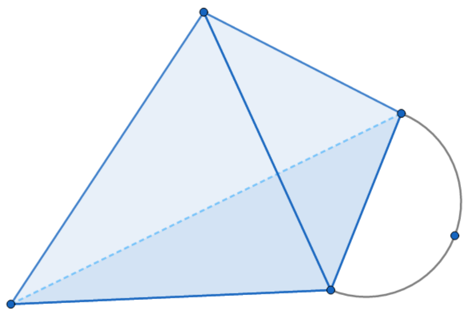


Figure 6.1: Simplicial complex of a sphere with an added handle.

The chains' groups are given by the linear span of the vector space basis:

$$C_0(X) = < 0,1,2,3,4 >$$
$$C_1(X) = < [0,1],[0,2],[0,3],[0,4],[1,2],[1,3],[2,3],[2,4] >$$
$$C_2(X) = < [0,1,2],[0,1,3],[0,2,3],[1,2,3] >$$

Now given (6.4) of ∂_k we can give a complete specification of the boundary map via the simplicial complex incidence matrix as follows.

For ∂_1, this would be

$$\partial_1 = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{c} \begin{array}{cccccccc} [0,1] & [0,2] & [0,3] & [0,4] & [1,2] & [1,3] & [2,3] & [2,4] \end{array} \\ \begin{pmatrix} -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

where the row labels are the basis for $C_0(X)$ and the column labels are the basis for $C_1(X)$. Similarly, ∂_2 is

$$\partial_2 = \begin{array}{c} \\ [0,1] \\ [0,2] \\ [0,3] \\ [0,4] \\ [1,2] \\ [1,3] \\ [2,3] \\ [2,4] \end{array} \begin{array}{c} \begin{array}{cccc} [0,1,2] & [0,1,3] & [0,2,3] & [1,2,3] \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

Now the composition of the two boundaries maps just corresponds to matrix multiplication, and the reader can verify (by hand, or via a linear algebra software) that indeed $\partial_1 \cdot \partial_2$ results in the zero matrix.

To compute the kernel of a linear map it suffices to solve the corresponding homogeneous system of linear equations. This usually implies reducing in echelon form by Gaussian elimination. Since reducing implies a change of basis, if we column reduce ∂_1 we have to take care of a corresponding change of basis when working with ∂_2, otherwise $\partial_1 \cdot \partial_2$ might not be the zero map[16].

$$T = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 \\ -1 & -1 & -1 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

[16]We can use the fact that elementary row operations are equivalent to multiplying with an elementary matrix on the left. Let D be a $m \times n$ boundary matrix. By the row reducing echelon form of the augmented transpose, in Matlab via z = rref([A' eye(n)])', we obtain the reduced matrix R = z(1:m,:), and T = z(m+1:end,:) gives us the transformation matrix.

Left multiplying with the boundary matrix ∂_1 gives us

$$\partial_1 T = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

The corresponding basis change for ∂_2, the inverse of T, is given by

$$T^{-1} = \begin{pmatrix} -1 & -1 & -1 & -1 & -0 & -0 & -0 & -0 \\ 1 & 0 & 0 & 0 & -1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and therefore the reduced form of ∂_2 with the compatible change of basis is

$$T^{-1}\partial_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore by inspecting the matrix pivots we get

$$\frac{\langle 5,6,7,8\rangle}{\langle 5,6,7\rangle} = \langle 8\rangle = \mathbb{Q}$$

Analogously we can compute the number of zero, and two dimensional holes, that is the homology group H_0, which rank corresponds to the number of connected components, and the homology group H_2.

6.2.3 Homology groups of a digital set

Visual processing is essentially discrete, the triangulation, therefore, cannot be assumed to be given by a homeomorphism from the simplicial complex to a continuous space. We need a mathematical theory that as topology is good for describing global properties, but of discrete or digital spaces.

Digital topology (Rosenfeld [146], Herman [67]) addresses the fundamental properties of binary object connectivity in two dimensional (and three dimensional) digital images. More broadly it allows to use the toolbox from standard topology (and homology) in the discrete realm of digital images. This in turn provides algorithms for computing interesting topological properties (like connectedness, the number of connected components, holes, thinning) in discrete sets. In particular, the theory of homology groups in n-dimensional digital images has been investigated in Arslan, Karaca, and Oztel [10] and Boxer, Karaca, and Oztel [16] from which we mainly borrow the following treatment.

Definition 6.8. Let n be any positive integer. An $n-xel\ q$ in an Euclidean n-space, $\mathbb{R}^n$, is a closed unit n-dimensional (hyper)cube $q \subset \mathbb{R}^n$ whose 2^nvertices have natural coordinates (more precisely, an n-xel in $\mathbb{R}^n$ is a cartesian product like $[i_1, i_1+1] \times [i_2, i_2+1] \times ... \times [i_n, i_n+1]$). We call pixel a 2-xel in $\mathbb{R}^2$. We define an n-dimensional binary image or nD-image, to be a finite set of $n-xels$ in $\mathbb{R}^n$

For simplifying the exposition we can define an image I as a subset of $\mathbb{Z}^n$, a lattice of points, with an $adjency$ relation. That is we represent a nD-image I as a finite n-dimensional array of 1's and 0's in which each 1 represents an n-xel in D and each 0 represents an n-xel that is not in D. Usually a variety of adjacency relations are used in the study of digital images. Being interested on the $2-D$ settings the following definition suffices:

There are various approaches to mimic standard topology in digital spaces, the most widely used in practice is via adjacency graphs.

Definition 6.9.
1. Two points $p, q \in \mathbb{Z}$ are called 2-adjacent if $|p-q| = 1$
2. Two points $p, q \in \mathbb{Z}^2$ are called 8-adjacent if they are distinct and differ by at most 1 in each coordinate
3. Two points $p, q \in \mathbb{Z}^2$ are called 4-adjacent if they are 8 adjacent and differ in exactly one coordinate

A well known results in digital topology is the necessity of defining two compatible adjacency relation for the 'foreground' and 'background' to avoid topological paradoxes (Rosenfeld [146]).

A digital picture is therefore commonly represented as a quadruple $\mathbb{Z}^n, k, \bar{k}, X$, where $n \in \mathbb{N}$, $X \subset \mathbb{Z}^n$ is the set of finite points, k represents an adjacency relation on X, and $\bar{k}$ represents an adjacency relation on $\mathbb{Z}^n \setminus X$.

Keeping in mind this distinction, and for the purposes of this thesis is sufficient to define an image stating only the adjacency relation on the foreground.

Definition 6.10. A 2 dimensional digital space, or a 2D digital image is a tuple (X,k), where $X \subset \mathbb{Z}^2$, and k is an adjacency relation on X.

We give the analogous definition of simplex and simplicial complex seen in the general case pinned down to a digital space as following:

Definition 6.11. Let $S \neq \emptyset \subset (X,k)$. We call each $\sigma \in S$ a simplex of the digital image (X,k) if the following holds:

1. If $\sigma, \sigma' \in S$ and $\sigma \neq \sigma'$, then σ, σ' are $k-$adjacent

2. If $\sigma \in S$ and $\emptyset \neq \tau \subset \sigma$ then $\tau \in S$

An n-simplex is thus a simplex S such that $|S| = m+1$. We call a nonempty proper subset of S a face of S

Definition 6.12. Let (X,k) be a finite collection of digital $m-$simplices, $0 \leq m \leq d$, for some $d \in \mathbb{N}$, if the following statements hold then (X,k) is called a finite digital simplicial complex:

1. If $P \in X$ then every face of P, P' is in X

2. If $P, Q \in X$ then $P \cap Q$ is either empty or a common face of P and Q

The dimension of a digital simplicial complex X is the biggest integer m such that X has an $m-$simplex.

The definitions of chain groups and chain complex carry over in the usual way (and has been proven in Arslan, Karaca, and Oztel [10]). Computing the homology of a digital space corresponds to compute the homology of the associated simplicial complex as in the general case.

6.2.4 Tolerance homology

Visual processing is essentially discrete, for this reason, and for computational simplicity, we resorted to digital topology. However, it remains to be answered the question of what attributes that belongs to a physically disconnected stimulus (such as a dot array) determine the perceptual connectivity[17]. Moreover, resorting to digital topology was based on the assumption that the lattice grid was given by retinal array. However more than the topographic adjacency, we wish to represent the functional and metabolic distance given by the neural activity. The difference is subtle, but important, whilst in the first case we see the grid lattice as an approximation of the retinal disposition, a move quite standard in the field of computer vision, in the latter we take the grid lattice to be an abstraction of the functional disposition induced by the neural activity. Both the *discretization problem* and the *perceptual connectivity problem* can be solved generalizing the notion of adjacency to that of tolerance. Indeed once the first problem is solved, it will become clear that the latter is just a natural extension.

The solution to the discretization problem is a way to construct a grid lattice that faithfully represents the external stimulus, by choosing an appropriate mesh resolution, that might be different from the retinal disposition. Intuitively any two points in the lattice represent two points in the stimulus that are perceptible different, and any finer resolution will be equivalent to a larger one, a move related to the psychophysical concept of just noticeable difference (JND) we already encountered in chapter 3. Formally the concept of JND has been investigated by Zeeman [184], defining tolerance spaces[18]. The idea, however, can be traced back to Poincaré[19] in introducing the representative space (*espace representatif*). In Poincaré [128] chapter 2 on Mathematical Magnitude and experience, Poincaré refers directly to Fechner's experiments.

> It has been observed, for example, that a weight A of 10 grams and a weight B of 11 grams produce identical sensations, that the weight B is just as indistinguishable from a weight C of 12 grams, but that the weight A is easily distinguished from the weight C. Thus the raw results of experience may be expressed by the following relations:
>
> $$A = B, B = C, A < C,$$

[17] With this, we ignore the possibility that all external stimuli might be discrete, and refer to disconnected stimuli as those that are 'evidently' not continuous.

[18] Although the presentation is sketchy. Sossinsky [156] provides a more detailed overview.

[19] Poincaré [128], Poincaré [127]

which may be regarded as the formula of the physical continuum[20]

In the space representative, therefore, sensations are collected in 'sets of similar sensations' (*ensemble de sensationes indiscernables*) or in Zeeman terminology, tolerance classes, consciously indistinguishable.

Definition 6.13 (Tolerance Space/physical continuum)**.** A tolerance space (X,ξ) is a set equipped with a tolerance relation ξ. We write briefly X_ξ for denoting a tolerance space.

Definition 6.14 (Tolerance relation)**.** A tolerance relation ξ on a set X is a relation $\xi \subseteq X \times X$ that is reflexive and symmetric.
We write briefly $x\xi y$ to replace the canonical $(x,y) \in \xi$ and read $a\xi b$ as a is indistinguishable from b.

The power of a tolerance relation stems from its simplicity. Indeed a tolerance relation is a natural generalization of an equivalence relation, by dropping the transitivity requirement. A simple move that allows to express the idea of 'resemblance' succinctly and precisely.

A few examples of tolerance spaces might help appreciating the definition.

1. Let X be a metric space and $\epsilon > 0$. Let's define ξ as $\{(x,x') \in X | d(x,x') < \epsilon\}$. The space (X,ξ) so obtained is a tolerance space.

2. Let V be the set of vertices of a simplicial complex, and ξ the relation 'the vertices x and y are in the some simplex'. The space (V,ξ) so obtained is a tolerance space.

3. The adjacency relation might be seen as a tolerance relation, reflexivity might be trivially imposed, defining 'adjacent or equal to', that is by taking the reflexive closure of the relation. Therefore a set X equipped with an adjacency relation is a tolerance space.

We might say that a tolerance relation acts like a 'glue' on the tolerance space, in a way similar as a topology acts as a 'glue' on a topological space [21]. In order to pin down this intuitive concept, we need some more definitions that will allow us to give a solution to the first problem.

[20] *But here is an intolerable discord with the principle of contradiction, and the need of stopping this has compelled us to invent the mathematical continuum. We will use again, in a different context, chapter 8, this idea.*

[21] It turns out that many of the things we can do with topological spaces, might be done with tolerance spaces. The reader interested on the similarities and differences between these two disciplines might refer to Poston [129].

Definition 6.15. Let (X,ξ), (Y,ξ), be tolerance spaces. A (ξ-continuous) map of tolerance spaces is any map $f : X \to Y$ preserving tolerance, that is such that

$$\forall x, x' \in X \quad x\xi x' \Rightarrow f(x)\xi f(x')$$

We denote the composition of tolerance maps and the identity map in the usual way, respectively given two tolerance maps f, g we write $f \circ g$ and id_X.

Definition 6.16. If f is a map of a set X into a tolerance space (Y,η), the induced tolerance denoted $f \star \xi$, arises in X according to the rule

$$x(f \star \eta)x' \Leftrightarrow f(x)\eta f(x')$$

In particular for an inclusion map $\iota : B \hookrightarrow Y$ the induced tolerance (or briefly *subtolerance*) $\iota \star \eta$ on B is simply denoted B_η.

Definition 6.17. If $A \subset X_\xi$ the the (1-fold) widening of A (also referred to as *neighborhood* or ξ-closure) is the set $\xi A = \{x \in X : \exists a \in A \quad x\xi a\}$. We define recursively the k-fold widening of $A : k\xi A = \xi(k-1)\xi A$

Definition 6.18. The doubled tolerance of X_ξ, denoted $X_{2\xi}$ is given by the rule

$$x(2\xi)x' \Leftrightarrow \exists y \in X \quad x\xi y, y\xi x'$$

In general the n-fold tolerance relation of X_ξ, denoted $X_{k\xi}$ is given by the rule

$$x(k\xi)x' \Leftrightarrow \exists y_1, ..., y_{k-1} \in X \quad x\xi y_1, \ldots, y_{k-1}\xi x'$$

Definition 6.19. A *skeleton* of the tolerance X_ξ is a subtolerance $\iota : A_\xi \hookrightarrow X_\xi$ for which there exists a map $r : X_\xi \to A_{2\xi}$ s.t. $r \circ i = id_A$ and $\forall x \in X \quad x\xi r(x)$.

We can now define precisely how the discretization problem might be solved. Let's assume for the sake of simplicity that the external input is a two dimensional configuration given by the luminance function, that is $X \subset \mathbb{R}^2$. Let moreover equip X with a tolerance relation, defined by the euclidean distance, as for example 1. above, representing the JND. Notice that, If A is a skeleton of X_ξ then $\xi A = X$, that is it approximates X. In particular, we would like A to be the 'discretization' of X.

As an example, suppose $X \subset \mathbb{R}^2$ is the unit disk in the plane, and A is the set of vertices of a square lattice of mesh h contained in X. Then for $\epsilon > \sqrt{2}h$, the set A is a skeleton of the tolerance X_ξ. (where the metric tolerance is defined as in example 1, and d is the euclidean distance).

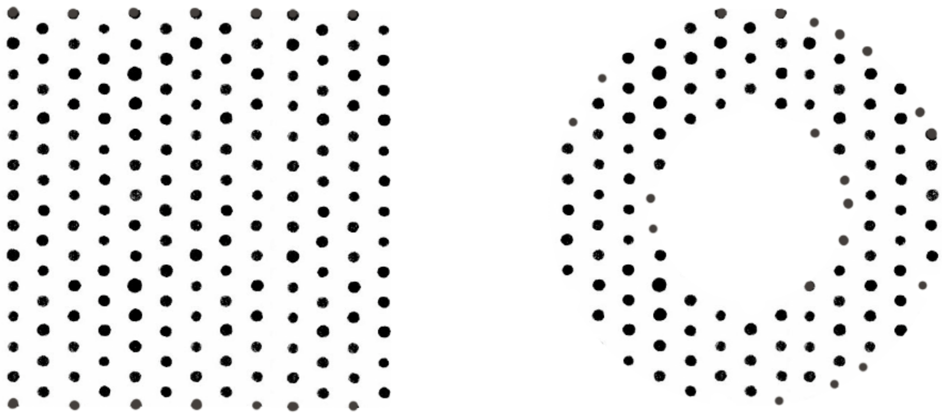


Figure 6.2: Topological global perception

The idea of a tolerance is however more powerful than just a mathematical expression of the concept of just noticeable difference and might be employed to perceptual organization in general, allowing us to solve the second problem. At the end of Buneman and Zeeman [18] the authors provide a suggestive image and hypothesis. (Figure 6.2: tolerance disk, and tolerance disk with a hole). While we clearly see the disk as disconnected, we can also perceive it as a whole. We can interpret this as if the visual system ignores details gaps within a certain tolerance. That is we might see the JND as the minimum tolerance ξ the functional architecture given by the skeleton A imposes, but also assume that larger tolerances might be computed by the neural system. Different tolerances might result in a 'Gestalt-like' perceptual instability, not too different from the perception of a Necker's cube for example. Any tolerance greater than the smaller will cover the stimulus with larger classes, and therefore with less connected components. The 'perceptual grouping problem', i.e. the fact that, given a visual stimulus with close enough items, the subject strongly underestimates its numerosity, might be addressed assuming the process that extracts the numerosity is sensible to this mechanism. For example if only the smallest and the largest tolerances are computed or affect numerosity discrimination, we might expect, for a visual array of n items, to perceive its numerosity as approximately $\frac{n+1}{2}$.

We leave the determination of how exactly the mechanism operates to a future time. We will end this chapter looking into a way to capture the global property of a tolerance space. The best way is again to resort to homology. Favorably tolerance spaces behave nicely and given a tolerance space (X, ξ) we can construct a simplicial complex and define the homology group as the homology group of this complex. The resulting homology behaves in a standard way (as proven in Sossinsky [156] and Poston [129, chapter 2]). In particular, given a tolerance space (X,ξ), it's usually associated the construction of a (Vietoris) simplicial complex consisting of all simplexes, where a simplex is a finite oriented subset of X all of whose points are within tolerance. In the case of the 'point cloud' given in Figure 6.2 we take the vertex points of the simplexes to be the points on the 'point cloud'.

Definition 6.20. Let X be a subset of a metric space and ξ the tolerance relation in example 1. We construct a simplicial complex in the following way inductively:

1. For each point $x \in X$, $\{x\}$ is a 0-simplex.
2. For each $x_1, x_2 \in X$, $\{x_1, x_2\}$ is a 1-simplex if $x_1 \xi x_2$
3. For $x_1, ..., x_n$, $\{x_1, ..., x_n\}$ is a $(n-1)$-simplex if all the points are within tolerance of each other.

According to this model distinghuishability is characterized by missing edges, faces, etc.. in the higher dimensional simplexes that make up the complex.

The problem of this definition is that we need extensive calculations to construct the simplicial complex [22]. This implies that computing the homology groups may become unwieldy even for a configuration with few dots. The assumption of discreteness of the perceptual input therefore may seem to suffer for a severe computational disadvantage.

However, with this construction we throw away the functional lattice structure defined by the tolerance skeleton. Being a tolerance relation a generalization of the adjacency relation used in digital topology allows us to apply the discussion of the preceding section in this special case. We can define a digital tolerance space as follows.

Definition 6.21. Let's denote the identity relation with τ_0, we define $\tau^{n+1} = \tau \circ \tau^n$ recursively for any integer n.
We call the pair (X,τ) a digital tolerance space.

We can therefore extend the definition of path-connectedness to account for the perceptual connectivity.

Definition 6.22. A digital tolerance space (X,τ) is connected if given $x, y \in X$ there exists a $n \in \mathbb{N}$ s.t. $x, y \in \tau^n$.

[22] It requires $\mathcal{O}(2^{|X|})$ time complexity to determine such a simplex if we directly implement a simple algorithm as suggested by the definition (namely brute force, by checking all possible simplexes).

To recast in digital topological terms x, y are nk-adjacent, that is there exists a k-adjacency path of length at most n connecting them.

We might associate nk-adjacency with an increased activation signal, or alternatively with a decreased threshold. The increased activation might be driven by top-down attentional processes, an hypothesis supported by some empirical evidence (Huang, Zhou, and Chen [71]) and in line with the holistic approach.

6.3 Summary

The term 'numerosity' is generally adopted as a mathematical uncommitted synonym of cardinality. In the discussion on natural numbers' representation (chapter 2) we argued that an informal definition might be safely used once its limitations are considered. The lack of a formal definition for 'numerosity' shows the other side of the medal: apply an informal characterization of a problem beyond the limits allowed prompts empty discussions. If numerosity is just cardinality, the definition assumes what it needs to be shown. What cognitively is referred to as an item, is the end stage of a delicate and complex process that can be disrupted by small modifications. Instead of defining the term numerosity directly for a broad class of phenomena, narrowing down the scope to a particular modality seems to us a promising strategy. Focusing on visual stimuli we suggest, in accordance with Kluth and Zetzsche [80], to define visual numerosity as the number of connected components in a configuration, the associated zeroth-Betti number. Computing the number of connected components in parallel by local operations is a feat in digital topology, and in computer vision. At the present time the proposed algorithms show a running time dependence on the number of items in the stimulus. We briefly introduced the reader to simplicial homology and give an example of how these algorithms might operate. In chapter 5 we have seen some indications towards a limit of the numerical representation. Therefore, for such small inputs the time complexity might be less of a concern. Although the discussion, and the title of this thesis, suggests a topological framework, visual perception is essential discrete. We, therefore, need a way to mimic topology in a discrete realm. Zeeman's [184] influential paper introduced tolerance spaces, set equipped with a tolerance relation, to solve this issue. Usually, simplicial complexes for tolerance spaces are constructed from point clouds, such that all points within tolerance are linked. If no topographic structure is known than this approach is sensible. However, the topographic order of the retinal neurons are preserved in the functional order in the higher visual area (cf. Toet [167]). For this reason digital topology might offer an upper bound on the complexity, once we consider the anatomical disposition of retinal receptors. Applying digital topology directly to the retinal grid is still expensive. For this reason, defining a tolerance skeleton over the functional activity allows us a more parsimonious way to triangulate a space. This layer of abstraction incorporates not only bottom up, but also top down activity, which modulation we hypothesized might be seen as the lattice mesh, allowing to solve the perceptual connectivity problem.

CHAPTER 7

COMPUTATIONAL MODELS OF VISUAL NUMEROSITY

The take home message of chapter 6 was that assessing whether a model genuinely computes numerosity requires checking the computed invariants. If the deviation from the invariance predicted matches the behavioral findings, and the model is equipped with numerosity detectors, then a strong case for a visual sense of number might be put forward. If, on the other scenario, the deviations from the invariance are predicted by a model that explains away numerosity coding, then it turns out to be hard to claim that the model is computing numerosity. This way of proceeding would be ideal. In the present chapter, we will therefore analyze two recent models that roughly compute the same invariant. The first model, Stoianov and Zorzi [160], supporting a visual sense of number, the second, Dakin et al. [26], explaining it away. On the other side Dehaene and Changeux's [35] model we analyzed in chapter 4 seems to be at odd with the invariance check.

7.1 Deahene & Changeaux, (1993)

We have already seen the architecture proposed by Dehaene and Changeux [35] in chapter 4. There we noted how a linear coding can be modeled in a plausible architecture. Here we are interested in the invariance principle the architecture can model.

7.1.1 Invariance principle

Importantly, The invariance is not achieved by the distributed computation of local features, instead an explicit normalization stage is assumed. Each object is seen as a blob represented by a dedicated detector. A certain size invariance can be attained in the one dimensional case, but it's quite hard to see how a blob matching system should attain a standardize form for various element shapes and spatial distribution in the two dimensional case. The interference effects (section 3.7) are thus quite hard to be accounted for.

7.2 Stoianov & Zorzi, (2012)

Stoianov and Zorzi [160] trained a Deep Belief Network (DBN) to reconstruct the input, given by binary images comprising of rectangular objects of different sizes. They were able to show that visual numerosity emerges as a statistical property of those images, without any preprocessing normalization mechanism, nor any information about numerosity during the training phase. There has been quite a lot of research in comparing (Sparse) Restricted Boltzmann machines[1] to the neural coding in vision (Lee, Ekanadham, and Ng [88], Bhand et al. [14]). Stoianov and Zorzi's [160] Deep Belief network, therefore, inherits the same neural plausibility. Moreover, for stressing the biologically plausibility, the greedy pre-training scheme lacked a back propagation fine tuning typical in the standard deep networks used in machine learning. However, from a developmental point of view, the pre training scheme appears quite unfeasible, where all numbers are given in a bunch. A more cognitively plausible solution would require a learning process in which random numerosity samples are given at a time. Such a solution doesn't seem technically out of reach and could, in principle, be used to explain the progressive sharpening of the Weber fraction observed in developmental numerical psychology. Interestingly the numerosity detectors where found only in the higher layer and the activation pattern reflected the monotonic coding. As we have seen in chapter 4, this alone is able to explain distance and size effects, and it might be used as the input representation to generate a number line coding (as we have seen in Verguts and Fias [179] model). To sound a note of caution, we remind the reader that deep learning models have high capacity and adapt to data statistics, it is therefore interesting to see whether the model trained with natural images is affected in its ability to learn to represent numerosity. At the

[1] Restricted Boltzman Machines (RBM) are the building block of DBNs.

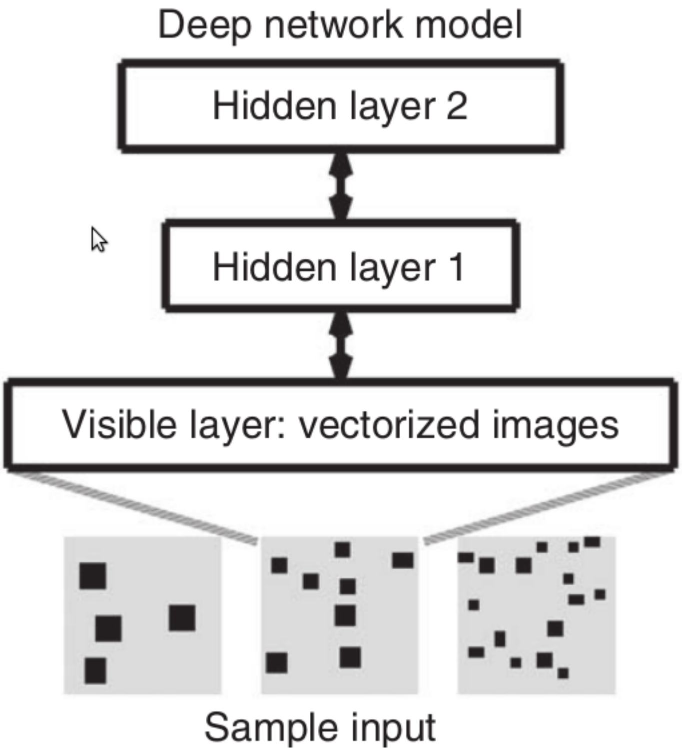


Figure 7.1: Inputs example and architecture used

present stage we will ignore this issue and instead look at the invariance principle the numerosity units are sensitive to.

7.2.1 The Model

The input is given by 52100, 30×30 pixels binary images containing from 1 to 32 randomly placed non overlapping rectangular shapes (Figure 7.1 bottom. On the on-line support material of this thesis are available Matlab and Python scripts to generate the Dataset as described in Stoianov and Zorzi [160] supplementary information[2]). The architecture used by the authors is a parallel implementation of Geoffrey Hinton's original code and it is freely available in their website[3]. Our implementation will soon be added to the GitHub page associated to this thesis[4].

The network architecture (Figure 7.1) comprises one visible layer, in which the

[2] Available at https://github.com/bramacchino/numberSense/tree/master/inputs/sz2012.

[3] http://ccnl.psy.unipd.it/research/deeplearning.

[4] At the time of writing I'm still unable to replicate their results fully. The invariance property is therefore, only hypothesized on the basis of the network description. As soon as I'll be able to test the invariance property the code will be added.

vectorized input is clamped, and two hierarchical organized hidden layers. In particular, the architecture might be seen as an auto-encoder consisting of two stacked RBMs. Each RBM is formed by a visible and a hidden layer of binary units. The units in the hidden layer fire with a probability that is the logistic function of the weighted input. The input layer of the first RBM comprises 900 binary units fully connected to the hidden layer of 80 binary units, that represents the visible layer of the second RBM with an hidden layer of 400 units. The output layer represents a dimensionality reduced version of the input layer.

The network is trained to maximize the product of probabilities assigned to the training set (i.e. to generate the sensory data), equivalently to minimize the average negative log-likelihood. This in turn is achieved by minimizing the weights (and biases). The minimization is achieved via (stochastic) gradient descent. The derivative gives us two terms, called the positive and the negative gradient. The first depends on observation whilst the latter depends only on the model. Learning is therefore achieved via Contrastive Divergence (CD) (Hinton, Osindero, and Teh [69]). Given an input vector v_i^+ , first the feature detectors h_j^+ are activated (positive phase). Starting from stochastically selected binary states of the feature detectors (using their state h_j^+ as a probability to turn them on), CD then infers an input vector v_i used in turn to reactivate the features detectors h_j^- ("negative" phase). The weights w_{ij} are updated with a small learning fraction η of the difference between input-output correlations measured in the positive and the negative phases:

$$\delta w_{ij} = \eta(v_i^+ h_j^+ - v_i^- h_j^-)$$

7.2.2 Invariance principle

In Stoianov and Zorzi [160] supplementary informations the authors provide a mathematical description of the learned model that help us in assessing the invariance principle. Most of the first hidden layer (HL1) units are center-surround detectors that uniformly cover the image space. The first layer consists of linear operations (2D Gaussian filters, sigma = 2, and spatial integration) followed by a non linear operation (a standard logistic function, f)

$$O_{ij} = f(\sum W'_{ij} I + 1)$$

The numerosity detectors found in the second hidden layer (HL2) are spatially selective as well (2D Gaussian filters, sigma=10). They receive positive input from HL1 units, and inhibition from HL1 units that were found to encode cumulative area (c). That is numerosity detectors are represented by the subtraction of the cumulative area from the low-pass-filtered features.

$$N_{kl} = \sum W_{kl} O - c$$

$$c = log(1 + \frac{\sum I}{c_{max}})$$

If we assume the first term to be related to the contour, then the invariance is achieved via a trade-off between objects' contour and area. For simple shapes (e.g. the rectangles of different sizes used to train the network) the deviation from invariance is minimal. However, for shapes with multiple edges, the deviation from the proposed invariance might be consistent. Whether this matches human behavior has not yet been experimentally addressed. A striking difference emerges with respect to Dehaene and Changeux's [35] model. Whilst in the model of Dehaene and Changeux [35] an intermediate layer of object locations was assumed, in this architecture the invariance is computed, from local features, without any explicit normalization stage.

We conclude the invariance analysis comparing it to our discussion in chapter 6. We hypothesized that the tolerance skeleton, from which the computation of the number of connected components is carried over, covers both top-down and bottom-up connections. The model of Stoianov and Zorzi [160] is based entirely on bottom-up features to compute the numerosity of a stimulus, whilst top down connections were used only during the learning stage. If it turns out that the systematic dependence on object's shape doesn't mirror human behavior, the model might still be accurate for what concerns the forward computations.

7.3 Dakin & Morgan, (2011-2014)

In this model numerosity perception is achieved as a by-product of texture processing. It's assumed that density (spacing of objects) and numerosity share the same underlying mechanism: filter-based texture computations (pooled high and low spatial-frequency filter). That is relative numerosity become a type of texture discrimination. Peculiarly the model doesn't need any contribution from a location code, that has we have seen in chapter 6 is hypothesized by the computational approach to visual perception[5]. Recall from section 4.2 that the models of Dehaene and Changeux [35] and Verguts and Fias [179] assume the input to be given by a location map, that we briefly suggest in chapter 5

[5]The authors argue that if a location code were available, the behavioral findings concerning the fact that a change in patch size disrupt our density and numerosity perception, would be pretty outlandish.

might be tentatively associated with the activity of LIP neurons[6]. On the other side a saliency map is not required by the global holistic approach (or topological account) to visual perception, and the adoption of spatial filtering is compatible with the computation of pseudo-topological properties described in Barth, Ferraro, and Zetzsche [11]. In this respect, the model is similar to the trained architecture of Stoianov and Zorzi [160]. However, here the filter's responses are used directly, without an explicit layer of numerosity detectors.

7.3.1 The model

Density ($d = N/A$) is estimated using the relative response of low and high spatial frequencies (SFs). Dakin et al. [26] and Morgan et al. [107] assume that high spatial frequencies are largely determined by the number of objects, whilst low SFs are largely determined by the cumulative area of the items in the stimulus.

Spatial frequencies sizes correspond in physiological terms to the size of the receptive fields. The SFs filters response are therefore estimated by convolving stimuli with Laplacian of Gaussian, center-surround filters, constructed from the combination of a Gaussian filter and a second derivative. Laplacian-of-Gaussian filters are tuned to high (s = 2 pixels) or lower (s = 8 pixels) spatial frequencies.

$$LoG(x,y) = \Delta^2 G(x,y) = \frac{1}{\pi s^4}(1 - \frac{x^2+y^2}{2s^2}) exp(-\frac{x^2+y^2}{2s^2})$$

The filter response is achieved by pooling the convolution across all image locations

$$R_s = \sum_{x,y} |LoG(x,y) \times |I||$$

By definition the pooled high-frequency output depends on the length of the object contours in the stimulus. That Rhi $\propto$ N, therefore is based on the fact that small filters generate isolated responses to individual elements, if the elements are 'small enough'. Analogously $Rlo \propto A$ follows from the fact that large filters responds to clusters of elements, with the same characteristics.

The response ratio C is thus seen as a correlate of density and number.

$$C = 2^{\gamma_s} \frac{R_{hi}}{R_{lo}}$$

[6] Although as reported in Knops et al. [81], it is still unclear how it might operate beyond the subitizing range.

Where 2^{γ_s} is a multiplicative Gaussian noise term ($s = 0.1$).

Denser stimuli is selected on the basis of

$$d_{a,b} = \frac{C_a}{C_b}$$

The behavioral observations in Dakin et al. [26] and Morgan et al. [107] suggest that R_{hi} cannot be directly accessed by the visual system. In particular mismatching the patch size has an effect on performance, and increasing contrast doesn't increase number estimation. The low spatial filters are used as a proxy for area in computing density, therefore an explicit weighting for degree of size mismatch is applied to recover the high spatial frequency component, based on the ratio of the low SFs response from the two stimuli.

$$N_{a,b} = (2^{\gamma_s} \frac{aR_{lo}}{bR_{lo}})^{2^{\gamma'_s}} d_{a,b}$$

7.3.2 Invariance principle

If there was no noise in the internal representation, then numerosity is attributed to the high frequency filters. Adding noise implies that the contribution of the low spatial filters is not canceled out, and Dakin et al. [26], Morgan et al. [107] consider this contribution as providing a moderate normalization for size. R_{hi}, the pooled response of high frequency low pass filters, is dependent on the items edges. In their report, the author tested the model with object of approximately equal size. However, for objects with large size mismatch and different shapes, the aggregate contour overestimates the number of objects. Ross [147] suggests that grouping by size precedes the operation of mechanisms that estimates numbers. Here, the estimation of number is only indirect, therefore the suggestion cannot directly be implemented, but as for Stoianov and Zorzi's [160] network, only bottom up processing are modeled in this architecture.

Although the influence on size was wanted by the authors to account for their data, resorting mainly on the contour length, without any compensation mechanism, seems a too coarse estimation. If we assume a stronger influence of the low-frequency filters, than the model invariance is based on a trade-off between contours and area as in Stoianov and Zorzi's [160] network.

Importantly, although the invariance computed by the model is roughly the same, numerosity estimation is computed directly, without any intermediate layer of numerosity detectors.

7.4 Summary

A cursory reading of the mathematical cognition literature might give the impression that numerosity is a perceptual feature which representations are neurons, or population of neurons, in the lowest level of the representations' hierarchy. We have seen on various occasions that this hypothesis is not without difficulties. Indeed, this thesis fits within the current debate on whether it is accurate to speak of a true sense of number (cf. section 3.1). As a matter of fact, numerosity might be either a cognitive process emerging from the combination of various capacities, or can be explained away as a by-product of non numerical perceptual features. Given the definition of numerosity as the number of connected components in a configuration allows us to analyze the models' predictions with respect to this invariance. If the deviation predicted by a model matches the subject performance, and the model's architecture comprises numerical representations as those reviewed in chapter 4, then we find that speaking of a true sense of number is not only a powerful metaphor. Stoianov and Zorzi's [160] architecture fits well within this situation. A Deep Belief Network trained to represent visual stimuli learned to represent numerosities via a thermometer encoding in the highest layer. No intermediate location map was needed, but the numerosity of the stimulus was estimated from low pass filters encoding the approximate contour and the cumulative area. Similar operations are those used in Dakin et al.'s [26] model. Here numerosity is assumed to be based on the recovered pooled response of high-frequency filters, with a moderate bias from size handled by low-frequency filters. If a stronger contribution from the low-frequency filters is assumed, however, this latter model might be seen as roughly computing the same invariance[7]. This implies that a stronger deviation from this invariance, as those proposed originally in Dakin et al. [26], results in a strong bias on size, whilst approaching the invariance results in a stronger shape bias, as is the case in the architecture proposed by Stoianov and Zorzi [160]. More behavioral studies are needed to assess human performances better, both with respect to the actual invariance the visual sense is sensible to, and to the deviations from it. However, Anobile, Cicchini, and Burr [8] and Zimmermann and Fink [185] suggest that both mechanisms might play a role, and that estimation of small and large numerosities (with a threshold around 32) are dissociated.

[7]That might be seen as the isoperimetric quotient $Q = 2\pi A/L^2$, where A is the area, and L is the contour length.

Chapter 8

Conclusions and further works

Ché perder tempo a chi più sa piu spiace.

Dante Alighieri, *La Divina Commedia*

Two main questions sit behind this thesis: is it correct to speak about an approximate number sense, and if it turns out to be the case, is it foundational for learning the concept of natural numbers and higher mathematical concepts? Indeed the curiosity towards 'numerical core systems', in particular, the approximate number system (ANS), was driven by the recognition of a certain agreement among the researchers, that the ANS plays a major role in higher mathematical capacities, although, has we have seen, they are divided about which role it plays. To recall chapter 2, 'Nativists' tend to see an ancillary role for the ANS, that is the ANS is representationally related to our understanding of numbers, but is neither foundational, nor necessary. 'Foundationalists' claim the ANS has a prominent role through lifetime. 'Developmentalists' maintain a paramount role for the ANS in higher mathematical concepts' acquisition, but only in order to learn, or in Carey [20]'s hypothesis, bootstrap, the natural numbers. In particular, with respect to the latter two positions, there is no agreement on how symbolic representations are learned. An effort, that of verbal counting, that takes many years to master, and absent in some cultures. Although the proposals differ, the underlying idea can be traced back to Poincaré: as we apply to the same objects, different systems, we are driven to reconcile the object's representations. In Poincaré [128, part

[0] For who knows most, him loss of time most grieves.

1: number and magnitude], this reconciliation is assumed to be driven by the *principle of contradiction*, and used to 'derive' the concept of mathematical continuum from the one of physical continuum. We already introduced the concept of physical continuum in chapter 6 when defining tolerance spaces. Briefly, starting from the Fechnerian experiment, we gain the knowledge that $A =_f B$, $B =_f C$ and $A < C$. The equality represented by the system f, however, is in disagreement with the 'standard equality' defined by another system a[1]. We are lead to reconcile the two notions, and this drove Poincaré to hypothesize that *the necessity of banishing the disagreement has compelled us to invent the mathematical continuum*[2]. The notion, that of mathematical continuum, is created by the mind, but were the empirical observations that provided the opportunity.

Young children struggling with learning natural numbers are like small mathematicians, trying to reconcile two or more different representations into a unique, consistent one. For example, when trying to learn the number 'nine', the child cannot depend only on the parallel individuation system (PIS) or the ANS, given that none of the two systems guarantee a meaning for it. The child is then driven to reconcile these representations. In section 2.3.1 we briefly considered how this can be achieved, without imposing any restriction on the process. This is just a first step, only to show that there is no principled impediment to reconcile the two representations. However, the fact that the representations provided by the two distinct core systems can be reconciled may not be guaranteed under cognitive plausible rules.

Why children are compelled to reconcile two, possibly unrelated representations? It's assumed, although it's not clear how, that the child is able to move from an implicit representation of the first three (or four) numbers encoded in the PIS, to an explicit one. It's moreover assumed that the ANS encodes approximate numerosities. Whether a true sense of numerosity, that is a meaningful representation also from the point of view of the animal, and not only from the observer standpoint, may be encoded within the brain is therefore a question of paramount importance. If numerosity estimation can be explained away as a byproduct of texture processing, for example, it is not clear how the child might extract any meaningful numerical information from it, before possessing the concept of number. The reconciliation hypothesis requires that the referent different vehicles represent is the same, and that the vehicles are 'cognitively compatible'. In the analysis

[1]These two systems might be roughly equated to system one and system two in the current debate on the psychology of reasoning.

[2]Poincaré [128, pag. 22]. By limiting finite resolutions we reach the mathematical continuum of the first order (the infinitude of the natural numbers), and then by reflection on the new contradictions that arise during the construction of the mathematical continuum of the first order, we are prompted to construct the mathematical continuum of the second order (that of the reals).

of the mathematical continuum given by Poincaré this was achieved by assuming that the physical equality, and the equality encoded in the principle of contradiction are perceived as instances of a general rule of equality that needs to be reconciled. In the case of numerical information, this means that the child knows that the information provided by the ANS is compatible with the one provided by the other system. If the information is not numerical, there is no need for reconciliation, and thus the account of how a number system emerges is at stake [3]. For assessing whether the information computed by the ANS is indeed numerical, an intuitive definition of numerosity doesn't suffice, and we were prompted to seek for a formal definition that captures the intuitions of the researchers. We defined numerosity as the number of connected components of a topological space, and noticed how an homological account may help in computing it. This definition is a mathematical idealization, since it is known, as reviewed in chapter 3, deviations from invariance are observed. Assessing the claim that the ANS is not only a metaphorical label for a process, but indeed a numerical perceptual mechanism, is thus a matter of constructing a model that predicts those deviations. If, for example, deviations are predicted from a model, like that proposed in Dakin et al. [26], in which numerosity perception is just a by-product of texture processing, then it would be hard to justify the claim of a true sense of number, and in turn, of the necessity the child is driven by to reconcile the different numerical representations. At the present time there is only a feeble evidence that a true sense of number exists, since deviations from invariance are reported in only few studies.

This gives us the opportunity to answer the second question posed in the introduction. There we claimed that philosophy of mathematics, and mathematical cognition are not communicating, and we were asking why it was so given the common problem they are facing.

On the psychological side, the lack of a firm answer to the nature of numerosity estimation, suggests more studies are needed, and that the mechanisms that give rise to numerical perception are still mainly unknown. It's just a matter of prudence not to resort to concepts that belongs to an usually much more abstract field, in order to avoid intellectual gimmicks, that would be mainly useless. There is also an historical concern, that philosophy of mathematics, deeply rooted in logic, cannot provide the right concepts for the cognitive scientist. This under appreciation, however, was mainly due to a misguided view of the vast varieties of logics, and the assumption that logic was just a

[3] Indeed, a general magnitude progression might suffice, for example we might hypothesize that the concept of succession is extracted from the intuition of time. In this case, however, the ANS has lost any privilege.

shorter word for classical first order logic [4].

On the philosophical side, the lack of collaboration might be traced back to Frege antypsychologism[5].

In more recent days, however, it's quite common too see philosophers of mathematics embracing a naturalistic approach and using psychological arguments. This use of results from cognitive science is often due to a will to avoid the Field-Benacerraf Dilemma: that is the problem of explaining abstract entities via 'naturalised' epistemology[6].

Benacerraf [13] notices that

> On a realist (i.e., standard) account of mathematical truth our explanation of how we know the basic postulates must be suitably connected with how we interpret the referential apparatus of the theory.

However,

> What is missing is precisely [...] an account of the link between our cognitive faculties and the objects known.

In particular he shows a certain skepticism that such an account might be given.

> If, for example, numbers are the kinds of entities they are normally taken to be, then the connection between the truth conditions for the statements of number theory and any relevant events connected with the people who are supposed to have knowledge cannot be made out.

One common way out is to commit to an internalist conception of meaning. That is instead of trying to glean truths about ethereal mathematical entities, we seek to explain how the concepts that underwrite our mathematical reasoning are constrained.

This move is usually based on 'intuitions', not too dissimilar from what we called in , following Darwin, 'fundamental intuitions'. Epistemic appeal to intuitions is sometimes seen as no different from perception and other times as worrisome as introspection [7].

[4] In the psychology of reasoning this line of argument has been debunked in Stenning and Van Lambalgen [159].

[5] In the introduction to 'Grundlagen der Arithmetic', Frege [49], formulated a severe and very influential criticism against the use of psychological methods in the philosophy of mathematics.

[6] See Benacerraf [13] and Hanna [60, chap. 6] for a broad overview of the Dilemma and a Neo Kantian structuralist proposal, much in line with the present discussion.

[7] Paradigmatic of the first position is Kripke's [84] view. *Of course, some philosophers think that something having intuitive content is very inconclusive evidence in favor of it. I think it is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking.*

In general ['s [[]pag. 2]williamson2008philosophy position is maintained, according to which

> Although there are real methodological differences between philosophy and the other sciences, as actually practiced, they are less deep than is often supposed. In particular, so called intuitions are simply [armchair] judgments (or dispositions to [armchair] judgment); neither their content nor the cognitive basis on which they are made need be distinctively philosophical.

Notwithstanding, the move is instantiated with only a superficial reference to cognitive faculties. More often than not, to some 'folk psychology'[8].

The basis for a collaboration between cognitive scientists and philosophers seems therefore, not out of reach. Intellectual acrobatics is not required from the psychologists, but well defined mathematical formulations of the core capacities are essential. For what concern philosophers, a greater attention to the core concepts studied by the psychologists, instead of a psychological flavored, yet ungrounded, reference to intuitions, is suggested [9].

8.1 Future works and works in progress

8.1.1 Development

As we repeatedly noticed it's not easy to assess whether numerosity perception can be explained away as merely a derived effect of other features. We hope that the proposed definition of numerosity might help in resolving the issue. This strategy requires assessing the empirical question to which topological invariant numerosity perception is sensible to, and how the computational models predict the deviation from the invariance. Moreover, analyzing data require data and too few are publicly available, more importantly for our concerns, only few behavioral experiments are aimed at assessing the invariance properties, and the deviation from it. As a sort of 'structured procrastination' I'm planning to spend some time on *Stimulus.js*, briefly presented in **??**, and building up

[8]For example Shapiro [155] via pattern recognition.

[9]I've been too optimistic in this remark. Recently, it has come to my attention, that for some philosophers, the possibility that one's philosophy might be falsified by empirical ground is unacceptable. A common joke regarding mathematicians and philosopher, argue that the latter are less expensive then the former, given that they need only paper and pen and no trash bin. Philosopher deeply absorbed in metaphysical questions, the remark continues, might be too far removed to be interested in psychological observations. If this is a general feeling, then it is much better to leave psychology to the psychologist and philosophers to speak among each other. Trash bins are still useful to the rest of us.

some experiments especially targeted at these purposes. As a long term plan, merging with jsPsych and creating a graphical user interface is foreseen.

For what concerns the computational models, there is also some work that needs to be done. The models in this thesis are coded in a fast prototyping style, and cannot be considered production ready. Moreover, they haven't been coded in a unitary framework. Whilst there is a plethora of libraries targeted at the machine learning community, there are practically no tools of comparable size aimed at the computational neuroscience needs. However, machine learning demands are not exactly those of cognitive modeling. For this reason, coding the models in TensorLayer seems a viable solution. It is abstract enough not too require recoding standard algorithms, and flexible enough to resort directly to TensorFlow when needed. The plan is thus to code the models in the numerical cognitive literature in TensorLayer and naturally releasing them with an open source license. This implies that depending on how it will be received it can turn out to be a personal Zen, or a useful endeavor.

8.1.2 Theory

One of the main reasons that drove this thesis was the curiosity to assess what the 'approximate number sense' was, whether it is legitimate to call it in this way (that is whether the concept is representational[10], or it makes sense only from our external perspective), and how it can be studied. To show how the proposed definition might have an algorithmic counterpart we resorted to simplicial homology, where the simplicial complex was assumed to be extracted from a lattice of points standing for the brain activity. We freed ourselves from stronger claims with respect to the functional lattice, by adopting tolerance spaces, however this construction remains too abstract for simulation purposes. It's interesting, moreover, to borrow a tool from topological data analysis, persistent homology, to compute the persistent Betti numbers and compare this result with empirical data.

The curiosity towards the ANS, was moreover driven by the recognition of a certain agreement, among researchers in mathematical cognition, that the ANS plays a major role in higher mathematical capacities. We argue that philosophers of mathematics (and we include logicians in the category) and cognitive scientists should collaborate toward this, and more broad, endeavors. Toward this direction, together we two other PhD candidates we are planning to develop a formal theory of the functional architecture and

[10]In the sense of representation discussed in chapter 2.

learnability of natural numbers concepts providing as input a formalization of the core representations.

BIBLIOGRAPHY

[1] Akin, Omer and Chase, William. “Quantification of three-dimensional structures”. In: *Journal of Experimental Psychology: Human Perception and Performance* 4.3 (1978), p. 397.

[2] Allik, Jüri, Tuulmets, Tiia, and Vos, Piet G. “Size invariance in visual number discrimination”. In: *Psychological research* 53.4 (1991), pp. 290–295.

[3] Alvarez, George A and Franconeri, Steven L. “How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism”. In: *Journal of vision* 7.13 (2007), pp. 14–14.

[4] Anderson, JA et al. “Experiments with representation in neural networks: Object motion, speech, and arithmetic”. In: *Synergetics of cognition*. Springer, 1990, pp. 54–69.

[5] Anderson, James A. “Learning arithmetic with a neural network: Seven times seven is about fifty.” In: *An invitation to cognitive science* 4 (1998), pp. 255–300.

[6] Anderson, James A, Spoehr, Kathryn T, and Bennett, David J. “A study in numerical perversity: Teaching arithmetic to a neural network”. In: *Neural networks for knowledge representation and inference* 311 (1994), p. 335.

[7] Anobile, Giovanni, Cicchini, Guido Marco, and Burr, David C. “Number as a primary perceptual attribute: A review”. In: *Perception* 45.1-2 (2016), pp. 5–31.

[8] Anobile, Giovanni, Cicchini, Guido Marco, and Burr, David C. “Separate Mechanisms for Perception of Numerosity and Density”. In: *Psychological Science* 25.1 (2013), pp. 265–270.

[9] Arrighi, Roberto, Togoli, Irene, and Burr, David C. “A generalized sense of number”. In: *Proceedings of the Royal Society of London B: Biological Sciences* 281.1797 (2014), p. 20141791.

[10] Arslan, H, Karaca, I, and Oztel, A. “Homology groups of n-dimensional digital images”. In: *XXI Turkish National Mathematics Symposium*. 2008, pp. 1–13.

[11] Barth, Erhardt, Ferraro, Mario, and Zetzsche, Christoph. "Global topological properties of images derived from local curvature features". In: *Visual Form 2001*. Springer, 2001, pp. 285–294.

[12] Bello, PF, Bridewell, W, and Wasylyshyn, C. "Attentive and pre-attentive processes in multiple object tracking: a computational investigation". In: *Proceedings of the thirty-eighth annual meeting of the cognitive science society. Philadelphia, PA*. 2016, pp. 1127–1139.

[13] Benacerraf, Paul. "Mathematical truth". In: *The Journal of Philosophy* 70.19 (1973), pp. 661–679.

[14] Bhand, Maneesh et al. "Unsupervised learning models of primary cortical receptive fields and receptive field plasticity". In: *Advances in neural information processing systems*. 2011, pp. 1971–1979.

[15] Bloom, Paul. "Generativity within language and other cognitive domains". In: *Cognition* 51.2 (1994), pp. 177–189.

[16] Boxer, Laurence, Karaca, Ismet, and Oztel, A. "Topological invariants in digital images". In: *Journal of Mathematical Sciences: Advances and Applications* 11.2 (2011), pp. 109–140.

[17] Bruner, Jerome S. *Beyond the information given: Studies in the psychology of knowing*. WW Norton, 1973.

[18] Buneman, OP and Zeeman, Erik Christopher. *Tolerance spaces and the brain*. Mathematics Institute, University of Warwick, 1967.

[19] Burr, David and Ross, John. "A visual sense of number". In: *Journal of vision* 9.8 (2008), pp. 12–15.

[20] Carey, Susan. "Cognitive foundations of arithmetic: Evolution and ontogenisis". In: *Mind & Language* 16.1 (2001), pp. 37–55.

[21] Carey, Susan. *The Origin of Concepts*. Oxford University Press, 2009.

[22] Chen, Li and Rong, Yongwu. "Digital topological method for computing genus and the Betti numbers". In: *Topology and its Applications* 157.12 (2010), pp. 1931–1936.

[23] Chen, Lin. "The topological approach to perceptual organization". In: *Visual Cognition* 12.4 (2005), pp. 553–637.

[24] Cicchini, Guido Marco, Anobile, Giovanni, and Burr, David C. "Spontaneous perception of numerosity in humans". In: *Nature Communications* 7 (2016).

[25] Cohen Kadosh, Roi et al. "Specialization in the human brain: the case of numbers". In: *Frontiers in Human Neuroscience* 5 (2011), p. 62.

[26] Dakin, Steven C et al. "A common visual metric for approximate number and density". In: *Proceedings of the National Academy of Sciences* 108.49 (2011), pp. 19552–19557.

[27] Davis, Hank and Pérusse, Rachelle. "Numerical competence in animals: Definitional issues, current evidence, and a new research agenda". In: *Behavioral and Brain Sciences* 11.04 (1988), pp. 561–579.

[28] Dayan, Peter and Abbott, LF. "Theoretical neuroscience: computational and mathematical modeling of neural systems". In: *Journal of Cognitive Neuroscience* 15.1 (2003), pp. 154–155.

[29] Dehaene, Stanislas. "Evolution of human cortical circuits for reading and arithmetic: The ‚Äúneuronal recycling‚Äù hypothesis". In: *From monkey brain to human brain, ed. S. Dehaene, J.-R. Duhamel, MD Hauser & G. Rizolatti* (2005), pp. 133–57.

[30] Dehaene, Stanislas. "Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation". In: *Attention & performance XXII. Sensori-motor foundations of higher cognition* (2007). Ed. by Rossetti, P. Haggard & Y., pp. 527–574.

[31] Dehaene, Stanislas. "The neural basis of the Weber–Fechner law: a logarithmic mental number line". In: *Trends in cognitive sciences* 7.4 (2003), pp. 145–147.

[32] Dehaene, Stanislas. *The number sense: How the mind creates mathematics*. Oxford University Press, 2011.

[33] Dehaene, Stanislas. "Varieties of numerical abilities". In: *Cognition* 44.1-2 (1992), pp. 1–42.

[34] Dehaene, Stanislas, Bossini, Serge, and Giraux, Pascal. "The mental representation of parity and number magnitude". In: *Journal of Experimental Psychology: General* 122.3 (1993), p. 371.

[35] Dehaene, Stanislas and Changeux, Jean-Pierre. "Development of elementary numerical abilities: A neuronal model". In: *Cognitive Neuroscience, Journal of* 5.4 (1993), pp. 390–407.

[36] Dehaene, Stanislas and Cohen, Laurent. “Cultural recycling of cortical maps”. In: *Neuron* 56.2 (2007), pp. 384–398.

[37] Dennett, Daniel C. *Consciousness explained*. Penguin UK, 1991.

[38] Durgin, Frank H. “Texture density adaptation and the perceived numerosity and distribution of texture”. In: *Journal of Experimental Psychology: Human perception and performance* 21.1 (1995), p. 149.

[39] Eliasmith, Chris. “A new perspective on representational problems”. In: *Journal of Cognitive Science* 6.97 (2005), p. 123.

[40] Eliasmith, Chris. “How Neurons Mean: A Neurocomputational Theory of Representational Content”. PhD thesis. Washington University, 2000.

[41] Eliasmith, Chris and Anderson, Charles H. “Neural engineering”. In: *Massachusetts Institue of Technology* (2003).

[42] Feigenson, Lisa and Carey, Susan. “On the limits of infants’ quantification of small object arrays”. In: *Cognition* 97.3 (2005), pp. 295–313.

[43] Feigenson, Lisa, Carey, Susan, and Hauser, Marc. “The representations underlying infants’ choice of more: Object files versus analog magnitudes”. In: *Psychological Science* 13.2 (2002), pp. 150–156.

[44] Fias, Wim. “The importance of magnitude information in numerical processing: Evidence from the SNARC effect”. In: *Mathematical cognition* 2.1 (1996), pp. 95–110.

[45] Fodor, Jerry A and Pylyshyn, Zenon W. “Connectionism and cognitive architecture: A critical analysis”. In: *Cognition* 28.1 (1988), pp. 3–71.

[46] Franconeri, SL, Bemis, DK, and Alvarez, GA. “Number estimation relies on a set of segmented objects”. In: *Cognition* 113.1 (2009), pp. 1–13.

[47] Franconeri, SL, Jonathan, SV, and Scimeca, JM. “Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity”. In: *Psychological science* (2010).

[48] Frank, Michael C et al. “Number as a cognitive technology: Evidence from Pirahã language and cognition”. In: *Cognition* 108.3 (2008), pp. 819–824.

[49] Frege, Gottlob. *Die Grundlagen der Arithmetik, translated by JL Austin as The Foundations of Arithmetic*. 1959.

[50] Gallistel, Charles R and Gelman, Rochel. “Non-verbal numerical cognition: From reals to integers”. In: *Trends in cognitive sciences* 4.2 (2000), pp. 59–65.

[51] Gallistel, Charles R and Gelman, Rochel. “Preverbal and verbal counting and computation”. In: *Cognition* 44.1 (1992), pp. 43–74.

[52] Garson, James. “Connectionism”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Zalta, Edward N. Winter 2016. Metaphysics Research Lab, Stanford University, 2016.

[53] Gelman, Rochel. “The case for continuity”. In: *Behavioral and Brain Sciences* 34.03 (2011), pp. 127–128.

[54] Gelman, Rochel and Butterworth, Brian. “Number and language: how are they related?” In: *Trends in cognitive sciences* 9.1 (2005), pp. 6–10.

[55] Gelman, Rochel and Gallistel, Charles R. *The child's understanding of number*. Harvard University Press, 1986.

[56] Gliksman, Yarden, Weinbach, Noam, and Henik, Avishai. “Alerting cues enhance the subitizing process”. In: *Acta Psychologica* 170 (2016), pp. 139–145.

[57] Göbel, Silke M et al. “Children‚Äôs arithmetic development: It is number knowledge, not the approximate number sense, that counts”. In: *Psychological science* 25.3 (2014), pp. 789–798.

[58] Gottlieb, Jacqueline. “From thought to action: the parietal cortex as a bridge between perception, action, and cognition”. In: *Neuron* 53.1 (2007), pp. 9–16.

[59] Graziano, Mario. “Numerical cognition and philosophy of mathematics. Dehaene‚Äôs (neuro) intuitionism and the relevance of language”. In: *Rivista Italiana di Filosofia del Linguaggio* 2 (2015).

[60] Hanna, Robert. *Cognition, Content, and the a Priori: A Study in the Philosophy of Mind and Knowledge*. Oxford University Press, 2015.

[61] Hatcher, Allen. “Algebraic topology”. In: *Cambridge UP, Cambridge* 606.9 (2002).

[62] Hauser, Marc D, Chomsky, Noam, and Fitch, W Tecumseh. “The faculty of language: What is it, who has it, and how did it evolve?” In: *science* 298.5598 (2002), pp. 1569–1579.

[63] He, Li-Feng, Chao, Yu-Yan, and Suzuki, Kenji. “An algorithm for connected-component labeling, hole labeling and Euler number computing”. In: *Journal of Computer Science and Technology* 28.3 (2013), pp. 468–478.

[64] He, Lixia et al. “Connectedness affects dot numerosity judgment: Implications for configural processing”. In: *Psychonomic Bulletin & Review* 16.3 (2009), pp. 509–517.

[65] He, Lixia et al. “Topology-defined units in numerosity perception”. In: *Proceedings of the National Academy of Sciences* 112.41 (2015), E5647–E5655.

[66] Hecht, Selig. “The visual discrimination of intensity and the Weber-Fechner law”. In: *The Journal of general physiology* 7.2 (1924), pp. 235–267.

[67] Herman, Gabor T. “Oriented surfaces in digital spaces”. In: *CVGIP: Graphical Models and Image Processing* 55.5 (1993), pp. 381–396.

[68] Heyer, Ken den and Briand, Kevin. “Priming Single Digit Numbers: Automatic Spreading Activation Dissipates as a Function of Semantic Distance”. In: *The American Journal of Psychology* 99.3 (1986), pp. 315–340.

[69] Hinton, Geoffrey E, Osindero, Simon, and Teh, Yee-Whye. “A fast learning algorithm for deep belief nets”. In: *Neural computation* 18.7 (2006), pp. 1527–1554.

[70] Holloway, Ian D et al. “Semantic and perceptual processing of number symbols: evidence from a cross-linguistic fMRI adaptation study”. In: *Journal of cognitive neuroscience* 25.3 (2013), pp. 388–400.

[71] Huang, Y., Zhou, T., and Chen, L. “The precedence of topological change over top-down attention in masked priming”. In: *Journal of Vision* 11.12 (2011), pp. 9–9.

[72] Hurewitz, Felicia, Gelman, Rochel, and Schnitzer, Brian. “Sometimes area counts more than number”. In: *Proceedings of the National Academy of Sciences* 103.51 (2006), pp. 19599–19604.

[73] Im, Hee Yeon, Zhong, Sheng-hua, and Halberda, Justin. “Grouping by proximity and the visual impression of approximate number in random dot arrays”. In: *Vision research* 126 (2016), pp. 291–307.

[74] Izard, Véronique et al. “Exact equality and successor function: Two key concepts on the path towards understanding exact numbers”. In: *Philosophical Psychology* 21.4 (2008), pp. 491–505.

[75] Jevons, W Stanley. “The power of numerical discrimination”. In: *Nature* 3.67 (1871), pp. 281–282.

[76] Kahneman, Daniel, Treisman, Anne, and Gibbs, Brian J. "The reviewing of object files: Object-specific integration of information". In: *Cognitive psychology* 24.2 (1992), pp. 175–219.

[77] Kaminski, Jennifer. "Young Children's Understanding of the Successor Function." In: *CogSci*. 2015.

[78] Kaufman, Edna L et al. "The discrimination of visual number". In: *The American journal of psychology* 62.4 (1949), pp. 498–525.

[79] Kirjakovski, Atanas and Matsumoto, Eriko. "Numerosity underestimation in sets with illusory contours". In: *Vision research* 122 (2016), pp. 34–42.

[80] Kluth, Tobias and Zetzsche, Christoph. "Numerosity as a topological invariant". In: *Journal of vision* 16.3 (2016), pp. 30–30.

[81] Knops, André et al. "A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration". In: *Journal of Neuroscience* 34.30 (2014), pp. 9857–9866.

[82] Kohonen, Teuvo. "The self-organizing map". In: *Neurocomputing* 21.1 (1998), pp. 1–6.

[83] Krajcsi, Attila, Szabó, Eszter, and Mórocz, István Ákos. "Subitizing is sensitive to the arrangement of objects". In: *Experimental psychology* (2013).

[84] Kripke, Saul A. "Naming and necessity". In: *Semantics of natural language*. Springer, 1972, pp. 253–355.

[85] Krysztofiak, Wojciech. "Representational structures of arithmetical thinking: part I". In: *Axiomathes* 26.1 (2016), pp. 1–40.

[86] Le Corre, Mathieu and Carey, Susan. "One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles". In: *Cognition* 105.2 (2007), pp. 395–438.

[87] Le Mouel, Charlotte and Pouget, Alexandre. "Computations with approximate neural representations of number".

[88] Lee, Honglak, Ekanadham, Chaitanya, and Ng, Andrew Y. "Sparse deep belief net model for visual area V2". In: *Advances in neural information processing systems*. 2008, pp. 873–880.

[89] Leslie, Alan M, Gelman, Rochel, and Gallistel, CR. “The generative basis of natural number concepts”. In: *Trends in cognitive sciences* 12.6 (2008), pp. 213–218.

[90] Leslie, Alan M, Gelman, Rochel, and Gallistel, CR. “Where integers come from”. In: *Festschrift for Jean Mandler* (2007).

[91] Libertus, Melissa E, Feigenson, Lisa, and Halberda, Justin. “Is approximate number precision a stable predictor of math ability?” In: *Learning and individual differences* 25 (2013), pp. 126–133.

[92] Lindskog, Marcus, Winman, Anders, and Poom, Leo. “Individual differences in nonverbal number skills predict math anxiety”. In: *Cognition* 159 (2017), pp. 156–162.

[93] Locke, John. *An essay concerning human understanding*. T. Tegg and Son, 1836.

[94] Luuk, Erkki. “Natural language–no infinity and probably no recursion”. In: *Proceedings of the Cognitive Science Society*. Vol. 34. 34. 2012.

[95] Lycan, William G. *Logical form in natural language*. Bradford Books, 1984.

[96] Lyons, Ian M, Ansari, Daniel, and Beilock, Sian L. “Qualitatively different coding of symbolic and nonsymbolic numbers in the human brain”. In: *Human brain mapping* 36.2 (2015), pp. 475–488.

[97] Lyons, Ian M, Ansari, Daniel, and Beilock, Sian L. “Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent.” In: *Journal of Experimental Psychology: General* 141.4 (2012), p. 635.

[98] Macmillan, Neil A and Creelman, C Douglas. *Detection theory: A user’s guide*. Psychology press, 2004.

[99] Majka, Piotr et al. “3D Brain Atlas Reconstructor Service,ÄîOnline Repository of Three-Dimensional Models of Brain Structures”. In: *Neuroinformatics* 11.4 (2013), pp. 507–518.

[100] Mandler, George and Shebo, Billie J. “Subitizing: An analysis of its component processes”. In: *Journal of Experimental Psychology: General* 111.1 (1982), p. 1.

[101] Marr, David. “Vision: A computational investigation into the human representation and processing of visual information, henry holt and co”. In: *Inc., New York, NY* 2 (1982), pp. 4–2.

[102] Mayberry, JP. “Natural Numbers and Infinitesimals”. In: *Logic, Mathematics, Philosophy, Vintage Enthusiasms*. Springer, 2011, pp. 255–286.

[103] Mazurek, Mark E et al. “A role for neural integrators in perceptual decision making”. In: *Cerebral cortex* 13.11 (2003), pp. 1257–1269.

[104] Mazzocco, Michèle MM, Feigenson, Lisa, and Halberda, Justin. “Preschoolers’ precision of the approximate number system predicts later school mathematics performance”. In: *PLoS one* 6.9 (2011), e23749.

[105] Minsky, Marvin. “Minds are simply what brains do”. 1997.

[106] Minsky, Marvin and Papert, Seymour. *Perceptrons*. MIT press, 1969.

[107] Morgan, Michael J et al. “A texture-processing model of the ‘visual sense of number’”. In: *Proceedings of the Royal Society of London B: Biological Sciences* 281.1790 (2014), p. 20141137.

[108] Moyer, Robert S and Landauer, Thomas K. “Time required for judgements of numerical inequality”. In: (1967).

[109] Newell, Allen. *Unified theories of cognition*. Harvard University Press, 1994.

[110] Nieder, Andreas. “Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices”. In: *Proceedings of the National Academy of Sciences* 109.29 (2012), pp. 11860–11865.

[111] Nieder, Andreas and Dehaene, Stanislas. “Representation of number in the brain”. In: *Annual review of neuroscience* 32 (2009), pp. 185–208.

[112] Nieder, Andreas, Freedman, David J, and Miller, Earl K. “Representation of the quantity of visual items in the primate prefrontal cortex”. In: *Science* 297.5587 (2002), pp. 1708–1711.

[113] Nieder, Andreas and Merten, Katharina. “A labeled-line code for small and large numerosities in the monkey prefrontal cortex”. In: *Journal of Neuroscience* 27.22 (2007), pp. 5986–5993.

[114] Nieder, Andreas and Miller, Earl K. “A parieto-frontal network for visual numerical information in the monkey”. In: *Proceedings of the National Academy of Sciences of the United States of America* 101.19 (2004), pp. 7457–7462.

[115] Nieder, Andreas and Miller, Earl K. “Coding of cognitive magnitude: Compressed scaling of numerical information in the primate prefrontal cortex”. In: *Neuron* 37.1 (2003), pp. 149–157.

[116] Palmer, Stephen and Rock, Irvin. "Rethinking perceptual organization: The role of uniform connectedness". In: *Psychonomic bulletin & review* 1.1 (1994), pp. 29–55.

[117] Parker, Andrew J and Newsome, William T. "Sense and the single neuron: probing the physiology of perception". In: *Annual review of neuroscience* 21.1 (1998), pp. 227–277.

[118] Peterson, Scott A and Simon, Tony J. "Computational evidence for the subitizing phenomenon as an emergent property of the human cognitive architecture". In: *Cognitive Science* 24.1 (2000), pp. 93–122.

[119] Piantadosi, Steven T, Tenenbaum, Joshua B, and Goodman, Noah D. "Bootstrapping in a language of thought: A formal model of numerical concept learning". In: *Cognition* 123.2 (2012), pp. 199–217.

[120] Piazza, Manuela et al. "A magnitude code common to numerosities and number symbols in human intraparietal cortex". In: *Neuron* 53.2 (2007), pp. 293–305.

[121] Piazza, Manuela et al. "Are subitizing and counting implemented as separate or functionally overlapping processes?" In: *Neuroimage* 15.2 (2002), pp. 435–446.

[122] Piazza, Manuela et al. "Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia". In: *Cognition* 116.1 (2010), pp. 33–41.

[123] Piazza, Manuela et al. "Single-trial classification of parallel pre-attentive and serial attentive processes using functional magnetic resonance imaging". In: *Proceedings of the Royal Society of London B: Biological Sciences* 270.1521 (2003), pp. 1237–1245.

[124] Piazza, Manuela et al. "Subitizing reflects visuo-spatial object individuation capacity". In: *Cognition* 121.1 (2011), pp. 147–153.

[125] Piazza, Manuela et al. "Tuning curves for approximate numerosity in the human intraparietal sulcus". In: *Neuron* 44.3 (2004), pp. 547–555.

[126] Plaisier, Myrthe A, Van Polanen, Vonne, and Kappers, Astrid ML. "The role of connectedness in haptic object perception". In: *Scientific Reports* 7 (2017).

[127] Poincaré, Henri. *Dernières pensées*. Flammarion, 1913.

[128] Poincaré, Henri. *Science and hypothesis*. Science Press, 1905.

[129] Poston, Tim. "Fuzzy geometry". PhD thesis. University of Warwick, 1971.

[130] Purushothaman, Gopathy and Bradley, David C. “Neural population code for fine perceptual decisions in area MT”. In: *Nature neuroscience* 8.1 (2005), pp. 99–106.

[131] Pylyshyn, Zenon W. “Visual indexes, preconceptual objects, and situated vision”. In: *Cognition* 80.1 (2001), pp. 127–158.

[132] Railo, Henry et al. “The role of attention in subitizing”. In: *Cognition* 107.1 (2008), pp. 82–104.

[133] Raphael, Sabine and Morgan, Michael J. “The computation of relative numerosity, size and density”. In: *Vision research* 124 (2016), pp. 15–23.

[134] Reveley, Colin et al. “Three-dimensional digital template atlas of the macaque brain”. In: *Cerebral Cortex* (2016).

[135] Revkin, Susannah K et al. “Does subitizing reflect numerical estimation?” In: *Psychological Science* 19.6 (2008), pp. 607–614.

[136] Reynvoet, Bert, Brysbaert, Marc, and Fias, Wim. “Semantic priming in number naming”. In: *The Quarterly Journal of Experimental Psychology Section A* 55.4 (2002), pp. 1127–1139.

[137] Rips, Lance J. “How many is a zillion? Sources of number distortion”. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 39.4 (2013), p. 1257.

[138] Rips, Lance J, Asmuth, Jennifer, and Bloomfield, Amber. “Can statistical learning bootstrap the integers?” In: *Cognition* 128.3 (2013), pp. 320–330.

[139] Rips, Lance J, Asmuth, Jennifer, and Bloomfield, Amber. “From numerical concepts to concepts of number”. In: *Behavioral and Brain Sciences* 31.06 (2008), pp. 623–642.

[140] Rips, Lance J, Asmuth, Jennifer, and Bloomfield, Amber. “Giving the boot to the bootstrap: How not to learn the natural numbers”. In: *Cognition* 101.3 (2006), B51–B60.

[141] Rips, Lance J and Hespos, Susan J. “Rebooting the bootstrap argument: Two puzzles for bootstrap theories of concept development”. In: *Behavioral and Brain Sciences* 34.03 (2011), pp. 145–146.

[142] Rock, Irvin and Palmer, Stephen. “Gestalt psychology”. In: *Sci Am* 263 (1990), pp. 84–90.

[143] Roelfsema, Pieter R. “Cortical algorithms for perceptual grouping”. In: *Annu. Rev. Neurosci.* 29 (2006), pp. 203–227.

[144] Roelfsema, Pieter R and Houtkamp, Roos. “Incremental grouping of image elements in vision”. In: *Attention, Perception, & Psychophysics* 73.8 (2011), pp. 2542–2572.

[145] Roitman, Jamie D, Brannon, Elizabeth M, and Platt, Michael L. “Monotonic coding of numerosity in macaque lateral intraparietal area”. In: *PLoS biology* 5.8 (2007), e208.

[146] Rosenfeld, Azriel. “Digital topology”. In: *American Mathematical Monthly* (1979), pp. 621–630.

[147] Ross, John. “Visual discrimination of number without counting”. In: *Perception* 32.7 (2003), pp. 867–870.

[148] Ross, John and Burr, David C. “Vision senses number directly”. In: *Journal of Vision* 10.2 (2010), pp. 10–10.

[149] Santens, Seppe et al. “Number processing pathways in human parietal cortex”. In: *Cerebral Cortex* 20.1 (2010), pp. 77–88.

[150] Sasanguie, Delphine et al. “The approximate number system is not predictive for symbolic number processing in kindergarteners”. In: *The Quarterly Journal of Experimental Psychology* 67.2 (2014), pp. 271–280.

[151] Sathian, Krish et al. “Neural evidence linking visual object enumeration and attention”. In: *Journal of Cognitive Neuroscience* 11.1 (1999), pp. 36–51.

[152] Sawamura, Hiromasa, Shima, Keisetsu, and Tanji, Jun. “Numerical representation for action in the parietal cortex of the monkey”. In: *Nature* 415.6874 (2002), pp. 918–922.

[153] Sengupta, Rakesh, Surampudi, Bapi Raju, and Melcher, David. “A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network”. In: *Brain research* 1582 (2014), pp. 114–124.

[154] Shaki, Samuel and Fischer, Martin H. “Reading space into numbers: a cross-linguistic comparison of the SNARC effect”. In: *Cognition* 108.2 (2008), pp. 590–599.

[155] Shapiro, Stewart. *Philosophy of mathematics: Structure and ontology*. Oxford University Press on Demand, 1997.

[156] Sossinsky, Alexei B. "Tolerance space theory and some applications". In: *Acta Applicandae Mathematica* 5.2 (1986), pp. 137–167.

[157] Spelke, Elizabeth S. "Principles of Object Perception". In: *Cognitive Science* 14.1 (1990), pp. 29–56.

[158] Spelke, Elizabeth S. "Quinian bootstrapping or Fodorian combination? Core and constructed knowledge of number". In: *Behavioral and Brain Sciences* 34.03 (2011), pp. 149–150.

[159] Stenning, Keith and Van Lambalgen, Michiel. *Human reasoning and cognitive science*. MIT Press, 2012.

[160] Stoianov, Ivilin and Zorzi, Marco. "Emergence of a'visual number sense'in hierarchical generative models". In: *Nature neuroscience* 15.2 (2012), pp. 194–196.

[161] Strang, Gilbert. *Introduction to linear algebra*. Vol. 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[162] Strang, Gilbert. *Linear algebra and its application, 4th edition*. Thomson, Brooks/Cole, 2006.

[163] Stuss, Donald T and Knight, Robert T. *Principles of frontal lobe function*. Oxford University Press, 2002.

[164] Testolin, Alberto et al. "Deep unsupervised learning on a desktop PC: a primer for cognitive scientists". In: *Frontiers in psychology* 4 (2013), p. 251.

[165] Thurstone, Louis L. "A law of comparative judgment". In: *Psychological review* 34.4 (1927), p. 273.

[166] Tibber, Marc, Greenwood, JA, and Dakin, SC. "Psychophysical evidence for a common metric underlying number and density discrimination". In: *J Vis* 10.11.11 (2011), p. 1205.

[167] Toet, Alexander. "Visual perception of spatial order". PhD thesis. 1987.

[168] Tokita, Midori and Ishiguchi, Akira. "How might the discrepancy in the effects of perceptual variables on numerosity judgment be reconciled?" In: *Attention, Perception, & Psychophysics* 72.7 (2010), pp. 1839–1853.

[169] Treisman, Anne and Gelade, Garry. "A feature-integration theory of attention". In: *Cognitive Psychology* 12.1 (1980), pp. 97–136.

[170] Treisman, Anne and Souther, Janet. “Search asymmetry: A diagnostic for preattentive processing of separable features”. In: *Journal of Experimental Psychology: General* 114.3 (1985), p. 285.

[171] Trick, Lana M and Pylyshyn, Zenon W. “Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision.” In: *Psychological review* 101.1 (1994), p. 80.

[172] Tudusciuc, Oana and Nieder, Andreas. “Comparison of length judgments and the Müller-Lyer illusion in monkeys and humans”. In: *Experimental brain research* 207.3-4 (2010), pp. 221–231.

[173] Tudusciuc, Oana and Nieder, Andreas. “Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex”. In: *Proceedings of the National Academy of Sciences* 104.36 (2007), pp. 14513–14518.

[174] Väänänen, Jouko. “Second-order logic and foundations of mathematics”. In: *Bulletin of Symbolic Logic* 7.04 (2001), pp. 504–520.

[175] Van Der Does, Jaap and Van Lambalgen, Michiel. “A logic of vision”. In: *Linguistics and Philosophy* 23.1 (2000), pp. 1–92.

[176] Van Oeffelen, Michiel P and Vos, Peter G. “A probabilistic model for the discrimination of visual number”. In: *Perception & Psychophysics* 32.2 (1982), pp. 163–170.

[177] Van Opstal, Filip and Verguts, Tom. “The origins of the numerical distance effect: the same–different task”. In: *Journal of Cognitive Psychology* 23.1 (2011), pp. 112–120.

[178] Verguts, Tom and Fias, Wim. “Interacting neighbors: A connectionist model of retrieval in single-digit multiplication”. In: *Memory & cognition* 33.1 (2005), pp. 1–16.

[179] Verguts, Tom and Fias, Wim. “Representation of number in animals and humans: a neural model”. In: *Journal of Cognitive Neuroscience* 16.9 (2004), pp. 1493–1504.

[180] Vetter, Petra, Butterworth, Brian, and Bahrami, Bahador. “Modulating attentional load affects numerosity estimation: evidence against a pre-attentive subitizing mechanism”. In: *PLoS One* 3.9 (2008), e3269.

[181] Von Glasersfeld, Ernst. “Subitizing: The role of figural patterns in the development of numerical concepts”. In: *Archives de Psychologie* (1982).

[182] Vuilleumier, Patrik O and Rafal, Robert D. “A systematic study of visual extinction: Between-and within-field deficits of attention in hemispatial neglect”. In: *Brain* 123.6 (2000), pp. 1263–1279.

[183] Walsh, Vincent. “A theory of magnitude: common cortical metrics of time, space and quantity”. In: *Trends in cognitive sciences* 7.11 (2003), pp. 483–488.

[184] Zeeman, E Christopher. “Topology of the brain and visual perception”. In: *Topology of 3-manifolds and related topics*. Prentice-Hall, 1961, pp. 240–256.

[185] Zimmermann, Eckart and Fink, Gereon R. “Numerosity perception after size adaptation”. In: *Scientific Reports* 6.1 (2016).

[186] Zorzi, M and Butterworth, B. “A computational model of number comparison”. In: (1999).

[187] Zorzi, M and Butterworth, B. “On the representation of number concepts”. In: (1997).