SAPIENZA
UNIVERSITÀ DI ROMA

Dottorato in Psicologia e Scienza Cognitiva

Faculty of Medicine and Psychology

Department of Psychology

Ph.D. Course in Cognitive Psychology and Science

32nd Cycle

Perceptual Grouping by Proximity and Orientation Bias:

Experimental and Modelling Investigations

Ph.D. Candidate:

Antonino Esposito

Supervisor:

Prof. Antonino Raffone

Co-Supervisors:

Prof. Marta Olivetti Belardinelli

Dr. Luca Simione

ABSTRACT

Grouping by proximity is the principle of perceptual organization by which the elements of a visual scene which are closer in space tend to be perceived as a coherent ensemble. Research into this topic makes substantial use of the class of stimuli known as dot lattices. The Pure Distance Law (Kubovy et al., 1998) predicts that the probability of grouping by proximity in these stimuli only depends on the relative inter-dot distance between competing organizations. Despite much effort to explain how grouping by proximity is shaped by the basic organization of visual stimuli, its neural mechanisms are still under debate. Moreover, previous studies reported that grouping in dot lattices also occurs according to an orientation bias, by which these stimuli are perceived along a preferred orientation (vertical), regardless of what predicted by the Pure Distance Law. The aim of this thesis is to shed light on the functional and neural mechanisms characterizing grouping by proximity in dot lattices, as well as the trade-off between proximity- and orientation-based grouping. Study 1 investigates the role of high-level visual working memory (VWM) in promoting for the shift between grouping by proximity and orientation bias. Both the quantity (load) and the quality (content) of the information stored in VWM shape online grouping for dot lattices. Study 2 presents a neural network model simulating the dynamics occurring between low- and high-level processing stages during dot lattices perception. The degree of synchrony between the units at low-level module has a key role in accounting for grouping by proximity. Overall, our results show that high-level (Study 1) and low-level (Study 2) operations contribute in parallel to the emergence of grouping by proximity, as well as to its reciprocity with orientation-based grouping.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

Chapter 1

Chapter 2

Chapter 3

Introduction

Given the limited resources of the visual system, humans and other species are constantly required to implement efficient perceptual strategies to account for the heterogeneity and complexity of the external world. To have sufficient chances of survival, they need to localize salient stimuli, even under conditions of reduced visibility or camouflage. To accomplish this goal, it is imperative to integrate the features belonging to a relevant object into a unitary and meaningful percept while, at the same time, segregating them from features that belong to non-relevant objects. Both processes are required to correctly identify objects, since different stimulus features project to separate sub-regions of the retinal mosaic (Ahnelt, 1998).

The mechanisms through which the visual system organizes the complex pattern of sensory stimulation into a spatio-temporal coherent ensemble are collectively known as *perceptual grouping*. The word “spatio-temporal” here denotes that perceptual grouping organizes, not only static patterns of stimulation, but also ones that move or change over time, including changes caused by movements of the eyes. A clear example is provided by the *phi phenomenon* (Wertheimer, 1912), by which the same single dot appearing at different locations over time is perceived as being in continuous motion.

Wertheimer (1923) was the first to describe a set of principles by which grouping processes occur. These are nowadays referred to as the classical principles of perceptual grouping. Some of them are illustrated in Figure I. Wertheimer showed that altering the distance between adjacent dots, such that some became closer than others, induced the observers to group the closest ones together as unitary structures. This first principle based on the spacing between separate elements was thus named as *proximity* (Figure IA). According to the principle of *similarity*, elements sharing some features are more strongly integrated in a whole than different stimuli (Figure IB). The principle of *common fate* posits that elements moving in the same direction tend to be perceived as belonging together (Figure IC). Panels D and E of Figure I show the principles of *good continuation* and *closure*, by which grouping emerges between elements aligned along a common direction or converging towards a closer shape, respectively. In more recent years, other principles have been added to the classical ones (for an exhaustive dissertation on classical and contemporary grouping principles see Brooks, 2015; Wagemans et al., 2012).

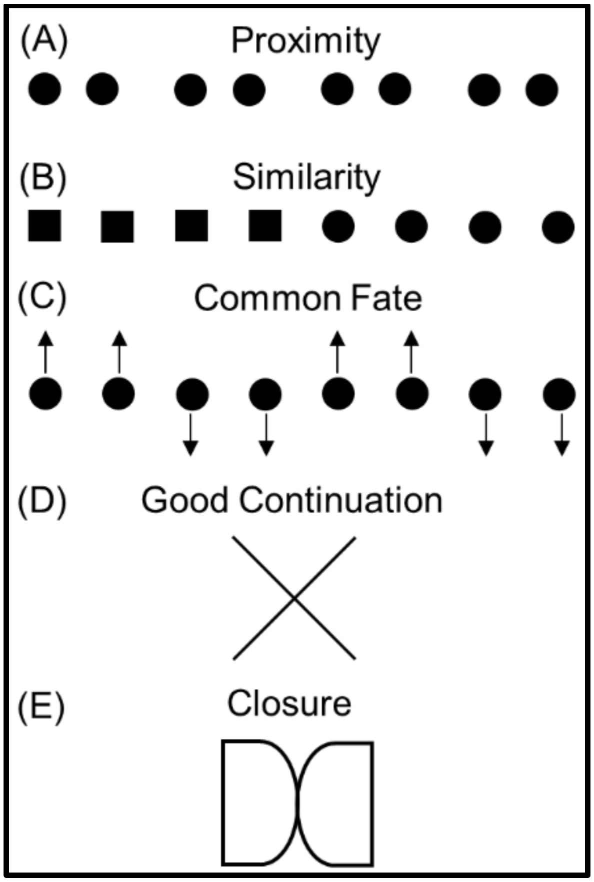


Figure I: Classical grouping principles individuated by Wertheimer (1923). (A). *Example of grouping by proximity. The dots tend to be grouped into four separate pairs with the shortest inter-dot distance in between.* (B). *Example of grouping by similarity. The elements of the array tend to* *be perceived as two groups depending on the shape.* (C). *Example of grouping by common fate. Dots moving in the same direction tend to be grouped together.* (D). *Example of grouping by good continuation. Observers prefer to perceive this patter as two straight lines intersecting in the middle, rather than as two 90° angles with touching vertices.* (E). *Example of grouping by closure. Elements composing closed patterns tend to be grouped together.*

Perceptual grouping has traditionally been considered as a set of operations taking place at early stages of visual hierarchy, namely in a preattentive fashion (Palmer, 2002a; Palmer et al., 2003; Peterson & Kimchi, 2014). These operations are meant to generate proto-representations of an object in the visual field, over which subsequent operations are applied, which would lead to a final percept. However, in more recent years, this assumption has been challenged by the evidence that several factors, such as past experience, expectations and attentional manipulations, can strongly shape the occurrence and the timing of grouping processes (see Peterson & Kimchi, 2014). This corpus of research contributed to the view of perceptual grouping as a highly complex process which arises from the interplay between bottom-up (early and automatic) and top-down (later and voluntary) processes.

The present thesis aims to deepen our understanding of the functional and neural mechanisms of perceptual grouping, particularly highlighting the role of the reciprocal relationship between low-level and high-level processing stages. In doing so, I specifically take into account the principle of grouping by proximity, even though many of the conclusions to which this work leads, might be extended to other grouping principles as well.

Chapter 1 introduces the issue of the objective measurement of perceptual grouping. The main focus is on psychophysical experiments aimed at quantitatively assessing the strength of grouping by proximity. The stimuli that most of these studies employed (i.e. the *dot lattices*) are extensively described. The chapter also introduces the issue of perceptual bias overruling grouping by proximity in dot lattices. Perceptual bias is understood as the top-down influence over low-level visual areas, reflecting the state of the observers.

Chapter 2 presents a behavioral experiment aimed at investigating how external stimulus factors, such as proximity, and internal factors, such as perceptual bias, shape perceptual grouping process. This study investigates the role of *visual working memory* (VWM) modulation in determining whether grouping process for ambiguous dot lattices occurs according to proximity principle or according to orientation bias. A dual-task paradigm is employed, where grouping evaluation is embedded in-between the encoding and the retrieval stages of a change-detection task, in which VWM is loaded with ungrouped information.

Chapter 3 introduces a dynamic network model accounting for the perception of grouping by proximity based on its quantitative measurement. In the model, grouping by proximity is achieved through the interplay of feedforward and lateral connections between low- and high-level modules. By replicating some of the experimental results exposed in Chapters 1 and 2, the model simulations offer a biologically plausible explanation of the neural mechanisms underlying grouping by proximity, as well as orientation bias interference.

The thesis ends with a general discussion of the results outlined in the two studies, and suggest possible developments for further research.

Chapter 1

Quantitative Assessments of Grouping by Proximity

1.1 The issue of Grouping Measurement

The long history of research investigating the operations governing perceptual grouping can be traced back to Wertheimer’s seminal paper (1923). His work was crucial in laying the foundations for our understating of how the perceptual organization of visual scenes is extracted and processed by our perceptual system. The method that Wertheimer (1923) used to illustrate classical principles of perceptual grouping was based on visual demonstration of sketches illustrating the principle to his audience. Thus, this approach fully relied on the qualitative experience of observers. This method may suggest which components of a visual scene are typically bound together as a whole, but did not provide any indication of the strength of the grouping process (Peterson & Kimchi, 2013; Wagemans et al., 2012a).

To assess this issue, a quantitative approach is needed (Palmer, 2002b). As pointed out by Palmer and Beck (2007), the distinction between qualitative and quantitative approaches to the study of perceptual grouping relates to the distinction between subjective and objective methods in perceptual studies. In subjective perceptual tasks, correct responses rely on the current experience of the observers and, as such, susceptible to distortion. In objective perceptual tasks, correct responses are defined on the physical parameters of the stimuli, and, as such, are observer-independent. In their perspective, quantitative estimation of grouping is best achieved when perceptual experience is unrelated to the subjective report of the participants.

This section provides an overview about such quantitative accounts to the investigation of perceptual grouping. Although several researches exposed here takes into account several principles of grouping, we will focus on proximity principle. As for many fields of psychology, research investigating grouping by proximity has used both indirect and direct methods, whose attributes are separately addressed in following .

1.2 Indirect Measures of Grouping by Proximity

One convenient strategy to overcome demand characteristics in perceptual research is to use tasks providing an indirect quantification of the processes under investigation (Palmer & Beck, 2007; Palmer & Nelson, 2000, Schmidt & Schmidt, 2013, Wagemans, 2018). The word “indirect” refers to the fact that perceptual judgments made by the observers do not reflect any task specification. In other words, participants are unaware about which dimension of visual stimuli represents the actual focus of the experiment (e.g. Gogel, 1976; Reingold & Merikle, 1988).

One of the most popular method to indirectly asses grouping by proximity is the repetition discrimination paradigm (RDT; Beck & Palmer, 2002; Palmer & Beck, 2007; Vickery, 2008, Vickery & Jiang, 2009). Participants are showed with a string of alternating shapes (i.e. squares and circles). At one point, one of the shapes is repeated instead of being alternated with the other. This constitute the target pair, which participants have to detect as quickly as possible. The distance between the shapes in the target pair and the distance between the target pair and the other elements of the string are orthogonally varied. The results show that participants are faster in detecting the target pair when repeating shapes were both closer each other and farther from other adjacent shapes (within-group condition). This procedure has provided comparable results with other grouping factors, such as similarity, common region, and connectedness (Palmer & Beck, 2007, Experiments 3-5).

A more recent paradigm has been developed by Wei et al. (2018). Their task is based on the well-known *tilt aftereffect* (TAE) typically found in adaptation studies (Dragoi et al., 2000; 2001; Fritsche et al., 2017; Jin et al., 2005; Webster, 2011). TAE refers to the fact that the perceived orientation of a stimulus (test) is altered after participants are exposed to a previous oriented stimulus (adaptor). In one of the conditions, adaptors are arrays of dots presenting a clear proximity organization. In this task, participants are required to respond to the orientation of the test stimulus, ignoring the previously presented adaptor. The results show that adaptors are still effective in inducing TAEs for subsequent test stimuli, with largest TAE implying a stronger proximity organization for the adaptors.

These two examples clearly illustrate that indirect methods configures themselves as a valid approach to quantify grouping by proximity, as well as other principles and their interaction (Wei et al., 2018, Experiment 3). The fact that participants are not aware of which dimension of perceptual experience is actually under investigation makes this type of tasks less susceptible to demand characteristics, preventing response biases. However, the disadvantage of not relying on the perceptual experience is that the performance in these tasks could reflect perceptual and cognitive processes other than grouping. In other words, it is not guaranteed that indirect methods to perceptual grouping measures what they are designed for (Brooks, 2015: Palmer & Beck, 2007, Schmidt & Schmidt, 2013). To overcome these limits, other scholars prefer investigating grouping phenomena through direct methods.

1.3 Direct Measures of Grouping by Proximity

Direct methods in grouping quantitative research involve presenting stimuli to the participants under multiple viewing conditions and collecting the reports about what they actually see (Wagemans, 2018). Although ruled by high experimental rigor, such approach derives from the qualitative demonstrations of early Gestaltists, and, as such, is based on the subjectivity of the observers. For this reason, a similar approach is referred to as *experimental phenomenology* (e.g. Albertazzi, 2015; Sinico, 2008), or, alternatively, as *phenomenological psychophysics* (Kubovy & Gephstein, 2003; Strother et al., 2003).

As argued by Kubovy and Gephstein (2003), the emphasis put on the phenomenal aspect indicates "A characteristic of or belonging to reality as perceived rather than as independent of mind.". This pits against idiosyncratic subjectivity, which strictly relates to the individuality of the perceiver, without being characterized by the sensory stimulation. Thus, phenomenology does not compromise the objectivity of psychological measurement, since it still is anchored to the appearance of a physical entity.

The application of phenomenological psychophysics to perceptual grouping has classically made use of a peculiar class of stimuli, knows as dot lattices. Given the importance of such stimuli for empirical section of the present work, their perceptual organization is extensively described.

1.3.1 The geometry of Dot Lattices

A dot lattice can be defined as a collection of dots on a plane, holding two fundamental proprieties: 1) the dots are discrete (i.e., they can be distinguished as singular elements); and 2) the dots entirely span the plane (i.e., the separation between them is not excessive).

The basic structure of a dot lattice is illustrated in Figure 1.1. Considering one dot **O** in the plane, we can define a distance ***a*** from its closest neighbor (**A**) and a distance ***b*** from its second closest neighbor (**B**). These distances coincide with the translation vectors **a** (with norm ||**a**||), and **b** (with norm ||**b**|| ≥ ||**a**||), respectively. The angle between these two vectors (γ) has been demonstrated to lie within the interval 60° ≤ γ ≤ 90° (Kubovy, 1994). **a**, **b** and γ together form the fundamental structure of a dot lattice, namely the *basic parallelogram*, with **a** and **b** being the shortest and longest sides, respectively. The basic parallelogram defines the spatial relationship between each quartet of dots in the stimulus, and, thus, the lattice itself (Figure 1.1).

The diagonals of the parallelogram are indicated as the vectors **c** and **d** and are parallel to the vectors having norm ||**a** - **b**|| and ||**a** + **b**||, respectively. In this way, all the possible directions between one dot and its eight surrounding dots are specified by these four vectors, which, are constrained to the relation ||**a**|| ≤ ||**b**|| ≤ ||**c**|| ≤ ||**d**||.

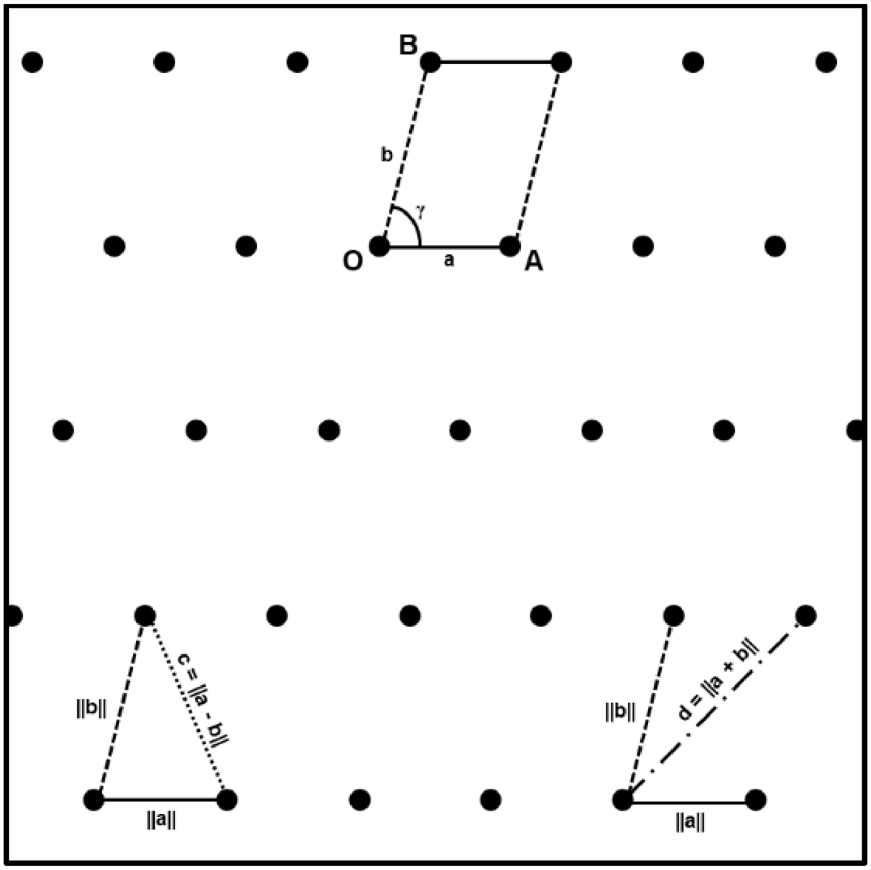


Figure 1.1: Basic organization of a dot lattice. The vector connecting **O** *and* **A** *is the shortest inter-dot distance and is typically indicated as* **a** *(with norm* ||**a**||*). The vector* **b** *(with norm* ||**b**|| ≥ ||**a**||*) connects* **O** *with* **B** *and, in this example, is the second shortest inter-dot distance. The angle these two vectors form is indicated by* γ*. Together,* **a**, **b**, *and* γ *constitute the basic parallelogram, which determines the spacing between four dots. The diagonals of the basic parallelogram are indicated as the vectors* **c** *and* **d**, *which have norm* ||**a** - **b**|| *and* ||**a** + **b**||, *respectively.*

Since **c** and **d** can be always computed starting from **a**, **b** and γ, these are the parameters useful to characterize a lattice. Keeping **a** constant, Bravais (1949), and subsequently Kubovy (1994), formulated the taxonomy of dot lattices by systematically varying b and γ. These authors individuated six classes of dot lattices, each of which presents different symmetry proprieties:

square, hexagonal, rhombic, oblique, rectangular, and centered rectangular (Figure 1.2). As we can see by Figure 1.2A, as ||**b**|| changes relative to ||**a**||, different types of lattices are defined. Such difference in length between inter-dot distances ***b*** and ***a*** (||**b**||/||**a**||) is known as *aspect ratio* (AR) (Kubovy, 1994; Kubovy & Wagemans, 1995). For example, squared dot lattices are characterized by AR = 1.0, which means that ||**b**|| is equal to ||**a**||, and thus ***a*** and ***b*** distances are characterized by the same inter-dot spacing. As the AR increases, the difference between a and b increases and a becomes the shortest inter-dot distance. For AR > 1.0 and γ fixed at 90°, dot lattices are classified as rectangular. Thus, squared dot lattices can be considered as a specific subclass of rectangular dot lattices, presenting AR = 1.0.

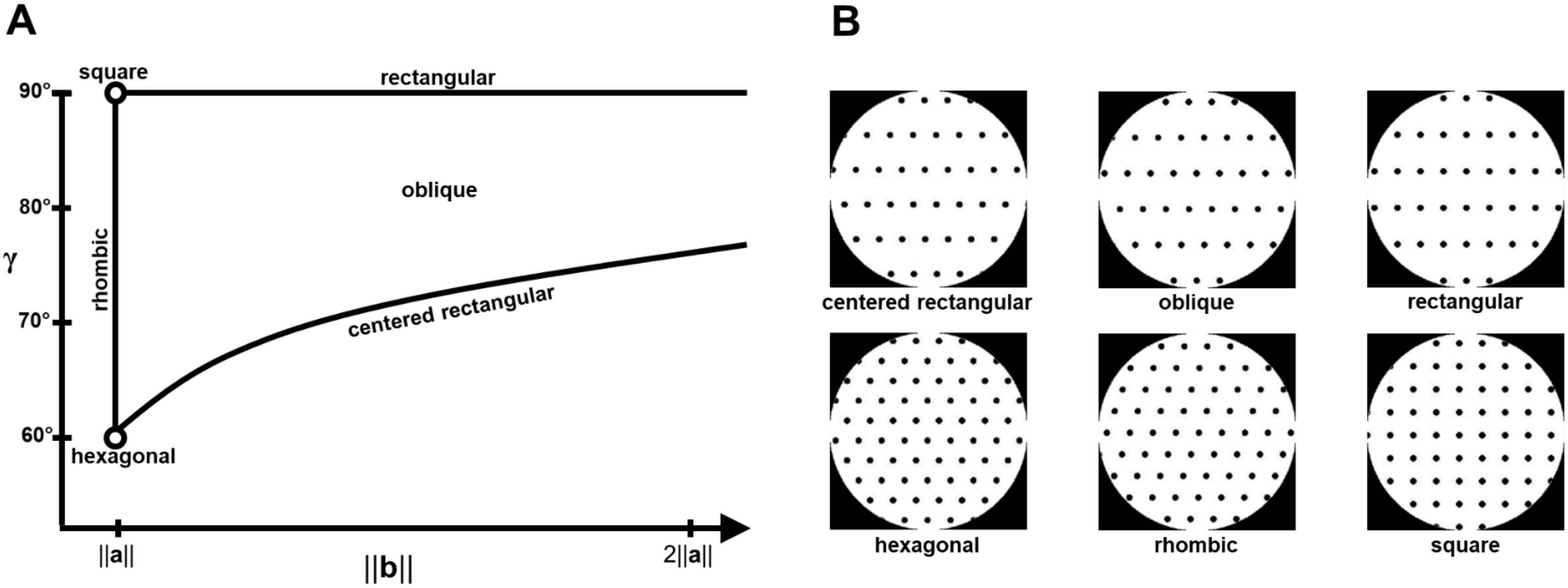


Figure 1.2: Taxonomy of dot lattices. (A). *The six classes of dot lattices are obtained by keeping constant the norm of vector* **a** *and varying* ||**b**|| *and* γ *accordingly. The ratio between vectors* **a** *and* **b** *is known as the aspect ratio* (AR = **b**/**a**)*, and it is an important parameter in determining how a dot lattice is perceived.* (B). *Illustration of one exemplar for each class. Note the circular configuration, which is typically how dot lattices are presented during experiments.*

Dot lattices are typically displayed in a circular configuration as in Figure 1.2B, where the dots can be perceived as aligned in a certain orientation. These orientations parallel the four distance vectors **a**, **b**, **c**, and **d**, and, therefore, henceforth we will refer to them as *a*, *b*, *c*, and *d*. Thus, grouping by proximity in these stimuli occurs when they are perceived as parallel strips of dots oriented along *a*. Notice that at AR = 1.0, inter-dot distance along orientation *b* is equal to the one along orientation *a*, and, thus, *b* defines grouping by proximity as well. Moreover, hexagonal dot lattices are specified by the condition $\|\mathbf{a}\| = \|\mathbf{b}\| = \|\mathbf{c}\|$ and, thus, in these stimuli grouping by proximity occurs along each of these orientations.

During a typical dot lattices experiment, the stimuli are presented in different orientations of *a*. Such orientations are specified by the angle between *a* and the horizontal diameter of the circular configuration of a dot lattice (θ), which ultimately defines the overall tilt of the stimulus (Figure 1.3).

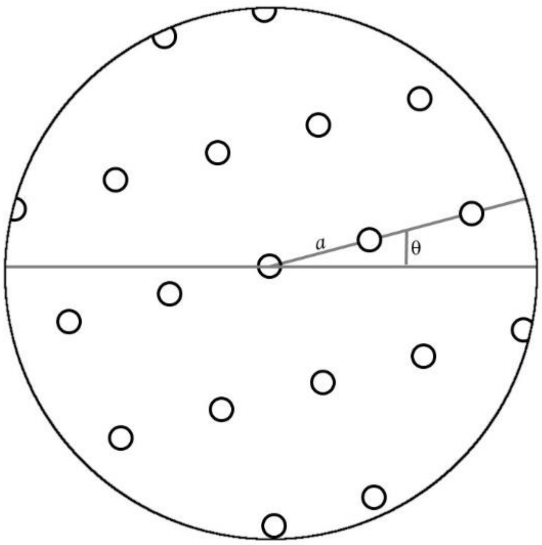


Figure 1.3: Example of oriented dot lattice. The angle θ *between a and the horizontal diameter of circular configuration defines the overall tilt of a dot lattice. In this exemplar,* AR = 1.2 *and* γ = 79°.

A last relevant feature of dot lattices is given by the fact that by increasing ||**a**||, the dot density of a finite dimension lattice is reduced, and vice versa. In other words, ||**a**|| represents the spatial scale of the stimulus.

1.3.2 The Pure Distance Law

In one of the first studies systematically varying all the parameters of dot lattices outlined above was conducted by Kubovy and Wagemans (1995). They presented participants with several types of dot lattices sampled the two-parameter stimulus space (Figure 1.2A). Stimuli lasted for 300 ms, after which participants were forced to report their perceived orientation. The frequencies for each orientation were then treated as an estimate of the strength of grouping by proximity and were plotted as a function of the relative inter-dot distance. The authors found that the probability to group the dots along one orientation *v* exponentially decreased with increasing its distance (***v***) relative to the shortest inter-dot distance (***a***):

$$\frac{p(\boldsymbol{v})}{p(\boldsymbol{a})} = e^{-\alpha\left(\frac{||\mathbf{v}||}{||\mathbf{a}||} - 1\right)} \qquad [1.1]$$

where the term ||**v**||/||**a**|| represents the length of the distance between two dots along orientation *v* scaled by the shortest inter-dot distance ***a*** and α (> 0) is referred to as *attraction constant* and defines the individual sensitivity to proximity (Figure 1.4). This function was defined *attraction function* because it states that the tendency to group together two dots (attraction) diminishes with increasing the distance between them.

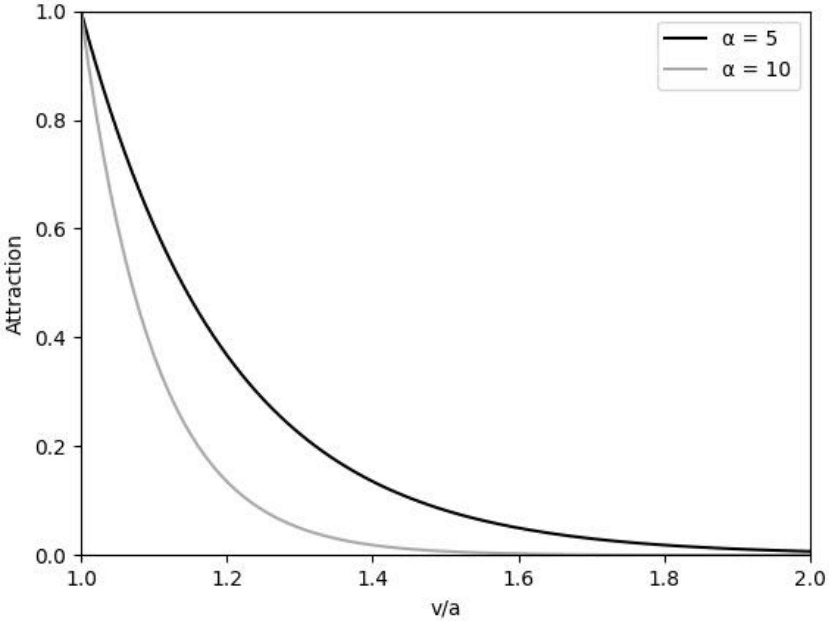
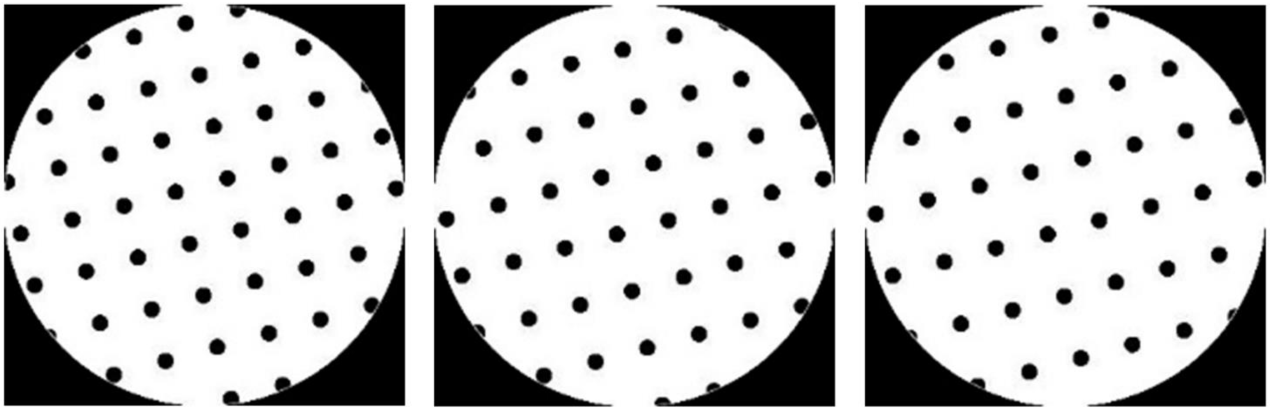


Figure 1.4: Attraction function for different values of α*. The tendency to group the dots along an orientation v other than a diminishes as the inter-dot distance along v* (**v**/**a**) *increases. In* Equation 1.1, α *defines an individual parameter of sensitivity to proximity. Thus, the higher* α*, the higher observers tend to group the dots according to proximity principle (i.e. attraction along v is reduced).*

In other words, attraction function predicts that the probability to group the dots by proximity decreases as the inter-dot distance along orientations other than *a* diminishes. Since the structural organization of dot lattices imposes that $||\mathbf{b}|| \leq ||\mathbf{c}|| \leq ||\mathbf{d}||$, reducing the distance along *b* also increases the attraction along *c* and *d*. In this sense, the probability of grouping by proximity in dot lattices depends on their AR (Figure 1.5). Recall that when AR = 1.0, inter-dot distances along *a* and *b* have the same length. In this case, the attraction function predicts that the perception of the dot lattices along these orientations is equiprobable. Thus, since both interpretations are equally likely, stimuli at AR = 1.0 present a higher degree of perceptual ambiguity (Kubovy, 1994). This model was therefore denominated *Pure Distance Law*, since, the only stimulus parameter predicting the occurrence of grouping by proximity is the relative distance between the single elements (Kubovy et al., 1998, but see Bleumers et al., 2008, for a discussion of how the scale of peripherally presented lattices affects the grouping).



AR = 1.0 AR = 1.1 AR = 1.2

Figure 1.5: Probability of grouping by proximity in dot lattices at different values of AR. *The probability to group the dots along a increases with the* AR. *At* AR = 1.0, **a** *and* **b** *have the same length and the attraction function predicts the same probability of grouping along both distances. Since the different stimulus interpretations are equally likely, dot lattices at* AR = 1.0 *present the highest level of perceptual ambiguity.*

The phenomenological psychophysics approach has been applied, among others, to investigate the interaction between proximity and similarity (Kubovy & Gephstein, 2000; Kubovy & van den Berg, 2008, Luna & Montoro, 2011). The aim of these studies was to characterize how the strength of grouping by proximity varies with the similarity between the elements of the stimuli, and vice versa. This has been pursued by combining the two factors in the stimuli such that they were either in accordance or in competition. Proximity and similarity contributed independently to grouping processes, whereas their joint effect was given by the sum of both their effects in isolation. Similar additive effects have been found for the combination of proximity and collinearity (Claessens & Wagemans, 2005; 2008), and between proximity and curvature, at least as they were not in conflict (Strother & Kubovy, 2006; 2012). Interestingly, these latter studies found that curvature overruled proximity when grouping along the two dimensions competed. The authors interpreted these results in terms of the non-accidental character (Blusseau et al., 2016; Kayaert & Wagemans, 2010; Vogels et al., 2001; Witkin & Tenenbaum, 1983) of curvilinear patterns in the stimuli. Participants were more likely to group along the most curvilinear orientation because they considered it less likely to be a casual factor in the stimulus and, as such, it gained higher weight in the process of selection. This suggests that grouping by proximity can be modulated by a top-down factor representing acquired knowledge about external contingences. This information provided a bias on visual processes responsible for perceptual grouping, which overruled the stimulus support (Gephstein & Kubovy, 2005) brought by the proximity of the dots.

1.4 Orientation Bias in Dot Lattices

Numerous studies on visual discrimination of oriented stimuli found evidence for an orientation bias, by which participants are more likely to perceive the stimuli in a preferred direction (e.g. Appelle, 1972; Girshick et al., 2011; Nasr & Tootell, 2012). Using simple configurations of stimuli, such as oriented gratings, these studies have reported evidence for an *oblique effect*. That is, participants are better in discriminating cardinal orientations (i.e. horizontal and vertical) than oblique ones (but see Essock et al., 2003; Hansen & Essok, 2004, for different results using broadband natural scenes). Single cell recordings as well as neuroimaging studies have linked such effect to the intrinsic properties neural populations tuned to cardinal orientations, such as their higher distribution within V1 and their narrower receptive fields (Chapman & Bonhoeffer, 1998; Furmanski & Engel, 2000; Li et al., 2003).

In dot lattices, several studies provided evidence that grouping operations are accompanied by a bias for the vertical orientation (Bleumers et al., 2008; Claessens & Wagemans, 2008; Nikolaev et al., 2016). In these studies, participants presented with rectangular dot lattices were more likely to group the dots along the vertical orientation, regardless of what was predicted by the Pure Distance Law function. Thus, the preference for vertical orientation during dot lattices dicrimination is accompanied by a reduction of proximity sensitivity (α parameter of the attraction function). However, the opposite could also be true: random fluctuations in the α parameter during some trials may render participants more susceptible to the influence of vertical bias (Bleumers et al., 2008).

Gephstein & Kubovy (2005) proposed that orientation bias during dot lattice perception may reflect a situation in which the amount of information brought by the stimulus to the perceptual system is low. In that case, the final decision relative to the orientation of the lattices is taken on the basis of the internal state of the observer.

Physiological and neuroimaging studies link these forms of perceptual bias to the top-down feedbacks from fronto-temporal networks over primary and secondary visual areas (V1 and V2), reflecting inferential (Frith & Dolan, 1997; Hohwy, 2017; Kok et al., 2013; Summersfield & de Lange, 2014; Wang et al., 2013), or associative (Albright, 2012) processes that may help people to deal with perceptual ambiguity. Accordingly, vertical preference in rectangular dot lattices was stronger for multistable stimuli at AR = 1.0, whereas it decreases as the stimuli become less ambiguous (Nikolaev et al., 2016). Grouping in dot lattices may therefore be intended as the contraposition of two forces: the current stimulus and stimulus expectations (Nikolaev et al., 2008).

Several studies on dot lattices considered preference for vertical orientation a noise factor, which could be balanced by presenting rotated versions of the stimuli, such that the angle θ always appeared to be rotated to a different degree (e.g. Kubovy & Wagemans, 1995; Kubovy et al., 1998). However, other studies found that its occurrence was modulated by several contextual factors, such as the recent visual experience (Gephstein & Kubovy, 2005) and perceptual learning (Nikolaev et al., 2016).

This suggests that orientation bias can exert systematic effects on dot lattices perception, which can alter the fate of ongoing grouping process. In the next sections, I will present a phenomenological psychophysics experiment aimed to shed light on the trade-off between grouping by proximity (as prospected by Pure Distance Law) and orientation bias in rectangular dot lattices.

Chapter 2

Top-Down Modulations on Grouping by Proximity and Orientation Bias

2.1 Introduction

The Pure Distance Law (Kubovy et al., 1998) assumes that the probability to group the dots in a lattice along one orientation purely depends on the relative distance between the dots along such orientation. In this sense, the emergence of perceptual organization in those stimuli is entirely determined by their bottom-up layout. However, top-down factors have been reported to alter grouping processes, overshadowing the distance between the dots. Strother and Kubovy (2006; 2012) used rectangular dot lattices, with different degrees of curvature along *a* and *b* orientations. They found that participants are more likely to group the lattices according to the highly curved direction, regardless of what predicted by the AR of the stimuli. The authors explained these results in terms of non-accidental character (Blusseau et al., 2016; Kayaert & Wagemans, 2010; Vogels et al., 2001; Witkin & Tenenbaum, 1983) of the curvilinear pattern of the stimuli. In other words, the curvature is judged as less likely to occur by chance in that context and, as such, it gained higher weight in the process of selection. This result suggests that grouping by proximity can be modulated by a top-down mechanism that biases the representation of one feature at the expense of the relative distances between the dots.

Similarly, proximity has been observed to be challenged by the relative preference for vertical orientation, a phenomenon termed orientation bias (Bleumers et al., 2008; Claessens & Wagemans, 2008; Gephstein & Kubovy, 2005; Nikolaev et al., 2016).

Using rectangular dot lattices at different levels of AR (i.e. stimulus ambiguity), Nikolaev et al. (2016) observed sequences of trials, in which orientations close to the vertical were preferred, regardless of proximity. The effect was stronger for stimuli at the highest level of ambiguity, whereas it weakened as the AR of the lattices increased. Such episodes of orientation bias tended to disappear as the experiment went by, together with as increase in grouping based on proximity. The authors attributed the presence of orientation bias to visual working memory (VWM). They proposed that, during the early stages of the experiment, participants maintain in VWM a representation of past trials in order to adjust their judgments. When VWM is occupied by those previous representations, a biased response tuned towards vertical orientation is given. This explanation is in line with the view that during perception VWM maintains coarse representations of the stimuli (proto-representations), which are gradually updated with more detailed information by object-based attention (Gao, Gao et al., 2011; Gao, Ding et al., 2013). This process would result either in perception based on stimulus features or in perception based on bias, depending on whether VWM resources are available at hand. Thus, accordingly, whether proximity or orientation control grouping of the dots depends on the load of the information hold in VWM.

To investigate top-down effects of VWM in determining the trade-off between proximity and orientation bias in perceptual grouping, we use a classical dual task procedure, where a visual discrimination task is embedded within the encoding and the test stages of a change detection task. For the former task, participants are shown with dot lattices with different degrees of ambiguity and have to report which orientation in the stimuli they perceive as the most prominent. For the latter task, participants are presented with four oriented items (i.e. Gabor patches) and are asked to retain them for a subsequent comparison. Gabor patches are presented in two conditions of complexity, which differently engaged VWM resources. In the low-load condition, all Gabor patches are in the same orientation, whereas in the high-load condition, they differ.

If the amount of information occupying VWM alters grouping processes for dot lattices, we expect to find higher rate of proximity-based responses in low-load than in high-load condition. If such reduction of proximity sensitivity is promoted by orientation bias, we expect this effect to coincide with an increase of vertical responses in high-load condition compared to low-load condition, regardless of the AR of the stimuli.

However, it could be that, instead of VWM load, the factor shaping how dot lattices are grouped is VWM content. In this case, the key factor in determining the trade-off between proximity-based and orientation-based grouping would be the quality of the information in VWM, rather than the quantity. For example, when participants maintain in VWM a representation of preceding vertical information, it could influence whether the dots are grouped according to the proximity of according to the orientation in a newly presented stimulus. Thus, the report of participants would be determined by VWM content.

The effect of VWM content on the perception of newly presented stimuli has been demonstrated (Kang et al., 2011; Pan et al., 2014; Salahub & Emrich, 2016; Scocchia et al., 2013, Teng & Kravitz, 2019, Experiment 3; van Leeuwen et al., 1988). In these studies, participants are typically asked to hold a stimulus or a single feature (e.g. orientation) in their VWM while performing a discrimination task. These studies show that perceptual performance improves when the items hold in VWM and the to-be-discriminated stimuli share some commonality compared to when they differ (congruence effect).

Thus, we check for the possible influence of VWM content. In both load conditions, half of the trials require to hold in VWM items with orientations similar to *a*, the orientation predicted by proximity for the dot lattice (congruent content); and others require to hold in VWM items with orthogonal orientations (incongruent content). Beside VWM load, we also expect congruency to affect grouping by proximity in the dot lattices. Specifically, proximity-based responses are expected to be enhanced in the congruent compared to the incongruent condition (congruence effect). Moreover, if VWM content can also promote for orientation-based grouping of dot lattices, a stronger congruence effect for vertical than for horizontal VWM content should emerge.

2.2 Method

2.2.1 Participants

Thirty-three participants took part to the experiment (17 female, age = 26 ± 6). They all had normal or corrected to normal vision. Half of the data were collected at KU Leuven where participants were recompensed with 2 course credits for taking part to the experiment. The other half of the data were collected at Sapienza University of Rome, where participants voluntary joined the study. The whole procedure was in accordance to the Declaration of Helsinki and was approved by the Ethical Committee of KU Leuven.

2.2.2 Stimuli

We used two sets of stimuli, one for each of the tasks involved in our dual task procedure. For the visual discrimination task, we used rectangular dot lattices, with AR values varying between 1.0, 1.1 and 1.2. As discussed in Chapter 1, when the stimuli presented AR = 1.0, **a** and **b** distance vectors have equal norm, and, thus, according to the Pure Distance Law, grouping by proximity can occur along both *a* and *b* orientations and the stimuli have higher level of ambiguity. As the AR of the lattices increases, grouping by proximity more strongly tunes towards orientation *a* and the stimuli become less ambiguous.

Similar to previous experiments (Nikolaev et al., 2008; 2010; 2016), dot lattices were arranged within a bivariate isotropic Gaussian distribution, to reduce possible impact of the edges of the screen. This created a circular configuration of supra-threshold dots with a diameter of 11.3 degrees of visual angle (dva). In each dot lattice, the dots had a diameter of 0.46 dva and the inter-dot distance at AR = 1.0 was 0.65 dva. Dot lattices were not presented in cardinal orientations (i.e. $\theta \neq 0°$ and $\theta \neq 90°$), since grouping by proximity is reduced along such orientations (Claessens & Wagemans, 2008), and to minimize the risk of categorical effects on participants' perceptual judgments. Instead, we used values of θ such that the stimuli appeared close to cardinal orientations (Figure 2.1): when θ was at 22.5° or 157.5°, *a* was closer to the horizontal axis (horizontal dot lattices); whereas when θ was 67.5° or 112.5°, *a* was closer to the vertical axis (vertical dot lattices).

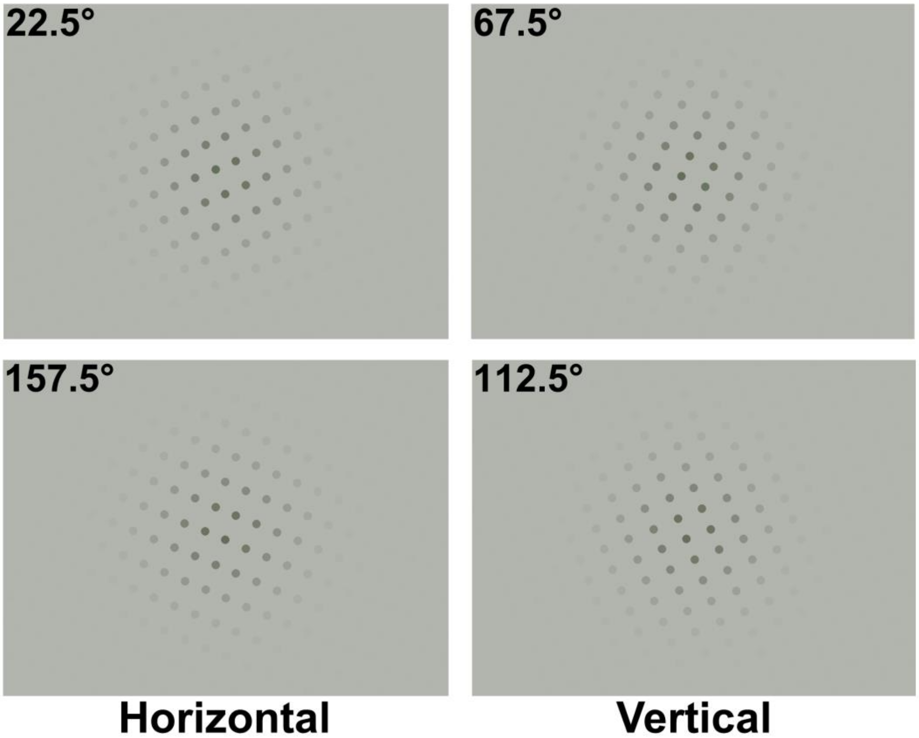


Figure 2.1: Examples of dot lattices at AR = 1.2, we employed in our experiment. The orientation of the stimuli is defined according to the angle θ *between the horizontal diameter of the circular configuration and the vector* **a**. *The stimuli tuned at 22.5° and 157.5° are closer to the horizontal axis, and, thus, they are referred to as "horizontal dot lattices", whereas the ones tuned at 67.5° and 112.5° are closer to the vertical axis, and, thus, they are referred to as "vertical dot lattices".*

For the change detection task, we presented memory arrays formed by Gabor patches embedded in a Gaussian envelope. The visible area of the patches covered 3 dva and their spatial frequency was 1.5 cycles per degree. They were placed at four equally distant locations from the fixation point (a circle of 0.32 dva of diameter at the center of the screen). The distance between the fixation point and the center of each Gabor patch was 2.5 dva. To limit for possible interference arising from the global configuration of a memory array, in half of the presentations the whole array was tilted by 22.5° clockwise from the horizontal, and in the other half it was tilted by 22.5° counterclockwise from the horizontal (Figure 2.2A).

In low-load trials, the four Gabor patches all had the same orientation, whereas, in high-load trials, the four Gabor patches presented the same orientation two by two, with equally oriented stimuli that were either horizontally or vertically aligned (Figure 2.2B). In the high-load condition, Gabor patches subtended to three possible organizations: 1) a “horizontal memory array”, in which two Gabor patches were oriented at 22.5° and the other two at 157.5°; 2) a “vertical memory array”, in which two Gabor patches were oriented at 67.5° and the other two at 112.5°; 3) a “mixed memory array”, in which two Gabor patches were presented in one of the horizontal orientations (22.5° or 157.5°) and the other two in one of vertical orientation (67.5° or 112.5°). Thus, depending on the orientation of the following dot lattice, we obtained a congruent condition when a horizontal memory array was followed by a horizontal dot lattice or when a vertical memory array was followed by a vertical dot lattice; an incongruent condition when a horizontal memory array was followed by a vertical dot lattice or when a vertical memory array was followed by a horizontal dot lattice; or a mixed condition when mixed memory arrays were presented. This latter provided a neutral baseline for comparison with the congruent and the incongruent conditions. In the low-load condition, since all the stimuli within the memory array were in the same orientation, only vertical and horizontal memory arrays were possible, thus defining only congruent and incongruent conditions.

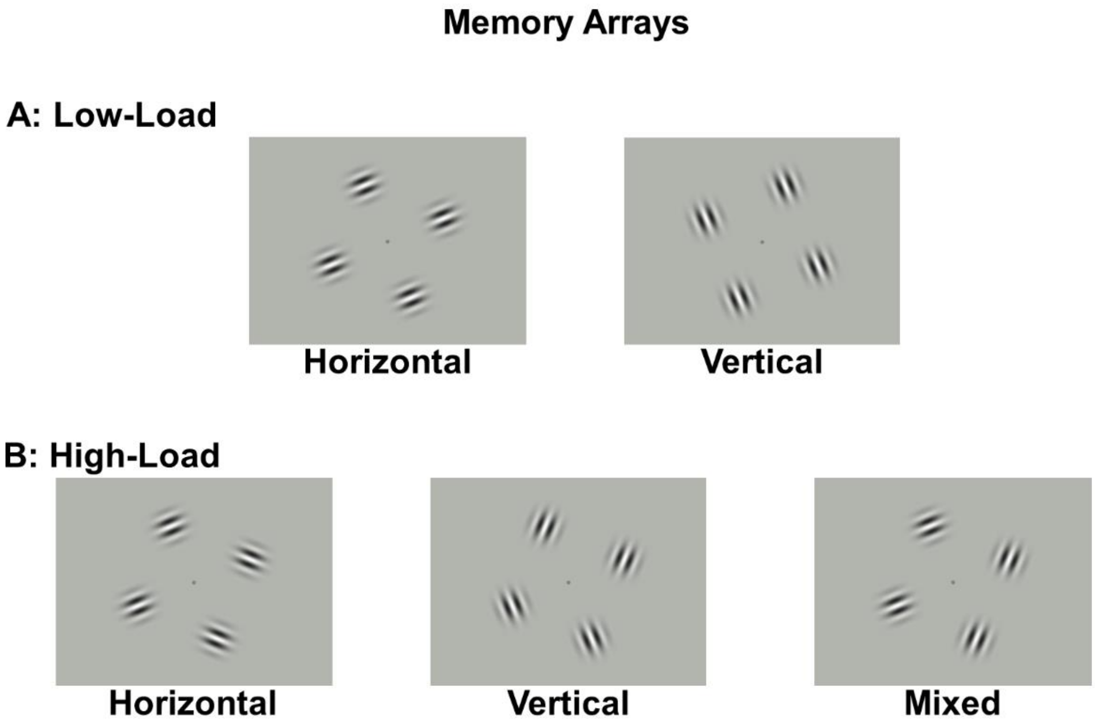


Figure 2.2: Memory arrays used for the different conditions of change detection task. Each array consists of a squared configuration of Gabor patches, rotated 22.5° clockwise or counterclockwise from the horizontal, with each stimulus being equidistant from the central fixation point. (A). *In the low-load condition, all Gabor patches are in the same orientation: either horizontal (left) (22.5° and 157.5° rotations) or vertical (right) (67.5° and 112.5° rotations).* (B). *In the high-load condition, Gabor patches are in the same orientation two by two. Same orientation pairs can be either vertically or horizontally aligned. With respect to patch orientation, we distinguish: horizontal (with one pair of stimuli being at 22.5° and the other one being at 157.5°), vertical (with one pair of stimuli being at 67.5° and the other one being at 112.5°) and mixed (with all possible combinations of horizontal and vertical pairs).*

2.2.3 Procedure

The experiment consisted in a visual discrimination task sandwiched in-between the two displays of a change detection task. Tasks were performed on a 23’’ LCD monitor (refresh rate of 60 Hz), placed at 70 cm from the participants. The whole procedure was administered using PsychoPy3 software (Pierce, 2007). Trials consisted of the following displays: fixation point, memory array, mask array, dot lattice, response probe, and test array (Figure 2.3). Memory array and test array referred to the change detection task, whereas dot lattice and response probe referred to the visual discrimination task.

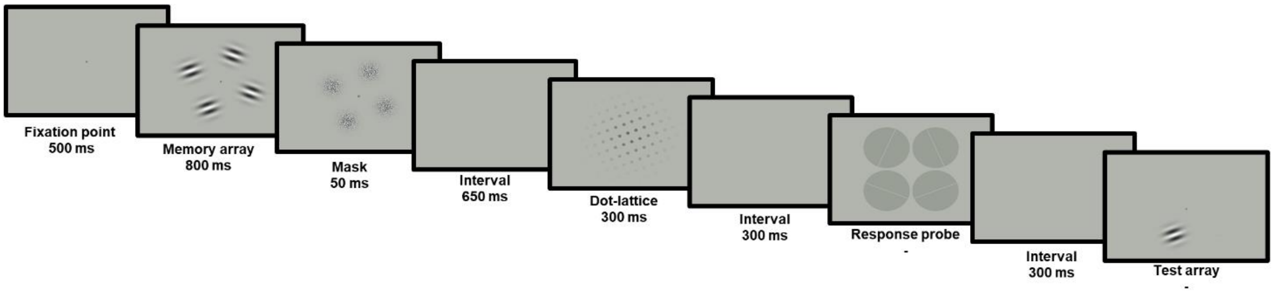


Figure 2.3: Single trial routine in our experiment. The four Gabor patches in the memory array could be either all in the same orientation (low-load condition) or, as in this case, in the same orientation two by two (high-load condition). After the mask and a 650 ms interval, participants performed the perceptual judgment task. A dot lattice was shown for 300 ms and followed by the response probe, which lasted until participants reported an orientation. Once completing the visual discrimination task, a test array with only one Gabor patch was shown and participants responded whether its orientation had changed between the two displays. Also the test array lasted on the screen until participants provided a response.

Trials began with 500 ms of fixation point, followed by 800 ms presentation of the memory array, which could appear either in low-load or in high-load condition. To prevent for the retinal persistence of the memory array (e.g. Sligte et al. 2008), Gabor patches were followed by a 50 ms mask. The mask consisted in four two-dimensional squared sine-wave gratings of 3 dva with a noisy texture embedded in a Gaussian envelope. Each grating was superimposed on one of the Gabor patches of the memory array. Then, after 650 ms of blank screen, participants saw the dot lattice for 300 ms. After additional 300 ms blank screen, the response probe was displayed until response. The response probe consisted of four circles (response icons), each half-cut by a diameter parallel to one of the four possible orientations of the dot lattice (i.e. *a*, *b*, *c* or *d*). For each of the four orientations of the dot lattices, four different response probes were obtained by counterclockwise rotating the response icons to one of four possible locations on the screen (i.e. upper-left, lower-left, upper-right, lower-right, respectively). Participants reported the orientation of the dot lattice by pressing with the right hand one of four keys from the numerical keyboard ("4", "1", "5", or "2"). These keys spatially corresponded to the locations of the response icons on the screen during each trial (i.e. "4" for top-left response icon, "1" for down-left response icon, "5" for top-right response icon, and "2" for down-right response icon). If participants did not perceive a prominent orientation in the dot lattice, they were asked to guess one. The visual discrimination report was followed by a 300 ms blank screen. Finally, the test array appeared and lasted until a response was provided. For the test array, only one of the four Gabor patches in the memory array was presented. Participants indicated whether the orientation between the test Gabor patch had changed between the two displays. In half of the trials, its orientation remained the same, whereas, in the other half, it was different. The new orientation was selected among the remaining three orientations, each of which occurring the same number of trials. Participants responded to the test array with the keys “1” and “2” of the numerical keyboard. The association between keys and responses was counterbalanced across participants. Before the new trial, there was a 500 ms inter-trial interval. In order to familiarize with the whole procedure, before starting the experiment, participants completed 12 practice trials. At this stage, for each response icon, the corresponding key was printed on the screen. The practice trials were equal for all participants but presented in a random order. Participants could ask for more practice trials until they felt ready to proceed with the experiment. To avoid that participants exclusively focused on one task, our instructions emphasized that the change detection and the visual discrimination were equally important and that they had to perform both at their best.

The experiment consisted of 768 trials, divided into 8 blocks. Participants were allowed to do short breaks in-between the blocks. The number of high-load and low-load trials, as well as the number of trials for each AR and orientation of the dot lattices, was the same for each block. Within a block, the configuration of the memory array as well as the congruence between memory array and dot lattices was balanced for both high-load and low-load trials. The location of response icons were balanced between the blocks, and their order was randomized for each participant to prevent for an effect of the response probe configurations. The whole experiment lasted approximately 90 minutes.

2.3 Results

Based on the binomial distribution, we set the chance-level threshold for the change detection at 54% of correct responses. Six participants did not reach such threshold and were excluded from the analyses. In addition, one participant was excluded because of software errors during the experiment. This led to a final sample of 26 participants. This sample size was sufficient to obtain reliable estimates of the effects under investigation: sample computation using the software G-Power with the assumptions of medium effect sizes (η_p^2 = .08) and high statistic power (β = 0.9) established that the minimum sample size required was 21 participants.

Since orientations along *c* and *d* were rarely reported (9.8% of total responses for *c* and 9.3% of total responses for *d*), we restricted our analyses to the trials reporting *a* and *b* responses.

A preliminary check determined that no difference between the sample collected at KU Leuven and the one collected at Sapienza University of Rome emerged for both the change detection task and the perceptual judgment task. Thus, these subgroups were merged for all next analyses.

We first compared the percentages of correct responses (accuracy) to the change detection task between the two VWM conditions (low vs. high). Mean accuracy was lower in high-load than in low-load condition ($t_{(25)}$ = -9.4, $p < .001$, η_p^2 = .78), meaning that our manipulation of VWM load had been effective (Figure 2.4).

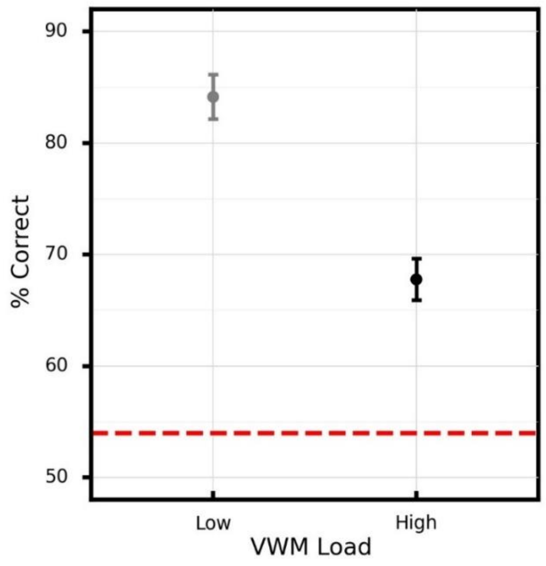


Figure 2.4: Performance to the change detection task for the two VWM load conditions. Dots represent average percentages correct and bars represent standard errors. The red dashed line represents chance level accuracy .

For all next analyses, we used a within-subjects ANOVA with Greenhouse-Geisser correction (ε) of p-values when the sphericity of factors with more than one degree of freedom was violated and Bonferroni correction for post-hoc pairwise comparisons.

We checked for an orientation bias for vertical orientation using a 2x2 repeated-measures ANOVA on response probability using response type (a vs. b) and response orientation (vertical vs. horizontal) as factors (Figure 2.5). Similarly to the study of Nikolaev et al., 2016, we found higher probability of a than b responses ($F_{(1,25)} = 38.85$, $p < .001$, $\eta_p^2 = .6$) and higher probability of vertical responses than horizontal ones ($F_{(1,25)} = 6.95$, $p < .05$, $\eta_p^2 = .22$). These findings are consistent with both the Pure Distance Law of grouping by proximity (Kubovy et al., 1998, Kubovy & Wagemans, 1995) and the existence of a bias for vertical orientation in the grouping of rectangular dot lattices (Bleumers et al., 2008; Claessens et al., 2008; Nikolaev et al., 2016).

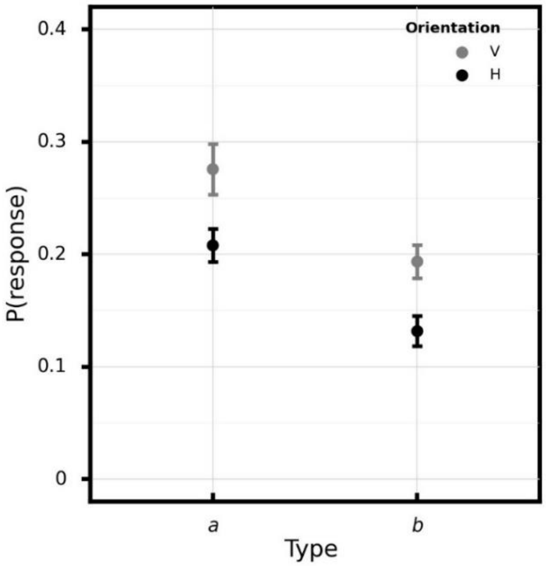


Figure 2.5: The probability of a and b responses separated for vertical (V) and horizontal (H) response orientations. Dots represent average response probability and bars represent standard errors.

Based on these observations, we computed a proximity-based grouping index, P(*a*-*b*), by subtracting the proportions of *a* responses to the proportions of *b* responses (regardless of their orientation), and an orientation-based grouping index, V(v-h), by subtracting the proportions of vertical responses to the proportions of horizontal ones (regardless of *a*- or *b* type). These indexes were then used as dependent variables for further analyses. Notice that P(*a*-*b*) and V(v-h) are not mutually exclusive. Both indexes enhance when a vertical *a* is chosen. In this case, however, we cannot determine whether this occurs because it is the shortest distance or because it is the preferred orientation. Similarly, they both decreases if a horizontal *b* is chosen. Horizontal *a* and vertical *b* have an opposite effect on these indexes. In former case, the enhancement of P(*a*-*b*) co-occurs with a decrease in V(v-h), and vice versa for the latter case. Thus, using the combination of both indexes, we can determine whether grouping in different conditions is mostly driven by proximity or by orientation bias. Specifically, if an increase in V(v-h) co-occurs to a decrease in P(*a*-*b*), this would reflect an increase of orientation-based grouping at the expense of proximity-based grouping. However, if both V(v-h) and P(*a*-*b*) increase, this implies a net shift from horizontal *b* responses to vertical *a* responses. If the VWM content is vertical, such a result may indicate a congruence effect. However, this does not necessarily imply an increase in orientation bias.

To test how VWM load affects grouping by proximity, we used a 2x3 ANOVA on the proximity-based grouping index P(*a*-*b*), with VWM load (low vs. high) and AR (1.0 vs. 1.1 vs. 1.2) as factors (Figure 2.6). In line with our predictions, we found P(*a*-*b*) to be higher in low-load than in high-load condition ($F_{(1,25)} = 14.08$, $p = .001$, $\eta_p^2 = .36$). Moreover, proximity-based grouping increased as the dot lattices became less ambiguous ($F_{(2,50)} = 38.59$, $p < .001$, $\eta_p^2 = .61$, $\varepsilon = .56$), as expected from the Pure Distance Law of grouping by proximity (Kubovy et al., 1998, Kubovy & Wagemans, 1995). We found an interaction between the two factors ($F_{(2,50)} = 7.8$, $p = .001$, $\eta_p^2 = .24$). Pairwise comparison showed that at AR = 1.0, P(*a*-*b*) was equal between low- and high-load conditions, whereas the difference between conditions emerged for AR = 1.1 and AR = 1.2, ($t_{(25)} = 4.09$, $p < .001$, $\eta_p^2 = .4$ and $t_{(25)} = 4.13$, $p < .001$, $\eta_p^2 = .4$, respectively), This finding was not surprising, since at AR = 1.0, no net preference for *a* over *b* responses is expected, so that the proximity index approaches to zero in both load conditions.

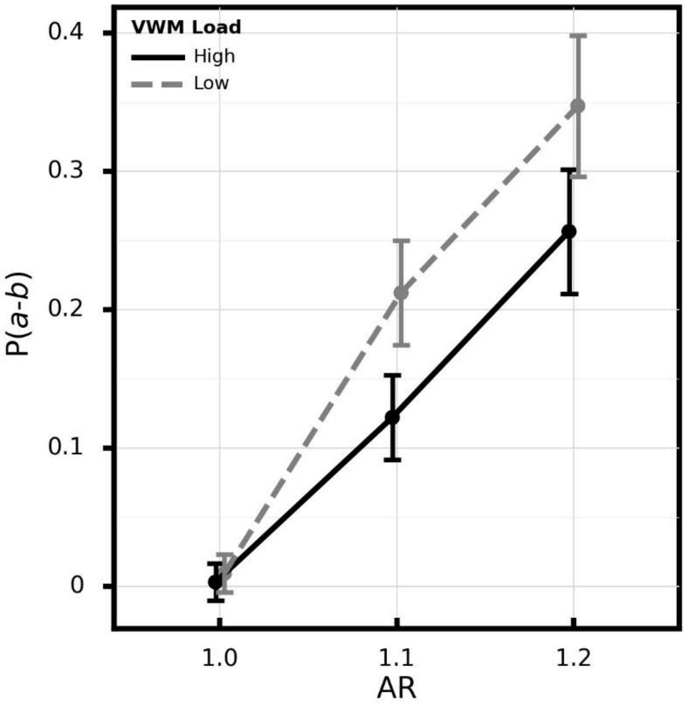


Figure 2.6: Proximity-based grouping indexes P(*a*-*b*) *in two VWM load conditions for three levels of* AR. *Notice that when the* AR *is* 1.0 *(i.e.* **a** = **b***), the two responses are equiprobable. As the stimuli become less ambiguous, proximity-based responses increase. Such increases are larger in the low-load condition than in the high-load condition. Dots represent average proximity index values and bars represent standard errors.*

Same ANOVA over V(v-h) showed that preference for the vertical orientation increased in high-load condition with respect to low-load condition ($F_{(1,25)} = 6.79$, $p < .05$, $\eta_p^2 = .21$), and was higher for more ambiguous dot lattices ($F_{(2,50)} = 7.98$, $p < .01$, $\eta_p^2 = .24$, $\varepsilon = .72$). No interaction between VWM load and AR effects was observed (Figure 2.7).

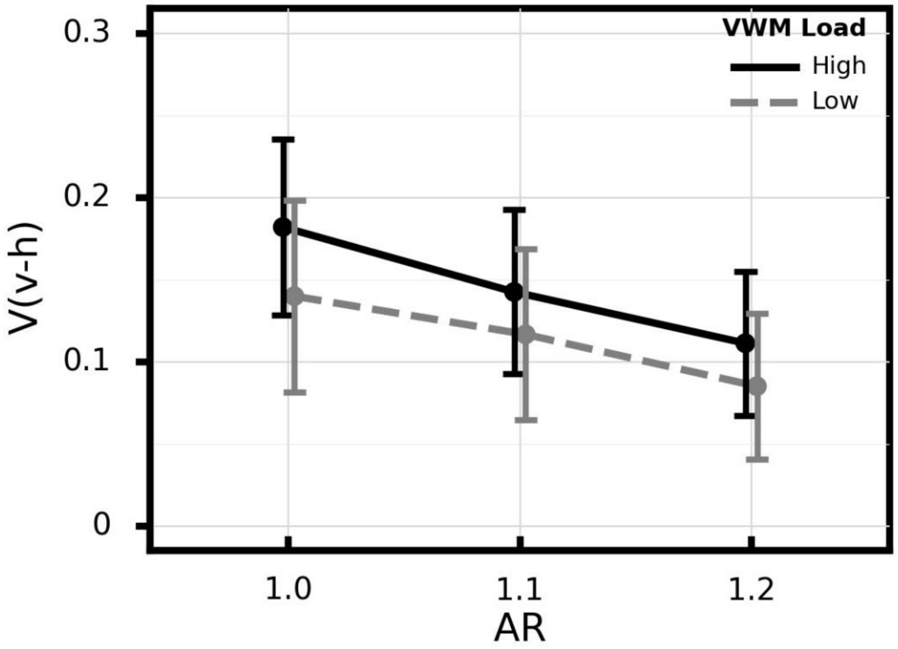


Figure 2.7: Orientation bias V(v-h) *in the two VWM conditions and for the three levels of* AR*. The preference for vertical orientation is higher in high-load condition, regardless of the dot lattices ambiguity. Dots represent average vertical preference and bars represent standard errors.*

The overall opposite trends of P(*a*-*b*) and V(v-h) in the two VWM load conditions imply that they result from an increase in orientation-based grouping with memory load. However, some differences in P(*a*-*b*) and V(v-h) between conditions seem to emerge. In particular, we saw that proximity-based grouping is unaffected by VWM when AR = 1.0, whereas higher VWM load enhances orientation-based grouping indistinctly for all the levels of AR. Recall that at AR = 1.0, proximity-based grouping is defined both along *a* and *b* orientations, because the equivalence between ***a*** and ***b*** for these stimuli. Thus, this discrepancy does not prevent us from concluding that

VWM load is involved in the shift between proximity- and orientation-based grouping of dot lattices.

To investigate the role of VWM content, we ran separate ANOVAs on P(*a*-*b*) and V(v-h), taking the orientation of the memory arrays into account. Since the possible memory array orientations differed between low- and high-load conditions, we analyzed these two separately.

We ran a 3-way repeated-measures ANOVA on P(*a*-*b*) for low-load condition, with memory array orientation (vertical vs. horizontal), dot lattice orientation (vertical vs. horizontal) and AR (1.0 vs. 1.1 vs. 1.2) as factors (Figure 2.8). P(*a*-*b*) was higher for vertical dot lattices than for horizontal ones ($F_{(1,25)} = 5.1$, $p < .05$, $\eta_p^2 = .17$) and increased with the AR of the stimuli ($F_{(2,50)} = 34.34$, $p < .001$, $\eta_p^2 = .58$, $\varepsilon = .62$). We found an interaction between memory array orientation and dot lattice orientation ($F_{(1,25)} = 16.64$, $p < .001$, $\eta_p^2 = .4$). For vertical dot lattices, P(*a*-*b*) was higher after vertical memory arrays than after horizontal ones ($t_{(25)} = 3.48$, $p < .005$, $\eta_p^2 = .32$), whereas the opposite was true for horizontal dot lattices ($t_{(25)} = 4.29$, $p < .001$, $\eta_p^2 = .42$).

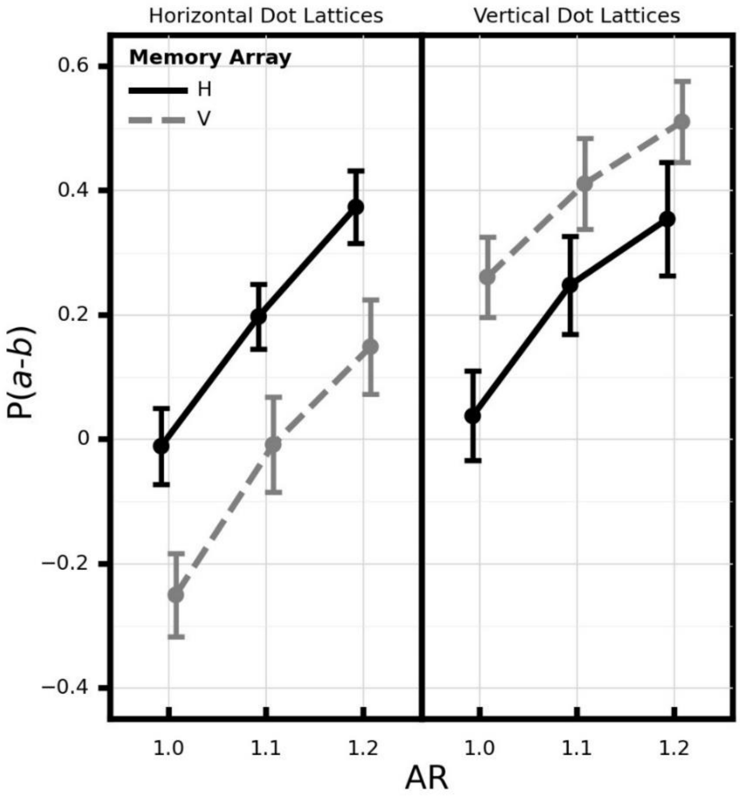


Figure 2.8: Proximity-based grouping index P(*a*-*b*) *after vertical and horizontal memory arrays in the low-load condition. Proximity-based grouping is higher for vertical dot lattices after vertical memory arrays and for horizontal dot lattices after horizontal memory arrays. Dots represent average proximity-based grouping index and bars represent standard errors. Abbreviations: H: horizontal; V: vertical.*

These results suggest a congruence effect promoted by VWM content. However, from Figure 2.8 it is evident that there is no difference in P(*a*-*b*) between vertical and horizontal dot lattices after horizontal memory arrays. In other words, the congruence effect observed for horizontal dot lattices would be due to vertical memory arrays increasing *b* responses, instead of an increase of *a* responses after horizontal VWM content. This is also evident for stimuli at AR = 1.0, which after horizontal memory arrays approached zero, whereas after vertical memory arrays tuned towards vertical orientation.

This claim was confirmed by the same ANOVA performed on orientation-based response index V(v-h). We found the main effect of memory array orientation ($F_{(1,25)}$ = 16.64, p < .001, η_p^2 = .4): V(v-h) was higher overall with vertical than with horizontal memory arrays (Figure 2.9). This result shows that vertical memory arrays increased both *a* responses for vertical dot lattices and *b* responses for horizontal dot lattices. In other words, the preference for vertical orientation was strengthened for both dot lattice orientations after vertical memory arrays. This pattern was consistent across all the values of AR. This analysis also showed an interaction between AR and dot lattice orientation ($F_{(2,50)}$ = 34.34, p < .001, η_p^2 = .58, ε = .62), meaning that at AR = 1.0, the preference for vertical grouping was the same in vertical and horizontal dot lattices, whereas it diminished with AR for horizontal stimuli ($t_{(25)}$ = 5.65, p < .001, η_p^2 = .56 and $t_{(25)}$ = 6.8, p < .001, η_p^2 = .65; for AR = 1.1 and AR = 1.2, respectively). This trend implies an increase of *a* responses with AR in both dot lattices orientation, whereas there was no difference between *a* and *b* responses at AR = 1.0, in line with what prospected by the Pure Distance Law (Kubovy et al., 1998).

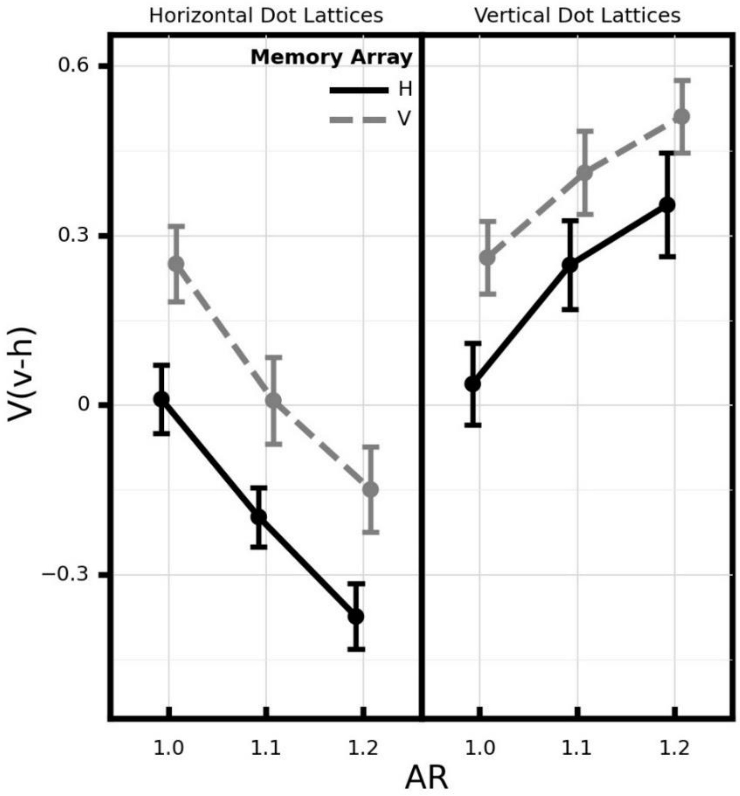


Figure 2.9: Orientation-based grouping index V(v-h) *after vertical and horizontal memory arrays in the low-load condition. Orientation-based grouping increases when vertical items are kept in VWM. Dots represent average vertical preference and bars represent standard errors. Abbreviations: H: horizontal; V: vertical.*

For high-load condition, we performed a similar ANOVA as before on P(*a*-*b*), but now with three levels of memory array orientation (vertical vs. horizontal vs. mixed). This ANOVA kept the main effects of AR ($F_{(2,50)}$ = 35.35, $p < .001$, η_p^2 = .59, ε = .58), dot lattice orientation ($F_{(1,25)}$ = 9.26, $p <$ .005, η_p^2 = .27), and the interaction between memory array orientation and dot lattice orientation ($F_{(2,50)}$ = 8.57, $p < .001$, η_p^2 = .26). Proximity-based grouping for horizontal dot lattices was higher after horizontal than after vertical ($t_{(25)}$ = 3, $p < .01$, η_p^2 = .26) and mixed ($t_{(25)}$ = 3.13, $p < .05$, η_p^2

= .28) memory arrays, as well as higher after mixed than after vertical ones ($t_{(25)}$ = 2.58, $p < .05$, η_p^2 = .21). On the other hand, there was no difference in P(*a*-*b*) between the memory arrays for vertical dot lattices (Figure 2.10).

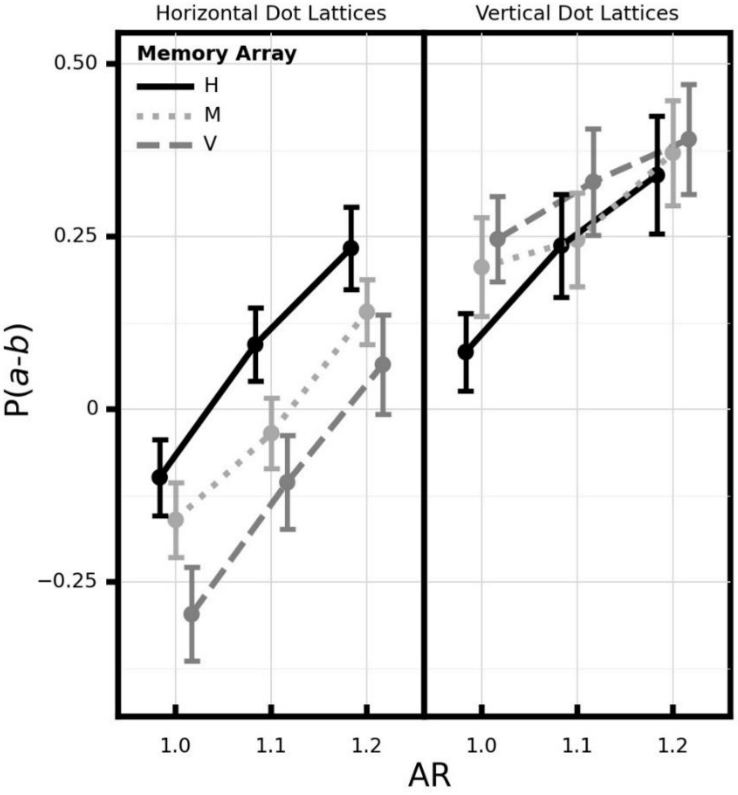


Figure 2.10: Proximity-based grouping index P(*a*-*b*) *after vertical and horizontal memory arrays in the high-load condition. Proximity-based grouping in vertical dot lattices does not depend on VWM content, whereas it is reduced for horizontal dot lattices after vertical memory arrays. Dots represent average proximity index and bars represent standard errors. Abbreviations: H: horizontal; M: mixed; V: vertical.*

These results resemble the ones found for the low-load condition. Consistently to Figure 2.8, in Figure 2.10 the congruence effect for horizontal dot lattices is attributable to the effect of vertical memory arrays, favoring vertical *b* over horizontal *a* responses. The main contrast with the low-load condition is the absence of memory array effects for vertical dot lattices.

Again, the pattern observed for P(*a*-*b*) was confirmed by the same ANOVA as before, using V(v-h) as dependent variable (Figure 2.11). As can be observed in Figure 2.9, the ANOVA reproduced a main effect of memory array orientation ($F_{(2,50)} = 8.94$, $p < .005$, $\eta_p^2 = .26$, $\varepsilon = .63$) and an interaction between AR and dot lattice orientation ($F_{(2,50)} = 35.64$, $p < .001$, $\eta_p^2 = .59$). Vertical and mixed memory arrays produced a larger number of vertical responses than horizontal ones ($t_{(25)} = 3.22$, $p < .05$, $\eta_p^2 = .29$ and $t_{(25)} = 2.96$, $p < .05$, $\eta_p^2 = .26$, respectively). We found an interaction between memory array orientation and dot lattice orientation ($F_{(2,50)} = 3.6$, $p < .05$, $\eta_p^2 = .13$). In horizontal dot lattices, orientation-based grouping increased after vertical memory arrays relative to horizontal ($t_{(25)} = 3.44$, $p < .01$, $\eta_p^2 = .32$) and mixed ($t_{(25)} = 2.58$, $p < .05$, $\eta_p^2 = .21$) ones, and following mixed memory arrays relative to horizontal ones ($t_{(25)} = 3.13$, $p < .05$, $\eta_p^2 = .28$). This further supports the finding that orientation-based grouping increases for horizontal dot lattices after vertical VWM content. As previously seen for P(*a*-*b*), V(v-h) for vertical dot-lattices was unaffected by memory array orientation. Such an interaction was not observed in the corresponding analysis for low-load condition, where both vertical and horizontal dot lattices showed a generalized enhancement of V(v-h) after vertical memory arrays.

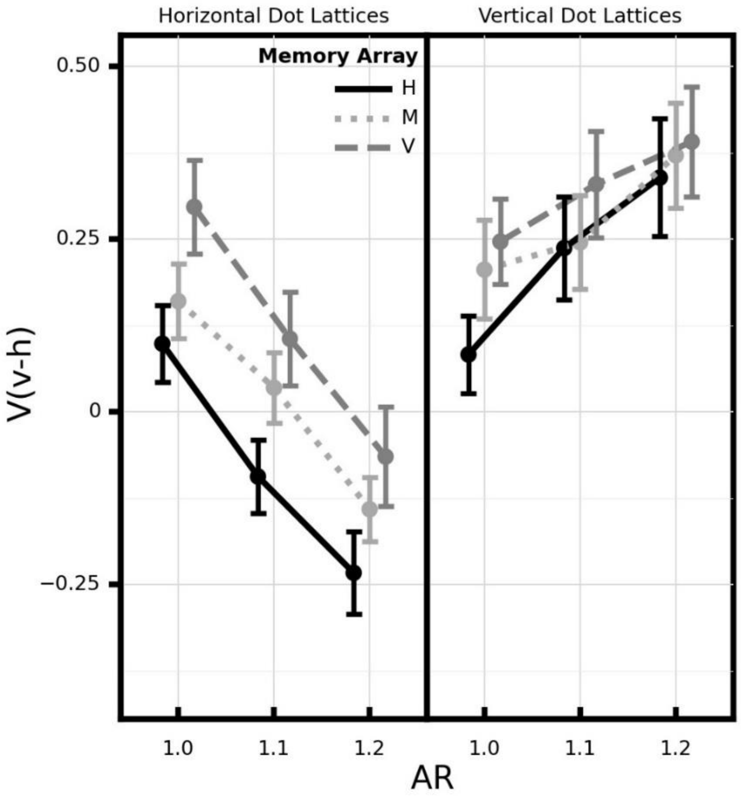


Figure 2.11: Orientation-based grouping index V(v-h) *after vertical and horizontal memory arrays in the high-load condition. Orientation-based grouping in vertical dot lattices is unaffected VWM content, whereas it increases for horizontal dot lattices after vertical memory arrays. Dots represent average proximity index and bars represent standard errors. Abbreviations: H: horizontal; M: mixed; V: vertical.*

Taken together, these analyses suggest that both P(*a*-*b*) and V(v-h) are highly sensitive to vertical VWM content. In other words, when VWM content is consistent with the orientation bias, we find a higher tendency for orientation-based grouping of the dot lattices. When the stimuli are actually vertical, this manifests as enhanced *a* responses, whereas when they are horizontal this manifests as enhanced *b* responses. This produces a congruence effect in perceiving both vertical and horizontal dot lattices, which, however, is asymmetric. This is clearly evident in Figure 2.8, where horizontal VWM content leaves proximity-based responses between horizontal and vertical dot lattices invariant. This effect seems to dissipate in the high-load condition when vertical dot lattices are presented. In this case, therefore, vertical VWM content would not promote vertical *a* responses, compared to horizontal and mixed VWM contents. This can be considered a ceiling effect due to the higher presence of orientation bias in this condition, which equates vertical responses among the different types of VWM contents.

Overall, the results of this study are in line with the hypothesis that proximity-based grouping depends, at least in part, on VWM top-down modulation on lower level perceptual areas. When high VWM load prevents this modulation, or when VWM content matches with the biased orientation, the perceptual system is more likely to group the dots in accordance to the orientation bias. This enhancement in preference for vertical orientation is accompanied by a drop in proximity-based grouping. Thus, both VWM load and content shape how the grouping process in dot lattices unfolds. To address this complex pattern of results within a unitary and coherent framework, we propose and evaluate a model of the role of VWM in perceptual grouping.

2.4 VWM Gating of Perceptual Grouping

The effect of VWM load on grouping are consistent with studies suggesting that VWM promotes for the active maintenance of intermediate representations of percepts (proto-representations) until visual processing is completed (Gao, Gao et al., 2011; Gao, Ding et al., 2013; Roelfsema, 2005). In our case, such proto-representation may underlie to the proximity organization of the dot lattices. This mechanism would be refined by VWM content, which would promote proximity-based grouping at the expense of orientation-based grouping when it is inconsistent with biased orientation.

Based on our observations, we suggest that whether grouping in dot lattices is proximity- or orientation-based depends on the competition between the orientation bias and the proximity proto-representation of the stimulus. In high-load condition, given the smaller amount of VWM capacity available to maintain proximity proto-representations, these would be weaker than in the low-load condition. This weakening would render the perceptual system more subjected to the orientation bias. The role of VWM content would be to gate which one amongst proto-representation and orientation bias gets to determine grouping. Specifically, when consistent with orientation bias (i.e. it is vertical), the content promotes orientation-based grouping, whereas when it is inconsistent (i.e. it is horizontal), it promotes proximity-based grouping.

To evaluate this hypothesis, we ran separate additive linear models for the low- and high-load conditions, using P(*a*-*b*) and V(v-h) as criteria and the following conditions as categorical predictors[1]:

- V_h: VWM content is vertical and the proto-representation is horizontal;

- V_v: VWM content is vertical and the proto-representation is vertical;

- H_h: VWM content is horizontal and the proto-representation is horizontal;

- H_v: VWM content is horizontal and the proto-representation is vertical.

The regression model for low-load condition on P(*a*-*b*) was significant ($F_{(3,100)} = 7.65$, $p < .001$, $R^2 = .19$). For the single comparisons, we used H_h as the control condition. P(*a*-*b*) decreased in

---

[1] In this notation, VWM content is expressed in upper-case and the proto-representation is expressed in lower-case.

V_h with respect to H_h ($t_{(100)}$ = -2.47, $p < .05$), meaning that vertical VWM content reduced the impact of proximity proto-representation on grouping process. There was no difference in P(*a-b*) between H_v and H_h. Since VWM content did not coincide with the biased orientation, grouping in these conditions was driven by their proximity proto-representations, which would equate between vertical and horizontal dot lattices. Finally, P(*a-b*) was higher in V_v than in H_h ($t_{(100)}$ = 2.3, $p < .05$), meaning that *a* responses increased for vertical congruent stimuli compared to horizontal congruent ones. Notice that in the V_v condition the proximity proto-representation is consistent with the combination of VWM content and orientation bias. Thus, the proto-representation provides for an additive effect in perceiving the vertical organization of the lattices, enhancing vertical *a* responses.

The same model on V(v-h) ($F_{(3,100)}$ = 15.11, $p < .001$, R^2 = .31) confirmed that the decrease in proximity-based grouping observed in V_h co-occurred with an increase in orientation-based grouping ($t_{(100)}$ = 2.47, $p < .05$), favored by vertical VWM content.

The regression model on P(*a-b*) for the high-load condition was also significant ($F_{(3,100)}$ = 9.61, $p < .001$, R^2 = .22). Proximity-based grouping was lower in V_h than in H_h ($t_{(100)}$ = -2.2, $p < .05$), reflecting again a decrease in horizontal *a* responses when VWM content was consistent to orientation bias. However, with respect to the effect observed in the low-load condition, this time the slope of the linear model is reduced (β = -.189 and β = -.223, for high- and low-load conditions, respectively). This implies that in condition H_h the number of horizontal responses decreased between low- and high-load conditions. This is in line with the idea of proximity proto-representation weakening due to the higher amount of resources occupied in VWM. The comparison between H_v and H_h found no difference in P(*a-b*) score, implying that in both conditions grouping is driven by the proximity proto-representation of the stimuli. However, in this case, the slope of the effect was higher than in the low-load condition (β = -.143 and β = -.027, for high- and low-load conditions, respectively). Again, this is consistent with the idea of proximity proto-representation weakening, which would decrease *a* responses in both conditions. The reduction for H_v, in fact, would be supplied by orientation bias, which would slightly enhance vertical responses for high-load condition. Proximity-based grouping in high-load condition also equated between V_v and H_v conditions. In this case, given the weakening of proximity proto-representations, grouping is driven by orientation bias, which promotes for the same amount of vertical responses in both conditions. Thus, the boost in *a* responses for V_v found in low-load condition would be absent because of the weakening of VWM proximity proto-representations.

Coherently with latter analysis, the regression on V(v-h) score was significant ($F_{(3,100)} = 7.83$, $p < .001$, $R^2 = .19$), and confirmed the increase in orientation-based grouping in V_h with respect to H_h condition ($t_{(100)} = 2.2$, $p < .05$).

To summarize, these additive effects confirm our model that grouping in dot lattices is constantly shaped by both proximity proto-representations and orientation bias. Whether grouping in a stimulus depends on proximity or on orientation bias is a matter of the quantity (load) and the quality (content) of the information held in VWM. In this sense, VWM gates grouping process.

2.5 Discussion

Previous studies adopting a phenomenal psychophysics approach to investigate grouping by proximity in rectangular dot lattices provided evidence of an orientation bias. Accordingly, participants were more likely to perceive those stimuli as vertical, even when *a* happened to be orthogonal to vertical orientation (Bleumers et al., 2008; Claessens & Wagemans, 2008; Gephstein

& Kubovy, 2005; Nikolaev et al. 2016). Thus, in these studies orientation-based grouping co-occurred with a reduction in proximity sensitivity.

By comparing the temporal dynamics of proximity-based and orientation-based responses, Nikolaev et al. (2016) attributed orientation-based grouping episodes to VWM load. They proposed that when information about previous trials was updated in VWM, it engaged the central capacity of the perceptual system (Raffone et al., 2104; Simione et al., 2012), inducing observers to respond based to the preferred orientation, rather by the organization invoked by the stimulus. Orientation-based grouping reduced in later blocks of the experiment and the authors attributed this effect to perceptual learning. Thus, in this case perceptual learning was governed by a decline in orientation bias, rather than by a boosting in proximity sensitivity (α parameter in Equation 1.1).

We directly tested this proposal by systematically varying the information load in VWM and comparing grouping processes between the conditions. Such scenario was obtained by using a dual task procedure in which participants evaluated dot lattices while engaged in a simultaneous change detection task. Accordingly, we found that proximity-based grouping is lower in the high-load condition than in the low-load condition. This pattern is paralleled by an increase in orientation-based grouping. Thus, accordingly to the results of Nikolaev et al. (2016), the decrease in grouping by proximity found in the high-load condition would be due to higher orientation bias, rather than to reduced stimulus discriminability. This conclusion is further supported by the finding that the effect of orientation bias on grouping in the high-load condition persists even for highly discriminable stimuli (Figure 2.7). Thus, as observed by Gephstein and Kubovy (2005), orientation bias does not implicate a mere weakening of stimulus support but rather an actual adjustment of default system preference.

We extended our investigation by raising the question of whether the trade-off between proximity and orientation bias in perceptual grouping was influenced by the quality of information represented in VWM (Kang et al., 2011; Pan et al., 2014; Salahub & Emrich, 2016; Scocchia et al., 2013, Teng & Kravitz, 2019). The results showed that visual discrimination is shaped by VWM content, with vertical VWM traces effectively promoting for orientation-based grouping.

To account for this pattern of results we propose a model of VWM role in perceptual grouping by proximity. In this model, VWM is the key factor determining whether perceptual grouping of dot lattices occurs in accordance to proximity or in accordance to orientation bias. Thus, in line with previous theories of perceptual grouping (Roelfsema 2005; Ullman, 1984), we consider VWM as an integral part of the grouping process. We propose that the grouping process depends on the competition between a proto-representation conditioned by the stimulus (i.e. proximity) and the intrinsic bias for vertical orientation. This competition is regulated by both VWM load and VWM content.

High VWM load reduces the chances of forming a stimulus-based proto-representation, promoting orientation-based grouping. At the same time, when VWM is fully available for visual discrimination, the strength of the proto-representation overcomes the impact of orientation bias, at least for non-ambiguous stimuli. VWM content, on the other hand, would provide higher weight to either proximity proto-representation or orientation bias, thus further impacting the outcome of the competition.

Our model considers VWM and perception as two intertwined processes. Rather than a mere storage of perceptual outputs, VWM is conceived as dynamically involved in ongoing perceptual processes (Gao, Gao et al., 2011; Gao, Ding et al., 2013; Roelfsema, 2005). This view is also in line with studies showing that VWM is sensitive to the basic perceptual organization of stimuli

(Gao et al., 2016; Jiang et al., 2004; Quinlan & Cohen 2012; van Lamsweerde et al., 2015; Woodman et al., 2003).

Our results suggest that VWM affects perceptual grouping in dot lattices by reducing orientation bias. Orientation bias, as other forms of perceptual bias, is thought to depend on top-down feedbacks on low-level visual areas. These reflect prior knowledge, probabilistic inferences and beliefs (Albright, 2012; Frith & Dolan, 1997; Hohwy, 2017; Kok et al., 2013; Summersfield & de Lange, 2014; Wang et al., 2013), which help to deal with perceptual ambiguity. Another form of top-down perceptual bias found during perceptual grouping of dot lattices refers to non-accidentalness of stimulus features (Strother & Kubovy 2006; 2012). In this case, participants were more likely to group the dots along the curved orientation, regardless of proximity organization, because they judged as less likely for curvature to be a casual feature of the stimuli. Thus, previous knowledge about external contingencies impacts again the way a stimulus is processed.

A future follow-up of this study may investigate whether VWM processes govern the trade-off between proximity- and curvature-based grouping. If the results were consistent between the experiments, higher proximity-based grouping in the low-load condition will be accompanied by higher curvature-based grouping in high-load condition, as well as more curvature-based grouping after curved information happens to be represented in VWM. A similar pattern of results would indicate that VWM modulates the activity on low-level areas, contrasting alternative top-down signals promoted by several types of perceptual biases. This would be in accordance with EEG and fMRI observations that VWM operations and visual perception processes rely on the same neural substrates and that VWM representations are maintained for several seconds in early visual areas, such as V1 and V2 (Bisley et al., 2004; Ester et al., 2009; Harrison & Tong, 2009; Motter, 1994; Serences et al., 2009; Teng & Kravitz, 2019; van Kerkoerle et al., 2017).

A limitation of the present study is the fact that our change detection procedure did not involve any concomitant articulatory suppression. Thus, one may argue that participants may have used verbal representations instead of purely relying on VWM. It should be noted, however, that the memory arrays we employed could be hardly verbalized and this is effective in preventing the verbal representation strategy to overcome the visual one (Sense et al., 2017).

Further studies are needed to address how VWM gates the trade-off between proximity and perceptual biases. However, our results suggest that grouping in dot lattice relies on parallel processing performed at multiple stages of cognitive hierarchy, ranging from early visual areas up to high-level prefrontal areas involved in perceptual bias and VWM operations.

Next chapter will present a dynamic neural model accounting for the mechanisms of grouping by proximity and orientation bias in perceptual grouping within a biologically plausible architecture.

Chapter 3

A Dynamic Model of Grouping by Proximity and Orientation Bias

3.1 Introduction

One of the key aspects of the perceptual system is the ability to create stable representations of external objects (Maia & Cleeremans, 2005; van Leeuwen et al., 1997; Wagemans et al., 2012b). Stability refers to the robustness of these representations against perturbation, and it is particularly clear-cut in the case of perceptually non-ambiguous stimuli. This supposes that being presented with the same object under different viewing conditions does not alter its perception. For example, Kubovy and Wagemans (1998) demonstrated that grouping by proximity in dot lattices is not affected when either the spatial and the temporal scale are changed.

On the other hand, such resistance to interference is counterbalanced by a certain degree of flexibility, which produces different outcomes during the perception of identical stimuli in distinct presentations. This is the case, for example, of rectangular dot lattices with AR = 1.0, where one can alternate between *a* and *b* organizations across separate trials. Similarly, during prolonged inspection of ambiguous stimuli, observers tend to switch between the possible interpretations of a percept (Wagemans et al., 2012b). Such alternation between settled perceptual organizations seems mostly to occur without voluntary control (Kleinschmidt et al., 1998; Toppino, 2003; Toppino & Long, 2005).

The interplay between stability and flexibility in visual processing is well accounted by dynamical models. Perceptual stability in these models is achieved via the convergence to a permanent state of activation, known as *attractor* (e.g. Hopfield, 1982). When an attractor is approached, the state of a system reaches equilibrium and does no longer change in response to the inputs received. In neural terms, an attractor can be thought as a state of sustained firing of a group of neurons (assembly) subtending to a specific representation (Raffone & van Leeuwen, 2003).

At the same time, to assure perceptual flexibility, these systems have to incorporate mechanisms to escape from attractors. One convenient strategy is to provide the system with internal stochastic noise (van Leeuwen et al., 1997; Wagemans et al., 2012b). The assumption of intrinsic noise is in line with psychophysical approaches to perception based on signal detection theory (e.g. Mueller & Weidemann, 2008). In other words, noise alters the way in which a signal is filtered by sensory channels and may induce a drift from one attractor to another. Importantly, this trade-off between attractors is spontaneously accomplished by the system. This feature highlights a main characteristic of dynamical models, namely *self-organization*, which is also a fundamental propriety of Gestalt emergence (Pomerantz & Portillo, 2011).

Grouping processes in the brain have been proposed to occur from the mutual excitation of neurons coding for the elements of a group (Herzog, 2018; Roelfsema & Houtkamp, 2011). This mechanism would segregate grouped components of a visual scene from non-grouped information. A possible neural code by which the activity of assemblies subtending to a group is differentiated from the activity of other assemblies is based on the synchrony of their firing (Luck & Beach, 1998; Singer, 1999; von der Malsburg, 1981). Singer (1999) proposed that the synchrony in neuronal activity may represent a plausible mechanism for all the processes characterized by a flexible grouping of visual scenes. Synchronous activity at early perceptual stages has a stronger impact on further processing stages. This would be due to the saliency gained by synchronous presynaptic spikes, which have been found to elicit higher postsynaptic responses both at the local level of the visual system (Dan et al., 1998; Freiwald et al., 2001) and at a more global level of different brain areas (Fries, 2005; 2015). In this sense, postsynaptic neurons would behave as *coincidence detectors* (Abeles, 1991), which read out afferent activity and enhance their responsiveness whether incoming signals are temporally correlated (Raffone & van Leeuwen, 2003; Raffone & Wolters, 2001).

Previous dynamic models accounting for grouping between parts of a visual scene into a coherent organization consisted of coupled oscillators (Grossberg & Grunewald, 1997; Grossberg & Somers; 1991; Raffone & van Leeuwen, 2001; Raffone & van Leeuwen, 2003; Raffone & Wolters, 2001; van Leeuwen et al. 1997). In these models, synchronized activity is understood in terms of coherence of oscillatory activity between neurons coding for to-be-grouped information.

Here, we propose a dynamical network model simulating the processes underlying grouping by proximity and orientation bias. Similarly to previous modeling work, we posit that perceptual organization emerges from the coherent oscillatory activity between the units in low-level perceptual modules. Synchrony is achieved via lateral recurring connections between neurons coding for the same orientation (Eckorn et al., 1990; Engel et al., 1990). The differential pattern of coupled activity is captured by higher levels modules and ultimately leads to an output which can be interpreted as a stable representation of an external stimulus. Our model investigates the dynamics occurring in the brain when an observer is presented with a rectangular dot lattice and is asked to judge which orientation is the most prominent. Attractors in our model are intended as the possible orientations along which a dot lattice can be perceived (*a*, *b*, *c*, and *d*). The model reaches an attractor as one of the four possible orientation is selected via a winner-takes-all dynamics. During each simulated trial, the model autonomously converges towards a winning attractor via a self-organized process driven by the intra- and inter-modules pattern of connectivity.

To account for the stability of grouping by proximity, the strength of the connections between neurons coding for each orientation is scaled according to the attraction function (Kubovy et al., 1998; Kubovy & Wagemans, 1995). This reinforce lateral connections in way that is proportional to the inter-dot distances. On the other side, to account for the flexibility observed for dot lattices perception, we provide our model with internal stochastic noise, which allows for the trial by trial variability of the winning attractor.

We test the model under different conditions of AR and proximity sensitivity (α parameter in Equation 1.1). We also implemented a possible mechanism for the vertical preference discussed in previous chapters.

3.2 Model Architecture and Functional Logic

The model consists of four interacting modules (Figure 3.1), each of which computes a different operation in order to achieve the emergence of perceptual organization in the system. These modules can be split into two distinct levels of perceptual hierarchy. The two lower-level modules mimic the functionality of primary visual cortices, whereas the two higher-level modules are involved in advanced stages of visual processing as well as in perceptual judgment. The lower-level modules are the *perceptual layer* and the *coincidence detectors* module. The higher-level modules are the *accumulators* module and the *decision units* module. Following a functional description of the modules is provided.

- *Perceptual Layer* (PL). This module receives the external excitatory input signaling the presence of an external stimulus and its function can be linked to V1. PL consists of a 64 neuronal assemblies, arranged into a configuration of eight rows and eight columns. In PL, each dot of a simulated dot lattice is mapped into a single assembly in a topographical fashion. These assemblies are composed by four units coding for the same dot in the framework of a specific orientation (*a*, *b*, *c*, and *d*). This is in line with the orientation selectivity of neuronal receptive fields in V1 (e.g. Hubel & Wiesel, 1968). For simplicity, we posit that *a* is always horizontal, *b* is always vertical and the diagonal orientations, *c* and *d*, parallel the distance vectors with norm $||\mathbf{a} - \mathbf{b}||$ and $||\mathbf{a} + \mathbf{b}||$, respectively.

Units within each assembly receive excitatory connections from the units coding for neighboring dots along the same orientation in other assemblies (Gilbert et al., 1996; Stettler et al., 2002). Given the fact that only one interpretation of the stimulus can perceived at time, units also receive inhibitory connections from the units coding for different orientations of a dot within the same assembly. This is in line with cross-orientation models of V1 selectivity (Brouwer & Heeger, 2011; Morrone et al., 1982). In other words, cooperative inter-assembly and competitive intra-assembly dynamics take place at this stage. Finally, during each iteration of the simulated process, units of PL are fed by a randomly distributed noise, which impacts the effectiveness of the sensory input. In this sense, noise can be considered as an intrinsic propriety of the sensory channels, in line with what postulated by signal detection theory (e.g. Mueller & Weidemann, 2008).

- *Coincidence Detectors* (CD) *module*. CD module is involved in the readout of the synchronous activity between the units in previous module (Raffone & van Leeuwen, 2003; Raffone & Wolters, 2001).

  For the units representing the dots along orientation *a*, CD module consists of eight units. Each of these units codes the synchrony between the neurons lying on the same row (i.e. horizontally aligned). Therefore, each row of the PL is mapped into a single unit in CD module. Similarly, for the units coding for orientation *b*, the eight CD neurons process the synchrony between neurons lying on the same column (i.e. vertically aligned).

On the other hand, the synchrony between the units representing oblique orientations (*c* and *d*) is read out by seven CD units each. This is because we limited the number of diagonals taken into account within PL.

The information concerning the synchrony at local PL level is then fed as input signal to the higher-level modules.

- *Accumulators* (ACC) *module*. This module consists of four units, each of which is representative for an orientation. Accumulators are a class of neurons which has been widely employed in simulating the processes underlying perceptual judgment (Forstmann et al., 2016; Ratcliff et al., 2016). The activity of accumulators can be represented as the gathering of perceptual evidence computed in different modules, until a threshold is reached. Typically, the overcoming of the threshold coincides with perceptual decision-making.

  Here, we assume a different role for accumulators, which are involved in gathering evidence of the synchrony in lower-level units. In line with what theorized by Singer (1999), the more the units in PL are synchronous the more the rate of accumulators activity drifts. When a threshold for a specific orientation is reached, a stronger input relative to that orientation is passed to the decision units module. Since the orientations compete for the selection, the four accumulator mutually inhibit each other. For the simulations taking into account orientation bias, we provide an additional excitatory input to the accumulator coding for vertical orientation from a not explicitly modelled module. This influences the gathering perceptual evidence from previous stages as well as the activity supplied to the higher module.

- *Decision Units* (DU) *module*. This module consists of four neurons, each of which receives inputs from a single accumulator. When a DU is activated, the system enters an attractor and a specific orientation in the dot lattice is finally selected. Therefore, this module resembles the functioning of the high-level processing stages, such as dorso-lateral prefrontal cortex (DLPFC) during task requiring perceptual discrimination judgments among several alternatives (Heekeren et al, 2008).

  Since only one orientation can be active at time, we provide strong inhibitory connections between these units. Thus, when one of these units becomes active, it shuts off all the concurring neurons in a winner-takes-all fashion.

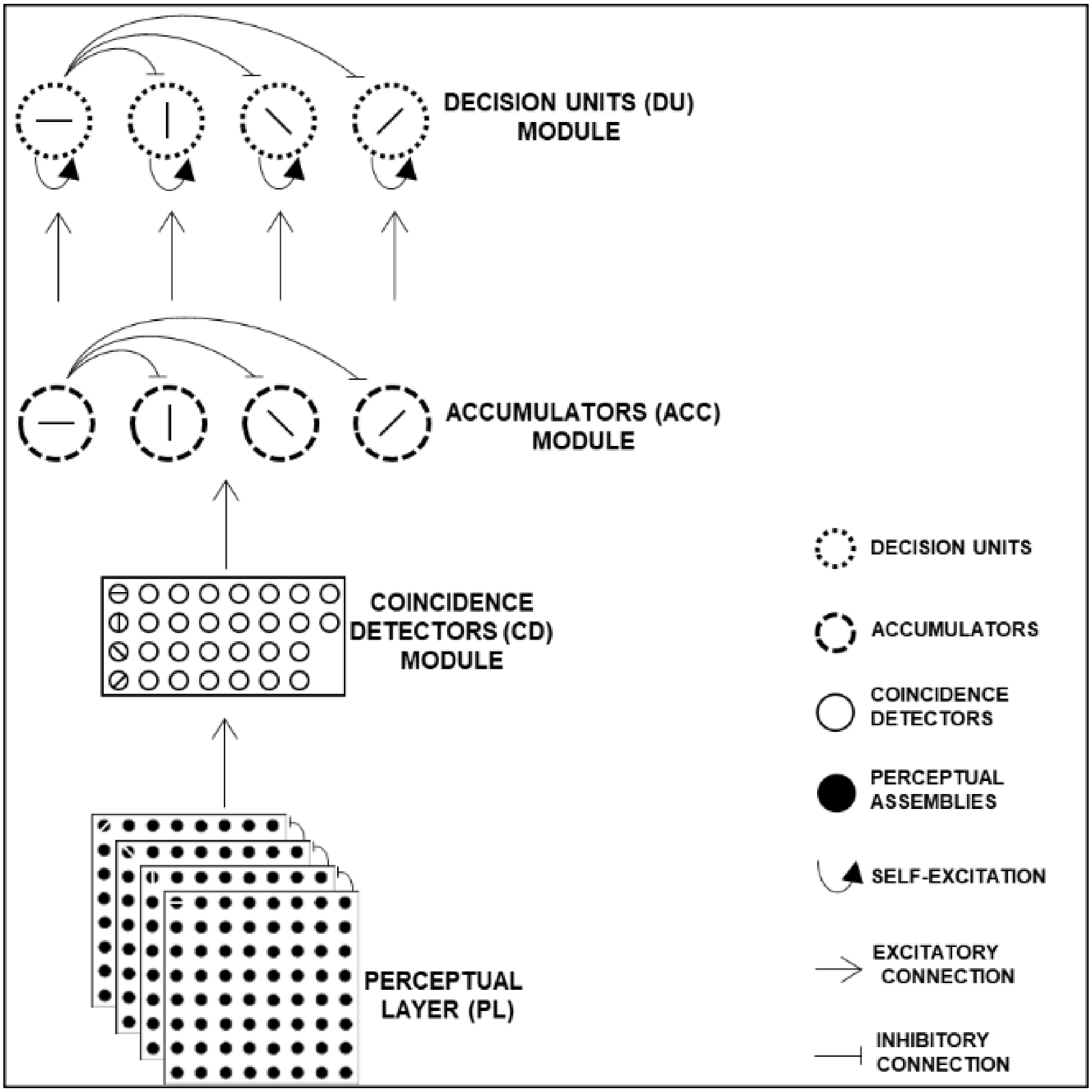


Figure 3.1: Architecture of the model. Perceptual layer (PL) and coincidence detectors (CD) module simulate the functioning of low-level perceptual areas, whereas accumulator (ACC) module and decision units (DU) model account for the operations performed at higher levels of perceptual hierarchy. Lateral inhibitory connections are represented within the modules, whereas excitatory feedforward connections are represented between the modules. In the figure, different assemblies of PL are represented as the units occupying the same position across the depth dimension. Excitatory connections between units in neighbor assemblies are not shown here.

3.3 Model Units

For each modules of the network, we model different neuronal populations with specific activation dynamics.

Units in PL are composed by the coupling between a *fast shunting excitatory neurons* and a *slow additive inhibitory neurons* (Grossberg & Grunewald, 1997; Grossberg & Somers; 1991; Figure 3.2A). The activity of fast excitatory neurons can be thought as a phasic cellular response to an external stimulus, whereas the activity of slow inhibitory neurons resembles a tonic hyperpolarizing current on such neurons. The different temporal dynamics between fast and slow components is responsible of their behavior as *relaxation oscillators* (Figure 3.2B).

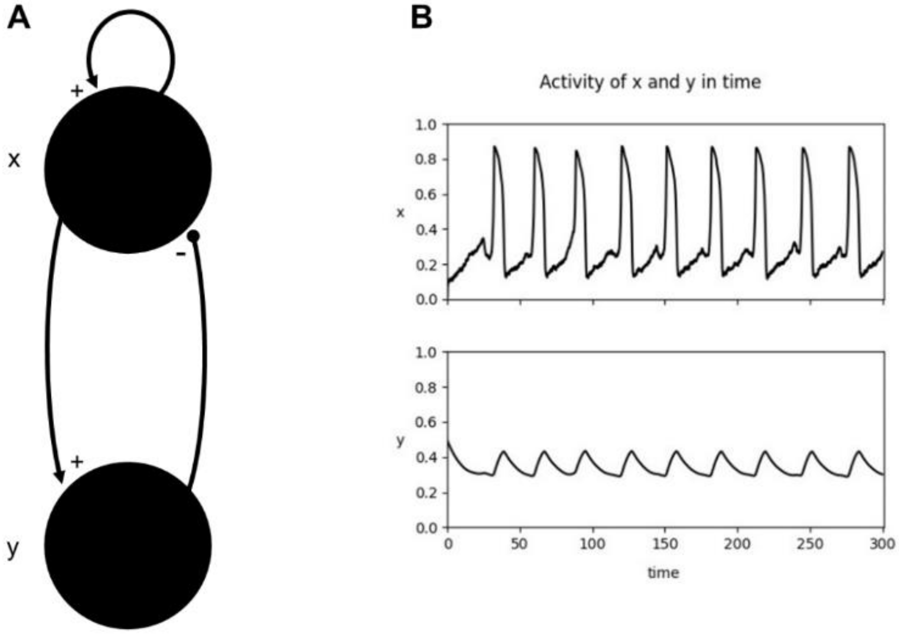


Figure 3.2: Illustration of oscillators implemented in PL. (A). *The units are composed by a fast excitatory component* (x) *coupled with a slow inhibitory component* (y). (B). *The slower rise and decay of the inhibitory subunit induces the oscillatory activity in the excitatory subunit.*

The activation of these neurons is specified by the following first-order differential equations:

$$\frac{dx_i}{dt} = -Ax_i + (B - x_i)\left[Cf_a(x_i) + I^{neigh} + I^{stim} + I^{noise}\right] - Dx_i f_a(y_i) - x_i I^{inh} \qquad [3.1]$$

$$\frac{dy_i}{dt} = E\left(x_i - y_i\right). \qquad [3.2]$$

The terms x_i and y_i represent the activity of the *i*th fast and slow unit, respectively, and could be physiologically interpreted as the average voltage of a neural population at time t (Ellias & Grossberg, 1975). At time $t = 0$, values for x_i and y_i are randomly assigned from the ranges [0, 0.15] and [0.15, 0.55], respectively. In Equation 3.1, the term $-A$ is the decay rate parameter, whereas the terms $(B - x_i)$ and $-Dx_i$ are the shunting factors controlling for the total excitiatory and total inhibitory input, respectively. Total excitation is provided by the terms $Cf_a(x_i)$, I^{neigh}, I^{stim}, and I^{noise}, representing non-linear self-excitation, excitatory input from units coding for neighboring dots along the same orientation, external input produced by dot lattice exposure, and stochastic noise of the channels, respectively. Total inhibition is given by $f_a(y_i)$ and I^{inh}, which are the non-linear tonic input from the slow coupled unit and the intra-assembly inhibition from units coding for the same dot along another orientation. In Equation 3.2, the term E designates the rise and the decay of the slow unit, thus controlling the rate of oscillatory activity (Figure 3.2B). A, B, C, D, E, and I^{stim} are constant parameters of the model and their value is indicated in the Appendix. During each iteration, I^{noise} assumes a random value in the range [0, 0.1625]. I^{neigh} and I^{inh} are specified by the following equations:

$$I_{i,k}^{neigh} = \left[f_a(x_{i-1,k}) + f_a(x_{i+1,k})\right]W_k^{exc} \qquad [3.3]$$

$$W_k^{exc} = mg + m(1-g)e^{-\alpha\left(\frac{k}{a} - 1\right)} \qquad [3.4]$$

$$I_i^{inh} = \sum_{\substack{k=1 \\ \forall\, k \neq i}}^{N} f_b(x_{i,k}) W_{PL}^{inh} \qquad [3.5]$$

Equation 3.3 shows that each unit receives excitatory non-linear inputs from its two closest neighbors along *k*th orientation. This mechanism has been shown to effectively synchronize the activity between directly and indirectly connected fast oscillators (Grossberg & Somers, 1991). The incoming excitation is scaled by the term W_k^{exc}, which corresponds to the synaptic strength between *i*th unit and its neighbor units along *k*th orientation. Equation 3.4 describes how W_k^{exc} is computed and its last term is the attraction function of Pure Distance Law (Kubovy et al. 1998; Kubovy & Wagemans, 1995). Thus, the efficacy of lateral excitatory connections in synchronizing the activity of one orientation-selective population in PL exponentially decreases with the corresponding inter-dot distance. In Equation 3.4, *m* defines the maximum value that W_k^{exc} can assume, whereas g controls the difference between the values of W_k^{exc} for each orientation. Equation 3.5 describes the intra-assembly inhibition. It can be thought as a signal from a not explicitly modelled interneuron integrating the non-linear activity of alternative fast units, scaled by the strength of inhibitory synapsis, W_{PL}^{inh}. The values for parameter *m*, *g* and W_{PL}^{inh} are provided in the Appendix.

At each iteration, the synchrony of oscillatory activity between same orientation-selective fast units in the PL is read out by the CD, according to:

$$\frac{dcd_{i,k}}{dt} = \tau \left[\frac{1}{N} \sum_{i=1}^{N} (x_{i,k}) - cd_{i,k} \right] (1 - \tau) \qquad [3.6]$$

In Equation 3.6, τ represents the time window by which two oscillations can be considered synchronous and can be linked to a neuronal membrane time constant (Raffone & van Leeuwen, 2001). In other words, it controls the advantage for synchronous activity in affecting higher-level processing.

The read out of the synchrony made by CD is then fed to the ACC, whose temporal evolution is governed by the following equations:

$$\frac{dact_k}{dt} = -Fact_k + \left[(B - act_k)\left(\frac{1}{N}\sum_{i=1}^{N} f_c(cd_{i,k}) + bias\right) - I_k^{inh} act_k\right] G \qquad [3.7]$$

$$I_k^{inh} = \sum_{\substack{o=1 \\ \forall\, o \neq k}}^{N} f_d(act_o) W_{ACC}^{inh} \qquad [3.8]$$

Equation 3.7 describes the activity (act_k) of ACC based on the non-linear input from CD. Similarly to Equation 3.1, the terms $-F$ and $(B - act_k)$ are the decay rate and the shunting scaling factor for excitatory input, respectively. The term $bias$ is set to 0.035 for the ACC coding for b (vertical) orientation, during the simulations reproducing the effects of orientation bias; otherwise it is set to 0. The parameter G defines the rise of the units, that is the speed by which information about the synchrony in PL is gathered. Equation 3.8 computes the lateral inhibition among ACC and is based on the same principles of Equation 3.5.

In DU module, each unit receives information relative to the synchrony for a specific orientation processed in previous levels. At this stage, unit activation is governed by the following equations:

$$\frac{ddec_k}{dt} = -Adec_k + (B - dec_k)[Uact_k + Cf_e(dec_k)] - dec_k I_k^{inh} \qquad [3.9]$$

$$I_k^{inh} = \sum_{\substack{o=1 \\ \forall\, o \neq k}}^{N} f_g(dec_o) W_{DEC}^{inh} \qquad [3.10]$$

By comparing Equations 3.9 and 3.1, we can see how the activity dynamics of DU is similar to the one described for the fast shunting units used to model the PL. The difference is that these neurons are not coupled to any slow additive unit. This prevents oscillatory activity at this stage, such that when an orientation is selected, a single unit stabilizes its activity (i.e. an attractor is entered). Fundamental to such mechanism is the lateral inhibition computed by Equation 3.10, where the parameter W_{DEC}^{inh} is higher than the synaptic weights for inhibition provided in previous layers. The parameter U in Equation 3.9 is an amplification factor for the activity feeding from the ACC.

Finally, $f(x)$ is a sigmoid function converting neuronal activity into non-linear signals, given by:

$$f(x) = \frac{x^n}{Q^n + x^n} \qquad [3.11]$$

With Q determining the threshold (i.e. the value where the function starts to rise) and n being the gain (i.e. the steepness) of the sigmoid.

3.4 Simulations

Our model is used to simulate grouping process with and without orientation bias influence. For both scenarios, we provide the model with different AR (1.0, 1.1, and 1.2), and different values of the α parameter in Equation 3.4 (6.72 and 9.145). In accordance with what outlined in previous chapters, this allows us to investigate the behavior of the model in different conditions of perceptual ambiguity and sensitivity to proximity. The values for α have been established on the basis of the results of Kubovy and Wagemans (1995). They found that α among the participants oscillated in the range [6.72, 11.57]. Accordingly, we use as lower α condition the minimum value these authors found in their experiment, and compute the value for higher α condition by adding to 6.72 the difference between the highest and the lowest α found by Kubovy and Wagemans (1995).

Thus, we provide the network with 2(bias: present vs. absent) x 2(α: 6.72 vs. 9.145) x 3(AR: 1.0 vs. 1.1 vs. 1.2) conditions. We test 50 randomly initialized networks (*seeds*), for 30 replications (i.e. trials) per condition. Each trial lasts until either an orientation is selected (i.e. an attractor is entered) or a maximum of 3000 iterations is reached. During each iteration, differential equations are solved using the Euler's method, with a time-step of 0.1. This corresponds to 300 ms of simulated time, in line with the exposure timing typically adopted in dot lattices experiments.

3.5 Model Behavior

In this section, we show how the behavior of one seed in a single trial changes with AR. The simulations reported here have been performed using α = 6.72 and without the additional parameter accounting for orientation bias.

Figure 3.3 illustrates the behavior of the model at AR = 1.0. Figure 3.3A show the activity of eight fast excitatory units of PL coding for the different orientations (*a*, *b*, *c*, and *d*). In each subplots of Figure 3.3A, oscillatory activity represents the temporal evolution of a single unit coding for a dot on the same straight line. Each unit receives excitatory connections from the two closest neighbors along the same orientation and sends inhibitory connection to the complementary units coding for alternative orientations. The efficacy of such excitatory connections in synchronizing their activity depends on their weight, which, based on Equations 3.4, depends on the attraction function.

The activity of the units coding for *a* and *b* presents a similar pattern of high synchronization, whereas the activity of the units processing for *c* and *d* orientations is completely stochastic. Since at AR = 1.0, *a* and *b* orientations are characterized by the same inter-dot distance, which is shorter than the one subtending to *c* and *d*, this pattern is in accordance with the claim considering neural synchrony as a plausible marker for grouping by proximity (Singer, 1999).

The synchrony between the units in the PL is then captured by the CD. Figure 3.3B, displays the activity of four synchrony detectors, each of which processes the temporal coincidence between the eight collinear perceptual units of Figure 3.3A. Note how the amplitude of such units is higher for synchronous orientations (*a* and *b*) than for asynchronous ones (*c* and *d*). This produces a net differentiation for the activity of the four ACC (Figure 3.3C). In this specific trial, we observe a toss-up between *a* and *b* orientations, which ultimately leads to the selection of *a* by the DU (Figure 3.3D). When *a* is selected at this stage, the layer stabilizes its activity and the competitor orientations are shut-off via the lateral inhibition from the unit coding for the winning interpretation.

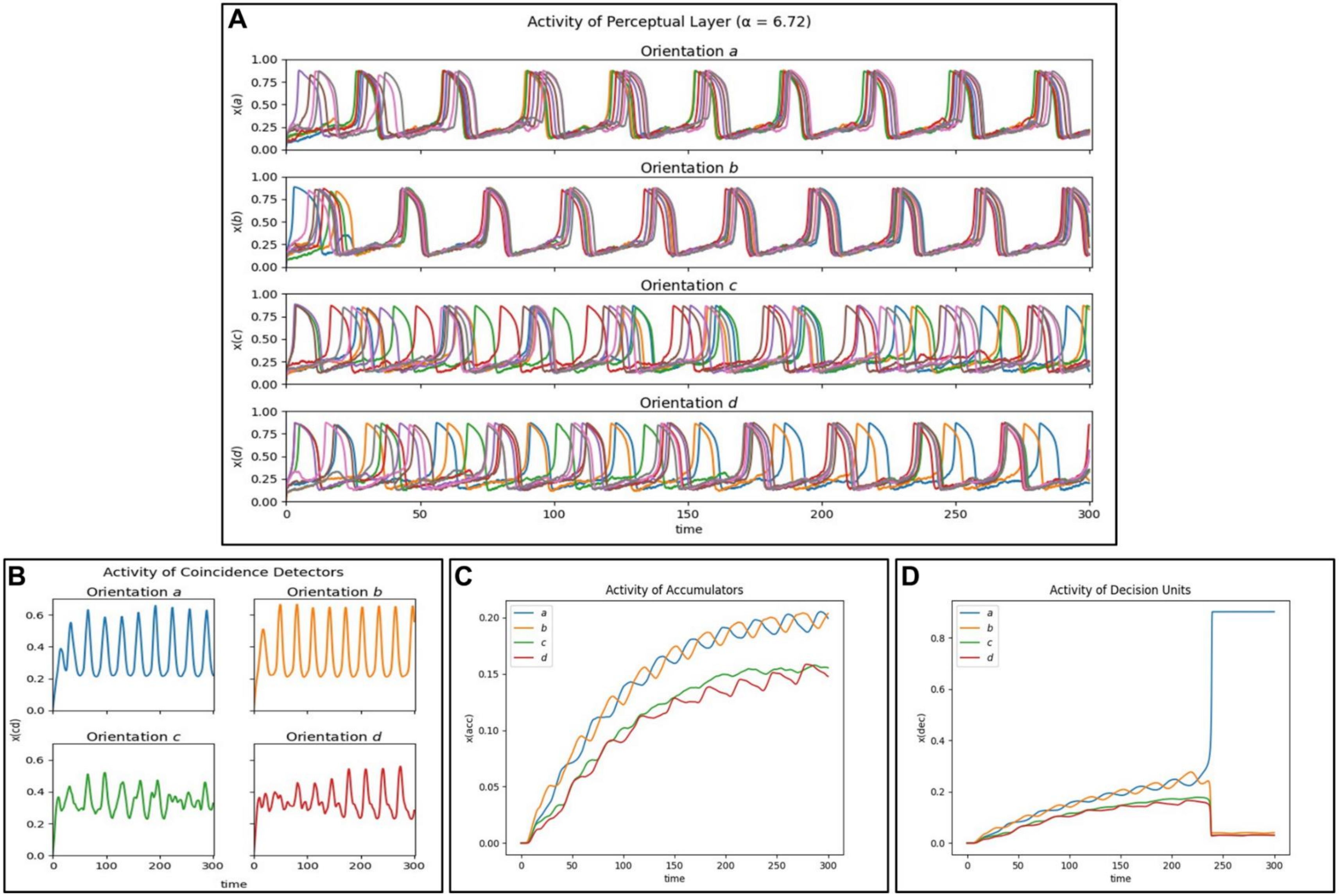


Figure 3.3: Simulation of a single trial with AR = 1.0, α = 6.72, *and no orientation bias parameter*. (A). *Perceptual units coding for a and b orientations show a similar pattern of synchrony, whereas the activity for c and d is stochastic.* (B). *The amplitude of* CD *is higher for the synchronous orientations (a and b) than for the asynchronous ones.* (C). *The same pattern is reflected at the level of* ACC, *which show a toss-up between a and b.* (D). *Once an orientation is selected, the activity relative to other orientations is shut-off via the lateral inhibitory connections between* DU.

Figure 3.4 reproduces a trial for a dot lattice at AR = 1.1. Figure 3.4A shows that *a* is the most synchronized orientation. Synchrony is still present among the units coding for *b*, although in a weaker form with respect to previous simulation. Again, stochastic activity is observed for the units representing *c* and *d*. The higher synchronous activity for the neurons coding for *a* and *b* is read out by the synchrony detectors (Figure 3.4B), which send this information to the respective accumulators. Although at the activity of CD show a clear predominance for orientation *a*, we again observe a close competition with *b* at the ACC level (Figure 3.4C). Such competition leads to the selection of *b* at the level of DU (Figure 3.4D). Therefore, even if the activity at lower layers provides a clear specification of what orientation should be selected based on proximity information, the recurrent dynamics between low- and high-level modules induces for a misinterpretation.

This pattern is in line with what predicted by the Pure Distance Law (Kubovy et al., 1998; Kubovy & Wagemans, 1995), by which grouping by proximity depends on both the AR and the sensitivity parameter α. In fact, when α is low and the AR still allows for a certain degree of perceptual ambiguity (as for AR = 1.1), the probability to report orientation *b* is still considerable.

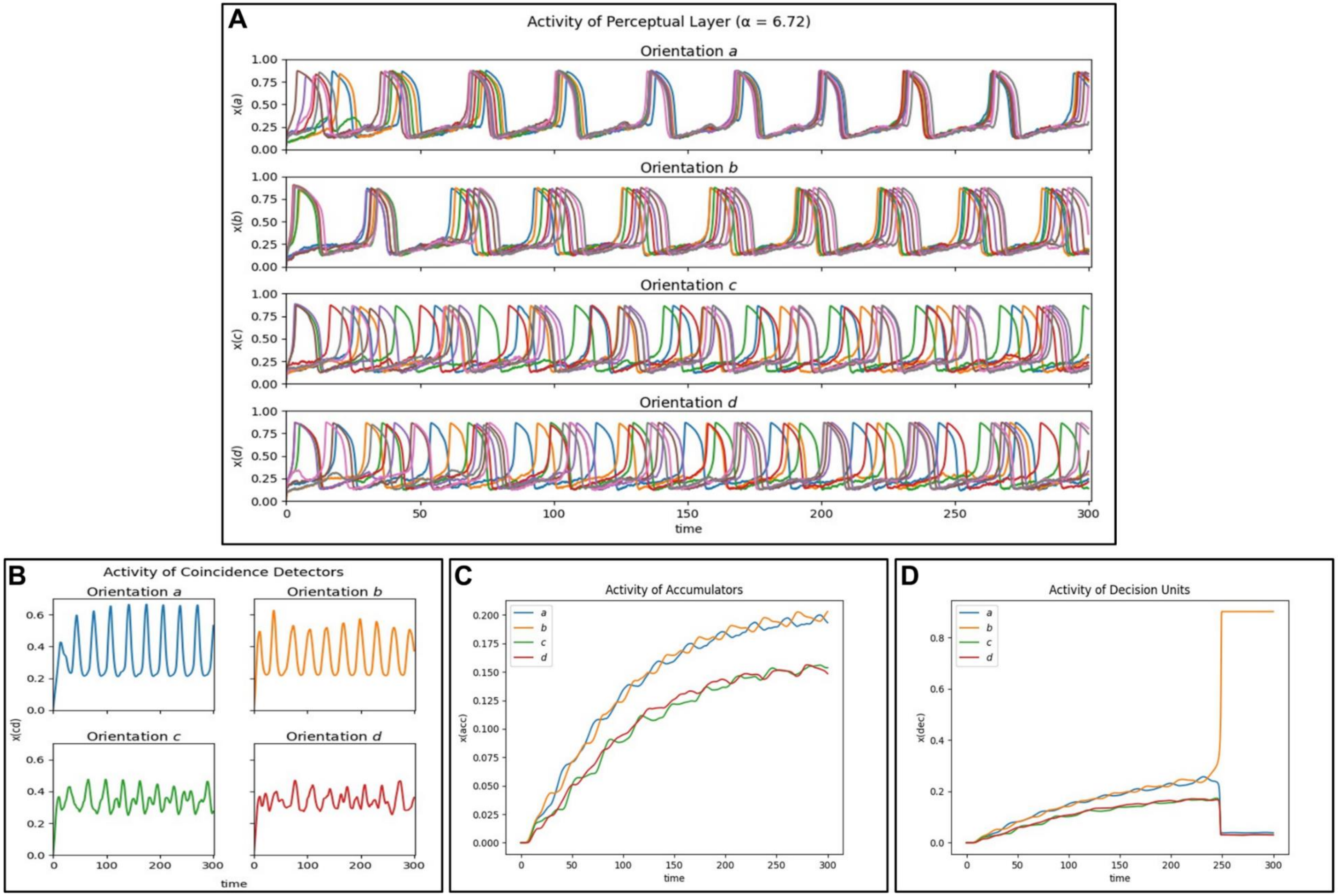


Figure 3.4: Simulation of a single trial with AR = 1.1, α = 6.72, *and no orientation bias parameter*. (A). *Synchrony in* PL *is higher between the units coding for a. A lower degree of synchronization is also observed between the units coding for b* (B). *The amplitude of* CD *is higher for the a than for other orientations. Note, however, that the* CD *coding for b still have considerable amplitude.* (C). *Despite the higher synchrony for a at lower stages,* ACC *activity still shows a toss-up between a and b.* (D). *The high activity of b at* ACC *level induces its selection at the stage of* DU. *This can be considered as a misinterpretation of the stimulus.*

At AR = 1.2 (Figure 3.5), the attraction of the dots along *b* further reduces, and such trend is mirrored by the further desynchronization of the units within the PL with respect to preceding simulations(Figure 3.5A). In this case, both the amplitude of CD (Figure 3.5B) and the rise of ACC (Figure 3.5C) show a net preference for orientation *a*, which ultimately leads to its selection at the DU stage (Figure 3.5D).

These single-trial simulations illustrate how the model could account for the attraction function in dot lattices (Kubovy et al., 1998; Kubovy & Wagemans, 1995). The basic mechanism through which this is achieved is the gradual desynchronization for *b*, due to the increase of its relative inter-dot distance (AR). When the AR = 1.0, both *a* and *b* show the same level of synchronization and the selection of the model is stochastic. Moreover, the interplay between lateral within-module connections and feedforward between-module connections reproduces the occasional rise of orientations other than *a*, especially when the stimulus support (AR) and the sensitivity to proximity (α) are low.

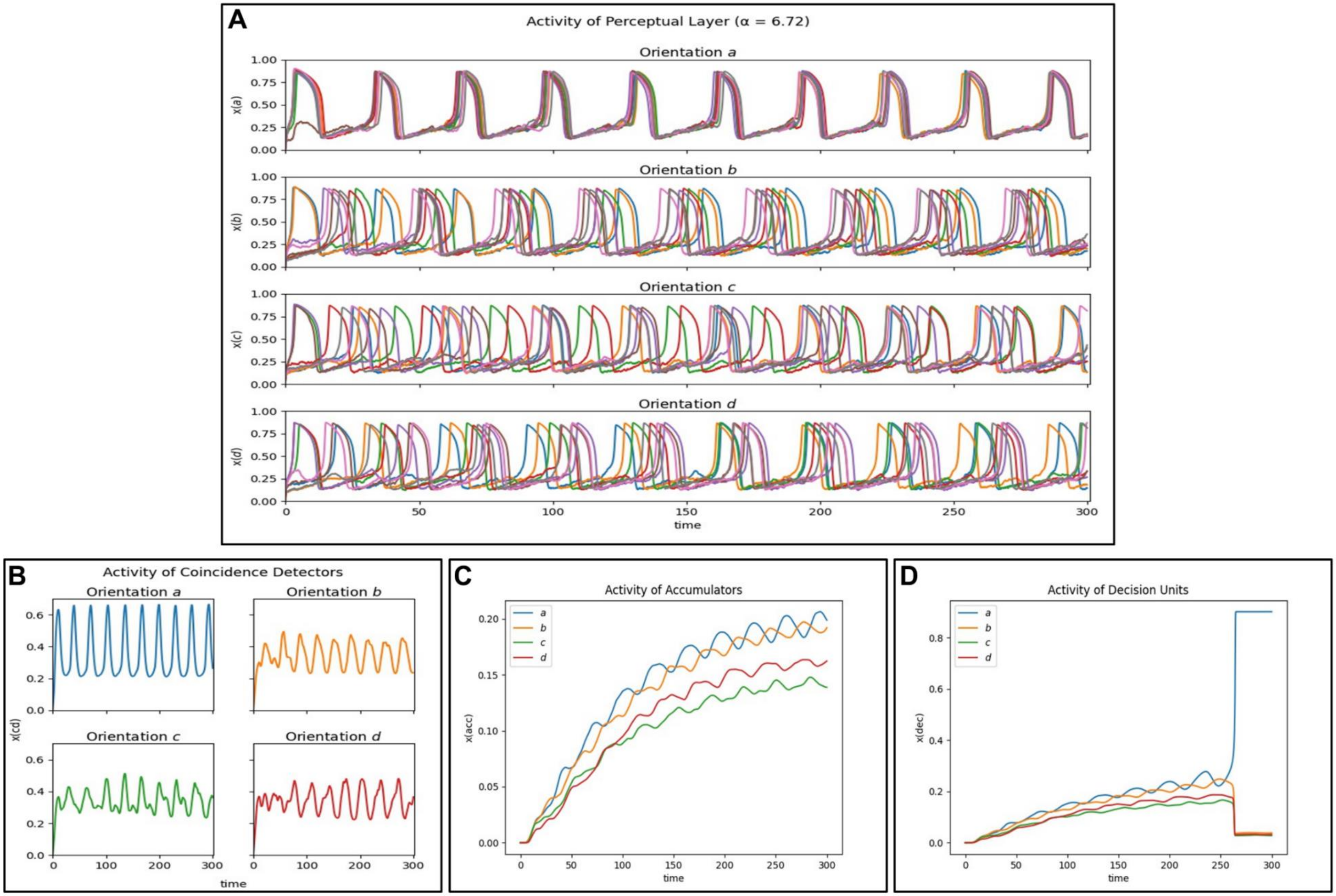


Figure 3.5: Simulation of a single trial with AR = 1.2, α = 6.72, *and no orientation bias parameter*. (A). *Synchrony between* PL *units coding for b further reduces, compared to previous simulations.* (B). *The amplitude of* CD *puts b on the same level of c and orientations.* (C). ACC *activity shows a superiority for a with respect to b. Thus, in this condition, the low synchrony for b in previous layers cannot be overcame at this stage.* (D). *Given the higher synchrony for a in previous modules, this orientation is selected by DU.*

3.6 Simulation Results

The results of all simulations run are summarized in Figure 3.6. The plots show the frequency of winning orientations across the simulated conditions. The first thing to note is that orientations *c* and *d* are never selected by the model. This is in line with previous studies with human observers (Bleumers et al., 2008; Claessens & Wagemans, 2005; 2008; Nikolaev et al., 2016; Strother & Kubovy, 2006) finding that the report of these orientations in rectangular dot lattices is negligible.

The upper panels of Figure 3.6 show the results of the simulations without any bias provided to vertical orientation (*b*). The model replicates the effect of both AR and α parameters of attraction function (Kubovy et al., 1998; Kubovy & Wagemans, 1995). For AR = 1.0, the frequency in selecting *a* and *b* equates, whereas grouping along orientation *a* increases with the AR of the simulated dot lattices. Concerning the effects of α, no difference emerges when the simulated dot lattices are at the extremes of the perceptual ambiguity continuum (i.e. AR = 1.0 and AR = 1.2). This pattern of results is expected, since the effect of α is to attenuate the synchrony between the units coding for orientations characterized by higher inter-dot distance. When AR = 1.0, *a* and *b* subtend to the same inter-dot distance and, thus, both these orientations are in accordance with dot proximity. On the other hand, at AR = 1.2, the difference in synchronous activity between the units coding for the two orientation is already high, and thus further reducing the synchrony for *b* provides a negligible effect. This suggests that grouping by proximity can be effectively promoted by highly unambiguous stimuli, even for low values of proximity sensitivity. On the other hand, the modulation of α is evident at AR = 1.1. In this case, AR alone does not promote for a clear differentiation of the activity relative to *a* and *b* in higher level module and, thus, the action of α in impacting the activity underlying orientation *b* produces visible effects.

The lower panels of Figure 3.6 show the effect of AR and α in the simulations accounting for orientation bias. Orientation-based grouping is more prominent when the stimuli are at the highest degree of ambiguity, whereas it dissolves for clearly unambiguous stimuli. This is in line with previous findings, which showed that orientation bias decreases with AR (Gephstein & Kubovy, 2005; Nikolaev et al., 2016). The ratio between *a* and *b* responses does not change for different values of α at AR = 1.0 and AR = 1.2. Again, this depends on the difference in synchrony promoted by AR. In particular, when AR = 1.0, the synchrony between the units coding for *b* is not reduced by α, and, thus, the additive bias parameter produces a similar effect across the two conditions. Conversely, at AR = 1.2, the difference in synchrony between *a* and *b* makes the model resistant to the additive effect of bias already at the lowest value of α. At AR = 1.1, we find that for low values of α the model is sensitive to orientation bias. This effect depends on the dynamics occurring in high-level modules, where the activity between *a* and *b* is not clearly differentiated. This promotes for the higher rate of orientation-based grouping during these trials. This trend is, however, counteracted by α, which effectively promotes for a shift towards proximity-based grouping as its value increases.

Overall, our simulations account for several findings in psychophysical literature about grouping by proximity and orientation bias. Specifically, our model well replicates the attraction function (Kubovy et al., 1998; Kubovy & Wagemans, 1995), as well as the higher impact of orientation bias for ambiguous stimuli. Moreover, the influence of α on grouping process is effective only at AR = 1.1, that is when the difference in synchrony between *a* and *b* is not large enough to be efficiently captured by higher level modules.

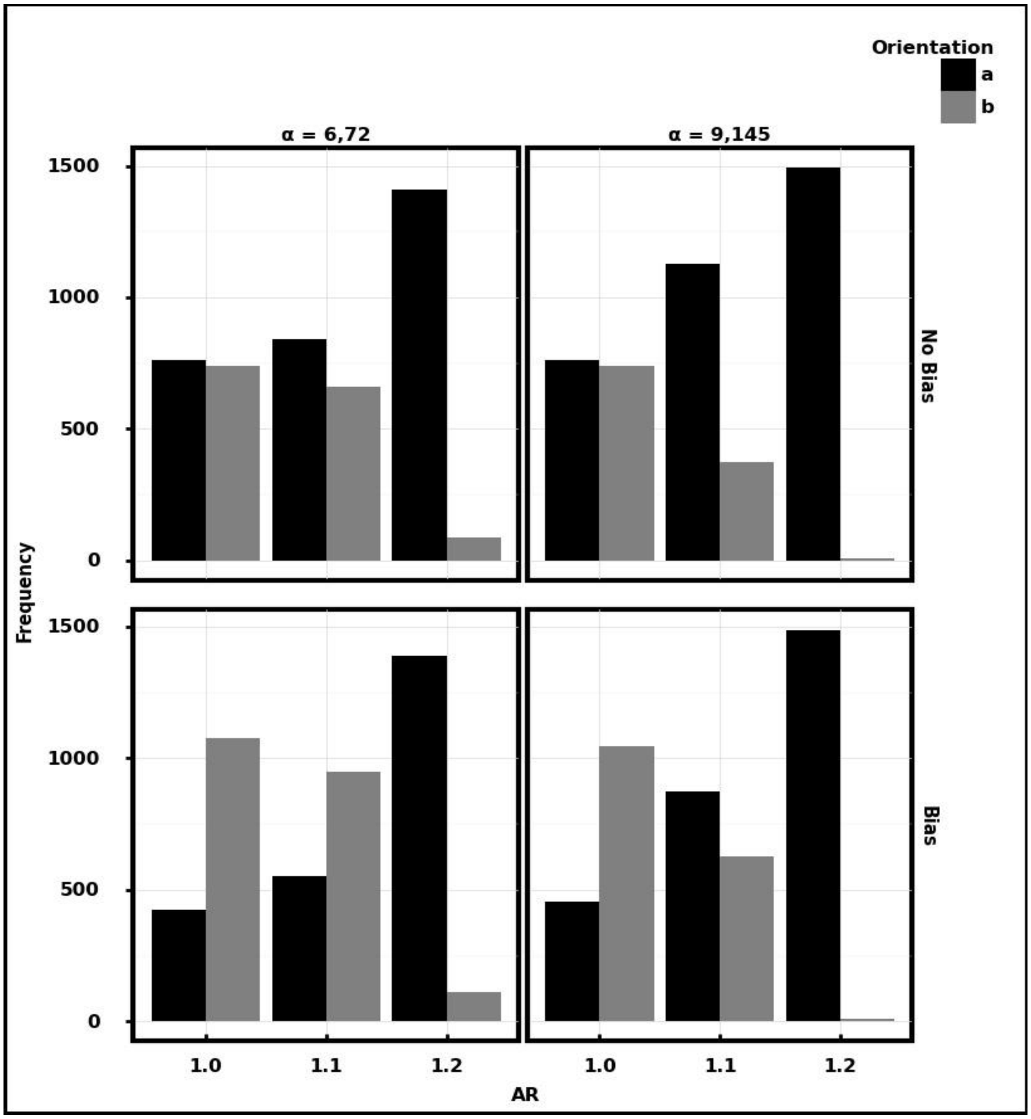


Figure 3.6: Simulation results for all the testing conditions averaged over 50 seeds. The simulations not providing for orientation bias show a clear effect of AR *in increasing the rate of proximity-based grouping. Notice how increasing* α *produces visible effects only at* AR = 1.1. *In the simulations accounting for bias, there is a clear inverse relationship, with orientation-based grouping being reduced by* AR. *Also in this condition, the modulation of* α *produces a drift in grouping process only at* AR = 1.1.

3.7 Discussion

In this section, we provided a dynamical model to investigate the neural mechanisms underlying grouping processes in dot lattices. Gestalt organization, such as grouping by proximity, is an emergent feature of our perceptual system (Pomerantz & Portillo, 2011), which is characterized by both stability and flexibility. Stability is assured in our model by the convergence to a fixed activation state (i.e. attractor) in the DU module. Once an attractor is settled on, activity in DU module is no longer perturbed by lower level inputs. This corresponds to the selection of one orientation from the stimulus, which persists for the remaining duration of the trial. Different attractors in our model represent all the possible interpretations along which the dots can be grouped. The possibility to switch between different attractors over the trials is assured by the stochastic noise at the level of sensory neurons. This mechanism ultimately provides grouping processes with a certain degree of flexibility. In our model, this balancing between stability and flexibility results from the self-organization dynamics. In other words, we do not provide any external teaching signal specifying for the ideal output, but the model is free to evolve towards a particular state in an unsupervised fashion. This is in line with the view that Gestalt phenomena spontaneously arise from self-organized processes in the brain (Wagemans et al., 2012b).

Critical to the emergence of grouping by proximity in our model is the interplay between competitive and cooperative connections occurring at every level of the model’s architecture. Competitive connections are ultimately responsible for the winner-take-all dynamic implemented at DU level, where only one attractor can be entered at time. On the other hand, cooperative connections play a critical role at low perceptual module, where they promote for synchronous activity between neighboring units.

Synchrony between units coding for the same orientation in PL module is the marker for how the sensory input is represented at higher processing modules (Luck & Beach, 1998; Singer, 1999; von der Malsburg, 1981). Ultimately, this is the substrate underlying grouping by proximity in our model: the higher the proximity along one orientation, the more the units coding for that orientation will synchronize their firing; increasing, thus, the chances to win the competition in DU module. The idea of synchrony between neurons coding for elements belonging to the same group as a possible substrate of Gestalt phenomena has successfully been applied in a previous model (van Leeuwen et al., 1997). The authors proposed a model composed by four layers of chaotic oscillators. Excitatory connections within and between the layers coding for different orientations were found to be effective in producing the emergence of grouping by proximity and by good continuation. However, with respect to our work, grouping by proximity was investigated as the shift between *a* and *b* organizations in rectangular dot lattice with AR = 1.0 over sustained stimulus exposure; without simulating the effects of different levels of stimulus ambiguity. Moreover, the previous model only implemented low level dynamics, without defining a large scale neural architecture accounting for subsequent steps of visual processing.

In our model, the higher synchrony for the orientations characterized by smaller inter-dot distance is accomplished via the scaling of the weight of excitatory connections in PL module according to the attraction function (Equation 3.4). This implies that synchrony depends on both AR and α, with higher values of both parameters reducing the synchrony for orientations other than *a*. This functional logic replicates several aspects of grouping by proximity in rectangular dot lattices (Bleumers et al., 2008; Claessens & Wagemans, 2005; 2008).

We find that grouping by proximity was more prominent for high AR, implying lower synchronization between the neurons coding for orientations characterized by higher inter-dot distances. When AR = 1.0 and the inter-dot distances along *a* and *b* equates, the same level of synchrony is observed between these two orientations. In this case, there is no difference in the selection between *a* and *b*. This pattern also explains why the simulations addressing for the orientation bias produces a higher increase for *b* selection at AR = 1.0 than at other AR conditions. In this case, orientation bias is an additive term to the already high synchrony for *b*, which boosts its activation in further modules with respect to *a*.

In our simulations, varying α only affects the grouping by proximity at AR = 1.1. Thus, α provides an additive effect to AR, which, however is effective only when the stimulus support alone is not sufficient to promote for a clear differentiation between *a* and *b*. Again, this pattern can be attributed to a weakening of *b*, rather than to an actual boost of proximity sensitivity. A similar mechanism also occurs in the simulations accounting for orientation bias, where higher α decreases the grouping along *b*, despite the additional bias term received.

A critical aspect of our model is that the input relative to the bias for vertical (*b*) orientation is fed at the level of accumulators. This apparently contrasts with a view of perceptual bias as top-down signal from fronto-temporal networks on early visual areas (Frith & Dolan, 1997; Hohwy, 2017; Kok et al., 2013; Summersfield & de Lange, 2014; Wang et al., 2013). However, in the context of our model, such top-down modulation would be translated by a higher synchrony for vertical (*b*) orientation, which ultimately would led to a higher saliency in further processing stages (Singer, 1999). Thus, even if our model does not appear to reproduce the precise neural dynamics underlying orientation bias, it produces an equivalent outcome.

It has been proposed that each dot lattice can be thought of as four different stimuli which compete for awareness, each of which being represented by a straight line of dots parallel to *a, b, c,* or *d* (Kubovy & van den Berg, 2008). In our model, this competition for conscious access is depicted in the DU module, which allows only one orientation to be active at a given time. Previous models dealing with selection and awareness of limited portions of visual stimulation (Raffone & Pantani, 2010; Simione et al., 2012) incorporate Block's distinction between *phenomenal* and *access consciousness* (Block, 1995; 2005). In these models, phenomenal consciousness is linked to the activity in lower level perceptual modules, where perceptual representations have broad accessibility (Block, 2007; Simione et al., 2019), whereas access consciousness is linked to the activity of higher-level stages, where few representations are available at a given time for cognitive manipulation (narrow accessibility). The gating of broad phenomenal representation into access consciousness is promoted by a *global workspace* (Dehaene et al., 2006; Dehaene & Naccache, 2001), which serially allows competing representations to gain access consciousness when they receive top-down attentional modulation. In line with both Raffone and Pantani (2010) and Simione et al., 2012, our model maintains the distinction between broad phenomenal accessibility in lower-level modules and narrow accessibility in the higher-level module, where only one orientation can be effectively represented at a time. However, we do not implement any attention-mediated gating signal to control for the winning representation, but the mechanism determining its selection is the bottom-up synchronization achieved at perceptual stages. This is in line with research suggesting that perceptual grouping emerges as a pre-attentive process (Kramer & Jacobson, 1991; Lamy et al., 2006; Luna et al., 2016; Montoro et al., 2014; Villalba-García et al., 2018).

Despite the wide range or psychophysical results accounted, our model in its current form present some limitations which should be addressed in future developments. A central feature of perceptual organization concerns the alternation of different interpretations over the sustained presentation of a stimulus. In other words, attractors are entered and escaped within the same trial, due to the stochastic dynamics of the system. This propriety is defined *metastability* (Kelso, 2012; Kelso et al., 1995) and the attractors are better understood as *fragile attractors* (van Leeuwen, 2007). However, in our model, such intermittent propriety of perceptual system is not rendered, and the switch between the attractors only occurs across different trials.

Another aspect that is not fully supported by human perception studies concerns the fact that, in our model, all the units in the PL receive the same amount of input signaling for a dot lattice. However, in a typical experiment, these stimuli are embedded in a circular configuration, and sometimes in a Gaussian envelope (confront stimuli in Figure 2.1), which may provide a different degree of sensory input between neurons coding for central dots and neurons coding for peripheral ones.

Despite these limitations, our model, already in its current form, offers a biologically plausible substrate of grouping processes. In doing so, it supports both the generation of representations proximity-based and based on orientation bias, which have been found to jointly participate during the perception of rectangular dot lattices.

Conclusion

The aim of the present work was to investigate the functional and neural mechanisms underlying the emergence of perceptual grouping, with a special focus on the trade-off between grouping by proximity and orientation bias in dot lattices perception. The preference for grouping dot lattices along the vertical orientation, regardless of the relative distance between the dots, has been already reported in previous psychophysical experiments (Bleumers et al., 2008; Claessens & Wagemans, 2008; Nikolaev et al., 2016). However, none of these studies provided a systematical investigation about the factors influencing this effect, as well as its interaction with grouping by proximity. The studies reported in this thesis provide the first attempt to characterize the conditions defining whether stimuli are grouped according to extrinsic factors, such as proximity, or according to intrinsic ones, such as orientation bias.

In Study 1, we investigate how the shift between grouping by proximity and orientation bias is shaped by high-level VWM. We considered both the effect of load and content in VWM. We found that both aspects play a key role in determining the perception of rectangular dot lattices. Increasing VWM load produces a net shift from proximity-based to orientation-based grouping. At the same time, we also observe a clear effect of vertical VWM representations in promoting for the biased perception of those stimuli. We propose that grouping process is governed by the competition between proximity proto-representations of the stimuli and the vertical preference imposed on lower-level stage by orientation bias. In our view, grouping processes in low-load condition are mostly driven by proximity proto-representations, whereas in high-load conditions, such representations are weakened and perception is largely under the control of orientation bias.

In line with this claim, we found no difference in low-load condition in the rate of *a* and *b* responses at AR = 1.0. In this case, when both orientations supported stimulus proximity (AR = 1.0), VWM proto-representations subtend to *a* in half of the cases and to *b* in the other half of the cases. This pattern is abolished as the AR increases and proximity proto-representations tune towards *a*. Moreover, we found that in high-load condition the preference for vertical orientation is equal across all the level of AR. This implies that, in such condition, proximity proto-representations have no impact on visual processing of dot lattices.

Model simulations performed in Study 2 are only partially in line with these results. Specifically, we found that the model without the additional bias term shows no difference between *a* and *b* responses at AR = 1.0, and a higher rate of *a* responses as the AR increased. This scenario may be ascribed to the low-load condition of Study 1, thus exhibiting a similar pattern of results across the studies. However, in simulations with the bias term provided (which mimics the high-load condition of Study 1), we found that at AR = 1.2 proximity was still preferred over orientation bias. This difference may be explained by the fact that in Study 1, due to the load provided to VWM, the generation of proximity proto-representation for stimuli at AR = 1.2 is effectively impeded, whereas in Study 2, despite the bias parameter, no mechanism interferes with the gradual building of *a* representation in DU module.

In our model, grouping by proximity emerges as the parallel activation of several processing stages of visual hierarchy. This scheme is in line with the view of a visual system as characterized by *intrinsic holism*. Intrinsic holism has been proposed to occur in the brain through the interplay between short range connections within early visual areas and top-down signals from higher-level processing stages (van Leeuwen, 2015a). Thus, holistic perception would be achieved through recurrent cycles of activity operating at multiple stages of stimulus processing. This supposes a visual system organized as a small-world network, where assemblies of units densely connected via short range connections provide for the coding of local features (e.g. single elements) and their coupling via long range connections allows the global structure to unfold (van Leeuwen, 2015b). This network structure is reproduced in our model, where long range connections between the modules effectively promote for the emergence of grouping.

Simulations with our model provide evidence that the synchrony between the units in perceptual layer has a key role in grouping by proximity. We can speculate that VWM and orientation bias may regulate grouping process by influencing such low-level synchronization. In particular, it could be hypothesized that proximity proto-representations, once established following stimulus presentation, provide the units coding for *a* with top-down feedbacks which further facilitate their synchronization. On the other hand, the effects of orientation bias on units coding for vertical orientations may be more pervasive. It has been proposed that top-down projections affect the spontaneous activity at early stages of visual hierarchy, altering the synaptic efficacy between neurons coding for features that reflect perceptual regularities of the environment (Harmelech & Malach, 2013), such as vertical orientation (Coppola et al., 1998). In this sense, orientation bias may induce a long lasting effect on the synchrony between orientation-tuned neurons in V1.

Future developments of Study 1 may address whether the modulation of VWM on grouping by proximity also occurs for other grouping principles. The reciprocal effect may also be investigated: namely, whether grouping by proximity is affected by other higher level processes, such as the stronger allocation of spatial attention towards the left visual hemifield (i.e. *pseudo-neglect*). It has been reported that detection accuracy is higher when stimuli are presented in the left visual hemifield compared to the right one (Learmonth et al., 2015a; 2015b; Thut et al., 2006) and grouping by proximity is impaired when dot lattices are shifted to the right of the fixation point

(Bleumers et al., 2008). Thus, one might expect proximity-based responses being higher whether the stimuli are presented leftward than rightward. This research would provide us with useful information about how perceptual system is organized, as well as about how the mechanisms subtending to different forms of perceptual bias affect low-level processing.

Concerning Study 2, successive versions of our model should account for a more biologically plausible implementation of the bias input at the level of perceptual units. For example, it could be simulated as external input to the neurons coding for vertical orientation within each assembly from a non-explicitly modelled module. In this way, we would have a direct test that the effect of orientation would promote for a further synchrony between vertical units. To simulate a scenario more similar to that of Study 1, new implementations of the model should account for the effect of VWM content in enhancing the selection of vertical (b) orientation. A possible way to replicate this dynamics may be obtained by acting on the decay rate parameter ($-A$) for vertical DU. Specifically, we could impose that, when a vertical interpretation (b) is selected in one trial, such activity decay is reduced. This would enhance the probability to reselect vertical orientation in next trial. Moreover, simulation results obtained for grouping by proximity are related to a subset to the dot lattices (see Figure 1.2). In fact, our model only deals with the classes of dot lattices constrained by the relation $\mathbf{a} \leq \mathbf{b} < \mathbf{c} = \mathbf{d}$. Thus, we do not account for grouping process in hexagonal and centered rectangular dot lattices, which are constrained by the relations $\mathbf{a} = \mathbf{b} = \mathbf{c} < \mathbf{d}$ and $\mathbf{a} < \mathbf{b} = \mathbf{c} < \mathbf{d}$, respectively. In the former case, we should expect the synchrony between a, b, and c to equate at the perceptual stage. In the latter case, we should expect a similar pattern of low-level synchrony across b and c, which, depending on both α and AR, may induce for misinterpretations of the stimuli. A last development for our model would be to take into account the time elapsed from the presentation of the stimulus ($t = 0$) and the time in which an attractor is entered. We could expect that similarly to grouping by proximity, decision time would be modulated by the attraction function parameters (α and AR), with reaction times for perceptual decision being slower as the values of AR and α decrease. This hypothesis seems supported by the diffusion models of reaction time (Bogacz et al., 2006; Forstmann et al., 2016; Ratcliff et al., 2016), which predict slower reaction times as the to-be-decided alternatives have the same weight of perceptual evidence.

Overall, both the studies reported here agree that proximity- and orientation-based grouping result from the relationship between low-level and high-level processing stages. Further studies are required to better characterize the role of low-level synchrony and high-level VWM proto-representations in the emergence of grouping by proximity. The studies in this thesis contribute to highlight such new research perspective, which hopefully will promote for a renewed interest in the investigation of the mechanisms other than stimulus parameters in determining perceptual grouping.

References

Abeles, M. (1991). Corticonics: Neural circuits of the cerebral cortex. Cambridge, UK: Cambridge University Press.

Ahnelt, P. K. (1998). The photoreceptor mosaic. *Eye, 12*(3), 531-540.

Albertazzi, L. (2015). Philosophical background: Phenomenology. In: J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization*. *Oxford Handbook of Perceptual Organization* (pp. 21-31). Oxford, UK: Oxford University Press.

Albright, T. D. (2012). On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron, 74*(2), 227-245.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The "oblique effect" in man and animals. *Psychological Bulletin, 78*(4), 266-278.

Beck, D. M., & Palmer, S. E. (2002). Top-down influences on perceptual grouping. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1071-1084.

Bisley, J. W., Zaksas, D., Droll, J. A., & Pasternak, T. (2004). Activity of neurons in cortical area MT during a memory for motion task. *Journal of Neurophysiology, 91*(1), 286-300.

Bleumers, L., De Graef, P., Verfaillie, K., & Wagemans, J. (2008). Eccentric grouping by proximity in multistable dot lattices. *Vision Research, 48*(2), 179-192.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences, 18*(2), 227-247.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences, 9*(2), 46-52.

Block, N. (2007). Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences*, *30*(5-6), 481-499.

Blusseau, S., Carboni, A., Maiche, A., Morel, J. M., & von Gioi, R. G. (2016). Measuring the visual salience of alignments by their non-accidentalness. *Vision Research*, *126*, 192-206.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700-765.

Bravais, A. (1949). *Crystallographic studies: On the systems formed by points regularly distributed on a plane or in space* (A. J. Shaler, Trans.). New York, NY: Crystallographic Society of America. (Original work published 1850).

Brooks, J. L. (2015) Traditional and new principles of perceptual grouping. In: J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization*. *Oxford Handbook of Perceptual Organization* (pp. 57-87). Oxford, UK: Oxford University Press.

Brouwer, G. J., & Heeger, D. J. (2011). Cross-orientation suppression in human visual cortex. *Journal of Neurophysiology*, *106*(5), 2108-2119.

Chapman, B., & Bonhoeffer, T. (1998). Overrepresentation of horizontal and vertical orientation preferences in developing ferret area 17. *Proceedings of the National Academy of Sciences*, *95*(5), 2609-2614.

Claessens, P. M., & Wagemans, J. (2005). Perceptual grouping in Gabor lattices: Proximity and alignment. *Perception & Psychophysics*, *67*(8), 1446-1459.

Claessens, P. M., & Wagemans, J. (2008). A Bayesian framework for cue integration in multistable grouping: Proximity, collinearity, and orientation priors in zigzag lattices. *Journal of Vision*, *8*(7), 33-33.

Coppola, D. M., Purves, H. R., McCoy, A. N., & Purves, D. (1998). The distribution of oriented contours in the real world. *Proceedings of the National Academy of Sciences*, *95*(7), 4002-4006.

Dan, Y., Alonso, J. M., Usrey, W. M., & Reid, R. C. (1998). Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nature Neuroscience*, *1*(6), 501-507.

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204-211.

Dehaene, S., & Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, *79*(1-2), 1-37.

Dragoi, V., Rivadulla, C., & Sur, M. (2001). Foci of orientation plasticity in visual cortex. *Nature*, *411*(6833), 80-86.

Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron*, *28*(1), 287-298.

Eckhorn, R., Reitboeck, H. J., Arndt, M. T., & Dicke, P. (1990). Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex. *Neural Computation*, *2*(3), 293-307.

Ellias, S. A., & Grossberg, S. (1975). Pattern formation, contrast control, and oscillations in the short term memory of shunting on-center off-surround networks. *Biological Cybernetics*, *20*(2), 69-98.

Engel, A. K., König, P., Gray, C. M., & Singer, W. (1990). Stimulus-dependent neuronal oscillations in cat visual cortex: Inter-columnar interaction as determined by cross-correlation analysis. *European Journal of Neuroscience*, *2*(7), 588-606.

Essock, E. A., DeFord, J. K., Hansen, B. C., & Sinai, M. J. (2003). Oblique stimuli are seen best (not worst!) in naturalistic broad-band stimuli: A horizontal effect. *Vision Research*, *43*(12), 1329-1335.

Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially global representations in human primary visual cortex during working memory maintenance. *Journal of Neuroscience*, *29*(48), 15258-15265.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, *67*(1), 641-666.

Freiwald W.A., Kreiter A.K., & Singer W. (2001). Synchronization and assembly formation in the visual cortex. In: M.A.L. Nicolelis (Ed.), *Advances in Neural Population Coding, Progress in Brain Research, Volume 130* (pp. 111-140). Amsterdam, The Netherlands: Elsevier Science Publishers.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*(10), 474-480.

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron*, *88*(1), 220-235.

Frith, C., & Dolan, R. J. (1997). Brain mechanisms associated with top-down processes in perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1358), 1221-1230.

Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite effects of recent history on perception and decision. *Current Biology*, *27*(4), 590-595.

Furmanski, C. S., & Engel, S. A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, *3*(6), 535-536.

Gao, Z., Ding, X., Yang, T., Liang, J., & Shui, R. (2013). Coarse-to-fine construction for high-resolution representation in visual working memory. *PloS One*, *8*(2), e57913.

Gao, T., Gao, Z., Li, J., Sun, Z., & Shen, M. (2011). The perceptual root of object-based storage: An interactive model of perception and visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 37*(6), 1803-1823.

Gao, Z., Gao, Q., Tang, N., Shui, R., & Shen, M. (2016). Organization principles in visual working memory: Evidence from sequential stimulus display. *Cognition*, *146*, 277-288.

Gepshtein, S., & Kubovy, M. (2005). Stability and change in perception: spatial organization in temporal context. *Experimental Brain Research*, *160*(4), 487-495.

Gilbert, C. D., Das, A., Ito, M., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences, 93*(2), 615-622.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience, 14*(7), 926-932.

Gogel, W. C. (1976). An indirect method of measuring perceived distance from familiar size. *Perception & Psychophysics*, *20*(6), 419-429.

Grossberg, S., & Grunewald, A. (1997). Cortical synchronization and perceptual framing. *Journal of Cognitive Neuroscience, 9*(1), 117-132.

Grossberg, S., & Somers, D. (1991). Synchronized oscillations during cooperative feature linking in a cortical model of visual perception. *Neural Networks*, *4*(4), 453-466.

Hansen, B. C., & Essock, E. A. (2004). A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes. *Journal of Vision*, *4*(12), 5-5.

Harmelech, T., & Malach, R. (2013). Neurocognitive biases and the patterns of spontaneous correlations in the human cortex. *Trends in Cognitive Sciences*, *17*(12), 606-615.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*(7238), 632-635.

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews Neuroscience*, *9*(6), 467-479.

Herzog, M. H. (2018). Perceptual grouping. *Current Biology*, *28*(12), R687-R688.

Hohwy, J. (2017). Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. *Consciousness and Cognition*, *47*, 75-85.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, *79*(8), 2554-2558.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215-243.

Jiang, Y., Chun, M. M., & Olson, I. R. (2004). Perceptual grouping in change detection. *Perception & Psychophysics*, *66*(3), 446-453.

Jin, D. Z., Dragoi, V., Sur, M., & Seung, H. S. (2005). Tilt aftereffect and adaptation-induced changes in orientation tuning in visual cortex. *Journal of Neurophysiology*, *94*(6), 4038-4050.

Kang, M. S., Hong, S. W., Blake, R., & Woodman, G. F. (2011). Visual working memory contaminates perception. *Psychonomic Bulletin & Review*, *18*(5), 860-869.

Kayaert, G., & Wagemans, J. (2010). Infants and toddlers show enlarged visual sensitivity to nonaccidental compared with metric shape changes. *i-Perception*, *1*(3), 149-158.

Kelso, J. S. (2012). Multistability and metastability: understanding dynamic coordination in the brain. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1591), 906-918.

Kelso, J. A. S., Case, P., Holroyd, T., Horvath, E., Raczaszek, J., Tuller, B., & Ding, M. (1995). Multistability and metastability in perceptual and brain dynamics. In P. Kruse & M. Stadler (Eds.), *Ambiguity in Mind and Nature: Multistable Cognitive Phenomena* (pp. 159-184). Berlin, Germany: Springer.

Kleinschmidt, A., Büchel, C., Zeki, S., & Frackowiak, R. S. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings. Biological sciences*, *265*(1413), 2427-2433.

Kok, P., Brouwer, G. J., van Gerven, M. A., & de Lange, F. P. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience*, *33*(41), 16275-16284.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics, 50*(3), 267-284.

Kubovy, M. (1994). The perceptual organization of dot lattices. *Psychonomic Bulletin & Review*, *1*(2), 182-190.

Kubovy M., & Gepshtein S. (2000). Gestalt: From Phenomena to Laws. In: K.L. Boyer & S. Sarkar (Eds.) *Perceptual Organization for Artificial Vision Systems. The Kluwer International Series in Engineering and Computer Science, Vol 546*. Boston, MA: Springer.

Kubovy, M., & Gepshtein, S. (2003). Perceptual grouping in space and in space-time: An exercise in phenomenological psychophysics. In R. Kimchi, M. Behrmann, & C.R. Olson (Eds.), *Perceptual Organization in Vision: Behavioral and Neural Perspectives* (pp. 45-85). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kubovy, M., Holcombe, A. O., & Wagemans, J. (1998). On the lawfulness of grouping by proximity. *Cognitive Psychology*, *35*(1), 71-98.

Kubovy, M., & van den Berg, M. (2008). The whole is equal to the sum of its parts: A probabilistic model of grouping by proximity and similarity in regular patterns. *Psychological Review*, *115*(1), 131-154.

Kubovy, M., & Wagemans, J. (1995). Grouping by proximity and multistability in dot lattices: A quantitative Gestalt theory. *Psychological Science*, *6*(4), 225-234.

Lamy, D., Segal, H., & Ruderman, L. (2006). Grouping does not require attention. *Perception & Psychophysics*, *68*(1), 17-31.

Learmonth, G., Gallagher, A., Gibson, J., Thut, G., & Harvey, M. (2015a). Intra-and inter-task reliability of spatial attention measures in pseudoneglect. *PloS One*, *10*(9), e0138379.

Learmonth, G., Thut, G., Benwell, C. S., & Harvey, M. (2015b). The implications of state-dependent tDCS effects in aging: behavioural response is determined by baseline performance. *Neuropsychologia*, *74*, 108-119.

Li, B., Peterson, M. R., & Freeman, R. D. (2003). Oblique effect: a neural basis in the visual cortex. *Journal of Neurophysiology*, *90*(1), 204-217.

Luna, D., & Montoro, P. R. (2011). Interactions between intrinsic principles of similarity and proximity and extrinsic principle of common region in visual perception. *Perception, 40*(12), 1467-1477.

Luna, D., Villalba-García, C., Montoro, P. R., & Hinojosa, J. A. (2016). Dominance dynamics of competition between intrinsic and extrinsic grouping cues. *Acta Psychologica, 170*, 146-154.

Luck, S. J., & Beach, N. J. (1998). Visual attention and the binding problem: A neurophysiological perspective. In R. Wright (Ed.), *Visual Attention* (pp. 455-478). New York, USA: Oxford University Press.

Maia, T. V., & Cleeremans, A. (2005). Consciousness: Converging insights from connectionist modeling and neuroscience. *Trends in Cognitive Sciences*, *9*(8), 397-404.

Montoro, P. R., Luna, D., & Ortells, J. J. (2014). Subliminal Gestalt grouping: Evidence of perceptual grouping by proximity and similarity in absence of conscious perception. *Consciousness and Cognition*, *25*, 1-8.

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *216*(1204), 335-354.

Motter, B. C. (1994). Neural correlates of feature selective memory and pop-out in extrastriate area V4. *Journal of Neuroscience*, *14*(4), 2190-2199.

Mueller, S. T., & Weidemann, C. T. (2008). Decision noise: An explanation for observed violations of signal detection theory. *Psychonomic Bulletin & Review*, *15*(3), 465-494.

Nikolaev, A. R., Gepshtein, S., Kubovy, M., & van Leeuwen, C. (2008). Dissociation of early evoked cortical activity in perceptual grouping. *Experimental Brain Research, 186*(1), 107-122.

Nikolaev, A. R., Gepshtein, S., Gong, P., & Van Leeuwen, C. (2010). Duration of coherence intervals in electrical brain activity in perceptual organization. *Cerebral Cortex, 20*(2), 365-382.

Nikolaev, A. R., Gepshtein, S., & van Leeuwen, C. (2016). Intermittent regime of brain activity at the early, bias-guided stage of perceptual learning. *Journal of Vision*, *16*(14), 11-11.

Nasr, S., & Tootell, R. B. (2012). A cardinal orientation bias in scene-selective visual cortex. *Journal of Neuroscience, 32*(43), 14921-14926.

Palmer, S. E. (2002a). Perceptual grouping: It's later than you think. *Current Directions in Psychological Science, 11*(3), 101-106.

Palmer, S.E. (2002b). *Perceptual Organization in Vision*. In H. Pashler (Ed.), *Stevens' Handbook of Experimental Psychology, Third Edition. Volume 1: Sensation and Perception* (pp. 177-234). John Wiley & Sons, Inc., New York, USA.

Palmer, S. E., & Beck, D. M. (2007). The repetition discrimination task: An objective method for studying perceptual grouping. *Perception & Psychophysics*, *69*(1), 68-78.

Palmer, S. E., Brooks, J. L., & Nelson, R. (2003). When does grouping happen?. *Acta Psychologica*, *114*(3), 311-330.

Palmer, S. E., & Nelson, R. (2000). Late influences on perceptual grouping: Illusory contours. *Perception & Psychophysics, 62*(7), 1321-1331.

Pan, Y., Lin, B., Zhao, Y., & Soto, D. (2014). Working memory biasing of visual perception without awareness. *Attention, Perception, & Psychophysics*, *76*(7), 2051-2062.

Peterson, M. A., & Kimchi, R. (2013). *Perceptual organization in vision.* In D. Reisberg (Ed.), *Oxford Library of Psychology. The Oxford Handbook of Cognitive Psychology* (p. 9-31). Oxford University Press, Oxford, UK.

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8-13.

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: Toward a theory of basic Gestalts. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1331.

Quinlan, P. T., & Cohen, D. J. (2012). Grouping and binding in visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1432-1438.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception & Psychophysics*, *44*(6), 563-575.

Raffone, A., Srinivasan, N., & van Leeuwen, C. (2014). The interplay of attention and consciousness in visual search, attentional blink and working memory consolidation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1641), 20130215.

Raffone, A., & Van Leeuwen, C. (2001). Activation and coherence in memory processes: Revisiting the parallel distributed processing approach to retrieval. *Connection Science*, *13*(4), 349-382.

Raffone, A., & van Leeuwen, C. (2003). Dynamic synchronization and chaos in an associative neural network with multiple active memories. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *13*(3), 1090-1104.

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience*, *13*(6), 766-785.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*(4), 260-281.

Roelfsema, P. R. (2005). Elemental operations in vision. *Trends in Cognitive Sciences*, *9*(5), 226-233.

Roelfsema, P. R., & Houtkamp, R. (2011). Incremental grouping of image elements in vision. *Attention, Perception, & Psychophysics*, *73*(8), 2542-2572.

Salahub, C. M., & Emrich, S. M. (2016). Tuning perception: Visual working memory biases the quality of visual awareness. *Psychonomic Bulletin & Review*, *23*(6), 1854-1859.

Schmidt, F., & Schmidt, T. (2013). Grouping principles in direct competition. *Vision Research*, *88*, 9-21.

Scocchia, L., Cicchini, G. M., & Triesch, J. (2013). What's "up"? Working memory contents can bias orientation processing. *Vision Research*, *78*, 46-55.

Sense, F., Morey, C. C., Prince, M., Heathcote, A., & Morey, R. D. (2017). Opportunity for verbalization does not improve visual change detection performance: A state-trace analysis. *Behavior Research Methods*, *49*(3), 853-862.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, *20*(2), 207-214.

Simione, L., Di Pace, E., Chiarella, S. G., & Raffone, A. (2019). Visual attention modulates phenomenal consciousness: evidence from a change detection study. *Frontiers in Psychology*, *10*, 2150.

Simione, L., Raffone, A., Wolters, G., Salmas, P., Nakatani, C., Belardinelli, M. O., & van Leeuwen, C. (2012). ViSA: A neurodynamic model for visuo-spatial working memory, attentional blink, and conscious access. *Psychological Review*, *119*(4), 745-769.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations?. *Neuron*, *24*(1), 49-65.

Sinico, M. (2008). Demonstration in experimental phenomenology: How to bring out perceptual laws. *Theory & Psychology*, *18*(6), 853-863.

Sligte, I. G., Scholte, H. S., & Lamme, V. A. (2008). Are there multiple visual short-term memory stores?. *PloS One*, *3*(2), e1699.

Stettler, D. D., Das, A., Bennett, J., & Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron*, *36*(4), 739-750.

Strother, L., Van Valkenburg, D., & Kubovy, M. (2003). Toward a psychophysics of perceptual organization using multistable stimuli and phenomenal reports. *Axiomathes*, *13*(3-4), 283-302.

Strother, L., & Kubovy, M. (2006). On the surprising salience of curvature in grouping by proximity. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 226-234.

Strother, L., & Kubovy, M. (2012). Structural salience and the nonaccidentality of a gestalt. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 827-832.

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nature Reviews Neuroscience*, *15*(11), 745-756.

Teng, C., & Kravitz, D. J. (2019). Visual working memory directly alters perception. *Nature Human Behaviour, 3*(8), 827-836.

Thut, G., Nietzel, A., Brandt, S. A., & Pascual-Leone, A. (2006). α-Band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *Journal of Neuroscience*, *26*(37), 9494-9502.

Toppino, T. C. (2003). Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics*, *65*(8), 1285-1295.

Toppino T.C., Long G.M. (2005) Top-Down and Bottom-Up Processes in the Perception of Reversible Figures: Toward a Hybrid Model. In: N. Ohta, C.M. MacLeod, & B. Uttl (Eds.), *Dynamic Cognitive Processes* (pp. 37-58). Tokyo, Japan: Springer.

Ullman S. (1984). Visual routines. *Cognition*, *18*(1-3), 97-159.

van Kerkoerle, T., Self, M. W., & Roelfsema, P. R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nature Communications*, *8*(1), 1-14.

van Lamsweerde, A. E., Beck, M. R., & Johnson, J. S. (2016). Visual working memory organization is subject to top-down control. *Psychonomic Bulletin & Review*, *23*(4), 1181-1189.

van Leeuwen, C. (2007). What needs to emerge to make you conscious?. *Journal of Consciousness Studies*, *14*(1-2), 115-136.

van Leeuwen, C. (2015a). Hierarchical stages or emergence in perceptual integration?. In: J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization*. *Oxford Handbook of Perceptual Organization* (pp. 969-988). Oxford, UK: Oxford University Press.

van Leeuwen, C. (2015b). Cortical dynamics and oscillations: What controls what we see?. In: J. Wagemans (Ed.), *Oxford Handbook of Perceptual Organization*. *Oxford Handbook of Perceptual Organization* (pp. 989-1007). Oxford, UK: Oxford University Press.

van Leeuwen, C., Buffart, H., & van der Vegt, J. (1988). Sequence influence on the organization of meaningless serial stimuli: Economy after all. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 481-502.

van Leeuwen, C., Steyvers, M., & Nooter, M. (1997). Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology*, *41*(4), 319-344.

von der Malsburg, C. (1981). The correlation theory of brain function (Departmental Technical Report No. 81-2). Göttingen, Germany: Max- Planck-Institut für Biophysical Chemistry.

Vickery, T. J. (2008). Induced perceptual grouping. *Psychological Science*, *19*(7), 693-701.

Vickery, T. J., & Jiang, Y. V. (2009). Associative grouping: Perceptual grouping of shapes by association. *Attention, Perception, & Psychophysics*, *71*(4), 896-909.

Villalba-García, C., Santaniello, G., Luna, D., Montoro, P. R., & Hinojosa, J. A. (2018). Temporal brain dynamics of the competition between proximity and shape similarity grouping cues in vision. *Neuropsychologia*, *121*, 88-97.

Vogels, R., Biederman, I., Bar, M., & Lorincz, A. (2001). Inferior temporal neurons show greater sensitivity to nonaccidental than to metric shape differences. *Journal of Cognitive Neuroscience*, *13*(4), 444-453.

Wagemans, J. (2018). Perceptual Organization. In J. T. Wixted (Series Ed.) & J. Serences (Vol. Ed.), *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience:*

Vol. 2. Sensation, Perception & Attention (pp. 803-872). Hoboken, NJ: John Wiley & Sons, Inc.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012a). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin, 138*(6), 1172-1217.

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., & van Leeuwen, C. (2012b). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin, 138*(6), 1218-1252.

Wang, M., Arteaga, D., & He, B. J. (2013). Brain mechanisms for simple perception and bistable perception. *Proceedings of the National Academy of Sciences*, *110*(35), E3350-E3359.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision*, *11*(5), 3-3.

Wei, N., Zhou, T., & Chen, L. (2018). Objective measurement of gestalts: Quantifying grouping effect by tilt aftereffect. *Behavior Research Methods, 50*(3), 963-971.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, *61*, 161-265.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II. [Investigations in Gestalt Theory: II. Laws of organization in perceptual forms]. *Psychologische Forschung, 4*, 301-350.

Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and Machine Vision* (pp. 481 -543). New York, NY: Academic Press

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, *10*(1), 80-87.

Appendix

Model's Parameters

Table 1: Units parameters.

A	1
B	1
C	20
D	33.3
E	0.05
I^{stim}	0.3
m	0.5
g	0.125
τ	0.4
F	1/150
G	1/150
U	1.25
W_{PL}^{inh}	0.4
W_{ACC}^{inh}	0.7
W_{DEC}^{inh}	5

Table 2: Sigmoid function arguments.

	Q	n
f_a	0.9	4
f_b	0.95	4
f_c	0.4125	15
f_d	0.3	4
f_e	0.95	4
f_g	0.4	20