KU LEUVEN

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

Variability in multistable apparent motion perception

Charlotte Boeykens

Proefschrift aangeboden tot het verkrijgen van de graad van Doctor in de Psychologie

Promotor: Prof. Dr. Johan Wagemans

Copromotor: Dr. Pieter Moors

Onderzoekseenheid: Brein en Cognitie

2022

Summary

Visual perception is not merely the result of stimuli being processed by passive observers. Our perceptual system actively disambiguates and interprets the information it receives from our eyes. Multistable perception of ambiguous stimuli allows to make a clear distinction between stimuli and perception. Perceptual multistability occurs when one stimulus leads to multiple different percepts in the same observer. This way, we can study how a percept emerges without the confounding factor of stimulus change. In this PhD thesis, we use ambiguous stimuli to study variability (within and between individuals) in multistable perception. In the first three empirical chapters, a 'local-global' ambiguous motion stimulus, which enables both a 'local' and 'global' apparent motion percept, is employed. In the last empirical chapter, a motion quartet stimulus, enabling apparent motion percepts in different directions, is employed.

Although individual differences have often been considered in research, many inter-individual (between individuals) and intra-individual differences (within individuals) in the general population have been neglected when they are not the result of an a priori defined manipulation or demographic group. However, these individual differences and perceptual dynamics may be very informative and are therefore a major focus in this PhD thesis. Multistable perception is an excellent test case to study this variability in depth.

A fundamental question pertains to which factors determine the variability in multistable perception. These factors are often described in terms of their "position" on the different levels of the visual system (i.e. "low-level", "mid-level", and "high-level") and different empirical chapters in this PhD thesis examine the influence of different factors along these levels.

The perception of the 'local-global' ambiguous motion stimulus is characterized by a 'shift to global' over time and the underlying factors remain unclear. Therefore, the first empirical chapter, 'Relevance study', examines the role of mere exposure vs. "higher-level" relevance (i.e., the necessity of a percept to perform an auxiliary task) on the perception of this stimulus over time. In the second empirical chapter, 'EEG study', electrophysiological correlates of the perception of the ambiguous local-global motion stimulus are examined. We measure electro-encephalography (EEG) activity during perception of the ambiguous local-global motion stimulus vs. disambiguated versions of this stimulus over time. Moreover, we compare activity during perception of the 'local' and 'global' percept. The third empirical chapter of this thesis, 'CSF study', examines the relationship between the "low-level" measure of individual luminance contrast sensitivity and the perception of the 'local-global' ambiguous motion stimulus.

The first empirical studies allow us to control for variability caused by the stimulus. However, visual perception is the result of complex interactions between observer and stimulus characteristics. Therefore, in the last empirical chapter, 'MQ study', we explore the perception of multiple variations of a motion quartet stimulus and examine the interaction of these different stimulus characteristics with individual differences.

Our understanding of visual perception is still growing and it is important to note that the terms "low-" and "high-level" may be used to imply different meanings. The discussion chapter, 'Level with me', addresses the ambiguity of the terms "low-" vs. "high-level", and aims to illustrate where the distinction between different levels can be useful vs. where it can be contra-productive for developing a better understanding of visual perception.

Samenvatting

Visuele perceptie is niet louter het resultaat van de verwerking van stimuli door passieve waarnemers. Ons perceptueel systeem disambigueert en interpreteert actief de informatie die het van onze ogen ontvangt. Multistabiele perceptie van ambigue stimuli maakt het mogelijk om een duidelijk onderscheid te maken tussen stimuli en perceptie. Perceptuele multistabiliteit treedt op wanneer één stimulus tot meerdere verschillende percepten leidt bij dezelfde waarnemer. Op deze manier kunnen we bestuderen hoe een percept tot stand komt, zonder de verstorende invloed van een stimulusverandering. In dit proefschrift gebruiken we ambigue stimuli om variabiliteit (binnen en tussen individuen) in multistabiele perceptie te bestuderen. In de eerste drie empirische hoofdstukken wordt gebruik gemaakt van een 'local-global' ambigue bewegingsstimulus die zowel een 'lokale' als een 'globale' schijnbare bewegingsperceptie mogelijk maakt. In het laatste empirische hoofdstuk wordt een ‘motion quartet’ stimulus gebruikt die schijnbare bewegingspercepties in verschillende richtingen mogelijk maakt.

Hoewel in onderzoek vaak rekening wordt gehouden met individuele verschillen, zijn veel inter-individuele (tussen individuen) en intra-individuele verschillen (binnen individuen) in de algemene bevolking verwaarloosd wanneer zij niet het resultaat zijn van een a priori gedefinieerde manipulatie of demografische groep. Deze individuele verschillen en perceptuele dynamieken kunnen echter zeer informatief zijn en vormen daarom een belangrijk aandachtspunt in deze doctoraatsthesis. Multistabiele perceptie is een uitstekende test case om deze variabiliteit in de diepte te bestuderen.

Een fundamentele vraag is welke factoren variabiliteit in multistabiele perceptie bepalen. Deze factoren worden vaak beschreven in relatie tot hun "positie" op de verschillende niveaus van het visuele systeem (d.w.z. “low-level”, “mid-level”, en “high-level”) en verschillende empirische hoofdstukken in dit proefschrift onderzoeken de invloed van verschillende factoren op deze niveaus.

De perceptie van de 'local-global' stimulus wordt gekarakteriseerd door een ‘shift to global’ overheen de tijd waarvan de onderliggende factoren onduidelijk blijven. Daarom onderzoekt het eerste empirische hoofdstuk, ‘Relevance study’, de rol van loutere blootstelling versus “higher-level” relevantie (d.w.z. de noodzaak van het percept bij het uitoefenen van een taak) op de perceptie van deze stimulus overheen de tijd. In het tweede empirische hoofdstuk, 'EEG studie', worden elektrofysiologische correlaten van de perceptie van de 'local-global' stimulus onderzocht. We meten electro-encephalografie (EEG) activiteit tijdens de perceptie van de ambigue 'local-global' stimulus versus gedisambigueerde versies ervan overheen de tijd. Bovendien vergelijken we de activiteit voor het 'lokale' en 'globale' percept. Het derde empirische hoofdstuk, 'CSF study', onderzoekt de relatie tussen de “low-level” meting van individuele luminantie contrastgevoeligheid en de perceptie van de 'local-global' stimulus.

De eerste empirische studies stellen ons in staat om variabiliteit veroorzaakt door de stimulus te beperken. Visuele perceptie is echter het resultaat van complexe interacties tussen waarnemer en stimulus karakteristieken. Daarom onderzoeken we in het laatste empirische hoofdstuk, 'MQ study', de perceptie van meerdere variaties van een ‘motion quartet’ stimulus en de interactie van deze verschillende stimuluskarakteristieken met individuele verschillen.

Ons begrip van visuele perceptie groeit nog steeds en het is belangrijk om op te merken dat de termen “low-“ en “high-level” verschillende betekenissen kunnen hebben. Het discussie hoofdstuk, 'Level with me', gaat in op de ambiguïteit van de termen “low-“ versus “high-level”, en heeft als doel te illustreren waar het onderscheid tussen verschillende niveaus nuttig kan zijn versus waar het contra-productief kan zijn voor het ontwikkelen van een beter begrip van visuele perceptie.

Dankwoord

Ik heb mijn studies in Gent gedaan, waar de faculteit psychologie vooral gespecialiseerd is in onderzoek over taal. Doorheen die jaren - na een Bachelor topic te kiezen over de Navon taak, een stage over visueel werkgeheugen en een masterscriptie over spatiële cognitie en gezichtsherkenning - besefte ik dat mijn grootste interesse toch eerder bij visuele perceptie lag. Gelukkig was er ook de universiteit in Leuven, waar net wél vaak de focus bij visuele perceptie bleek te liggen. Ik heb gesolliciteerd bij Johan en ben erna officieel een overloper geworden. Bedankt Johan voor deze kans, om je vertrouwen in mij te leggen, om de toegewijde wetenschapper te zijn die je bent, en om ondanks een enorm drukke agenda altijd klaar te staan om advies te geven wanneer ik het nodig had, hoe laat ik er soms ook om vroeg.

Ik had aan het begin nog veel te leren over visuele perceptie en uiteraard ook over onderzoek in het algemeen. Niemand heeft me daarin zoveel gesteund en me dat proces van zo dichtbij zien meemaken als mijn copromotor. Pieter, aan het begin zei je dat er nog veel ups en downs zouden komen tijdens een doctoraat en de mijne heb je telkens goed weten te modereren. Wanneer ik naïef en laaiend enthousiast was over onderzoek stak jouw cynisme de kop op. Wanneer de nihilist in mij naar boven kwam en ik bedacht: “hoe meer ik weet, hoe minder ik weet”, wist je me ervan te overtuigen dat dat helemaal niet zo erg is. Bedankt voor alle waardevolle raad.

I would also like to thank Cees van Leeuwen, Andrey Nikolaev, and Jürgen Kornmeier for the useful scientific advice along various stages of the work in this PhD.

De voorbije jaren hebben mij gevormd tot wie ik vandaag ben, en daarvoor ben ik ook de vele collega’s dankbaar die daar deel van hebben uitgemaakt. Agna, Eline, Lisa, Hanne, Lore, Nathalie, Jaana, Christophe, Rudy, Kevin, en zo veel anderen, merci om altijd klaar te staan voor een leuke babbel en goede raad. To my dear colleagues living nearby in Brussels, Thiago, Claudia, Laurie-Anne, Chris, thank you for the fun times, the scientific advice and emotional support.

Ik moet toegeven dat ik soms heb getwijfeld of het me wel zou lukken om het proefschrift af te ronden. Dat ik er toch in geslaagd ben, heb ik te danken aan een collectie onwaarschijnlijk mooie mensen in mijn omgeving.

Ten eerste wil ik twee vrienden bedanken die me perfect konden bijstaan in de highs en lows van zowel het academische als het privéleven, dankjewel Jelle en dankjewel Sofie.

Dankjewel Véronique voor het plezier van een ideale woon/werkomgeving die me soms zelfs nostalgisch kan maken naar tijden van een covid-lockdown.

Alle lieve vrienden die mee wouden vieren wanneer ik een stap vooruit zette en waar ik mijn hart bij kon luchten wanneer ik niet meer vooruit geraakte, dankjewel Kim, Eliane, Arnaud, Simon, Hanne, Emmy, Hannah, Alicia, Laura, Maud en zoveel anderen.

Mijn familie verdient het meeste lof voor alle onvoorwaardelijke steun. Mama, papa, Steffi, Vitoria, Nandje, Fil, Erika, Herman, dankjewel om mijn grootste supporters en rotsen in de branding te zijn.

Tot slot, Jonas, je bent een held, dankjewel om me zo ongelooflijk te steunen en mij zelfs letterlijk met voedsel in leven te houden tijdens het schrijven van mijn thesis. Zonder jou had deze thesis er niet gelegen. Of toch zeker niet met een cover errond.

Leuven, 22 juni 2022.

Table of contents

p.5 1. General Introduction

p.5 1.1 Visual perception & Perceptual organization

p.7 1.2 Motion perception & Apparent motion perception

p.10 1.3 Multistable perception of ambiguous stimuli

p.20 1.4 Variability in multistable perception: individual differences

p.21 1.5 Origins of variability in perception

p.24 1.6 Factors determining variability in multistable perception

p.31 2. Relevance study

p.32 2.1 Introduction

p.36 2.2 Methods

p.41 2.3 Results

p.44 2.4 Discussion

p.47 2.5 Conclusion

p.53 3. EEG study

p.53 3.1 Part 1: Perception of the ambiguous vs. disambiguated 'local-global' stimulus over time

p.53 3.1.1 Introduction

p.57 3.1.2 Methods

p.65 3.1.3 Behavioral results and discussion

p.67 3.1.4 EEG results

p.75 3.1.5 Interim discussion

p.78 3.2 Part 2: Local vs. global perception of the ambiguous vs. disambiguated 'local-global' stimulus

p.78 3.2.1 Introduction

p.84 3.2.2 EEG results

p.92 3.2.3 Discussion

p.95 3.3 General conclusion

p.101 4. CSF study

p.101 4.1 Introduction

p.106 4.2 General Methods

p.108 4.3 Experiment 1

p.110 4.4 Experiment 2

p.116 4.5 General Discussion

p.125 5. Motion quartet study

p.126 5.1 Introduction

p.126 5.1.1 (Meta-)theoretical context

p.129 5.1.2 Empirical context

p.133 5.2 Spatial Scale Experiment

p.133 5.2.1 Methods

p.135 5.2.2 Results and discussion

p.141 5.3 Spatiotemporal Scale Experiment

p.141 5.3.1 Methods

p.142 5.3.2 Results and discussion

p.147 5.4 Spatial Location Experiment

p.148 5.4.1 Methods

p.149 5.4.2 Results and discussion

p.153 5.5 Sequential effects of percepts

p.154 5.5.1 Methods

p.157 5.5.2 Results and discussion

p.164 5.6 Conclusions

p.171 6. General Discussion

p.171 6.1 Summary and general conclusions

p.178 6.2 Level with me: A re-evaluation of the use of “low-“ and “high-level” in visual perception.

p.193 6.3 Some (meta-) theoretical observations and speculations

p.196 6.4 Looking back and looking forward

1. General Introduction

1.1 Visual perception and perceptual organization

Visual perception is the process of interpreting the environment by means of information available through our eyes. When we are looking at the world, it seems only natural that the world in our mind's eye is reflecting what is "out there" in the external world with the amount of depth and detail as we experience it with. However, our visual experience, or percept, is the result of a highly complex process and, in contrast to this coherent, meaningful percept, our visual system has a rather messy, obscure input to start from (i.e., light captured by our retinae and transformed into neuronal signals). Visual perception is not merely the result of stimuli being processed by passive observers. Our perceptual system actively disambiguates and interprets the information it receives from our eyes.

Of crucial importance for this interpretation is perceptual organization, the process of giving structure to the information provided by our eyes. The concept of perceptual organization was introduced in Gestalt psychology and is considered to be a "mid-level" process, where basic visual features (i.e., "low-level") are organized into coherent wholes, which are employed in further processing (i.e., "high-level") (Wagemans, 2018). Nevertheless, in principle, perceptual organization can occur at multiple levels and at each stage of visual processing, including feedback from further levels (Wagemans et al., 2012a). Moreover, "mid-level" characteristics alone are not sufficient to explain the entirety of perception and perceptual organization.

Perceptual organization implies a differentiation in 'globality' between 'local' and 'global' as it describes the structuring or grouping of 'local' parts into a relevant 'global' whole, or Gestalt. 'Globality' is a relative concept, implying that something can only be local with respect to something more global, and vice versa. For instance, a word on this page might be local with respect to the page, but it may be global with respect to a letter in the word. The local parts of a global whole are sometimes assumed to be singular (physical) entities when in fact they can be wholes (or Gestalts) as well. The local parts of a word on this page are made up from multiple letters, which in their turn consist of the perceptual grouping of multiple lines, which

in their turn consist of multiple points, etc. Therefore, we consider 'local' to be the *relative* parts and 'global' to be the *relative* whole in a hierarchical relationship.

The critical difference in globality is not necessarily spatial, but is most often intertwined with this. Yet, it is still not straightforward to define the critical difference between local and global. There is a wide range of stimuli and tasks that are assumed to be operationalizations of local vs. global processing (de Wit & Wagemans, 2015; Chamberlain et al., 2017). However, these stimuli and tasks may actually engage very different underlying processes, leading to inconsistency in the literature regarding the essence and implications of the terms local and global. The chapter on the 'EEG study' and the 'Level with me' part in the General Discussion focus on the difference between 'local' and 'global' in more depth.

In the beginning of the 20th century, the empiricist approach of structuralism (the dominant school of psychology at that time) studied perception in terms of additive associations of observable sensory units (analogous to elements in physics). In contrast, the Gestalt psychologist Wertheimer (1912) advocated for an approach to the study of perception which employed descriptions of conscious experience. He believed that sensory units are not part of our mental life and should therefore not be (exclusively) examined in order to advance our understanding of our mental life. Rather, our phenomenological, conscious experience is in varying degrees structured into wholes. These wholes, or Gestalts, are dominant in our perception, without the need to *perceive* the local elements they are made up from (Wertheimer, 1912; 1922; 1923). According to Gestalt psychology, our visual system arrives at the perception of such wholes by means of dynamic interactive processes, contrasting the linear approach of the empiricists at the time.

In Gestalt psychology, "perceptual organization" referred to all the ways in which a stimulus could be organized in our phenomenal experience, or conscious perception of visual stimuli. Therefore, 'local' and 'global' were initially defined in terms of characteristics of our conscious percept. However, when we do not exclusively consider our phenomenological experience, we can see that the process of structuring local elements into global wholes is omnipresent in the visual system, starting with the increasing receptive field sizes of neurons in visual cortex and extending to the interaction of visual information with multiple other information sources into functional networks that aid us in complex tasks.

Many of the original studies in the Gestalt tradition have failed to thoroughly specify or provide an explanation for their descriptive principles of perception. Nevertheless, their insights into our phenomenal experience and their conceptualization of perception as nonlinear dynamics retain their importance to this day. As our experimental methodologies advance, we start to uncover empirical evidence for some of their ideas and uncover that some seemingly contrasting ideas from the Gestaltist and empiricist approach might partially be united. For instance, Wertheimer (1923) stated that the most fundamental principle of perceptual organization is the law of Prägnanz, which implies that our percept will take on the simplest and most encompassing, good structure, or Gestalt, as possible (2012a). The principle of simplicity was originally contrasted with an approach based on likelihood (von Helmholz, 1867), which states that we tend to perceive what we have learned to be the most likely interpretation of a stimulus. However, more recent considerations of the simplicity principle have argued that the implications of the descriptive concepts of simplicity and likelihood can at least partially be reconciled in explanatory accounts of perception (2012a).

We believe that Gestalt psychology should be integrated into our contemporary approach in vision science in order to address the vital question (first formulated by Koffka, 1922): Why do things look as they do? Accordingly, the work in this PhD thesis has addressed research questions with a contemporary scientific approach as well as considerations from the Gestaltist perspective of perceptual organization.

1.2 Motion perception and apparent motion

A crucial feature of visual perception is the ability to experience stability and change. In order to successfully interact with our continuously evolving external world, we need to categorize what is stable in this world and successfully detect any changes that may be of interest to us. Consider the perception of motion, or a change in position over time. In traditional Gestalt psychology (Wertheimer, 1923), basic grouping principles describing the structure in the phenomenology of perceived wholes were defined. The Gestalt grouping principle of common fate (Wertheimer, 1923), later expanded to generalized common fate and synchrony (Alais, Blake & Lee, 1998), describes that parts which move in synchrony, or change simultaneously, tend to be grouped together in our perceptual experience.

Perceiving motion may seem straightforward to us, but is a critical challenge to our visual system. A moving object will result in an image with different positions on our retinae over time and space. This ambiguous information is then employed in further processing by our visual system. Determining which parts of a static image belong to one object, or "whole", already requires a substantial amount of visual processing. Determining which parts of multiple subsequent images belong to the same object that is changing its position over time adds another level of complexity. Moreover, in the absence of any movement in the external world, the images on our retinae will also change their position when we move around, or even when we only make small eye movements, or saccades. Therefore, our visual system should also succeed in determining whether a change in position on our retinae is due to a change in position of our eyes or a change in position of something in the external world. Importantly, the processing of static information from multiple points in time can give rise to motion perception, but motion processing in turn is also important to facilitate static image processing or perceptual organization. For instance, you might be able to distinguish an animal that is perfectly camouflaged in its environment only when it is moving around with respect to a static background.

Our visual system has mastered the task of motion perception almost perfectly, but there are instances where our percept does not match reality, or rather, where we can get a glimpse of the complexity of the processes required to achieve this seemingly perfect perception. For instance, the image on the cover of this thesis portrays the perception of motion in a static image. Obviously, no real motion is present in this image, yet our perceptual experience suggests that there is. The image is based on the peripheral drift illusion and research suggests that the combination of (micro-) saccades and differences in luminance causing temporal differences in motion processing are at the basis of the illusion (Faubert & Herbert, 1999; Kitaoka & Ashida, 2003).

When motion is perceived even though there is no real motion present in the external world, we refer to the phenomenon of apparent motion. In 1912, Max Wertheimer published a paper on the phenomenon of phi motion, a specific form of apparent motion (i.e., "pure" motion), which is recognized as the start of the Berlin school of Gestalt psychology (Wagemans et al., 2012). Closely related to this, beta apparent motion is the visual phenomenon where two separate stimuli presented in two successive static images, with or without a certain temporal

delay, are perceived as if they are one stimulus moving across space over time. Again, no real motion is occurring. Nevertheless, our visual system interprets this information as if there is real motion.

A good example of apparent motion is provided by the motion quartet stimulus (Figure 1). In the first frame, a dot is displayed in the top left corner and another in the bottom right corner. In the second frame, a dot is presented in the top right corner and another in the bottom left corner. When the two frames are presented successively, given certain temporal and spatial properties, observers experience a percept of motion even though they were presented with two static images.

We chose to work with apparent motion stimuli for this PhD thesis because they allow for both simple, well-controllable paradigms and the application to more complex interpretations and ecologically relevant stimuli (e.g., biological motion). Moreover, they enable a comparison between static and dynamic perceptual organization. Importantly, for this PhD thesis, the intention is not to fully understand motion perception. Rather, apparent motion is employed as a means to explore visual perception and its variability.

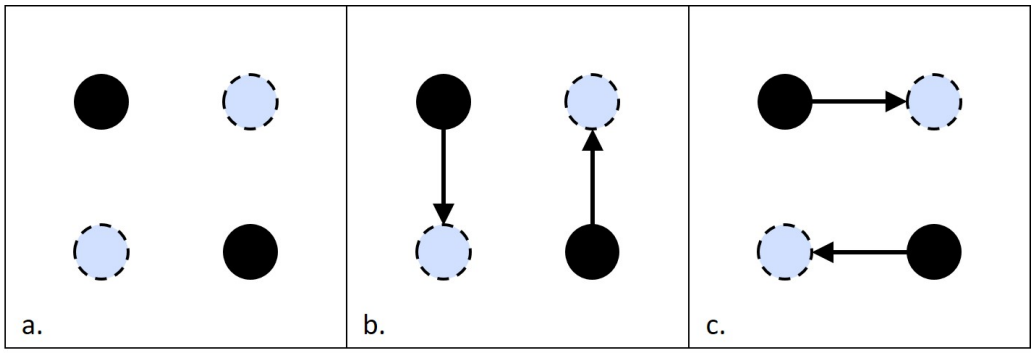


Figure 1. Motion Quartet. a: The ambiguous motion quartet (MQ) is depicted with the first frame in black and the subsequent frame in dashed lines. In an alternating presentation of these two frames, apparent motion can be perceived. **b:** Vertical motion percept of the ambiguous MQ. **c:** Horizontal motion percept of the ambiguous MQ. **d:** Ambiguous MQ in Spatial Scale and Spatiotemporal Scale experiment. **e:** Ambiguous MQ in Spatial Location experiment. An animated version can be found by following this link: https://osf.io/2yktz/?view_only=ef697c56758547858ca4d02c63b62013.

1.3 Multistable perception of ambiguous stimuli

Some perceptual phenomena may occur quasi-independently of the observer; it seems as if the stimulus ultimately determines the perceptual experience in all observers. Gestalt 'grouping principles' describe how we may predict perception by means of stimulus characteristics (e.g., grouping by proximity, grouping by similarity, grouping by common fate, etc.) (Wertheimer, 1923/2012; Wagemans et al., 2012). The traditional Gestalt 'grouping principles' are mainly descriptive, and their explanatory role is still unsatisfactory (Spillmann, 2009). Even though we can predict perceptual phenomena by means of stimulus characteristics in general, there is no such thing as a single relationship between a stimulus and its percept. As mentioned above, visual perception is not merely the result of stimuli being processed by passive observers; Our perceptual system disambiguates the information it receives from our eyes. In some instances, our percept does not seem to match reality and we can get some insight into the complexity of the processes required to enable visual perception.

Multistable perception of ambiguous stimuli allows to make a clear distinction between stimuli and perception. Perceptual multistability occurs when one and the same stimulus enables multiple different percepts. Schwarz et al. (2012) described the key features of multistability as follows: "(i) stimuli have more than one plausible perceptual organization; (ii) these organizations are not compatible with each other". Identical stimuli are perceived in substantially different ways between different observers or from one moment to the next within single observers (Kornmeier & Bach, 2012; Brascamp, Sterzer, Blake, & Knapen, 2018). Accordingly, ambiguous stimuli provide an interesting tool to examine perceptual organization. For instance, a Necker cube can be perceived both facing the upper right corner and facing the lower left corner, even though the physical image remains the same (Figure 2). Incidentally, the motion quartet stimulus is also a good example to demonstrate perceptual multistability. The apparent motion in the motion quartet may be perceived in two (or more) different ways. More specifically, one may perceive a horizontal (or clockwise) motion of the dots, or, one may perceive a vertical (or counter-clockwise) motion of the dots.

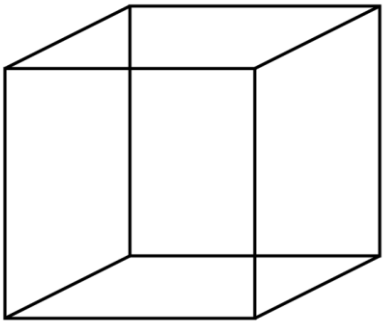


Figure 2. Ambiguous Necker cube.

Ambiguous stimuli have been widely employed throughout the history of vision science. The most promising quality of these stimuli is the fact that they enable us to study the formation of or a change in a percept without the confounding factor of a change in the stimulus. In Gestalt psychology, ambiguous figures were a compelling tool to demonstrate the dynamic, nonlinear nature of the processes underlying what we perceive. Brascamp et al. (2018) dissociate between two types of theoretical models on multistable perception: The dynamical systems account and the Bayesian principle of predictive coding. Although both types of models have important differences, they appear to agree that multistable perception is associated with rapid neural changes at the time of transition and gradual changes during stable periods of perception.

The majority of contemporary research on multistable perception has included neuronal correlates from functional magnetic resonance imaging (fMRI) measures. Whereas earlier explanations of multistable perception focused mostly on the processes within sensory areas, empirical findings of activity in “higher-level” brain areas (i.e., frontoparietal network) inspired later explanations to incorporate non-sensory and cognitive processes into their models of multistable perception (Leopold & Logothetis, 1999). The degree to which different levels of the visual system are causally involved in multistable perception and initiating perceptual switches is still debated to this day (Brascamp et al., 2018). However, as was also already pointed out by Long and Toppino (2004), most vision scientists will agree on the fact that a satisfactory account of perceptual multistability will involve a hybrid conceptualization of a series of interacting sensory (bottom-up) and cognitive (top-down) processes (Blake & Logothetis, 2002; Sterzer, Kleinschmidt, & Rees, 2009; Brascamp et al., 2018). As most research on the topic of perceptual multistability implicates different levels in a visual

hierarchy, it is important to point out the potential difficulties associated with multiple interpretations of the concepts of "low-" and "high-level", as will be discussed in the General Discussion.

The volume of research on multistable perception incorporating electro-encephalography (EEG) measures appears to be less extensive. Kornmeier and Bach (2012) reviewed available studies and proposed an integrative theory of multistable perception and its neural correlates in EEG measurements. In the chapter 'EEG study', EEG correlates of multistable perception (in particular, event-related potentials or ERP) will be discussed in more depth.

Irrespective of the methods employed, the main motivation for this field of research is the conviction that we can develop a better understanding of perception in general by means of studying perception of ambiguous stimuli. This way, we can study how a percept emerges, without the confounding factor of stimulus change, and extrapolate this knowledge about the perception of ambiguous stimuli to perception in general. However, one may ask whether multistable perception of ambiguous stimuli is a good proxy for perception in general. In a recent review on multistable perception, Brascamp and Shevell (2021) argued that implicit ambiguity in visual neural representations is pervasive and that resolving ambiguity is a fundamental process of vision. This implies that the processes we study in relation to ambiguous stimuli are indeed very closely related, if not identical, to processes of perception in general.

It is important to note that there is a crucial difference between stimuli that result in ambiguity because they enable different, but clear percepts (i.e., multistable perception), and stimuli that result in ambiguity because they do not enable any clear percept (i.e., what can be characterized as noisy stimuli). In the following, we will not engage in an analysis of the latter type of ambiguity, and will therefore refer to ambiguity that results in multistability (i.e., the stimulus is ambiguous and gives rise to multiple, clear percepts) as ambiguity, and refer to the absence of multistability (i.e., only one, clear percept is possible) as unambiguity in perception.

Let's assume, for the sake of argument, that indeed *every* stimulus gives rise to a retinal image which could be related to a multitude of different external physical "sources" and also gives rise to ambiguity in the neural representation and interpretation. We would then face the fact that, in contrast to ambiguity in neural representations, not every stimulus gives rise to

ambiguity in perception or variability in observable behavior in relation to perception. In fact, in order to successfully interact with the external world, the majority of stimuli should not give rise to variability in perception and associated behavior. This implies that, for these stimuli, somewhere along the transition from neural representations to perception and observable behavior in relation to perception, there is a change from ambiguity to unambiguity. This change to unambiguity may occur at different levels along this transition from neural representations to perception and behavior. In contrast, ambiguous stimuli would, along the transition from neural representations to perception and observable behavior in relation to perception, remain ambiguous.

A first type of ambiguous stimuli are stimuli that evoke intra-individual variability in the form of alternating percepts and are typically studied with stimulus characteristics at an observer's point of subjective equality (PSE). A defining aspect of ambiguous stimuli at the PSE is the fact that, just as the assumed ambiguity in neural representations, the perception and behavior associated to perception is ambiguous or variable as well.

A second type of ambiguous stimuli are stimuli that evoke inter-individual variability in percepts or intra-individual variability in the form of robust changes in percepts over time, which are typically studied with stimulus characteristics for which an observer has an unambiguous percept. In contrast to the first type of ambiguous stimuli, the behavior associated to the perception of ambiguous stimuli which are studied at unambiguous percepts is unambiguous, unlike the assumed ambiguity in some of the underlying neural representations. However, importantly, those unambiguous percepts and associated unambiguous behaviors are different across observers or within an observer over time.

In the following paragraphs, we aim to show that, despite the potential universality of ambiguity in neural representations, a dissociation between two types of ambiguous stimuli might still be valuable to enable a comparison of perception of ambiguous stimuli to perception in general. The most important difference between ambiguous and unambiguous stimuli is the amount of inter- or intra-individual variability associated with its perception. Depending on how this variability manifests, the perception of ambiguous stimuli can be compared to the perception of unambiguous stimuli in different ways. In order to make more precise comparisons of the perception of ambiguous and unambiguous stimuli, we would like

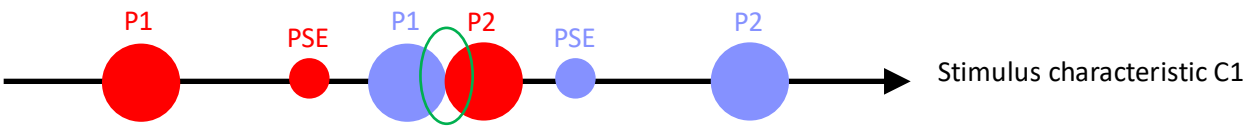
to distinguish different types ambiguous stimuli and their potential relationships to unambiguous stimuli in more depth.

For each stimulus **Si** (with stimulus characteristics **C(Si)** = C1, …, Cx) and observer **Oi** at a particular moment in time **Ti**, there are the corresponding potential percepts **P(Si,Oi,Ti)** = P1, …, Pn. For n > 1 and dependent on stimulus characteristics C(Si), there are point(s) of subjective equality (**PSE**) between those percepts. Furthermore, each of those potential percepts *may* correspond to a particular detectable behavior **B(P)**. The ambiguous stimuli employed in multistable perception can be categorized according to these concepts.

For instance, at time T1 an observer O1 is presented with the motion quartet stimulus S1, which has the stimulus characteristic C1 of vertical distance (i.e., the vertical distance between one dot in the first frame and another in the second frame). The two corresponding potential percepts P(S1,O1,T1) are the vertical apparent motion percept P1 and the horizontal apparent motion percept P2. P1 and P2 both correspond to a detectable behavior which could be, for instance, verbal identification of the apparent motion direction. Depending on the magnitude of C1, observer O1 will be more likely to experience P1 or P2. More specifically, the smaller the vertical distance compared to the horizontal distance, the more likely an observer is to experience vertical apparent motion. The PSE is the specific magnitude of C1 for which O1 will be equally likely to experience P1 and P2.

Another observer O2 at the same time T1 might perceive the same stimulus differently. For instance, the observer might have a more horizontal percept compared to O1. Therefore, the PSE for O2 might be located at a different magnitude of C1. The same situation may be applicable to the first observer O1, but at a different time T2, after the observer has experienced a shift in perception over time.

Then, consider a stimulus with characteristic C1 = Y units. Either a stimulus with stimulus characteristic C1 = PSE or C1 = Y will most likely be the stimulus characteristic that is used for experimentation with ambiguous stimuli, as it provides the highest variability for an observer at one point in time or across observers / within observers across time, respectively.



In the following visualizations, **PSE** represents a stimulus for which both percepts are equally probable and **P** represents a stimulus for which perception is unambiguous. Red colors depict one observer at a particular point in time. Blue colors depict another observer at the same point in time (to illustrate inter-individual variability), or the same observer at a different point in time (to illustrate intra-individual variability in the form of stable changes in percepts over time). Green circles point at the stimulus characteristic area of interest where the stimulus and the observer's perception of it is considered when we discuss the multistability or ambiguity. We will limit these abstract examples to two perceptual alternatives maximum: P(S,O,T) = P1,P2. The points depicted along these axis relate to the ambiguity associated to different perceptual states, and do not intend to represent the ambiguity of pre-perceptual neural processes/representations. A striped fill of the colors representing perceptual states indicates that it is unclear whether both percepts can still somehow be represented along one stimulus characteristic for one observer at one point in time (i.e., unclear whether this observer can change percepts if we change the stimulus characteristic).

Different types of ambiguous stimuli and their potential relationships to unambiguous stimuli:

1) **Ambiguous stimuli that typically evoke intra-individual variability in the form of alternating percepts.**

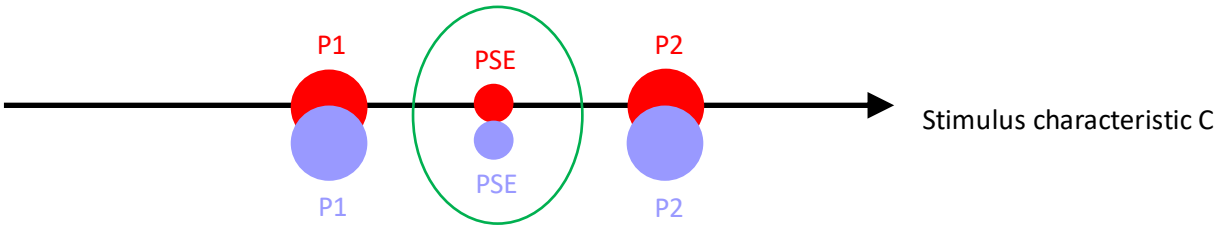
   These stimuli will enable the same multiple percepts P for different observers O and for different times T. Moreover, the possible percepts P will correspond to different detectable behaviors B(P) = B1,B2.

   The studies employing these kinds of stimuli will focus on a range of stimulus characteristics for which there is the most ambiguity around the PSE. This implies that, compared to the perception of unambiguous stimuli, there might be additional processes associated to the competition between different perceptual representations. As this process of competition between different perceptual representations should take place before the final percept is established, any subsequent processes will most likely be highly similar to the perception of unambiguous stimuli. In accordance with this, Kornmeier and Bach (2012) found that EEG ERP neural correlates of perception of ambiguous stimuli compared to

disambiguated versions is highly similar aside from a few differences in early components and a shift in timing of later components.

Moreover, the "position" of the PSE along stimulus characteristic C is interconnected with the "position" of either stimulus that enables unambiguous perception. Therefore, any finding about perception around the PSE (e.g., intra- or inter-individual variability in the perception of this stimulus) will not be independent of perception of either stimulus that enables unambiguous perception associated to this PSE.

Note that, in principle any unambiguous stimulus can be put on a similar scale as the one displayed here, as there is always a stimulus characteristic along which we can manipulate a stimulus until the percept changes into another. For instance, consider changing a square along the stimulus characteristic of 'roundedness' until it becomes a circle. However, the crucial difference between an unambiguous stimulus and an ambiguous stimulus in this scenario is the fact that the stimulus is employed around the PSE, or in its ambiguous form.



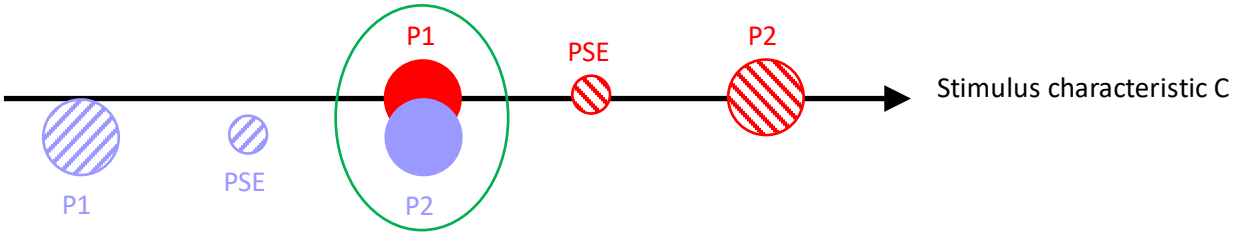
Example: Necker cube: We know both potential percepts exist and we are all able to switch between these different percepts relatively easily for the same stimulus characteristics. Extended viewing will evoke spontaneous alternations in percepts.

2) **Ambiguous stimuli that typically evoke inter-individual variability in percepts or intra-individual variability in the form of stable changes in percepts over time.**

These stimuli will enable one percept P(n=1) which is qualitatively different between different observers O or within one observer over time T. Moreover, different percepts P will correspond to detectable different behaviors B(P) = B1,B2.

Studies employing these stimuli will most likely focus on differing, but unambiguous perceptual states. As there are no additional processes associated to the competition between different perceptual representations, the main processes in the formation of a percept should be similar to unambiguous perception. This implies that we do not

expect to find any clear differences in neural correlates of perception between these kinds of ambiguous stimuli and perception in general. However, any differences in the qualities of the actual percepts may have an additional impact aside from differences due the formation of a percept.

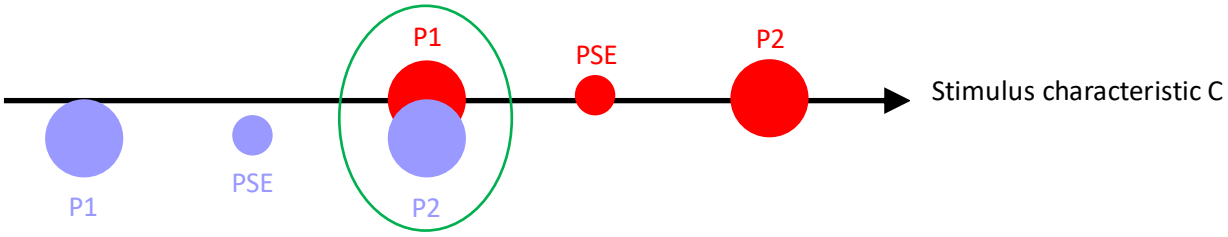


Example of inter-individual variability: The dress: We know that there are two different percepts across different observers. However, it is unclear whether the observers have the potential to experience the other percept when a particular stimulus characteristic is manipulated. Extended viewing will typically not evoke spontaneous alternations in percepts.

Example of intra-individual variability in the form of stable changes in percepts over time: Mooney images: When you see a Mooney image for the first time, you perceive a meaningless formation of shapes. However, once you have perceived the meaningful shape embedded in it, you are unable to "un-see" this. It is unclear whether both percepts can still somehow be represented along one stimulus characteristic. Extended viewing will typically not evoke spontaneous alternations in percepts.

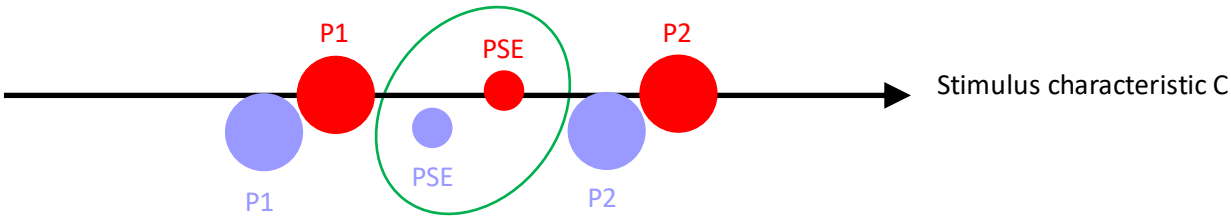
3) A scenario that is a combination of possibilities 1) and 2)

Implications for comparison to unambiguous stimuli depend on whether you are focusing on the PSE or the stimulus that enables unambiguous perception: see 1) and 2).



Example: local-global ambiguous motion stimulus. You know you can access both percepts if you change a stimulus characteristic C (e.g., between dot-pair distance). However, for the same stimulus characteristics, you are more likely to see either the

local or global percept compared to another individual at the same point in time, and the same holds true at one point in time compared to another point in time (due to the 'shift to global').



Example of more subtle difference: perception of the 'local-global' ambiguous motion stimulus that is employed in the first chapters of this PhD thesis.

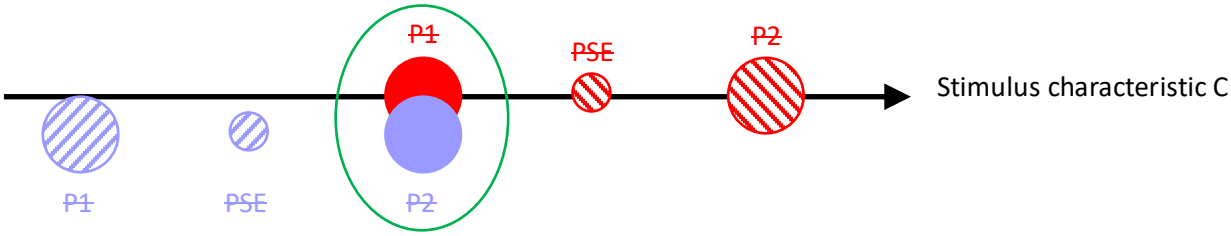
4) Undiagnosed ambiguous stimuli

Undiagnosed ambiguous stimuli can be either of the above three scenarios, but with the important difference that different percepts do not correspond to detectable different associated behaviors: B(P) = B1.

This fourth scenario implies that, in principle, there is no way for us to know whether a stimulus is ambiguous or unambiguous.

If there are no differences in detectable behavior between differing percepts, there may be hidden processes related to competition between different perceptual representations, or hidden processes related to differing perceptual states.

Nevertheless, having ambiguity in perceptual representations without having ambiguity in associated behaviors would be highly inefficient from an evolutionary perspective. We mainly include this theoretical possibility here because it demonstrates that multistability merely becomes noticeable to us for certain stimuli when the different percepts evoke an observable, differing behavior.



Example: Vannuscorps et al. (2021) reported a case study of a woman who perceives two-dimensional shapes as if they were rotated by 90, 180 degrees around their

center, mirrored across their own axes, or both. However, even though this woman had more difficulties with reading and some other perceptual abilities than average throughout her life, she was skillful at many other abilities involving visual perception (e.g., soccer and basketball) and her peculiar perception was identified only when she was already 15 years old.

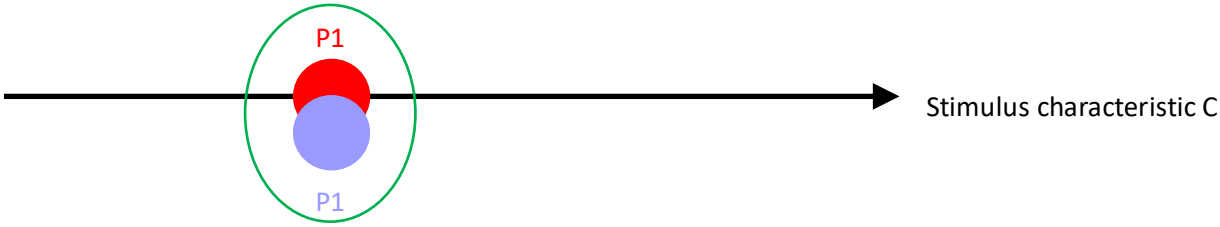
5) Unambiguous stimuli

Unambiguous stimuli evoke neither intra- or inter-individual variability in percepts.

These stimuli will enable one percept P(n=1) which is (close to) identical across observers O and time T. B(P) = B1.

Assumed to be "perception in general".

Note that, for any of these axes along a stimulus characteristic C, there exist a multitude of potential other axes along which any stimulus can change to enable multiple percepts and become ambiguous at some point.



We may conclude that the comparability of multistable perception to unambiguous or "normal" perception depends on the kind of multistability that is studied. An important consideration to be made is whether a particular ambiguous stimulus is ambiguous within one observer (scenario 1) vs. across observers or for an observer over time (scenario 2) vs. the combination of scenario 1 and 2. This will determine how certain findings in behavior or neural correlates associated to the perception of ambiguous stimuli can be extrapolated to perception in general. However, there is no doubt that a substantial set of the processes governing perception in both ambiguous and unambiguous stimuli are highly similar. At least, we can examine similarities between processes governing perception of ambiguous as well as unambiguous stimuli. Accordingly, we can exclude processes that are either purely associated to ambiguity or purely associated to stimulus differences and derive information about perceptual processes which should be common between these two forms of perception. In this PhD thesis, we use ambiguous stimuli to study variability in multistable perception. In the

first three empirical chapters, a 'local-global' ambiguous motion stimulus, which enables both a 'local' and 'global' apparent motion percept, is employed. In the last empirical chapter, a motion quartet stimulus, enabling apparent motion percepts in different directions, is employed.

1.4 Variability in multistable perception: individual differences

Already in 1957, Cronbach argued that the hypothesis testing of experimental psychology needed to be combined with research on individual differences. As we saw in the previous section, individual differences are a logical consequence of multistable perception. Multistable perception of ambiguous stimuli shows us that there is no such thing as a direct relationship between stimulus and percept. Perception is the result of a highly complex interaction of stimulus and observer properties, resulting in inter- (differences between individuals) and intra- (differences within individuals) individual variability in perception, and their interaction.

Most existing research in inter-individual differences in visual perception focuses on differences between individuals which are grouped based on a clinical diagnosis or a demographic feature. For instance, a vast amount of research looks into differences in visual perception between neurotypically developing individuals and individuals on the autism spectrum (Kaiser & Shiffrar, 2009; Van der Hallen et al., 2019), differences in visual perception between men and women (Vanston & Strother, 2017), etc. It is sensible to generate groups of individuals based on features other than visual perception as this enables researchers to correlate any differences in visual perception to a priori defined groups. Consequently, differences in visual perception between individuals which cannot be categorized into a priori groups are often ignored.

Inter-individual differences in visual perception in the general population, or within the groups that are employed to categorize different observers, are not considered as much. However, these individual differences may be very informative for our understanding of visual perception (Mollon, Bosten, Peterzell, & Webster, 2017). Kleinschmidt, Sterzer and Rees (2012) considered individual differences in multistable perception and stated that "perceptual

fluctuation rates are consistent for a given individual but vary considerably between individuals. There has been some evidence for a genetic basis and […] neuroanatomical substrates for individual variability". Moreover, neglecting the individual variability in visual perception can sometimes lead to conclusions which are based on the statistically average observer, but unrepresentative for any one observer in particular (Fisher, Medaglia, & Jeronimus, 2018). In this PhD thesis, inter-individual differences will be considered in the 'CSF study' chapter and more extensively in the 'MQ study' chapter.

Most existing research in intra-individual differences in visual perception has focused on developmental differences or perceptual learning. The former describe general perceptual differences across the lifespan of an individual, occurring for most or all individuals (Faubert, 2002). The latter describes changes in visual perception for one specific perceptual task training resulting in an increased performance in this task (Sagi, 2011). However, intra-individual differences in visual perception which are unrelated to general development, or differences which are related to general changes in appearance, rather than performance in one task, are not considered as much.

Some individuals may change in their perception over time without a clear developmental phase or clear task-related goal. These intra-individual differences are often dismissed as they were not the result of an a priori defined developmental phase or task manipulation. However, perception is continuously evolving and seemingly robust effects for one observer may change over time (Friston, & Kiebel, 2009; Koenderink, 2019; Molenaar, 2004; Hamaker, 2012). These perceptual dynamics can be very informative for our understanding of the visual system and, hence, they may be worth being examined further. In this PhD thesis, intra-individual perceptual dynamics will be considered in the 'Relevance study', 'EEG study' and 'MQ study' chapters.

1.5 Origins of variability in perception

Given the fact that Gestalt psychology put forward a set of 'grouping principles' that predict perception for all observers based on stimulus characteristics alone, it may seem plausible that they denied any inter- or intra-individual variability in visual perception. However, some Gestaltists have explicitly referred to the existence of individual differences in visual perception.

Metzger (1936/2006 p. 136) argued that previous experience and information other than stimulus characteristics can have a strong influence on perception when the simplicity, or 'Prägnanz', of the structure of a stimulus is low, therefore acknowledging the significance of previous experience in perception. However, understanding fundamental, universal laws of perceptual organization is considered to be more important, as they are present before any experience has taken place, and individual differences are seen as an additional variability to the fundamental laws that constitute common patterns in perception (Metzger, 1936, pp. 180-184). Koffka (1935, pp. 98-100) saw perceptual organization as a process that was constrained by the medium in which this organization takes place. Koffka defined this medium by external conditions, such as the proximal stimulus, and internal conditions, such as the observer's nervous system. This variability in internal conditions can then result in permanent as well as temporary individual differences. Accordingly, experience can be seen as a change in the internal conditions of the medium, changing the constitution of the observer's nervous system and therefore possibly changing perception (Koffka, 1935, pp. 104-105).

We can ask where or when variability in visual perception originates. When considering the origins of perception, we can distinguish between phylogenetic and ontogenetic origins. Phylogenetic origins concern all development throughout the evolution of a species. Ontogenetic origins concern all development throughout the lifespan of an individual. In general, phylogenetic origins should not result in inter- or intra-individual differences in perception, as these processed are independent of individual development or experience. Ontogenetic origins may result in development and experience that is common to essentially all individuals (e.g., essentially all individuals develop the ability to perceive animals), but may also result in development that is not common to all individuals. These ontogenetic differences may be innate (e.g. some individuals may have an innate higher sensitivity to luminance contrast for high spatial frequencies) or may manifest throughout the lifespan of an individual due to experience (e.g., some individuals may develop a higher sensitivity to luminance contrast of high spatial frequencies because of exposure to an environment with more high spatial frequency information). Phylogenetic origins and ontogenetic origins in the situation where development and experience is common give rise to the impression that the stimulus ultimately determines perception, independent of the observer. However, when considering ontogenetic origins in the situation where development and experience is not

common, we can see that the visual system of the observer can have an important role in determining perception.

We may also ask why there is individual variability in visual perception. I would like to allude to two theoretical accounts that have inter- and intra-individual variability in perceptual organization and perception as a logical consequence.

The first one entails that we *are not able to* have perceptual representations completely identical to the outside world because our visual system has limited input and resources. von Helmholtz (1867) introduced the concept of "unconscious inference" to describe how our perceptual system resolves ambiguities in visual input. The theory implies that human vision is incomplete and perception is the result of inference processes by the unconscious mind. According to von Helmholtz, the nervous system relies on information from prior experience, expectations, context, and motor activities to infer the *most likely* state of the outer world. Differences in the visual system or in visual experience would change the inference of the most likely state of the outer world, and may therefore result in inter- and intra- individual differences in visual perception.

The second one entails that we *do not need to* have perceptual representations completely identical to the outside world because this is not necessary to interact with it. Our visual system might rely on information other than stimulus input simply because it allows us to *successfully* interact with the environment. Indeed, some accounts (e.g., The 'Interface Theory of Perception'; Hoffman, Singh, & Prakash, 2015; Koenderink, 2019) have described perception as an adaptive interface that serves to guide useful actions, not to resemble truth. Accordingly, the formation of a percept from our basic visual experience should be shaped by what is useful. What is found to be useful by our visual system, however, may be variable within and between individuals, and it may be influenced by multiple information sources.

The two accounts differ in crucial aspects, yet they are not entirely mutually exclusive and may even strengthen each other in some aspects. For instance, a difference between both accounts is the notion that our visual system creates a percept that reflects the most likely state of the outer world in "unconscious inference", whereas the percept does not have to reflect the visual state of the outer world whatsoever in the 'Interface Theory of Perception'.

However, both accounts emphasize a limited amount of visual information and various other factors systematically influencing perception, giving rise to variability in visual perception.

1.6 Factors determining variability in multistable perception

Another fundamental question pertains to what the influencing factors of variability in multistable perception are. As Brascamp (2021) pointed out, ambiguity in representations of external stimuli is ubiquitous in visual perception. Since our visual system is always resolving ambiguity, everything that has an influence on perception may also have an influence on (individual) variability in multistable perception, albeit in different proportions depending on the particular ambiguities.

There are many different factors influencing visual perception, with examples ranging from the properties of an individual's neurons in visual cortex to task-related relevance of a particular percept. Brascamp et al. (2018) pointed out that multiple information sources influence multistable perception, by listing various examples such as stimulus evidence in accordance with a perceptual interpretation, ecological likelihood of a perceptual interpretation, information from other sensory modalities, the behavioral context and prior experience and expectations.

The factors influencing visual perception are often described in terms of their "position" on the different levels of the visual system: "low-level", "mid-level", and "high-level". It is important to note, however, that the borders between "low-", "mid-", and "high-level"vision may be rather arbitrary and flexible. Our understanding of the interaction of different levels in the visual system is still growing and, considering our current knowledge of interaction across the visual system, a textbook definition of distinct levels in a cortical hierarchy falls short (Wagemans, 2018). The different empirical studies in this PhD thesis are aimed at examining the influence of factors from different levels of the visual system on individual variability in visual perception.

The perception of the 'local-global' ambiguous motion stimulus is characterized by a 'shift to global' over time and the underlying factors causing this change in perception over time remain unclear. Therefore, *Chapter 2: 'Relevance study'* examines the role of mere exposure

vs. experienced “high-level” relevance in an auxiliary task on the perception of this stimulus over time.

Spillmann (2009, p. 1508) pointed out “the fruitfulness of pursuing complementary strategies by correlating phenomenological observations and psychophysical measurements with neurophysiological recordings … in the hope that a deeper understanding of visual perception may be obtained”. *Chapter 3: ‘EEG study’* examines electrophysiological correlates of perception of the ambiguous local-global motion stimulus vs. disambiguated versions of this stimulus over time.

Chapter 4: ‘CSF study’ examines the relationship between the “low-level” measure of individual luminance contrast sensitivity and the perception of a ‘local-global’ ambiguous motion stimulus which enables both a ‘local’ and ‘global’ apparent motion percept.

All of the above empirical studies explore the influence of various factors on multistable perception of a particular ambiguous motion stimulus. Accordingly, this allowed us to study perceptual variability aside from variability caused by stimulus characteristics. However, given the inherent complexity of visual perception, the influence of these factors causing individual variability may also interact with the influence of various stimulus characteristics. For instance, an observer may experience a ‘shift to global’ over time for large stimuli, but this may not be the case for small stimuli. Studying the influence of any factor on one particular stimulus in isolation may limit us in developing a deeper understanding of visual perception.

In *Chapter 5: ‘MQ study’*, we took a step back and explored the perception of multiple variations of a motion quartet stimulus and examined the interaction of the effect of these different stimulus characteristics with individual differences in the perception of these stimuli.

References

Alais, D., Blake, R., & Lee, S. H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience, 1*(2), 160-164.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience, 3*(1), 13-21. https://doi.org/10.1038/nrn701

Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable perception and the role of the frontoparietal cortex in perceptual inference. *Annual Review of Psychology, 69,* 77-103. https://doi.org/10.1146/annurev-psych-010417-085944

Brascamp, J. W., & Shevell, S. K. (2021). The certainty of ambiguity in visual neural representations. *Annual Review of Vision Science, 7*, 465-486. https://doi.org/10.1146/annurev-vision-100419-125929

Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research, 141*, 247-257. https://doi.org/10.1016/j.visres.2017.01.008

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*(11), 671-684.

de-Wit, L., & Wagemans, J. (2015). Individual differences in local and global perceptual organization. In J. Wagemans (Ed.), *The oxford handbook of perceptual organization* (pp. 601-620). Oxford, UK: Oxford University Press.

Faubert, J., & Herbert, A. M. (1999). The peripheral drift illusion: A motion illusion in the visual periphery. *Perception, 28*(5), 617-621. https://doi.org/10.1068/p2825

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 56*(3), 164-176. https://doi.org/10.1037/h0087394

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27), E6106-E6115. https://doi.org/10.1073/pnas.1711978115

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1521), 1211-1221. https://doi.org/10.1098/rstb.2008.0300

Hamaker, E. L. (2012). Why researchers should think" within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). The Guilford Press.

Hoffman, D. D., Singh, M., & Prakash, C. (2015). The interface theory of perception. *Psychonomic Bulletin & Review, 22*(6), 1480-1506. https://doi.org/10.3758/s13423-015-0890-8

Kaiser, M. D., & Shiffrar, M. (2009). The visual perception of motion by observers with autism spectrum disorders: A review and synthesis. *Psychonomic Bulletin & Review, 16*(5), 761-777. https://doi.org/10.3758/PBR.16.5.761

Kitaoka, A., & Ashida, H. (2003). Phenomenal characteristics of the peripheral drift illusion. *Vision, 15*(4), 261-262. https://doi.org/10.24636/vision.15.4_261

Kleinschmidt, A., Sterzer, P., & Rees, G. (2012). Variability of perceptual multistability: from brain state to individual trait. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1591), 988-1000. https://doi.org/10.1098/rstb.2011.0367

Koenderink, J. (2019). Vision, an optical user interface. *Perception, 48*(7), 545-601. https://doi.org/10.1177/0301006619853758

Koffka, K. (1922). Perception: an introduction to the Gestalt-Theorie. *Psychological Bulletin, 19*(10), 531-585.

Koffka, K. (1935). *Principles of Gestalt psychology*. London, England: Lund Humphries.

Kornmeier, J., & Bach, M. (2012). Ambiguous figures–what happens in the brain when perception changes but not the stimulus. *Frontiers in Human Neuroscience, 6*. https://doi.org/10.3389/fnhum.2012.00051

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences, 3*(7), 254-264. https://doi.org/10.1016/S1364-6613(99)01332-7

Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychological Bulletin, 130*(5), 748-768. https://doi.org/10.1037/0033-2909.130.5.748

Metzger, W. (1936). *Gesetze des Sehens*. Frankfurt am Main, Germany: Kramer. [Translation reprinted as "*Laws of seeing*" (L. Spillmann, M. Wertheimer, & S. Lehar, Trans.) (2006). Cambridge, MA: MIT Press.]

Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice?. *Vision Research, 141*, 4-15. https://doi.org/10.1016/j.visres.2017.11.001

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*(4), 201-218. https://doi.org/10.1207/s15366359mea0204_1

Sagi, D. (2011). Perceptual learning in vision research. *Vision Research, 51*(13), 1552-1566. https://doi.org/10.1016/j.visres.2010.10.019

Schwartz, J. L., Grimault, N., Hupé, J. M., Moore, B. C., & Pressnitzer, D. (2012). Multistability in perception: binding sensory modalities, an overview. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*(1591), 896-905. https://doi.org/10.1098/rstb.2011.0254

Spillmann, L. (2009). Phenomenology and neurophysiological correlations: Two approaches to perception research. *Vision Research, 49*(12), 1507-1521. https://doi.org/10.1016/j.visres.2009.02.022

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences, 13*(7), 310-318. https://doi.org/10.1016/j.tics.2009.04.006

Van der Hallen, R., Manning, C., Evers, K., & Wagemans, J. (2019). Global motion perception in autism spectrum disorder: a meta-analysis. *Journal of Autism and Developmental Disorders, 49*(12), 4901-4918. https://doi.org/10.1007/s10803-019-04194-8

Vannuscorps, G., Galaburda, A., & Caramazza, A. (2021). Shape-centered representations of bounded regions of space mediate the perception of objects. *Cognitive Neuropsychology*. https://doi.org/10.1080/02643294.2021.1960495

Vanston, J. E., & Strother, L. (2017). Sex differences in the human visual system. *Journal of Neuroscience Research, 95*(1-2), 617-625. https://doi.org/10.1002/jnr.23895

von Helmholtz, H. (1867). *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln (Vol. 9)*. Voss.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin, 138*(6), 1172-1217. https://doi.org/10.1037/a0029333

Wagemans, J. (Ed.). (2015). *The Oxford Handbook of Perceptual Organization*. Oxford, UK: Oxford University Press. https://doi.org/10.1002/9781119170174.epcn218

Wagemans, J. (2018). Perceptual organization. In Wixted J. T. (Series Ed.) & Serences J. (Vol. Ed.), *The Stevens' handbook of experimental psychology and cognitive neuroscience: Vol. 2. Sensation, perception & attention* (Chapter 18, pp. 803–872). Hoboken, NJ: Wiley. https://doi.org/10.1002/9781119170174.epcn218

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie, 61,* 161–265. [Translated as "Experimental studies on seeing motion". In L. Spillmann (Ed.), (2012). *On perceived motion and figural organization* (pp. 1-91). Cambridge, MA: MIT Press.]

Wertheimer, M. (1922). Untersuchungen zur Lehre von der Gestalt, I: Prinzipielle Bemerkungen. *Psychologische Forschung, 1*, 47-58. [Translated extract reprinted as "The general theoretical situation." In W. D. Ellis (Ed.), (1938). *A source book of Gestalt psychology* (pp. 12-16). London, U.K.: Routledge & Kegan Paul Ltd.]

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II. *Psychologische Forschung, 4*, 301–350. [Translated as "Investigations on Gestalt principles, II.". In L. Spillmann (Ed.), (2012). *On motion and figure-ground organization* (pp. 127-182). Cambridge, MA: MIT Press.]

2. Relevance study:

Does task relevance shape the 'shift to global' in ambiguous motion perception?

This study was published in Journal of Vision in 2019

Reference

Boeykens, C., Wagemans, J., & Moors, P. (2019). Does task relevance shape the 'shift to global' in ambiguous motion perception?. *Journal of Vision*, *19*(1), 8-8. https://doi.org/10.1167/19.1.8

Abstract

Perception can differ even when the stimulus information is the same. Previous studies have demonstrated the importance of experience and relevance on visual perception. We examined the influence of perceptual relevance in an auxiliary task on subsequent perception of an ambiguous stimulus. Observers were presented with an ambiguous motion stimulus that could either be perceived as rotating dot pairs ("local") or pulsating geometrical figures ("global"). Prolonged perception of this stimulus is characterized by a "shift to global", but it remained unclear whether this process is due to relevance of the global percept or mere exposure to the stimulus. During a relevance learning phase over 5 successive days, participants were divided into conditions determining the relevant percept in an auxiliary task: local, global, or none (active exposure). In a pre- and posttest, individual points of subjective equality between local and global percepts were measured. Results indicate that there is indeed a shift to global. Interestingly, auxiliary task relevance does not seem to modify this process.

2.1 Introduction

Visual perception is not merely the result of stimuli being represented in passive observers. Our perceptual system disambiguates and interprets the information it receives through our eyes. The perception of ambiguous stimuli provides a clear demonstration of the dissociation between stimulus and percept: Identical stimuli are perceived in substantially different ways between different observers (inter-individual differences) or from one moment to the next within single observers (intra-individual differences) (Necker, 1832; Leopold & Logothetis, 1999; Blake & Logothetis, 2002; Kornmeier & Bach, 2012; Nikolaev et al., 2016; Brascamp, Sterzer, Blake, & Knapen, 2018). For instance, a Necker cube can be perceived both facing the upper right corner and facing the lower left corner, even though the physical image remains the same. Perceptual organization describes how basic visual features are organized into more coherent units, which are employed in further processing (Wagemans, 2018). Understanding under which circumstances a stimulus is perceived one way or the other may guide our understanding of how sensory information is transformed into a perceptual representation.

Decades of research have been devoted to investigating what determines a perceptual outcome in the case of ambiguous stimuli. Differences in perception of ambiguous stimuli appear not to be merely coincidental, yet are systematically influenced by various factors. Brascamp et al. (2018) reviewed the literature on ambiguous perception, demonstrating the role of multiple information sources resolving perceptual ambiguities. One particularly important and complex source of information to resolve perceptual ambiguities is prior experience. Prior experience may influence perception in various ways, ranging from perceptual organization in the preceding seconds to implicitly and explicitly learned expectations (Denison, Piazza, & Silver, 2011; Di Luca, Ernst, & Backus, 2010; Piazza, Denison, & Silver, 2018; Schmack, Weilnhammer, Heinzle, Stephan, & Sterzer, 2016). Even in traditional Gestalt psychology, Wertheimer (1923/2012) described prior experience as an influencing factor in perceptual organization, in addition to isolated stimulus factors (see also Wagemans, 2018).

Prior experience may affect perception because it provides information about the relevance of sensory input to the observer. This way, our perceptual system is informed about which perceptual organization will allow successful interaction with the environment. Related to

this, the 'Interface Theory of Perception' (Hoffman, Singh, & Prakash, 2015; Koenderink, 2011) describes perception as an adaptive interface that serves to guide useful actions, not to resemble truth. Accordingly, the formation of a percept from our basic visual experience should be shaped by what is relevant. What is found to be relevant by our visual system, however, may be variable and influenced by multiple information sources. We may distinguish three levels of prior experience constituting perceptual relevance. "*Higher-level"* processes (e.g., previously experienced cognitive goals to perceive the Necker cube facing the upper right corner) may determine what is relevant to perceive. Alternatively, engagement with the same *"mid-level" percept* (e.g., previously perceiving the Necker cube facing the upper right corner) could determine its relevance. Or, *"low-level" exposure* to the same visual input (e.g., previously being presented with the Necker cube, irrespective of the percept) may be sufficient to determine its relevant percept.

Anstis and Kim (2011) studied perception of an ambiguous motion stimulus (Figure 1) that enabled a 'local' and a 'global' percept. The stimulus consists of multiple dot pairs placed around a common center. When the dot pairs change in phase position between two successive frames, one of two percepts can emerge. The local percept consists of multiple rotating dot pairs, whereas the global percept consists of two pulsating geometric forms. Such an ambiguous motion stimulus is particularly interesting since it allows us to control and manipulate its basic physical properties precisely, while still enabling us to study more complex perceptual interpretations (e.g., motion direction).

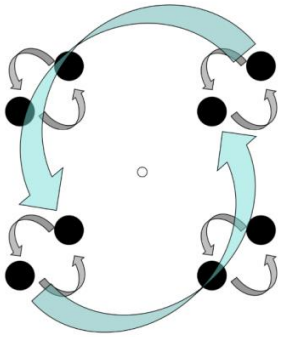


Figure 1. Ambiguous motion stimulus as employed by Anstis and Kim (2011). Small arrows indicate rotation of the dot-pairs and movement of the 'local' percept. Large arrows indicate movement of the 'global' percept. A similar animated version can be found under "stimuli" on https://osf.io/jryb9/.

Prolonged perception of the ambiguous motion stimulus was found to be characterized by a 'shift to global'. This implies that the longer an individual watches the stimulus, the more they

report to perceive it globally. However, it remains unclear whether the 'shift to global' occurs automatically over time, or is due to a "higher-level" attribution to the global percept. For instance, more relevance might be attributed to the global percept because it is initially more difficult to perceive and participants might try to achieve this in order to report both local and global percepts (as instructed). Accordingly, the global percept would become a "higher-level" cognitive goal. Anstis and Kim (2011) interpreted the 'shift to global' as a result of learning and adaptation in perceptual organization towards a global configuration. They concluded that "higher-level" attention and perceptual interpretation underlie the perceptual strength of global versus local percepts, in analogy to the "high-level" processes involved in perception of ambiguous motion transparency stimuli (McOwan & Johnston, 1996). However, since the manipulation of Anstis and Kim (2011) merely involved a stimulus induced change, it is difficult to determine exactly what is implied with the term "higher-level" and to exclude alternative explanations that do not rely on "high-level" perceptual processes.

Chopin and Mamassian (2010; 2011) examined the role of "high-level" processes in more detail and found that biases in perceptual organization can be updated on a short time scale based on relevance. They used an ambiguous motion transparency stimulus and had observers report the direction of motion that was perceived to be in front. In an auxiliary task, observers had to perform a visual search for a slowly moving dot in a similar stimulus. The preferred motion direction of the motion transparency task was altered by the direction of the slow dot in the visual search task. The authors concluded that the attribution of usefulness to one of the directions in the auxiliary task had an influence on spontaneous perception in the motion transparency task, even though this direction did not help them in that task. Harrison and Backus (2010) manipulated perceptual bias in rotating Necker cubes and showed that contextual information can bias subsequent perception of an ambiguous stimulus. Similarly, Feigin et al. (2017) found that previous behavioral relevance has a key role in biasing future perception. These findings suggest that "higher-level" processes, such as auxiliary task relevance, can affect "mid-level" perception. If this is the case, the 'shift to global' in the ambiguous motion stimulus employed by Anstis and Kim (2011) could be due to "higher-level" relevance of the global percept.

In this study, we similarly examined if and how relevance shapes perceptual organization. In particular, we asked whether the previously observed "shifttoglobal" in the stimulus used by

Anstis and Kim (2011) is due to relevance (i.e., akin to the findings of Chopin and Mamassian, 2010; 2011) or can be explained by mere exposure to the ambiguous stimulus. Observers were presented with an altered version of the ambiguous motion stimulus (Figure 2) employed by Anstis and Kim (2011) and were asked to report their spontaneous percept. In a first session, their point of subjective equality (PSE) between the local percept (pairs rotating) and the global percept (geometric figures pulsating) was measured. In a relevance learning phase, which took place across five successive days, observers were then divided into three conditions determining their relevant percept. This time scale was chosen because it was assumed to provide a more robust manipulation of perceptual relevance. In the *local condition*, participants performed a task that could only be solved by adopting a local percept. In the *global condition*, participants performed a task that could only be solved by adopting a global percept. Lastly, in the *exposure condition*, participants performed a task that could be solved without having to consistently adopt either percept. Accordingly, the "higher-level" relevance of either percept in the ambiguous motion stimulus was manipulated by means of a task. After the learning phase, participants' PSE was measured again. A 'shift to global' is reflected by an increase in the PSE.

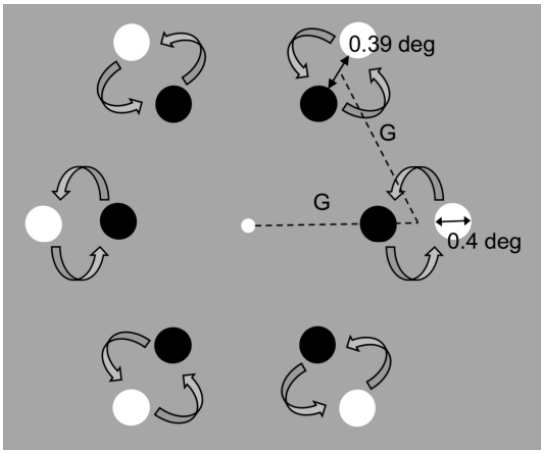


Figure 2. Ambiguous motion stimulus employed in the present study. Arrows indicate the movement of the dot pairs. Dot pairs were phase-shifted in comparison to Figure 1 in order to create a global percept with different characteristics (i.e., pulsating instead of shuffling). An animated version can be found under "stimuli" on https://osf.io/jryb9/.

If a "higher-level" relevance of the global percept or engagement with the global percept is necessary to elicit a 'shift to global', we should find a differential effect for task relevance of the global percept compared to the local percept. That is, the global condition should elicit

more global percepts whereas the local condition should elicit more local percepts. If a "higher-level" relevance of either percept or engagement with either percept is necessary to elicit a 'shift to global', we should find a differential effect for task relevance of either percept compared to the exposure task. That is, relative to the exposure condition, we expect more global percepts for either condition. If exposure to the same visual input, irrespective of the constructed percept, is sufficient to elicit a 'shift to global', we should not find any differential effect for task relevance.

2.2 Methods

Participants

30 university students (5 male, aged 18-24) participated in the study in return for course credit. The number of participants was based on previous studies employing an equivalent paradigm (Chopin & Mamassian, 2011; Harrison & Backus, 2010). All participants had normal or corrected-to-normal vision, provided informed consent, and were naïve with respect to the goals of the study. The study was approved by the Social and Societal Ethics Committee of KU Leuven (SMEC, approval code G 2017 05 834) and adhered to the tenets of the Declaration of Helsinki.

Apparatus

All stimuli were displayed on a CRT monitor (Sony GDM F520) with a refresh rate of 60 Hz, a diagonal length of 50 cm and a resolution of 1024 x 768 pixels. Participants were positioned in a dark room at a viewing distance of 57 cm from the monitor by means of a chin rest. Stimulus presentation and response registration was controlled by software programmed in Python using the PsychoPy library (Peirce, 2007).

Stimuli

The ambiguous motion stimulus (Figure 2) consisted of 6 pairs of dots displayed on a uniform grey background (luminance 17.9 cd/m^2) around a white fixation point. One dot of each pair was white (luminance 76.8 cd/m^2), whereas the other dot was black (luminance 0.2 cd/m^2). The random allocation of these colors to the dots was determined at the start of each trial.

The rationale for this color scheme is related to the relevance manipulation (see Procedure). The dots had a radius of 0.2 deg and were placed in pairs with an edge-to-edge distance of 0.39 deg. This distance may also be regarded as a phase shift of π rad around the pair-center. The 6 dot pairs are placed around a central fixation point with a distance G from the center. Since the dot pairs are placed in a hexagon, distance G is also equal to the distance between the dot pairs. The starting position of the dots varied according to the phase of their pair-centers relative to the common center. Accordingly, the position of a dot was not only related to the other dot in its pair by a phase shift around the pair-center, but also to other dots by the phase relation to the common center. Dots rotated around their pair-center at a speed of 2π rad/s for 30 frames, or 500 ms, per trial. This rotation was either clockwise or counter-clockwise, determined randomly at the start of each trial. At the end of each trial, the position of each dot pair differed from the starting position with a rotation of π rad around the pair-center.

Design and Procedure

A 2 x 3 design with within-subjects factor time (pre- and post-learning phase) and between-subjects factor relevance (local, global and exposure) was implemented. The experiment consisted of five separate sessions spread across five successive days. The first and last session lasted approximately 40 min, all other sessions lasted approximately 15 min. During the first session, a continuous version of the ambiguous motion stimulus was displayed shortly in order to assure that the participant could experience the difference between the 'local' and the 'global' percept. During the first and last session, the point of subjective equality (PSE) was measured, according to the procedure described below. During the relevance learning phase, taking place in all five sessions, participants performed the Relevance task, also described below. Participants were randomly assigned to a relevance condition for the learning phase, resulting in 10 participants per condition. The dependent measure of interest was the difference between the PSE pre- and post-learning, dependent upon the relevance condition. At the end of the experiment, participants were questioned about their knowledge of the experimental manipulation. We wanted to assure that participants did not figure out the purpose of the experiment, which would allow them to bias their responses. If participants

had a vague idea about the purpose of the experiment, we questioned whether they altered their responses according to their expectations instead of their actual percepts.

Point of subjective equality

In a pre- and post-learning test, the point of subjective equality (PSE) was measured for each participant. The PSE, or transition point, is the point (i.e., a particular stimulus configuration) at which both percepts of the ambiguous motion stimulus are equally likely. On each trial, participants were presented with the ambiguous motion stimulus for 500 ms, after which they had to indicate their percept by means of a button press (left arrow for local, right arrow for global, or vice versa). Responses were not permitted before the end of the trial and the stimulus remained in the ending position until a response was given. Between two successive trials, the fixation point was displayed for 1000 ms. The distance between the elements of the local percept, or between the dots of one dot pair, was kept constant at 0.39 deg. The distance G was manipulated on each trial according to an adaptive staircase procedure – Psi (Kontsevich & Tyler, 1999). We will refer to distance G at the PSE as G*. When G is larger than G*, an individual will perceive the ambiguous stimulus as local; when G is smaller than G*, an individual will perceive the ambiguous stimulus as global. The longer distance G* is, the longer the physical distance between the dot pairs can be for an individual to be able to see the global percept. A 'shift to global' may thus be quantified by an increase in distance G*. Three different Psi staircases, each consisting of 100 trials, were presented in an interleaved manner, resulting in 300 trials in total. Distance G* was computed by taking the mean of the estimates for G* of the three Psi staircases. Participants performed the Relevance task in all five sessions of the learning phase with a stimulus that has the distance G* measured in the beginning of their first session. This way, participants initially perform the Relevance task with an equal probability of seeing the local and the global percept, irrespective of which percept is relevant for the task. After the learning phase, at the end of the last session, the PSE was re-measured.

Relevance task

Participants performed the Relevance task during each session on five successive days. Each task session consisted of 300 trials divided into six blocks interspersed with short breaks. Crucially, the presented stimulus was identical in all tasks, except for participant-specific G* and a change in the color of the fixation point in the Exposure task. Moreover, in all tasks, participants received the instruction to maintain their fixation on the fixation point, but to spread their visual attention across the entire stimulus. Therefore, the only difference between conditions should be the manipulation of relevance of the percept by means of the task.

In the Local task, a local percept was necessary to perform the task. Participants had to indicate whether they saw a clockwise or counter-clockwise rotation of the dot pairs (Figure 3, top row). Crucially, the global percept was irrelevant to perform the Local task, and vice versa. As indicated in Figure 3, the same stimulus (rows) provides both a local and a global percept. However, as the figure shows, a correct response in the Global task corresponds to different correct responses in the Local task and vice versa.

In the Global task, a global percept was necessary to perform the task. Participants had to indicate whether they saw a growing-shrinking or shrinking-growing pulsating movement of the global geometric form (Figure 3, bottom row). During one trial, both options can be valid. However, because of the color scheme, only one option is correct. More specifically, participants were instructed to report the pulsating movement direction for either the black, or the white dots. The color to be reported was counter-balanced across participants.

In the Exposure task, neither a global nor local percept was necessary to perform the task. Participants had to report a color change (white to black) in the central fixation point, which occurred on 50% of trials. Consequently, both percepts were irrelevant. Nevertheless, participants were engaged in a task that was intended to be equally attention demanding as the other conditions.

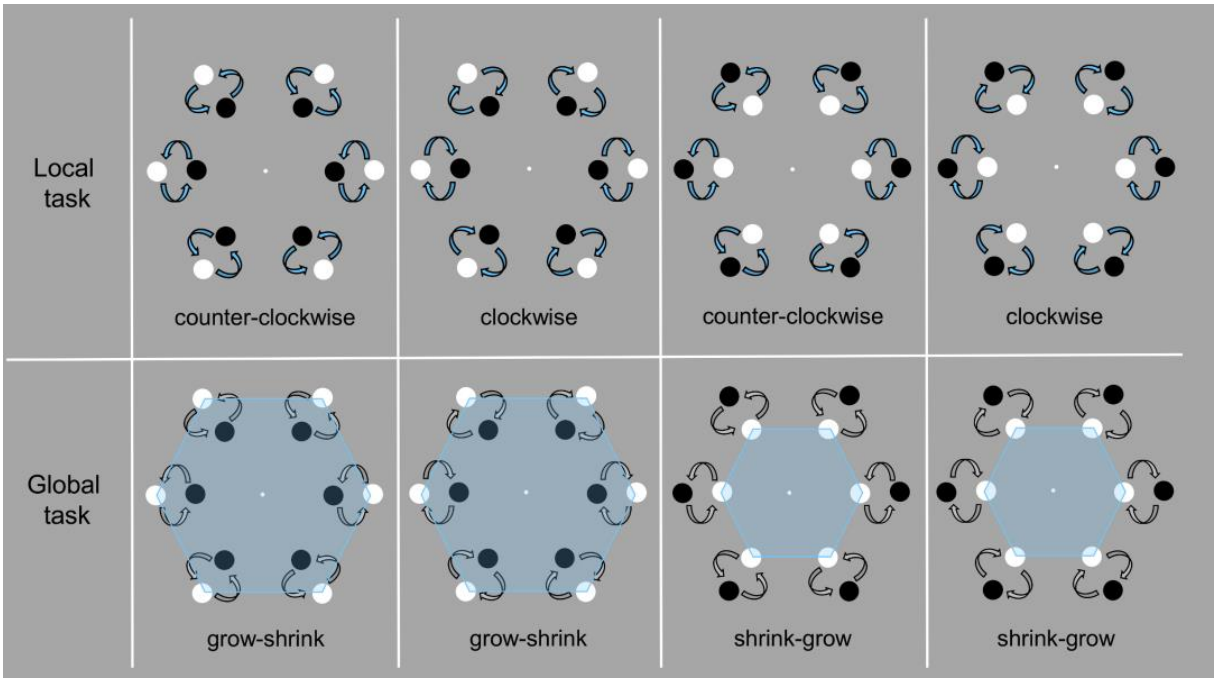


Figure 3. Local and global tasks employing ambiguous motion stimuli and examples with correct answers. In this example, white dots should be used for the global task. Dot color does not have any meaning in the local task. The same stimulus (columns) can elicit both a local and a global percept (percepts are indicated in blue). A change in correct answer for the global task does not necessarily imply a change in correct answer for the local task, and vice versa.

Data analysis

All data analysis was performed in R 3.4.1 (R Core Team, 2017). Statistical tests relied on model selection based on Bayes Factors using the BayesFactor package 0.9.12.2 (Morey & Rouder, 2015). In the Bayesian statistical framework, Bayes Factors quantify the consistency of the data with the predictions of a particular statistical model over another. Bayes Factors from 3 onwards were proposed as substantial evidence for one model over another by Jeffreys (1961). Fitted models included random intercepts for participants and used variables time (pre- and post-learning) and relevance (local, global and exposure) as predictors. Default priors as implemented in the package were used. This amounts to a "medium" scale ($r = 0.5$) for fixed effects and a "nuisance" scale for random effects ($r = 1$). These scales apply to the so-called g-priors which are placed on the effects in the ANOVA model (for a detailed description, see Rouder, Morey, Speckman, & Province, 2012). The data and analysis code are available under "data" and "analysis" on the Open Science Framework. (link: https://osf.io/jryb9/).

2.3 Results

The results of the Bayes Factor analysis indicated that there was strongest support for a model with only a main effect of time. Table 1 shows Bayes Factors of the model with the main effect of time compared to other models. The main effect of time indicates an overall change in perceptual organization after prolonged engagement with the ambiguous motion stimulus. G*, the distance between the dot pairs of the ambiguous stimulus at the point of subjective equality (PSE), was larger after the learning phase (*M* = 6.28 deg, *SD* = 3.04 deg) than before the learning phase (*M* = 5.36 deg, *SD* = 2.85 deg). There are substantial individual differences in this effect, but the overall trend is in the direction of a 'shift to global', indicating that the global percept was more dominant in most individuals post-learning. Figure 4A shows G* in the first and last session for all observers. Values of G* above the diagonal imply a 'shift to global' over time. Figure 4B displays the density of observers with a certain magnitude of G* in the first compared to the last session, with a change towards the right side of the plot indicating a 'shift to global'. Secondly, there is support for an additional main effect of relevance (a mere *BF* = 1.526 for a model without a main effect of relevance compared to a model with both main effects) indicating that there was a difference in G* depending on the assigned conditions of participants, irrespective of the learning phase. Participants assigned to the global condition (*M* = 6.80 deg, *SD* = 3.51 deg) had a larger G* than participants assigned to the exposure condition (*M* = 5.53 deg, *SD* = 2.61 deg) and the local condition (*M* = 5.14 deg, *SD* = 2.53 deg). Even though participants were assigned to relevance conditions randomly, the distribution of perceptual dominance was uneven over conditions by chance. Although the absence of a main effect of relevance would have been desirable, we do not consider it to be problematic here, as the critical effect was the interaction between time and relevance.

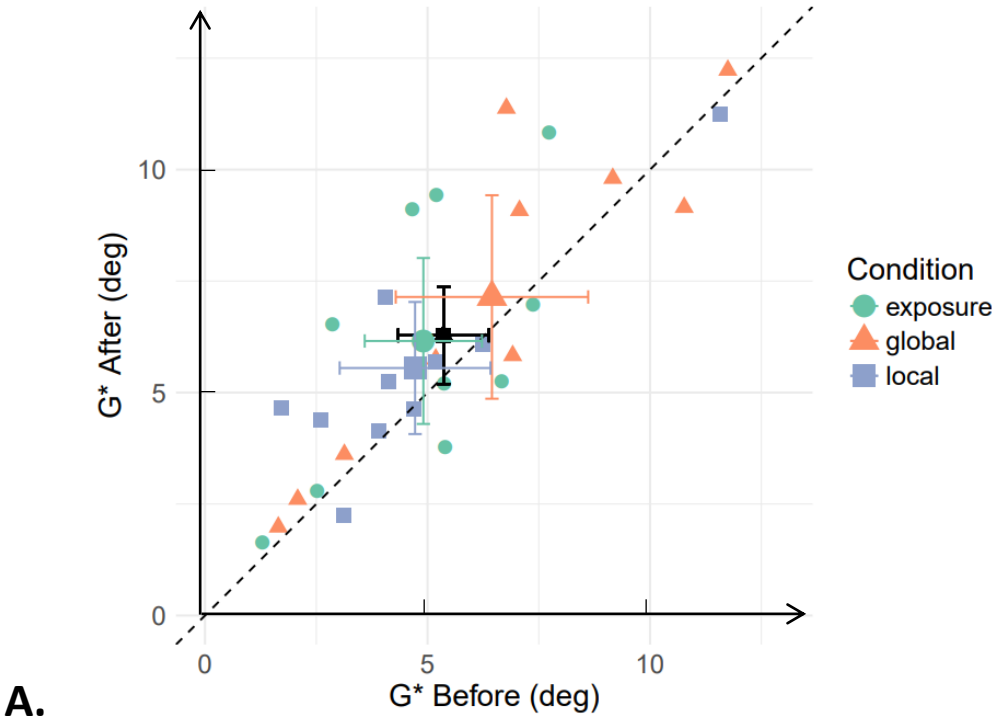
Table 1. *Fitted models and Bayes Factors (BF).*

Model	Bayes Factor
time + participant	1.000
time + relevance + participant	1.526
participant	4.865
time + relevance + time:relevance + participant	6.010
relevance + participant	7.939

Note. Reported BF correspond to the best fitting model (top row) with only main effects of time and participant, compared to each model. For instance, the Bayes factor of 4.865 displayed alongside the model with an effect of participant, indicates that the data is 4.865 times more consistent with the predictions of a model including effects of time and participant compared to a model including only an effect of participant. Or, in other words, the model including an effect of time as well as participant is 4.865 times more likely than the model only including an effect of participant.

As can be derived from Table 1, the model with only a main effect of time was supported substantially more by our analysis than a model with the interaction effect of interest (*BF* = 6.010 for a model without the interaction effect compared to a model including the interaction effect). This finding indicates that the manipulated relevance of either percept during the learning phase had no effect on the 'shift to global' of G*.

Responses to the questionnaire provided at the end of the last session indicated that almost none of the participants were aware of a link between the measurement of the PSE and the relevance manipulation during the learning phase. Importantly, none of the participants indicated that they had responded according to any reasoning other than their experienced percept.



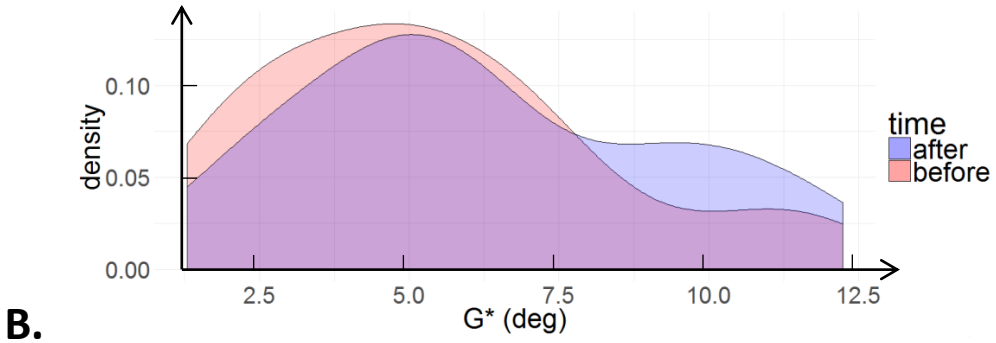


***Figure 4*. A.:** Individual measures of G*, the distance between the dot pairs of the ambiguous stimulus at the point of subjective equality, before and after the learning phase are plotted in a different color for each relevance condition. The mean of G* for each condition is indicated by a larger icon. Error bars are means +/- 1.96 times the standard error. The dotted diagonal line indicates no change in G* over time. Values above the diagonal indicate a 'shift to global'. The black symbol represents the mean shift to global over all individuals and conditions. **B.:** The density of observers with a certain magnitude of G* in the first ('before' in red) compared to the last ('after' in blue) session, with a change in magnitude along G* towards the right side of the plot indicates a 'shift to global'.

2.4 Discussion

Absence of a differential effect of auxiliary task relevance

In the current study, we examined the role of previous experience and relevance on perceptual organization. Relevance may manifest itself on various levels, ranging from a "high-level" cognitive goal to "low-level" repeated exposure. Based on previous findings of Chopin and Mamassian (2010; 2011), we predicted that "mid-level" perceptual organization in an ambiguous motion stimulus could be shaped by "higher-level" relevance triggered by an auxiliary task. Our results suggest that manipulating the relevance of a percept in an ambiguous motion stimulus by means of an auxiliary task does not affect its subsequent perception. This is in contrast to previous findings (Chopin & Mamassian, 2011; Harrison, & Backus, 2010; Feigin et al., 2017).

An explanation for these contradictory findings could be that the time scale of the manipulation of Chopin and Mamassian (2011) and Harrison and Backus (2010) was shorter than the time scale of the current experiment. In the current experiment, the influence of the relevance manipulation was examined over the course of a whole week, whereas Chopin and Mamassian (2011) obtained their results based on one experimental session. However, it is important to note that the last session in the current experiment consisted of the relevance task followed immediately by the measurement of the PSE. This implies that, in principle, the relevance manipulation needed only to last for one experimental session, not for an entire week. Another important difference between both studies is the alternation of perceptual report and the auxiliary task. Observers in the experiment of Chopin and Mamassian (2011) alternated between the perceptual report and the auxiliary task, whereas observers in our experiment performed both tasks separately. Therefore, the findings of Chopin and Mamassian (2011) may also have been due to a short-term attentional bias. Indeed, a study by Meng and Tong (2004) showed that attention can selectively bias perceptual organization in ambiguous stimuli such as the Necker cube, but not so strongly in binocular rivalry, where the ambiguity has been argued to be solved at a lower level (Tong, 2001). However, in a previous study, Chopin and Mamassian (2010) showed that the effect of task relevance is also present in binocular rivalry stimuli and argue against an account of attention for explaining

their results. Since the effect of task relevance was still present in the last block of their experiment where it was not relevant to the auxiliary task anymore, the effect is not merely due to a short-term bias. Moreover, in the experiment of Harrison et al. (2010) the perceptual report and auxiliary task were performed separately, as was the case in our experiment. Their reported effect was present on the next day and resisted reverse learning, eliminating the possibility of a short-term attentional bias. The absence of a differential effect of "higher-level" relevance on the perception of the ambiguous motion stimulus may also be due to the specific stimulus employed in this paradigm. Indeed, previous research has shown that the size of the effect of learned expectations on the perception of ambiguous stimuli varies widely across different stimuli (Meng & Tong, 2004; Schmack et al., 2016; Weilnhammer, Stuke, Sterzer, & Schmack, 2018). If our visual system adapts our perception to what is useful, the extent of the effect of "higher-level" relevance will naturally depend which range of factors influences usefulness in a particular stimulus.

A last consideration regarding the absence of an effect pertains to the sample size used in the current study, which was based on the range reported by previous studies. More specifically, 10 participants per condition may appear to lack sufficient statistical power to find an effect of auxiliary task relevance. Here, it is important to note that the dependent variable in experimental designs (in our case, the PSE) is usually a summary measure based on repeated measurements. Aggregating over these repeated measurements decreases measurement error, a factor not regularly taken into account in conducting power analyses. Consequently, in many psychophysical experiments with small sample sizes, statistical power is frequently much higher compared to studies where the dependent variable consists of a single measurement for each participant (Brysbaert & Stevens, 2018; Rouder & Haaf, 2018). Furthermore, our Bayes factor analysis indicates that the data are informative in the sense that they are more consistent with the absence rather than the presence of an effect.

'Shift to global'

Even though task relevance did not affect perceptual organization in our experiment, this does not necessarily suggest that relevance does not have an influence in any other way. As our results show, prior experience with the ambiguous motion stimulus affects its perception over time. The current experiment replicated earlier findings of a 'shift to global' (Anstis & Kim,

2011); The more observers see the ambiguous motion stimulus, the more they perceive it globally. Returning to the three levels of prior experience constituting perceptual relevance, the most probable candidate determining the 'shift to global' in our study would be *"low-level" exposure* to the same visual input. Since there was no differential effect for task relevance of either percept compared to the exposure task, "*higher-level" processes* (i.e. previously experienced cognitive goal to solve the global task) or any specific *"mid-level" percept* (i.e. previously experienced local or global percept) were not necessary to elicit a 'shift to global'. Moreover, since the 'shift to global' was present within one experimental session in the study of Anstis and Kim (2011), it is unlikely that it took as long as the time scale of our experiment (i.e. five short sessions spread across five successive days) to occur.

As was established previously, the absence of a differential effect of the relevance task on the 'shift to global' is most probably not due to the absence of a short-term attentional bias compared to previous studies (Chopin & Mamassian, 2011). However, attention may play a role in the 'shift to global' in general. In the exposure condition, no particular percept was necessary to perform the task. Nevertheless, a certain amount of attention to the stimulus was present in all conditions. Accordingly, the necessity of attention to the stimulus for a 'shift to global' remains an open question. Previous research has shown that attention influences perception of ambiguous stimuli (Meng & Tong, 2004). However, it is unclear why attention would bias the perception of ambiguous stimuli towards one particular perceptual interpretation (i.e. global) in the current study, irrespective of which perceptual interpretation the individual has previously payed attention to. Future research should clarify whether the 'shift to global' still occurs when attention is averted from the stimulus.

Another important question is whether the 'shift to global' occurs passively and purely stimulus-driven, versus actively and observer-driven. In the former case, the change in perception of the stimulus would occur in any observer, regardless of how the particular observer's visual system may interpret it. In the latter case, the change in perception of the stimulus is due to an observer-driven reinterpretation. As stated by Maloney et al. (2005, p. 3164): "We conclude that the visual system does not passively remember perceptual state: it analyzes recent perceptual history and attempts to predict what will come next. These predictions can alter what is seen". We may argue that the fact alone that a stimulus shows up in our environment more often might make it more relevant. More specifically, our visual

system may notice the reoccurrence of the stimulus and alter its perception to allow more efficient processing in the future. Accordingly, individuals may differ in the change of their perceptual dominance, depending on whether and how their visual system interprets the reoccurring stimulus. Indeed, there is individual variability in change in perception of the ambiguous motion stimulus (Figure 4).

Lastly, an essential issue is the role of globality in the perceptual organization of the ambiguous motion stimulus. The term globality here refers to the level of comprehensiveness and simplicity of the perceptual representation. For example, the percept of one diamond vs. four different bars in the bi-stable diamond (see example in de-Wit et al., 2012). In the study of Chopin and Mamassian (2011), both percepts of the ambiguous motion transparency stimulus were equal in globality. That is, the stimulus information was perceptually organized in an equal number of groups in both percepts. This was not the case for the ambiguous motion stimulus. As stated by Anstis and Kim (2011, p.11): "[The global percept] groups the maximum amount of evidence into the minimum number of perceptual hypotheses". Previous research has shown that effects of globality in a stimulus interact with effects of relevance of the stimulus (Poirel, Pineau, & Mellet, 2008). Therefore, it is possible that an effect of task relevance in the current experiment may have been overshadowed by an effect of globality. Our perceptual system might favor a global perceptual organization of a reoccurring stimulus because a configuration of multiple dots in two pulsating convex hexagons is a better Gestalt than six dot pairs moving coherently. Accordingly, it allows more efficient processing (Garner, 1974) and is stored more easily in memory (Attneave, 1955; Woodman, Vecera, & Luck, 2003). This rationale is in accordance with the notion that the purpose of perception is to be useful, not to resemble truth (Hoffman et al., 2015; Koenderink, 2011). The global perceptual organization might take some more time and effort when the stimulus is still unfamiliar compared to when it is more familiar and already represented, leading to an increase in global perceptual organization over time.

2.5 Conclusions

The results of our study imply that perception of the ambiguous motion stimulus used here (Figure 2) is not influenced by auxiliary task relevance, in contrast to previous studies where different ambiguous stimuli were used (Chopin & Mamassian, 2011: Harrison & Backus, 2010).

An effect of time showed that prolonged engagement with the ambiguous motion stimulus does result in a perceptual change - a 'shift to global' - irrespective of which percept was relevant in the auxiliary task during the relevance learning phase ('local', 'global' or neither). Although "low-level" exposure to the stimulus suffices to elicit the 'shift to global', we consider this to be an active process driven by the observer's visual system as well as by the stimulus. The necessity of attention to the ambiguous motion stimulus, as well as the role of globality of the percepts, remain open questions to be addressed in future research.

References

Anstis, S., & Kim, J. (2011). Local versus global perception of ambiguous motion displays. *Journal of Vision*, *11*(3): 13. https://doi.org/10.1167/11.3.13

Attneave, F. (1955). Symmetry, information, and memory for patterns. *The American Journal of Psychology*, *68*(2), 209-222.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13–21. https://doi.org/10.1038/nrn701

Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable Perception and the Role of the Frontoparietal Cortex in Perceptual Inference. *Annual Review of Psychology*, *69*(1), 77–103. https://doi.org/10.1146/annurev-psych-010417-085944

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: a tutorial. *Journal of Cognition*, *1*(1). https://doi.org/10.31234/osf.io/fahxc

Chopin, A., & Mamassian, P. (2010). Task Usefulness Affects Perception of Rivalrous Images. *Psychological Science*, *21*(12), 1886–1893. https://doi.org/10.1177/0956797610389190

Chopin, A., & Mamassian, P. (2011). Usefulness influences visual appearance in motion transparency depth rivalry. *Journal of Vision*, *11*(7): 18. https://doi.org/10.1167/11.7.18

de-Wit, L. H., Kubilius, J., Wagemans, J., & Op de Beeck, H. P. (2012). Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. *Journal of Vision*, *12*(11): 12. https://doi.org/10.1167/12.11.12

Denison, R. N., Piazza, E. A., & Silver, M. A. (2011). Predictive Context Influences Perceptual Selection during Binocular Rivalry. *Frontiers in Human Neuroscience*, *5*. https://doi.org/10.3389/fnhum.2011.00166

Di Luca, M., Ernst, M. O., & Backus, B. T. (2010). Learning to Use an Invisible Visual Signal for Perception. *Current Biology*, *20*(20), 1860–1863. https://doi.org/10.1016/j.cub.2010.09.047

Feigin, H., Baror, S., Bar, M., & Zaidel, A. (2017). Behaviorally relevant prior experience biases subsequent perception. *Journal of Vision*, *17*(10): 493. https://doi.org/10.1167/17.10.493

Garner, W. R. (1974). *The processing of information and structure*. Oxford, U.K.: Erlbaum.

Harrison, S. J., & Backus, B. T. (2010). Uninformative visual experience establishes long term perceptual bias. *Vision Research*, *50*(18), 1905–1911. https://doi.org/10.1016/j.visres.2010.06.013

Hoffman, D. D., Singh, M., & Prakash, C. (2015). The Interface Theory of Perception. *Psychonomic Bulletin & Review*, *22*(6), 1480–1506. https://doi.org/10.3758/s13423-015-0890-8

Jeffreys, H. (1961). *Theory of probability*, Oxford: Oxford University Press.

Koenderink, J. (2011). Vision as a user interface. *Human Vision and Electronic Imaging XVI* (Vol. 7865), 786504-786513. https://doi.org/10.1117/12.881671

Kornmeier, J., & Bach, M. (2012). Ambiguous figures – What happens in the brain when perception changes but not the stimulus. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00051

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729–2737. https://doi.org/10.1016/S0042-6989(98)00285-5

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254–264. https://doi.org/10.1016/S1364-6613(99)01332-7

Maloney, L. T., Martello, M. F. D., Sahm, C., & Spillmann, L. (2005). Past trials influence perception of ambiguous motion quartets through pattern completion. *Proceedings of the National Academy of Sciences of the USA*, *102*(8), 3164–3169. https://doi.org/10.1073/pnas.0407157102

McOwan, P. W., & Johnston, A. (1996). Motion transparency arises from perceptual grouping: evidence from luminance and contrast modulation motion displays. *Current Biology*, *6*(10), 1343–1346. https://doi.org/10.1016/S0960-9822(02)70722-7

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, *4*(7): 2. https://doi.org/10.1167/4.7.2

Morey, R. D., & Rouder, J. N. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-2.* From https://CRAN.R-project.org/package=BayesFactor

Necker, L. A. (1832). LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a

crystal or geometrical solid. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *1*(5), 329–337. https://doi.org/10.1080/14786443208647909

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. https://doi.org/10.1167/16.14.11

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Piazza, E. A., Denison, R. N., & Silver, M. A. (2018). Recent cross-modal statistical learning influences visual perceptual selection. *Journal of Vision*, *18*(3): 1. https://doi.org/10.1167/18.3.1

Poirel, N., Pineau, A., & Mellet, E. (2008). What does the nature of the stimuli tell us about the Global Precedence Effect? *Acta Psychologica*, *127*(1), 1–11. https://doi.org/10.1016/j.actpsy.2006.12.001

R Core Team. (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rouder, J. N., & Haaf, J. M. (2018). Power, dominance, and constraint: A note on the appeal of different design traditions. *Advances in Methods and Practices in Psychological Science*, *1*(1), 19-26. https://doi.org/10.1177/2515245917745058

Schmack, K., Weilnhammer, V., Heinzle, J., Stephan, K. E., & Sterzer, P. (2016). Learning What to See in a Changing World. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00263

Tong, F. (2001). Competing theories of binocular rivalry: A possible resolution. *Brain and Mind*, 2 (1), 55–83, https://doi.org/10.1023/A:1017942718744.

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* (Chapter 18, pp. 803-872). Hoboken, NJ: Wiley.

Weilnhammer, V. A., Stuke, H., Sterzer, P., & Schmack, K. (2018). The Neural Correlates of Hierarchical Predictions for Perceptual Decisions. *The Journal of Neuroscience*, *38*(21), 5008–5021. https://doi.org/10.1523/JNEUROSCI.2901-17.2018

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt, II. *Psychologische Forschung, 4,* 301-350. [Translated as “Investigations on Gestalt principles.” In L. Spillmann (Ed.), (2012). *Max Wertheimer. On perceived motion and figural organization* (pp. 127-182). Cambridge, MA: MIT Press.]

Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, *10*(1), 80–87. https://doi.org/10.3758/BF03196470

3. EEG study:

EEG correlates of perception of the ambiguous vs. disambiguated 'local-global' motion stimulus

3.1 Part 1: Perception of the ambiguous vs. disambiguated 'local-global' stimulus over time

3.1.1 Introduction

Visual perception is the result of a dynamic interaction process between stimulus and observer characteristics. There is no one-on-one-relationship between a stimulus and a percept. In order to make sense of the world, our perceptual system interprets the limited sensory information available and constructs rich percepts. This perceptual interpretation is not static, and may differ between individuals (Miller, & Saygin, 2013; Grzeczkowski et al., 2017; Takeuchi et al., 2017; Phillips, 2019) or within an individual over time (Gilbert, 1996; Molenaar, 2004; Karmarkar, & Dan, 2006; Friston, & Kiebel, 2009; Klink et al., 2010; Hamaker, 2012; Koenderink, 2019; Jia et al., 2020; Grasso et al., 2020).

Multistable perception of ambiguous stimuli provides a good illustration of the potential variability of perception: the same visual stimulus elicits more than one percept within an observer over time and/or between different observers (Kornmeier, & Bach, 2012; Brascamp, Sterzer, Blake, & Knapen, 2018). For instance, consider the Necker cube; One image enabling the percept of both a cube facing the upper right corner, and a cube facing the lower left corner (Figure 2 in the General Introduction of this PhD thesis). While the physical image remains the same, the observer's percept may switch from a cube facing the upper right corner to a cube facing the lower left corner. When examining what has changed from this one moment to the next, we may uncover processes that are specific to the formation of a percept, but unrelated to any changes in the stimulus itself.

Similarly, consider the ambiguous motion stimulus (Figure 1) which consists of multiple frames of multiple dots, creating a percept of motion when the frames are shown in rapid succession. This apparent motion may be perceived in two ways: the 'local' and 'global' percept. The 'local' motion percept consists of 6 pairs of dots positioned in a circle, and rotating around each other within the pair. The 'global' motion percept consists of 2 groups of 6 dots making up 2 circles or hexagons, and pulsating inwards and outwards. When perceiving the ambiguous motion stimulus for the first time, most observers experience a 'local' percept. However, once an observer has seen the 'global' percept as well, perception tends to oscillate between both percepts (Anstis, & Kim, 2011).

In a previous study (Boeykens, Wagemans, & Moors, 2019), we found that perception of the ambiguous motion stimulus was characterized by a 'shift to global' over time. Observers came to the lab during 5 successive days. On the first day, the point of subjective equality (PSE) between the local and global percept was measured by means of an adaptive psychophysical procedure. The PSE is a measure of stimulus characteristics for which the stimulus is maximally ambiguous for the observer (i.e. enabling an equal amount of 'local' and 'global' percepts). Subsequently, participants were assigned to one of 3 conditions and performed a short task by means of which the relevant percept of the ambiguous motion stimulus was manipulated. They were instructed to either indicate the local rotation direction (local condition), global pulsation direction (global condition), or a change in the color of the central fixation point (mere exposure condition). During the next 4 days, participants repeated this short task in the lab every day. At the end of the fifth and final day, we measured their point of subjective equality for a second time. We found that, for most observers, percepts had changed towards a more dominant global percept, irrespective of the condition they were in during the week. This indicates that, even though the stimulus remains identical, an observer's percept can change over time, seemingly automatically and irrespective of behavioral goals in an auxiliary task.

In this previous study, we intended to influence a change in perception by means of a manipulation of task relevance, which we extended over multiple short sessions in the course of 5 days in order to have a sufficiently strong effect. However, an earlier study (Anstis & Kim, 2011), as well as pilot testing in our lab, had indicated that the 'shift to global' can actually already occur over a much shorter time span (i.e., during one short experimental session)

compared to our study. Since this automatic change in perception is not dependent on the long duration of the behavioral manipulation, we decided that the time span for studying these dynamics in perception in the current study could be much shorter, focusing on changes which may occur during one experimental session.

Neurophysiological correlates of perception of the ambiguous motion stimulus over time may provide us with important insights into perceptual dynamics and the emergence of a percept in general. Even though it is known that there is a 'shift to global' (Anstis & Kim, 2011), the literature has yet to explore the perception of the ambiguous motion stimulus and its neural correlates over time. Moreover, the change in perception over time is sometimes corrected for in previous studies in order to ensure stability in the ambiguity of the stimulus; Zaretskaya et al. (2013) adjusted the speed of dot rotation after each part in their experiment in order to counter-act the 'shift to global'.

Previous neurophysiological research employing the ambiguous motion stimulus has predominantly looked into fMRI BOLD activity (Zaretskaya, Anstis & Bartels, 2013; Grassi, Zaretskaya & Bartels, 2016, Grassi, Zaretskaya & Bartels, 2017). These studies mainly focused on the difference in neural correlates between a 'local' percept and a 'global' percept. More specifically, they found that 'global' perception of the ambiguous motion stimulus is associated with a posterior parietal increase in activity and a decrease in early visual areas. Little electro-encephalography (EEG) research has been carried out with the ambiguous motion stimulus. One existing study (Zaretskaya & Bartels, 2015) examined perceptual reports during continuous presentation of the ambiguous motion stimulus and found that perception of the 'global' percept, which they referred to as the 'Gestalt' percept, was accompanied by reduced power of the EEG signal in the beta frequency band, which was source localized to the posterior parietal cortex.

Although fMRI research can provide some interesting insight into the spatial location of brain processes, its temporal resolution is rather low. In the current study, we employed EEG measures to study perception of the ambiguous motion stimulus over time. In order to do so, we first sought to determine which EEG correlates of the perception of the ambiguous motion stimulus at one static point in time might be interesting to explore. Subsequently, we could examine the dynamics of these EEG correlates over time. We started by considering the

neurophysiological measures that were previously studied with the same stimulus, as discussed above. However, even though analyzing perceptual reports during continuous presentation (akin to Zaretskaya & Bartels, 2015) can be an interesting tool to study neural correlates of perceptual switching, it lacks the temporal resolution needed to identify the dynamics of underlying brain processes.

Therefore, we chose to explore an EEG measure with a higher temporal resolution. Kornmeier and Bach (2012) developed a paradigm in which ambiguous stimuli are presented intermittently. They showed that, dependent on the timing of the inter-stimulus-interval (ISI), perceptual dynamics for intermitted presentation could be comparable to continuous presentation. In contrast to the classic continuous presentation paradigm, where the manual response indicating a change in percept is used as the time reference for the event-related potential (ERP) of the EEG signal, the time reference during intermittent presentation is stimulus onset. This enabled a much more precise temporal resolution, allowing the authors to uncover different components associated to the formation of a percept.

Kornmeier et al. (2016, 2020) found that perception of disambiguated stimuli compared to perception of various ambiguous stimuli is characterized by a clear difference in amplitude of two EEG event-related potentials (ERP) in fronto-central and parieto-central electrode positions. They named the ERP P200 and P400, with timings starting around 200 and 400 ms after stimulus-onset, respectively. They interpreted this effect as a difference in the uncertainty and ambiguity associated with the evaluation of a perceptual response. The stimuli employed in the studies by Kornmeier et al. (2016; 2020) were highly variable, ranging from a Necker cube to stroboscopic apparent motion (SAM) and smileys. Nevertheless, the 'uncertainty' effect they described remained very similar in timing and topography across these different stimulus types, leading them to conclude that the ERP reflects a "higher-level", abstract representation of the perceptual interpretation rather than "low-level" stimulus characteristics.

Therefore, we also predict to find an 'uncertainty effect' when comparing the ERP for the ambiguous motion stimulus to the ERP for disambiguated versions of this stimulus. The ambiguous motion stimulus is different from all previous stimuli tested by Kornmeier et al. (2020) in the sense that it is a moving stimulus, consisting of many more than 1 (or, in the case

of the SAM stimulus, 2) static images. Thus, it would be interesting to test whether this effect generalizes across this type of stimulus as well. Given that the ERP 'ambiguity/uncertainty effect' corresponds to the uncertainty of the perceptual outcome, the ERP uncertainty effect should be largest for maximally ambiguous stimuli (i.e., at the point of subjective equality between the possible percepts) compared to disambiguated stimulus variants. Accordingly, in the case of the ambiguous motion stimulus, the ERP uncertainty effect should be largest for stimuli with distance G* at the PSE compared to disambiguated motion stimuli. As previously mentioned, we know that when an observer has prolonged experience with / exposure to the ambiguous motion stimulus, the global percept will become more dominant. This implies that an identical ambiguous motion stimulus with distance G* that was maximally ambiguous (i.e., an equal number of 'local' and 'global' percepts) at the beginning of an experimental session, will be less ambiguous and predominantly perceived as 'global' by most observers towards the end of the experiment, due to the 'shift to global'. Therefore, we predict that towards the end of the experimental session, the ERP 'uncertainty effect' of the difference between perception of disambiguated stimulus variants and the perception of the now less ambiguous (i.e. more 'global') motion stimuli should be smaller.

3.1.2 Methods

Participants

25 university students (5 male, mean age = 21.24, *SD* = 2.96) participated in the study in return for course credit or monetary reward. The number of participants was based on the studies of Kornmeier et al. (Kornmeier, Wörner & Bach, 2016; Joos et al., 2020) employing an equivalent paradigm, in which 10 and 20 observers participated, respectively. Given that participants may need to be excluded from the analysis and a different stimulus is employed in this study, we decided to slightly increase the number of participants. 5 participants had to be excluded from the analysis because of various reasons (i.e. 2 participants were excluded before the EEG experiment started because they were unable to perceive the local and/or global percept, 1 participant took off her eyesight-correcting contacts in the middle of the experiment, and 2 participants did not report a sufficient amount of perceptual switches for the analysis based on a criterion of minimum perceptual switches per block, which was determined beforehand), leaving 20 participants (3 male, mean age = 20.8, *SD* = 2.84) to be

included in the analysis. All participants had normal or corrected-to-normal vision, provided informed consent, and were naïve with respect to the goals of the study. The study was approved by the Social and Societal Ethics Committee of KU Leuven (SMEC, approval code G-2018 03 1186) and adhered to the tenets of the Declaration of Helsinki.

Apparatus

All stimuli were displayed on an LCD monitor (Eizo FlexScan S2410W) with a refresh rate of 60 Hz, a diagonal length of 61.2 cm and a resolution of 1920 x 1200 pixels. Participants were positioned in a dark room and a chin rest was used to stabilize their head at a viewing distance of 57 cm from the monitor. Stimulus presentation and manual response registration was controlled through the experiment computer by software programmed in Python using the PsychoPy library (Peirce, 2007).

Stimuli

Ambiguous stimulus

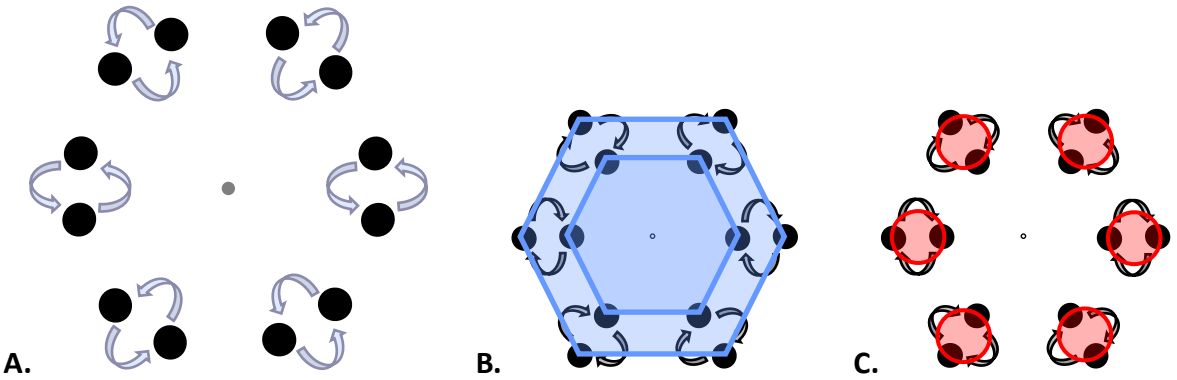


Figure 1. A. Static presentation of the ambiguous motion stimulus. Light blue arrows indicate the direction of the apparent motion. A video of a motion demonstration can be found in the files in the following link, where we recommend reading the wiki information to experience the 'local' and 'global' percepts of this stimulus: https://osf.io/du8fx/?view_only=e9a8eb6a4b294d20b1e9f911fe63833b. **B.** A representation of the global percept, consisting of two groups of six dots making up two circles or hexagons, pulsating inwards and outwards. **C.** A representation of the local percept, consisting of six rotating dot pairs.

The ambiguous motion stimulus (Figure 1) consists of 6 pairs of white dots (luminance 52.5 cd/m^2) displayed on a uniform grey background (luminance 11 cd/m^2) around a white fixation point. The dots have a radius of 0.2 deg and are placed in pairs with an edge-to-edge distance

of 0.39 deg. This distance may also be regarded as a phase shift of π rad around the pair-center. The 6 dot pairs are placed around a central fixation point with a distance G from the center. Since the dot pairs are placed in a hexagon, distance G is also equal to the distance between the dot pairs. The starting position of the dots varies according to the phase of their pair-centers relative to the common center. Accordingly, the position of a dot is not only related to the other dot in its pair by a phase shift around the pair-center, but also to other dots by the relation to the common center. Dots rotate around their pair-center at a speed of 2π rad/s for 30 frames, or 500 ms, per stimulus presentation. This rotation is either clockwise or counter-clockwise, determined randomly at the start of each trial. At the end of each stimulus presentation, the position of each dot pair differs from the starting position with a rotation of π rad around the pair-center.

During stimulus presentation, a local and a global percept are possible. The local percepts consists of 6 smaller groups of 2 dots, circling around in pairs. The global percept consists of 2 larger groups of 6 dots, shrinking and growing in the shape of a hexagon. The stimulus is ambiguous as multiple percepts (within- and between-individuals) of one stimulus are possible. The likeliness of either percept can be manipulated by means of changing stimulus characteristics. When distance G in the ambiguous motion stimulus is larger, a local percept becomes more likely and a global percept becomes less likely. Vice versa, when distance G is smaller, a global percept becomes more likely and a local percept becomes less likely.

Disambiguated stimuli

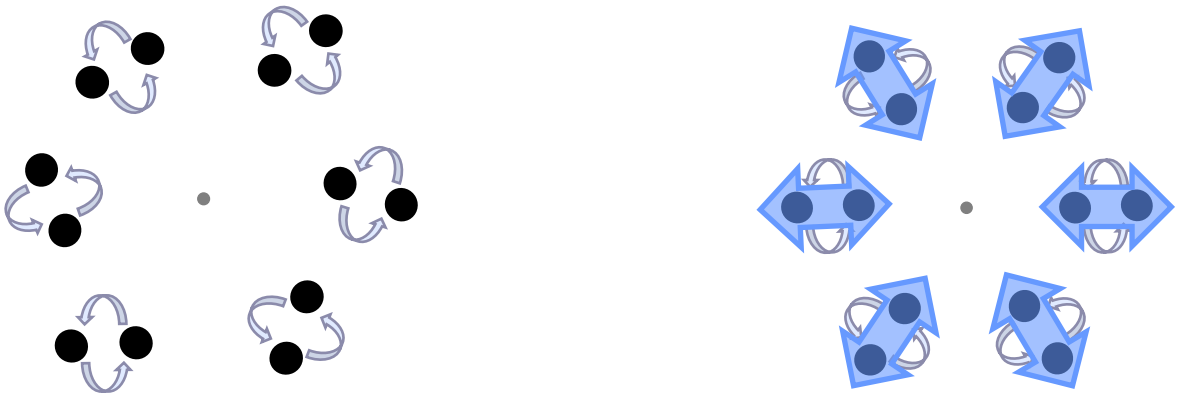


Figure 2. Static presentations of the local (left) and global (right) disambiguated versions of the 'local-global' motion stimulus. The creation of the local stimulus was enabled by a position and phase shift of the starting position of the dot pairs. The creation of the global stimulus was enabled by a dynamic addition and subtraction of the distance between the elements of a dot-pair (depicted with blue arrows), enforcing the perception of pulsation. Videos of motion demonstrations can be found in the files in this link: https://osf.io/hrn4y/?view_only=3267398b15d8407c8b645e6529228e28.

Disambiguated stimuli consist of a local and a global motion stimulus, corresponding to the local and global percept of the ambiguous motion stimulus, respectively. The disambiguated stimuli are identical to the ambiguous motion stimulus aside from certain manipulations influencing the probability of either percept. In the disambiguated local stimulus (Figure 2), the probability of a local percept is enhanced by means of adding random variation to the starting position and phase of each dot pair, maximizing local grouping cues and minimizing global grouping cues. In the disambiguated global stimulus (Figure 2), the probability of the global percept is enhanced by means of dynamically changing the distance between the dots in a dot pair during stimulus presentation, maximizing global grouping cues and minimizing local grouping cues.

Design and Procedure

When participants arrived in the lab, we started by showing them a short demonstration of the ambiguous motion stimulus in order to familiarize them with the 'local' and 'global' percepts. When a participant was unable to perceive either or both percepts, they were excluded from the further parts of the experiment.

As mentioned in the Stimulus section, distance G can be manipulated in order to change the likeliness of the local and global percept. We can estimate for which distance G, an observer reaches a point of subjective equality (PSE, or G*). At the PSE, both the local and global percept are equally likely for an observer at a particular point in time. Every observer's PSE, or G*, was measured at the beginning of the experiment. Subsequently, the EEG experiment was carried out with stimuli that had distance G = G*, in order to maximize stimulus ambiguity for each participant. After the EEG experiment, the individual PSE was measured again.

Measurement of PSE by means of adjustment

Participants reported their PSE by means of a method of adjustment in 20 trials, which took approximately 10 minutes in total. During each trial, the ambiguous motion stimulus was repeatedly presented for 500 ms, interleaved by a variable inter-stimulus-interval (ISI) between 450 and 650 ms, until a response is given. Distance G in the first stimulus presentation of each trial differed dependent on the parity of the trial number; G started from 1.08 deg in odd trials and started from 7.84 deg in even trials. Participants were instructed to slowly slide a bar on the bottom of the screen while monitoring their percept, thereby changing distance G (i.e. making G larger in odd trials and smaller in even trials). Stimulus presentation was repeated until a response, corresponding to a change in percept (i.e., global to local for odd trials, local to global for even trials), was given. When observers noticed that their percept had changed, they pressed the space bar and a new trial was initiated. The individual point of subjective equality, G*, between the local and global percept was estimated by taking the mean of all values of G corresponding to a perceptual change response.

EEG experiment

The EEG experimental design differentiated between presentations of ambiguous stimuli and disambiguated stimuli. Trials were divided into 4 blocks of 900 stimulus presentations. In half of the blocks, ambiguous stimuli were presented (i.e., A for ambiguous blocks), in the other half disambiguated stimuli were presented (i.e., D for disambiguated blocks). This resulted in the following two possible block orders (and trial types): A-D-A-D and D-A-D-A, counter-balanced across participants. Within each block, stimulus presentations with a duration of 500

ms were repeated, alternated by a variable ISI around 500 ms, until a response was given. In the ambiguous blocks, participants were presented with repetitions of the ambiguous stimulus. In the disambiguated blocks, participants were presented with alternations of the disambiguated local and global stimulus. The frequency of alternations of disambiguated stimuli was based on the average frequency of perceptual switches in a pilot experiment with added random variability. Participants were instructed to respond when they experienced a perceptual switch between two successive stimulus presentations (left arrow for Local-Global, right arrow for Global-Local, or vice versa) in the same way in the ambiguous and disambiguated blocks. When a perceptual switch response was given, the ISI was extended to a variable length between 1000 and 1300 ms. Participants could take a short break after each block and the EEG experiment part took approximately 1,5 h in total.

EEG recording

Electro-encephalogram (EEG) activity was recorded during the EEG experiment part of the session by means of a Geodesic Sensor Net with 256 Ag/AgCl electrodes. The recorded signal was amplified through a high input impedance Net Amps amplifier (EGI) and Net Station software on the EEG computer. An analog high-pass filter of 0.1 Hz and a 100-Hz low-pass filter were used. The EEG data were digitalized at a sampling rate of 250 Hz and impedance was kept below 50 kΩ. The activity in all channels was referenced to the vertex electrode (Cz) during recording. The EEG electrode net also recorded vertical and horizontal electro-oculograms (EOG). In order to assure a correct and precise timing estimation of all triggers sent from the experiment computer to the monitor and EEG system, a photodiode recording exact stimulus onset timing was set up in a small corner of the presentation monitor. The photodiode measured a change in luminance of a small square in the corner of the monitor that had a presentation onset which was synchronized in time with our stimulus presentation. The square was covered by the photodiode setup and not visible to participants.

Behavioral data analysis

Behavioral data analysis was carried out in R (version 4.1.0). We tested for a significant change of observers' G*, or PSE, measurements over the course of the experimental session by means of a paired samples t-test.

EEG data ERP analysis

Preprocessing

Preprocessing of raw EEG data was carried out with BrainVision Analyzer software (Brain Products, Gilching, Germany). EEG raw data were filtered with a low cutoff frequency of 0.5 Hz (and a time constant of 0.3, 48 dB/oct), a high cutoff of 30 Hz (48 dB/oct) and a Notch filter of 50 Hz. Channels located in the outer two rings, partially covering the neck of participants, as well as electrodes around the ears and cheeks were excluded from analysis in order to exclude motor artifacts, resulting in 156 remaining electrodes for all participants. A visual search for potential noisy channels per participant was carried out on the remaining electrodes. In the rare cases that a channel was excessively noisy, it was excluded and interpolated. Horizontal and vertical EOG (i.e., HEOG and VEOG) channels were created by means of a linear derivation of electrodes around the observer's eye. These HEOG and VEOG channels were employed later on in the independent component analysis (ICA), identifying eye movement artifacts. The continuous EEG signal was segmented into epochs of 800 ms, spanning from 100 ms before stimulus onset to 700 ms after stimulus onset.

Subsequently, ocular correction and identification of other artifacts was carried out by means of an ICA. Components relevant for vertical and horizontal ocular artifact activity were automatically defined by means of their correlation with the HEOG and VEOG channels, respectively. ICA components were rejected when the sum of squared correlations with the HEOG or the VEOG channel exceeded 0.2. Moreover, all components were visually inspected per participant and components with high levels of noise were manually set to zero. In a next step, artifact rejection was carried out by means of visual inspection as well as automatically marked epochs in which channels of interest (i.e., parieto-central electrode locations) had voltage steps larger than 50 µV/ms or differences larger than 100 µV in intervals of 100 ms.

Lastly, epochs were collapsed and averaged according to the experimental conditions of interest, activity was re-referenced to average activity of all remaining channels, and baseline correction was carried out relative to the 100 ms pre-stimulus interval. In order to avoid overlap of the effects of interest (i.e., uncertainty effect P200 and P400) with the P300 component, non-response trials were analyzed (akin to Kornmeier, 2016). Hence, the epochs

included in the statistical analysis were restricted to stimulus onsets after which there was no perceptual switch response.

Based on the studies of Kornmeier et al. (2016; 2020), 6 regions of interest and 2 timings of interest were defined for our statistical analysis.

Statistical analysis

The statistical analysis of the EEG data had a 2x2 design assessing the effect of ambiguity (ambiguous vs. disambiguated stimulus presentations) and part (beginning or first half vs. end or second half of the experiment).

In this statistical analysis, we aim to examine whether the uncertainty effect changes over time. However, in order to interpret any effect of part (beginning vs. end) in relation to ambiguity, there should also be an interaction effect with ambiguity (ambiguous vs. disambiguated stimulus presentations), as a general difference of EEG ERP components may be due to a wide variety of factors, such as fatigue, attention, etc. More specifically, we hope to find an interaction effect of part with ambiguity which indicates that a change in the ERP over time only occurs for ambiguous stimuli and does not occur for disambiguated stimuli. Accordingly, it is more likely that the difference in the ERP between the beginning and end of the experiment is related to the ambiguity of the stimulus, suggesting that the changes over time are not merely general or undefined.

Hypothesis driven statistical analysis was carried out in R (R Core Team). The averaged EEG segments per channel, participant and experimental condition were imported from exported preprocessed EEG files from Brain Vision Analyzer. Electrode regions of interest (ROI) and time windows of interest were selected based on the studies of Kornmeier et al. (2016), in which the P200 and P400 uncertainty effects were defined. The following eight spatial ROI were defined: FR, FL, Cz, Pz, CPL, CPR, OR and OL. Each ROI was computed by taking the average activity of 5 surrounding electrodes (Figure 4). For the P200 effect, we chose to focus on the fronto-central electrode sites (i.e., FR, FL, Cz, Pz) which were identified in the study of Kornmeier et al. (2016). For the P400 effect, we chose to focus on the centro-parietal electrodes (i.e., Cz, CPL, CPR, Pz), instead of central electrodes only, because the strongest effects for the stimulus employed in the study of Kornmeier et al. (2016), resembling our

stimulus –the motion quartet–the most, were also located over parieto-central electrode sites. In order to assess whether the effects of the P200 and P400 components are in accordance with early, "low-level" ERP effects, we also examined the P1 component by selecting time points between 80 and 120 ms post stimulus-onset over left and right occipital electrode sites (i.e., OR and OL). The P1 component is the first positive component of an event-related potential, occurring approximately 100 ms post stimulus-onset and is thought to reflect initial sensory and attentional processing of a stimulus (Luck, 2014). For the P200 analysis, we selected time points between 150 and 250 ms post stimulus-onset. For the P400 analysis, time points between 300 and 500 ms post stimulus-onset were selected. EEG segments were collapsed and averaged per ROI and the mean voltage was extracted per time window. We conducted an analysis of variance (ANOVA) separately per time window with 4 levels of ROI (i.e., Cz, Pz, FR and FL for the P200 effect and Cz, Pz, CPL and CPR for the P400 effect), 2 levels of ambiguity (i.e., ambiguous vs. disambiguated) and 2 levels of part (i.e., first vs. second part of the experimental session). For the estimated magnitudes of the effects in our statistical analysis, we employed generalized eta and omega squared effect sizes (ges) (Olejnik & Algina, 2003).

3.1.3 Behavioral Results & Discussion

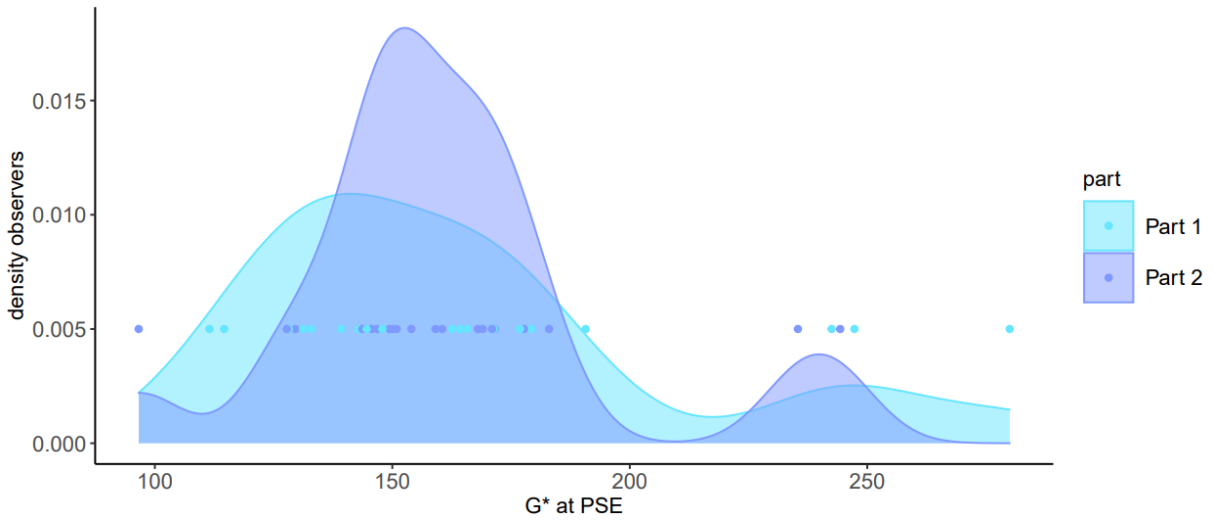


Figure 3. Individual measures of the point of subjective equality, or G*, before and after the EEG session.

Our results indicate that there was no significant difference between the PSE, or G*, at the beginning (*M* = 161.71, *SD* = 33.19) compared to the end (*M* = 165.095, *SD* = 45.42) of the experimental session; t(19) = -0.599, p = 0.556 (Figure 3). In other words, participants did not differ significantly in their percept over time and there was no significant 'shift to global', which is not in accordance with our expectations based on previous studies (Anstis & Kim, 2012; Boeykens, Wagemans & Moors, 2019). As our measure of the PSE was based on 20 trials of adjustment, instead of the more precise adaptive staircase procedure in our previous study (Boeykens et al., 2019), we decided to explore the behavioral data of the EEG experiment as well, in order to determine whether there were any clear signals of a 'shift to global'. We computed the number of non-response stimulus presentations after a perceptual switch was reported, which should be associated to the duration of a stable percept. We compared this approximation of the length of stable local and global perception across the two parts of the experiment and found that there was only a trend towards a 'shift to global' which was not statistically significant.

There may be various reasons explaining the absence of a behavioral 'shift to global' in the current experiment. A first consideration to be made is that the 'shift to global' may not be as robust as previously assumed. Nevertheless, as the 'shift to global' was relatively strong (*BF* = 4.865 for the best model including the effect of time compared to the best model without the effect of time) in our previous study and as we replicated it in pilot studies as well, it is important to note the differences between the current study and the previous literature.

As ambiguous stimulus blocks were alternated by disambiguated stimulus blocks, contrary to previous studies, it may be possible that the presentation and/or perception of disambiguated stimulus variants of the ambiguous motion stimulus counter-acted or reset the 'shift to global'. Secondly, the stimuli in our previous study (Boeykens, Wagemans & Moors, 2019) were different from the stimuli in the current study because of the color scheme employed for the relevance manipulation. More specifically, one dot in each dot pair was colored black in the previous study, whereas all dots were colored white in the current study. It may be possible that the 'shift to global' over time is enhanced by the stimulus cues favoring a global percept in our previous study. Moreover, the duration of one experimental session was much shorter than the total duration of 5 short sessions across 5 successive days in our previous study. The study of Anstis and Kim (2011) also identified the 'shift to global' and reported its

occurrence over the course of one experimental session. However, their stimuli were also slightly different from ours (i.e., continuous presentation and a variable number of dot pairs instead of 6 dot pairs). Therefore, it is likely that the 'shift to global' takes more time to occur for the currently employed stimulus and most of the observers in the current study. Results from a preceding pilot study for the current study did replicate the 'shift to global' and were obtained with alternated ambiguous and disambiguated blocks, the same stimuli (i.e., all dots were colored white) and over the same shorter time course. Nevertheless, a pilot study does not provide a reliable test as the sample size is too small.

Lastly, even though we observed a 'shift to global' in our previous study on average, there was also a substantial amount of inter-individual variability in the change of perception of the ambiguous motion stimulus. It is possible that the individual variability in the sample of observers in our current study was higher, making it more difficult to see a 'shift to global' for the majority of observers.

3.1.4 EEG results

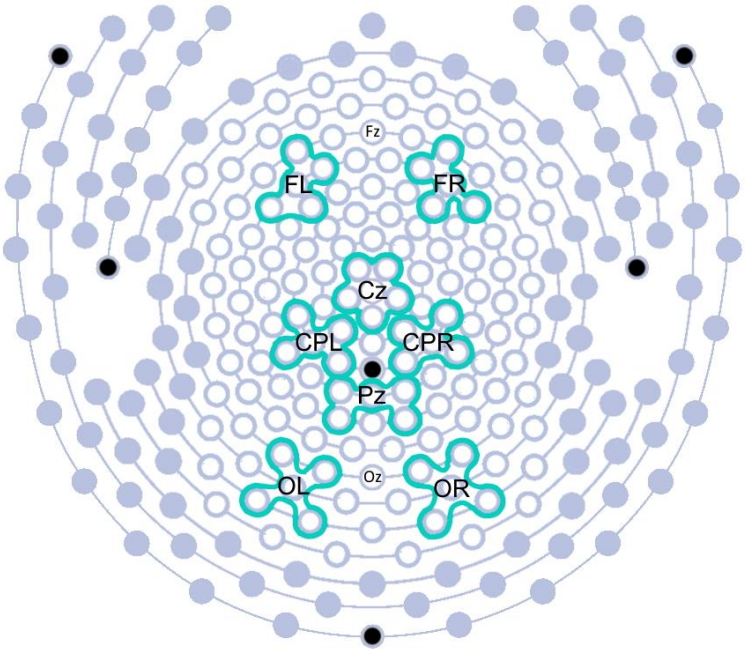


Figure 4. ROIs on the electrode net

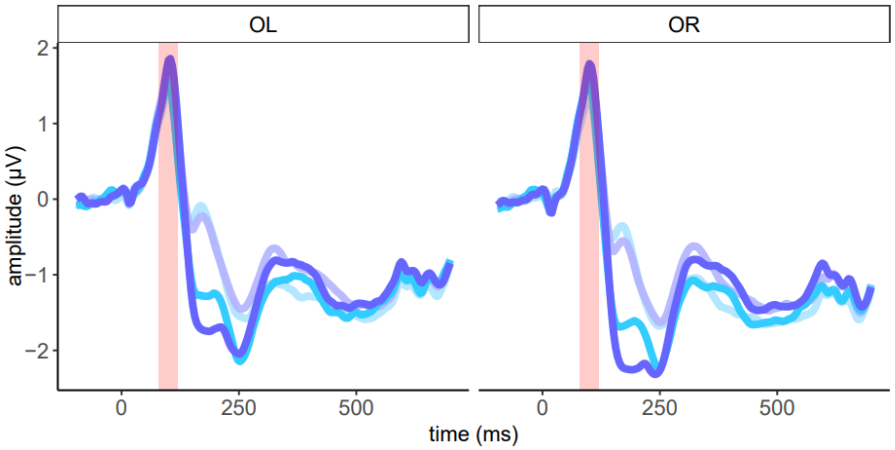


Figure 5. Event-Related Potential (ERP) traces for both percept and ambiguity conditions in 2 electrode ROIs for early visual components: occipital left (OL) and occipital right (OR). Vertical areas in pink indicate the time windows of interest for the P1 component.

We found no significant differences between the conditions of interest in the early visual component P1, indicating that the differences between the conditions of interest that are discussed in the time windows for the P200 and P400 effect are most probably not influenced by differences in stimulus characteristics or low-level visual processing.

P200

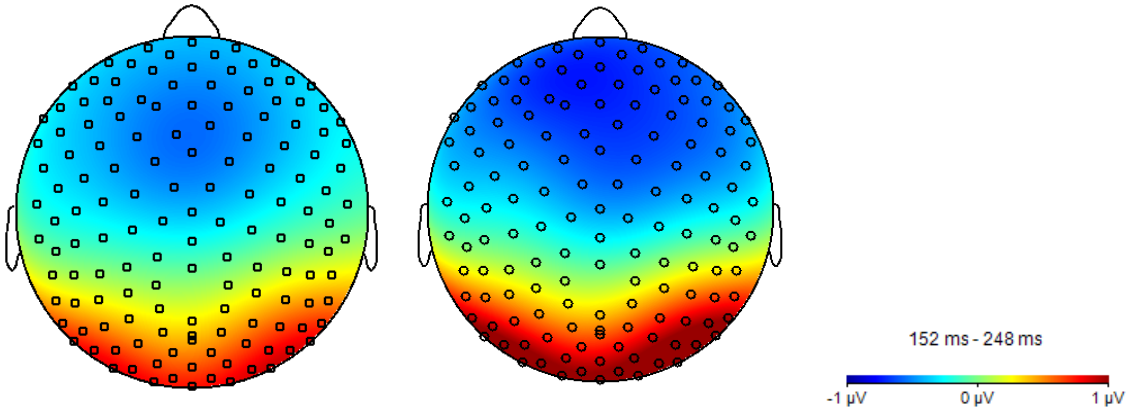


Figure 6. Topographic maps of difference waves for the P200 uncertainty effect in the time window of 150-250 ms post-stimulus-onset in the first (left) and second (right) part of the experiment. Voltages of disambiguated stimulus presentations were subtracted from ambiguous stimulus presentations.

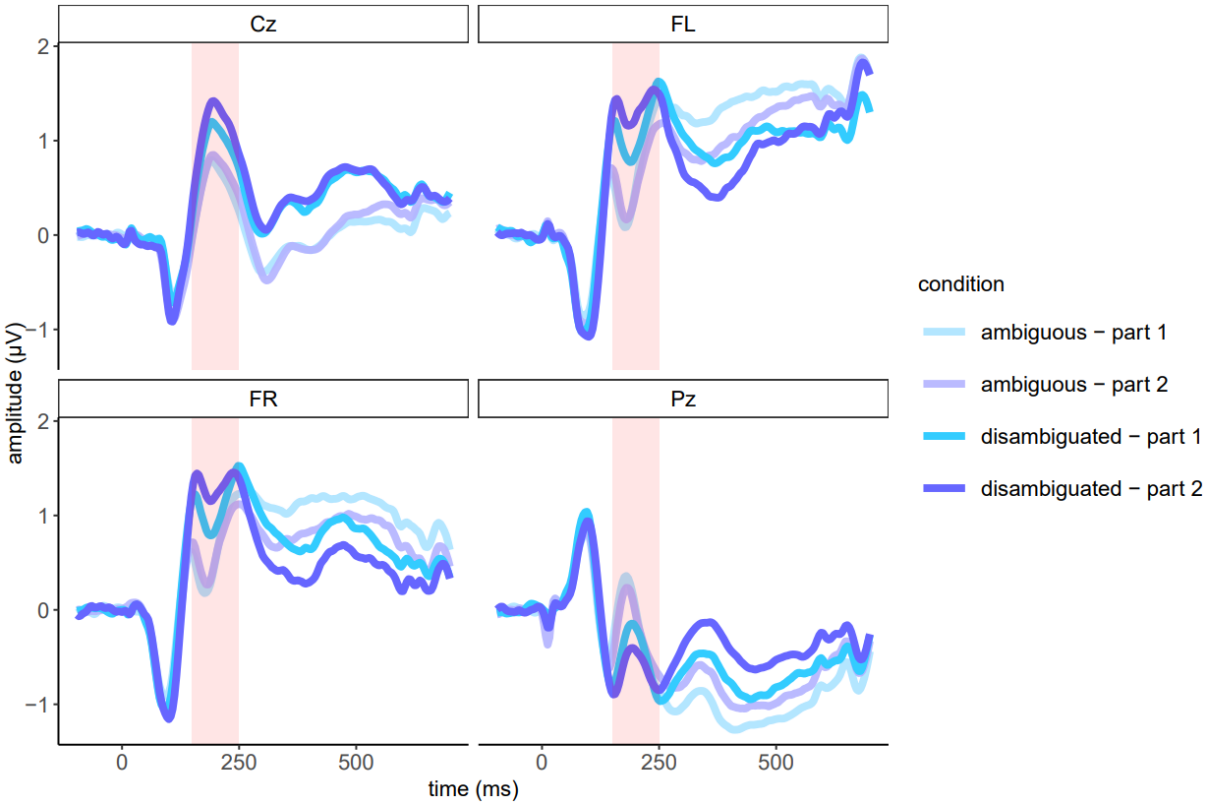
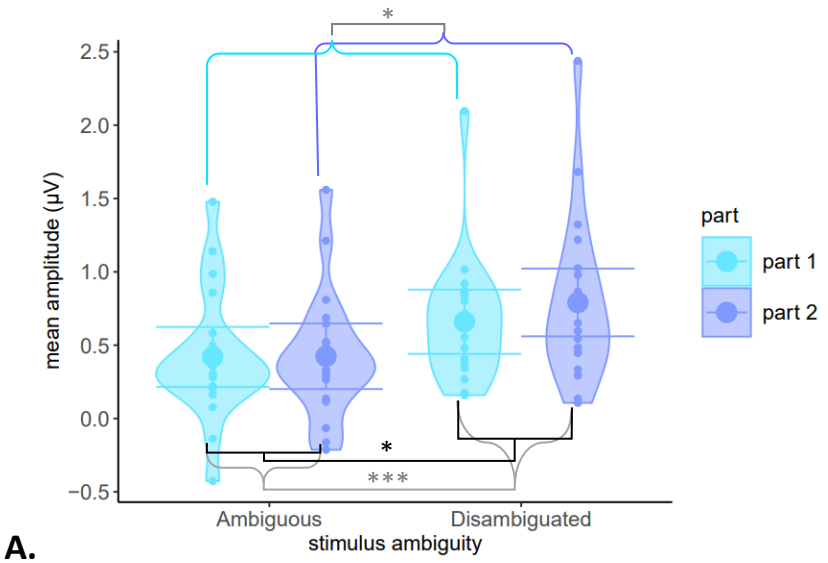


Figure 7. Event-Related Potential (ERP) traces for both parts and ambiguity of stimulus conditions in 4 electrode ROIs: Cz, FL, FR, and Pz. Vertical areas in pink indicate the time window of interest for the P200 effect.

Table 1. Summary of the ANOVA in the time window of the P200 effect.

Effect	df	MSE	F	ges	p value
Ambiguity	1, 19	0.20	37.03 ***	.024	<.001
Part	1, 19	0.08	4.46 *	.001	.048
ROI	1.52, 28.83	7.29	8.83 **	.247	.002
Ambiguity x Part	1, 19	0.07	4.63 *	.001	.045
Ambiguity x ROI	1.64, 31.11	0.33	21.94 ***	.038	<.001
Part x ROI	1.46, 27.72	0.26	1.61	.002	.219
Ambiguity x Part x ROI	1.71, 32.52	0.13	2.29	.002	.124



A.

B.

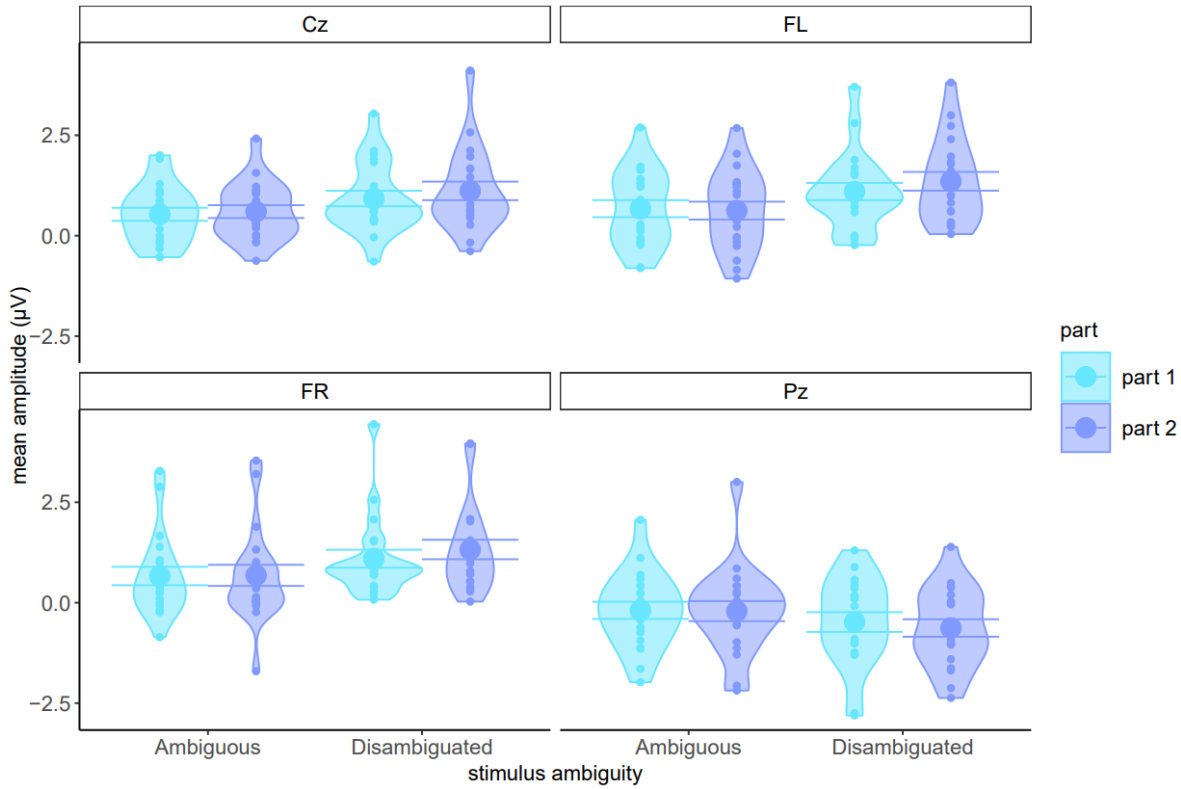


Figure 8. A. Individual mean amplitudes (in µV) in the time window between 150-250 ms post stimulus onset for all ambiguity and part conditions, averaged across ROIs. Large points depict the mean per condition and error bars depict one standard error around the mean. Results of statistical tests are shown in grey for main effects and in black for interaction effects. **B.** Individual mean amplitudes (in µV) in the time window between 150-250 ms post stimulus onset for all ambiguity and part conditions, for each ROI.

In the time window of interest for the P200 effect, between 150 and 250 ms post stimulus-onset, we observe a highly significant effect of ambiguity (i.e., the uncertainty effect): $F(1,19) = 37.03$, $p < .001$, with a small effect size (ges = 0.024) (Table 1, Figures 6, 7 and 8). More specifically, the mean amplitude was higher for disambiguated stimuli (M = 0.725 μV, SE = 0.108 μV) compared to ambiguous stimuli (M = 0.422 μV, SE = 0.094 μV).

In the same time window, the main effect of part of the experiment is significant as well ($F(1,19) = 4.46$, $p < .048$), but has a very small effect size (ges = 0.001). The mean amplitude for the first half of the experiment (M = 0.54 μV, SE = 0.092 μV) was significantly lower compared to second half of the experiment (M = 0.608 μV, SE = 0.108 μV). The implications of the main effects of ambiguity and part will be discussed below.

The main effect of ROI was also significant ($F(1.52, 28.83) = 8.83$, $p = .002$, corrected for sphericity by means of Greenhouse-Geisser), with a large effect size (ges = 0.247), indicating that the mean amplitude between 150 and 250 ms was significantly different between Cz, Pz, CPL and CPR. The amplitude was largest for FR (M = 0.941 μV, SE = 0.228 μV) and FL (M = 0.938 μV, SE = 0.208 μV), followed by Cz (M = 0.794 μV, SE = 0.179 μV), and Pz (M = -0.378 μV, SE = 0.223 μV).

Interestingly, the interaction effect of ambiguity and part of the experiment was also borderline significant: $F(1,19) = 4.63$, $p = .045$, with a very small effect size (ges = 0.001). Post-hoc tests revealed that difference between mean amplitude in ambiguous compared to disambiguated stimuli (i.e., the uncertainty effect) was significant in both parts of the experiment (all $p < .001$). However, the difference in mean amplitude between the first and second part of the experiment was significant only for disambiguated stimulus presentations ($p = .02$). This latter finding indicates that we did not find the interaction between ambiguity and part that we predicted (i.e., a difference in mean amplitude between the first and second part of the experiment which is larger for ambiguous stimulus presentations). The implications of the absence of the interaction effect of interest will be discussed below.

The interaction effect of ambiguity and ROI is highly significant, indicating that the difference in mean amplitude between ambiguous and disambiguated stimulus presentations is not the same across ROIs, $F(1.79, 34.02) = 28.36$, $p < .001$ (corrected for sphericity by means of Greenhouse-Geisser), with a small effect size (ges = 0.028). Post-hoc tests showed that the

difference between mean amplitude in ambiguous compared to disambiguated stimuli was significant in all ROIs (all $p < .001$). In Cz, FL and FR, the mean amplitude for disambiguated stimuli was significantly higher than for ambiguous stimuli, similar to the findings of Kornmeier et al. (2016). In contrast, in ROI Pz, the mean amplitude for ambiguous stimuli was significantly higher than for disambiguated stimuli. This indicates that the uncertainty effect differs in direction from more anterior (i.e., FR, FL and Cz) to more posterior (i.e., Pz) electrode positions.

P400

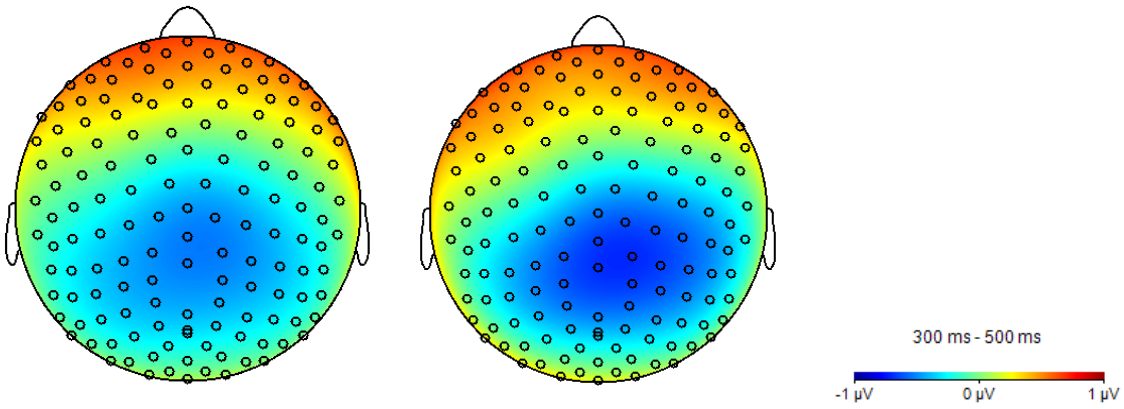


Figure 9. Topographic maps of difference waves for the P400 uncertainty effect in the time window of 300-500 ms post-stimulus-onset in the first (left) and second (right) part of the experiment. Voltages of disambiguated stimulus presentations were subtracted from ambiguous stimulus presentations.

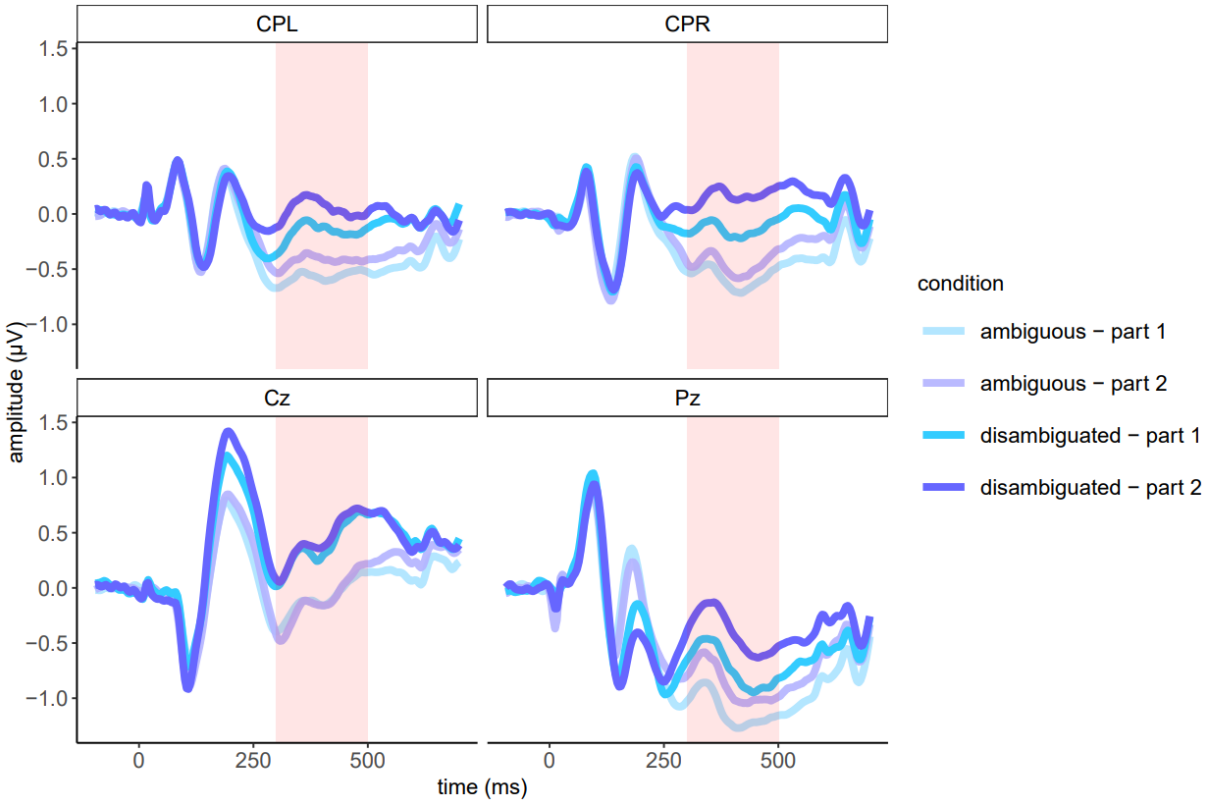
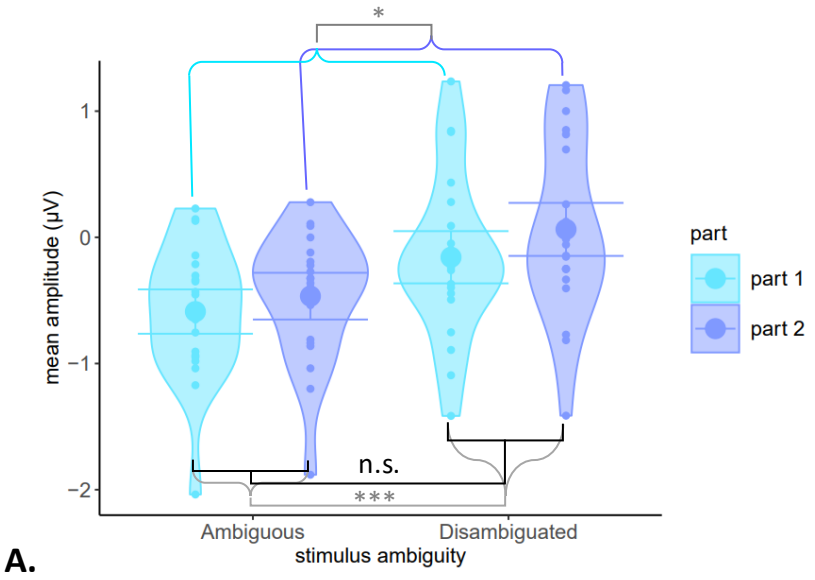


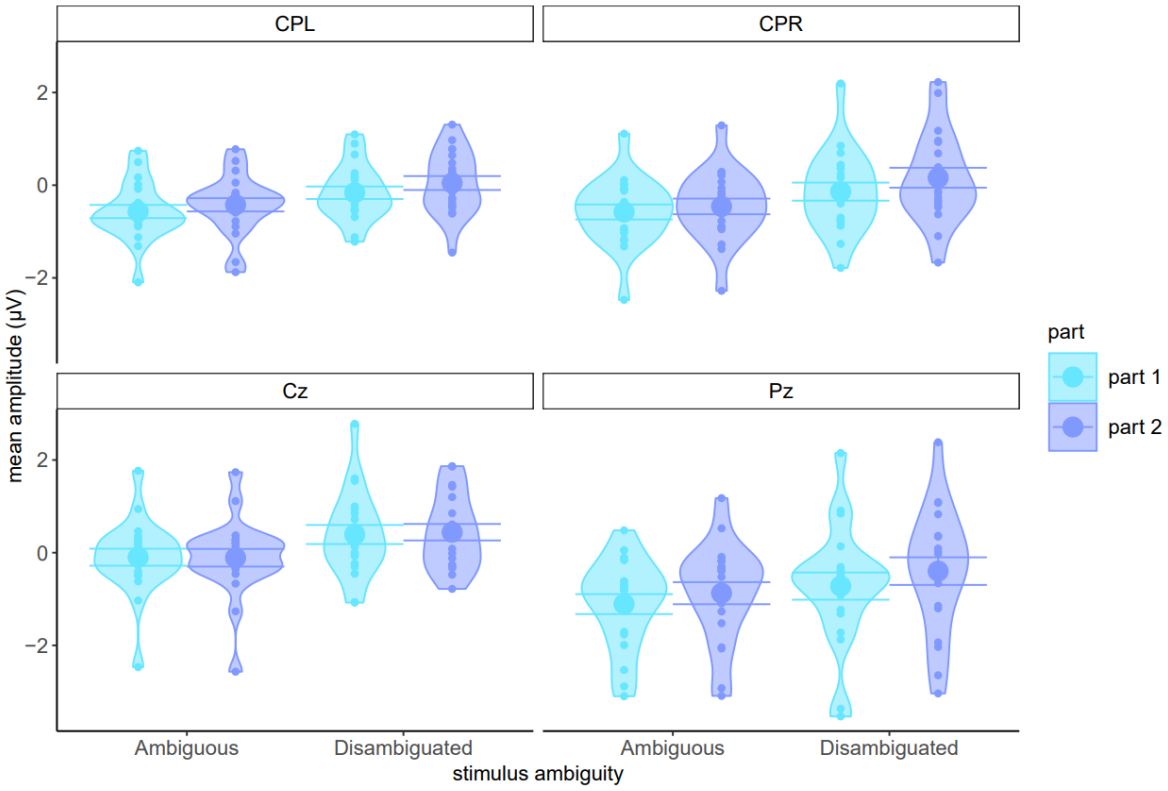
Figure 10. Event-Related Potential (ERP) traces for both parts and ambiguity of stimulus conditions in 4 electrode ROIs: Cz, CPL, CPR, and Pz. Vertical areas in pink indicate the time windows of interest for the P400 effect.

Table 2. Summary of the ANOVA in the time window of the P400 effect.

Effect	**df**	**MSE**	**F**	**ges**	**p value**
Ambiguity	1, 19	0.91	20.34 ***	.070	<.001
Part	1, 19	0.35	6.65 *	010	.018
ROI	1.43, 27.20	4.47	5.47 *	.125	.017
Ambiguity x Part	1, 19	0.12	1.66	<.001	.213
Ambiguity x ROI	2.11, 40.07	0.12	0.65	<.001	.536
Part x ROI	1.63, 30.96	0.10	4.28 *	.003	.029
Ambiguity x Part x ROI	1.78, 33.85	0.04	0.60	<.001	.535



A.



B.

Figure 11. A. Individual mean amplitudes (in μV) in the time window between 300-500 ms post stimulus onset for all ambiguity and part conditions, averaged across ROIs. Large points depict the mean per condition and error bars depict one standard error around the mean. Results of statistical tests are shown in grey for main effects and in black for interaction effects. **B.** Individual mean amplitudes (in μV) in the time window between 300-500 ms post stimulus onset for all ambiguity and part conditions, for each ROI.

In the time window of interest for the P400 effect, between 300 and 500 ms post stimulus-onset, the difference between the mean amplitude when observers were presented with disambiguated stimuli (*M* = -0.048 µV, *SE* = 0.145 µV) compared to when they observed ambiguous stimuli (*M* = -0.527 µV, *SE* = 0.115 µV) (i.e., the 'uncertainty effect') was highly significant; F(1,19) = 20.34, p < .001, with a medium effect size (ges = 0.070) (Table 2, Figures 9, 10 and 11). In the same time region, the main effect of part (or time) in the experiment was also significant, F(1,19) = 6.65, p = .018, with a small effect size (ges = 0.010). The mean amplitude in the second part of the experiment (*M* = -0.202 µV, *SE* = 0.124 µV) was significantly higher compared to the first part of the experiment (*M* = -0.373 µV, *SE* = 0.124 µV). Also in this time window, the interaction of ambiguity and part was not significant. The implications of the absence of the interaction effect between ambiguity and part, as well as the presence of the main effects of ambiguity and part will be discussed below.

The main effect of spatial ROI was significant (*F*(1.43, 27.20) = 4.47, *p* = .017, corrected for sphericity by means of Greenhouse-Geisser), with a medium effect size (ges = 0.125), which indicates that the mean amplitude was significantly different across our 4 selected ROIs: Cz, Pz, CPL and CPR. More specifically, the mean amplitude for Cz was highest, followed by CPR, CPL and Pz.

Finally, the interaction effect of part and ROI was significant (*F*(1.63, 30.96) = 4.28, *p* = 0.029, corrected for sphericity by means of Greenhouse-Geisser); The difference in mean amplitude between the first and second part of the experiment was different across ROIs. When inspecting the grand average ERP traces and violin plots of mean amplitudes, we can see that an effect of part appears to be present for the ROI of Pz, CPL and CPR, whereas it appears to be non-existent for the ROI of Cz. Indeed, post-hoc tests revealed that the difference in mean amplitude between the first and second part of the experiment was significant for all ROIs (*p* = .021 for Pz, *p* = .026 for CPL, *p* = .006 CPR), except for Cz.

3.1.5 Interim Discussion

In the time window of interest for the P200 effect, as well as in the time window for the P400 effect, we observed a highly significant effect of ambiguity (i.e., the uncertainty effect). This effect is equivalent to the ambiguity/uncertainty effect reported by Kornmeier et al. (2016)

and indicates that it is also present for (apparent) motion stimuli, which are an entirely different type of stimuli compared to the static stimuli employed in their previous research. The generalizability of the uncertainty effect across very different types of visual stimuli implies that it may indeed be associated to processes involved in later stages of visual perception, irrespective of the specific stimulus properties. This may be detection of the level of ambiguity involved in reaching a perceptual decision for the stimulus. Or, as Kornmeier et al. (2016) speculated, it may also involve meta-perceptual judgements about the certainty of the observer's perceptual decision, subsequent to this perceptual decision itself.

The main effect of part was significant for both the P200 and P400 time windows, indicating that the amplitude increases as the experiment continues. Interestingly, the mean amplitude of the P200 and P400 increases over time for both the ambiguous as the disambiguated stimuli. If a higher amplitude of the P200 and P400 components indeed corresponds to a lower uncertainty or ambiguity (Kornmeier, 2016), this would indicate that experience with the stimulus reduces the uncertainty, irrespective of whether the stimulus is ambiguous or disambiguated. This would put perceptual uncertainty in a different perspective, focusing not only on the (un-)ambiguity of a stimulus, but also on the frequency with which the observer has experienced this uncertainty in the past. Alternatively, if the uncertainty effect reflects unconscious neural processes, it could also reflect the frequency with which a stimulus has been associated to a percept by our visual system in the past.

However, as we highlighted in the Methods section, when we find a main effect of part in the absence of an interaction effect of part with ambiguity, it is important to consider whether the increase in amplitude over the course of the experiment is more general, in the sense of being unrelated to the uncertainty effect. If this were the case, we might expect to see an increase of amplitude over time across all time points of the ERP. However, as will be discussed further in the Results section, the main effect of time was not significant in the time window of the P200 effect. Furthermore, the increase of amplitude over time could be specific to this timing, yet still unrelated to the uncertainty effect. If so, we would expect to find similar findings in the literature covering other parieto-central components around 400 ms post-stimulus. To that end, we explored research addressing the evolution of the P300 component over the time course of an experimental session. Polich and Kok (1995) reviewed the literature concerning cognitive and biological determinants of P300 and concluded that amplitude

typically decreases as the session continues. A later study confirmed these findings (Sabeti et al., 2018). These conclusions describe the opposite trend compared to our data, suggesting that the increase of amplitude of the P400 component is independent of influences related to the P300 component over the time course of the experiment. This may be additional information demonstrating that the uncertainty effect is not entirely reducible to being a variant of the well-known P300 component (Kornmeier, 2016). However, even though the effect is restricted to a timing around 400 ms and we can exclude confounding effects over time related to the P300 component, we still cannot exclude every other possible confounding factor that might influence ERP amplitude over time. Therefore, we conclude that we are unable to draw a meaningful conclusion about the effect of experiment part, other than noting that it is not specific to ambiguous percepts only.

In the time window of the P200 effect, the interaction of ambiguity and part was significant, but did not entail the type of interaction that was of interest to us. In the time window of the P400 effect, the interaction of ambiguity and part was not significant. This interaction effect was of particular interest to us, as it could indicate that the uncertainty effect changes as the ambiguous stimuli become less ambiguous or uncertain over time (due to the 'shift to global'). We may conclude that the uncertainty effect does not reflect the perceptual ambiguity which may change over time as the 'shift to global' occurs. However, since the behavioral effect of a 'shift to global' was not found to be significant in the current study either, it is difficult to draw any conclusions. It may very well be that we can find a change in the uncertainty effect between a first and last part of the experiment if we were to increase the duration of the experiment and replicate the behavioral 'shift to global' as well. Nevertheless, as this is an EEG experiment and the duration of the EEG session during this experiment was already approximately 1.5 to 2 h, it may not be feasible to extend the session much longer because of practical reasons (i.e., the subject may become tired, the EEG cap may dry out, etc.). Future research could try to replicate this study over the course of multiple days, as we did in our previous experiment (Boeykens et al., 2019).

3.2 Part 2: Local vs. global perception of the ambiguous vs. disambiguated 'local-global' stimulus

3.2.1 Introduction

In the previous part of this chapter, we examined the evolution of the ERP 'uncertainty' effect in non-response trials over the time course of the experiment. However, we concluded that we are currently unable to interpret any main effect of experiment part, as we did not find a significant interaction effect of experiment part with ambiguity, and general changes in ERP amplitude over time may be caused by a wide variety of factors. Moreover, we did not find a significant effect of a change in percept over time in the behavioral data of this experiment. Nevertheless, aside from differences in the 'uncertainty' effect over time, there are also differences related to perception of the local percept vs. the global percept within both the ambiguous and disambiguated stimulus blocks, which were not considered thus far and will be discussed here.

The 'local-global' stimulus can be perceptually organized into two distinct percepts: 'local' perception of six rotating dot-pairs vs. 'global' perception of two pulsating hexagons. As mentioned in the General Introduction of this PhD thesis, despite the wealth of studies devoted to difference between local and global, it is still not straightforward to define the critical difference between 'local' and 'global'. We consider 'local' and 'global' to be *relative* concepts in a hierarchical relationship. The 'local' percept of rotating dot pairs can be considered local with respect to the 'global' percept of pulsating hexagons, but it may as well be considered global with respect to the percepts of each single dot within the dot-pairs.

Perceptual organization is a concept that has its roots in Gestalt psychology, where the focus of analysis is our phenomenological experience, or conscious perception of visual stimuli. Our visual experience is characterized by percepts that are already organized into coherent 'wholes', which are different from the mere sum of their 'parts'. Perceptual organization describes the structuring and organization of local 'parts' into global 'wholes', or Gestalts. These 'wholes' may be structured ('Gestaltet') up to a variable degree, and, accordingly, a local percept may be seen as a structured whole as well. However, a global percept is

considered to be structured up to a higher degree and makes up a better Gestalt compared to a local percept of the same stimulus (in the sense that an explanation of all the dot movements by two pulsating hexagons is easier than by six rotating dipoles). Even though the local percept of the 'parts' can be present in our phenomenological experience as well, the relatively global 'whole' dominates our perceptual experience most often (Wagemans et al., 2012).

According to Gestalt psychology, global percepts are favored in our perceptual experience because they are better Gestalts, implying that they have a better simplicity or 'Prägnanz'. Simplicity is not at all a straightforward characteristic to define, but some accounts have proposed that simplicity can be reflected by the number or length of descriptions necessary to define the percept, for instance in a coding language (Chater, 1996). In the current study, the 'local' percept would be described by 6 groups (of rotating dot pairs), whereas the 'global' percept can be described by 2 groups (of pulsating hexagons), making it a better Gestalt. Accordingly, the simplicity of a percept can be related to its efficiency and comprehensiveness in representing a stimulus. Aside from the hierarchical structure of their descriptive account of our phenomenal experience, Gestalt psychology proposed an explanatory account of the processes leading up to this phenomenal experience. According to them, perception is the result of a flexible and dynamic process that leads to the perceptual organization of stimuli, without the need for summation in successive levels of a hierarchical system, as per the empiricist approach.

Classic empiricists saw visual perception as the result of a series of summative operations carried out on measurable and/or sensory elements on different successive levels of analysis in a hierarchical visual system. They believed that there was no need to explore our phenomenological experience and focused on finding an explanatory account for visual perception that could be measured and adequately tested. Accordingly, 'local' and 'global' in this context are not considered to be characteristics of our perceptual experience, per se. Instead, the terms can be applied to the characteristics of a hierarchical visual system, where 'local' information on different levels is 'pooled' and summarized into more 'global' representations "higher up" in the visual system, irrespective of perceptual experience. In this interpretation, the terms 'local' and 'global' may sometimes be equated to, or at least intertwined with, different levels in a hierarchical visual system (i.e., low level vs. high level,

see also the section 'Level with me' in the General Discussion). Indeed, switches to more local and global perceptual organizations of ambiguous stimuli are associated with increases in activity at lower and higher stages of the visual system, respectively (Fang, Kersten & Murray, 2008; de-Wit, Kubilius, Wagemans & Op de Beeck, 2012). However, even though there are some commonalities between these concepts, it is important to note that they are not necessarily identical.

Similar to the Gestaltist approach, this perceptual process would lead up to percepts that are more comprehensive and efficient in representing the available stimulus information. Other characteristics that are commonly associated to the difference between a local and global perceptual organization are an increase in the level of abstraction and an increase in the spatial extent of the percept. Both of these concepts may sometimes be related, but are not necessarily equivalent to the difference between a local and global perceptual organization according to Gestalt psychology, whereas both concepts are more closely related to the interpretation of local and global according to the empiricist approach. Indeed, within a hierarchical visual system, more global perceptual interpretations increase in the level of abstraction and spatial extent when, for instance, information of multiple neuronal populations with small receptive fields is pooled together and averaged into information about a larger area of our visual field.

As can be seen from comparing the Gestaltist and empiricist approaches, which definition of perceptual organization or local vs. global you adopt critically depends on which theory of the visual processing system you adopt (de Wit et al, 2015). And as we have not reached a full understanding of the processes involved in visual perception, we cannot yet reach a consensus on the critical difference between local vs. global either. However, as argued by de Wit et al. (2015), these two approaches may possibly be reconciled if theoretic frameworks find a compromise between a focus on conscious perception and different levels of processing (cf. Reverse Hierarchy Theory by Hochstein & Ahissar, 2002). This suggestion will be elaborated in the General Discussion.

Nevertheless, the diversity in definitions of local and global percepts and perceptual processes has led to a wide range of stimuli and tasks that are assumed to be operationalizations of local vs. global processing (de Wit & Wagemans, 2015; Chamberlain et al., 2017) when in fact, these

stimuli and tasks may actually engage very different underlying processes. For instance, the embedded figures test or EFT (Witkin) and Navon type stimuli are generally seen as a typical measures of local vs. global perceptual processing. For both of these typical local-global stimuli, the terms local vs. global are associated to characteristics of certain percepts, bringing this type of interpretation closer to the approach in Gestalt psychology, where phenomenal experience was the main topic of interest. However, both stimuli and their associated tasks have important differences as well (Chamberlain et al., 2017). This diversity in stimuli and tasks has led to a low construct validity and inconsistency in the literature regarding the essence and implications of the terms local and global. Therefore, it is highly important to consider the implications of the local vs. global percept precisely in the stimulus and task at hand.

In the 'local-global' ambiguous motion stimulus employed in the current study, we may consider the local and global percepts as two different perceptual organizations or "interpretations" of the same stimulus. As Anstis and Kim (2011) noted, the local and global motion percepts can be seen as antagonistic solutions to the binding problem in this stimulus. In contrast to the EFT, the perception of the 'local' percept in the current stimulus is not necessarily facilitated by being able to embed it into the 'global' percept. In fact, the elements of the local percept need to be split up in order to create the global percept. The current stimulus may be more readily compared to other aspects of the EFT and Navon type stimuli, such as the fact that the local and global perceptual organizations result in percepts with a differing number of groups in each percept (i.e., less groups for the global percept) and a differing spatial scale of the groups in each percept (i.e., a larger spatial scale for the global percept). We could argue that the global percept constitutes a more efficient and comprehensive representation of the stimulus, as it groups the same amount of motion information into the minimum number of groups.

It is important to note, however, that efficiency in the perceptual representation does not necessarily correspond to efficiency in the perceptual process leading up to this representation. For instance, perceptually organizing a stimulus into a more comprehensive, global percept may initially take up more resources compared to a local perceptual organization. However, as the representation of the global percept is more efficient, our perceptual system may prefer this representation in the long term, leading to an increase in

the occurrence of the global percept and an increase in the ease with which this percept is organized over time as well.

As mentioned in the introduction of the first part of this chapter, several studies have examined the neural correlates of local vs. global perception of the 'local-global' ambiguous motion stimulus by means of fMRI BOLD activity and found that the global percept was associated with an increase in activity in posterior parietal areas and a decrease in activity in early visual areas (Zaretskaya, Anstis & Bartels, 2013; Grassi, Zaretskaya & Bartels, 2016, Grassi, Zaretskaya & Bartels, 2017). Zaretskaya and Bartels (2015) also carried out one EEG study in which they presented observers with continuous presentations of the 'local-global' stimulus and found that perception of the 'global' percept was associated to reduced posterior parietal beta-band (13-30 Hz) power. A remark concerning the interpretation by Zaretskaya et al. (2013; 2015; 2016; 2017) of both percepts in the 'local-global' stimulus that we would like to make, is the fact that they considered the global percept to be an "illusory" Gestalt, whereas the local percept was considered to be "veridical". Indeed, as we discussed in the current introduction, the global percept could be qualified as a better Gestalt compared to the local percept. However, it is important to indicate that a form of perceptual organization has taken place for the local percept as well.

Based on the research of Kornmeier et al. (2016), it is not straightforward to make a prediction of the differential magnitudes of the 'uncertainty' effect for global vs. local perception of the 'local-global' ambiguous motion stimulus. Following the interpretation of the local vs. global percept by Zaretskaya et al. (2015), the local percept may be more in line with the available stimulus information and may therefore be closer to "veridical" perception, which in turn is closer to an unambiguous or uni-stable form of perception. Therefore, we may expect to find a larger uncertainty effect for global compared to local perception of the ambiguous stimulus. For the disambiguated versions of the 'local-global' motion stimulus, however, we should not expect to find any differences in the uncertainty effect for the local vs. global percept.

EEG data and ERP analysis

Preprocessing

In order to assess the difference between the local and global perception of the 'local-global' stimulus, only response trials were analyzed in this analysis. Hence, the included epochs were restricted to stimulus onsets after which there was a perceptual switch response, in contrast to the first part of this chapter. Aside from this change, all preprocessing was identical to the first part of the experiment.

Statistical analysis

Here, the statistical design of the EEG data was a 2x2 design assessing the effect of ambiguity (ambiguous vs. disambiguated stimulus presentations) and perceptual switch (i.e., a switch to a local percept vs. global percept). Trials during which participants experienced and reported a perceptual switch were indicated by means of a marker corresponding to the experienced percept and stimulus type (i.e., L for a switch towards the local percept and G for the global percept; A for ambiguous stimulus presentations and D for disambiguated stimuli), resulting in the following two possible trial types: AL, DL, AG and DG.

All statistical analysis was identical to the first part of this chapter, except for the analysis of variance (ANOVA), which was computed separately per time window with 4 levels of ROI (i.e., FR, FL, Cz and Pz for the P200 effect and Cz, Pz, CPL and CPR for the P400 effect), 2 levels of ambiguity (i.e., ambiguous vs. disambiguated) and 2 levels of perceptual switch (i.e., a switch to a local percept vs. global percept).

3.2.2 Results

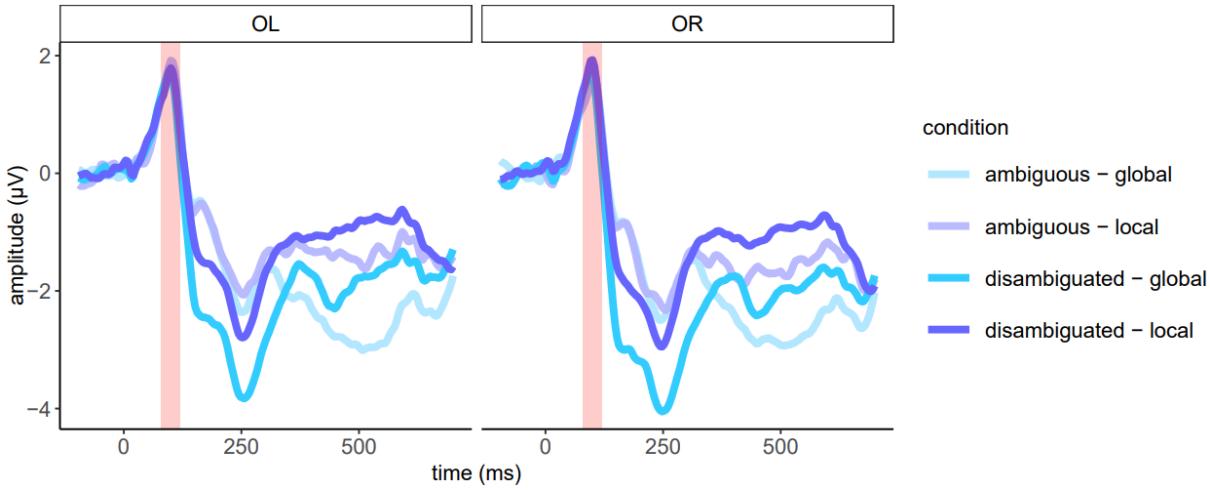


Figure 12. Event-Related Potential (ERP) traces for both percept and ambiguity conditions in 2 electrode ROIs for early visual components: OL and OR. Vertical areas in pink indicate the time windows of interest for the P1 component.

We found no significant differences between the conditions of interest in early visual component P1, indicating that the differences between the conditions of interest that are discussed in the time windows for the P200 and P400 effect are most probably not influenced by differences in stimulus characteristics or low level visual processing.

P200

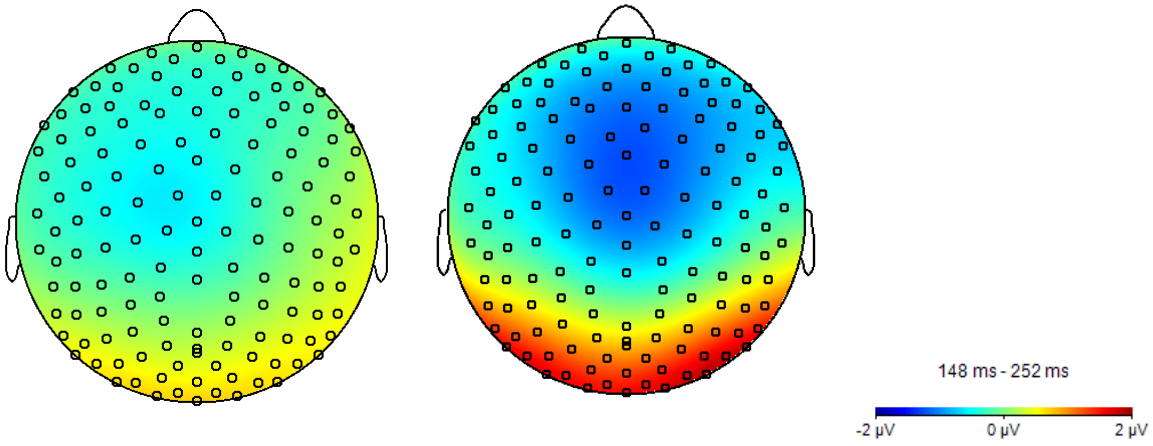


Figure 13. Topographic maps of difference waves for the P200 uncertainty effect in the time window of 150-250 ms post-stimulus-onset for trials during which the participant experienced a switch to the local (left) or global (right) percept. Voltages of disambiguated stimulus presentations were subtracted from ambiguous stimulus presentations.

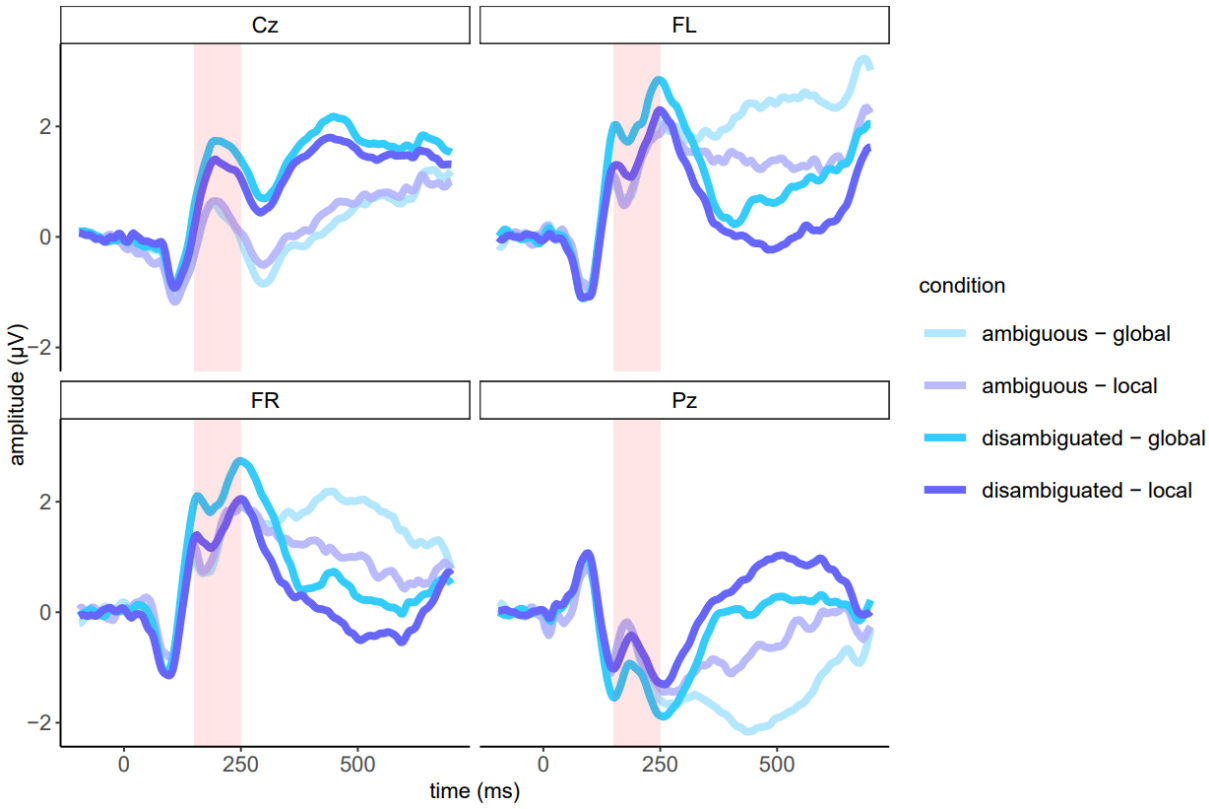
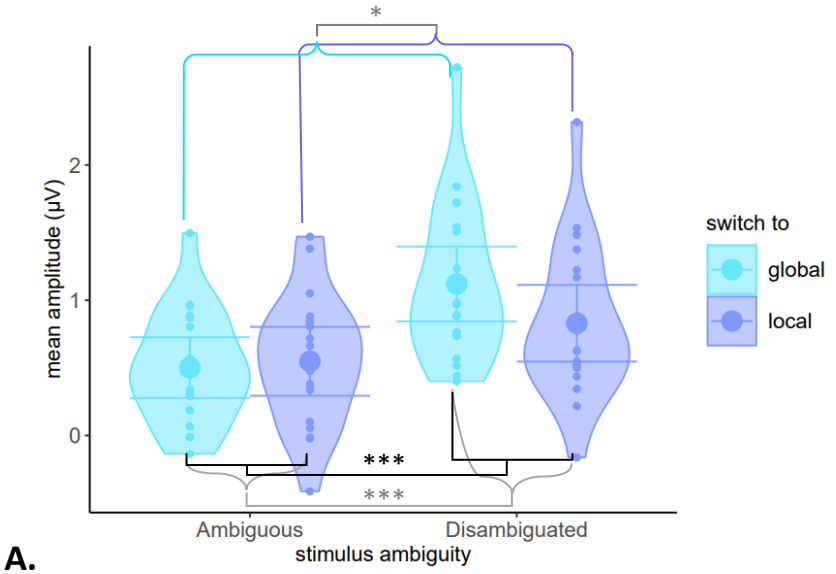


Figure 14. Event-Related Potential (ERP) traces for both percept and ambiguity conditions in 4 electrode ROIs: Cz, CPL, CPR, and Pz. Vertical areas in pink indicate the time windows of interest for the P200 and P400 effects.

Table 3. Summary of the ANOVA in the time window of the P200 effect.

Effect	df	MSE	F	ges	p value
Ambiguity	1, 19	0.42	38.38 ***	.037	<.001
Switch to percept	1, 19	0.20	6.09 *	.003	.023
ROI	1.66, 31.53	8.95	21.97 ***	.436	<.001
Ambiguity x Switch to percept	1, 19	0.12	19.75 ***	.005	<.001
Ambiguity x ROI	1.86, 35.38	0.84	10.55 ***	.038	<.001
Switch to percept x ROI	1.68, 31.85	0.52	5.60 *	.011	.011
Ambiguity x Switch to percept x ROI	1.58, 30.08	0.48	5.61 *	.010	.013



A.

B.

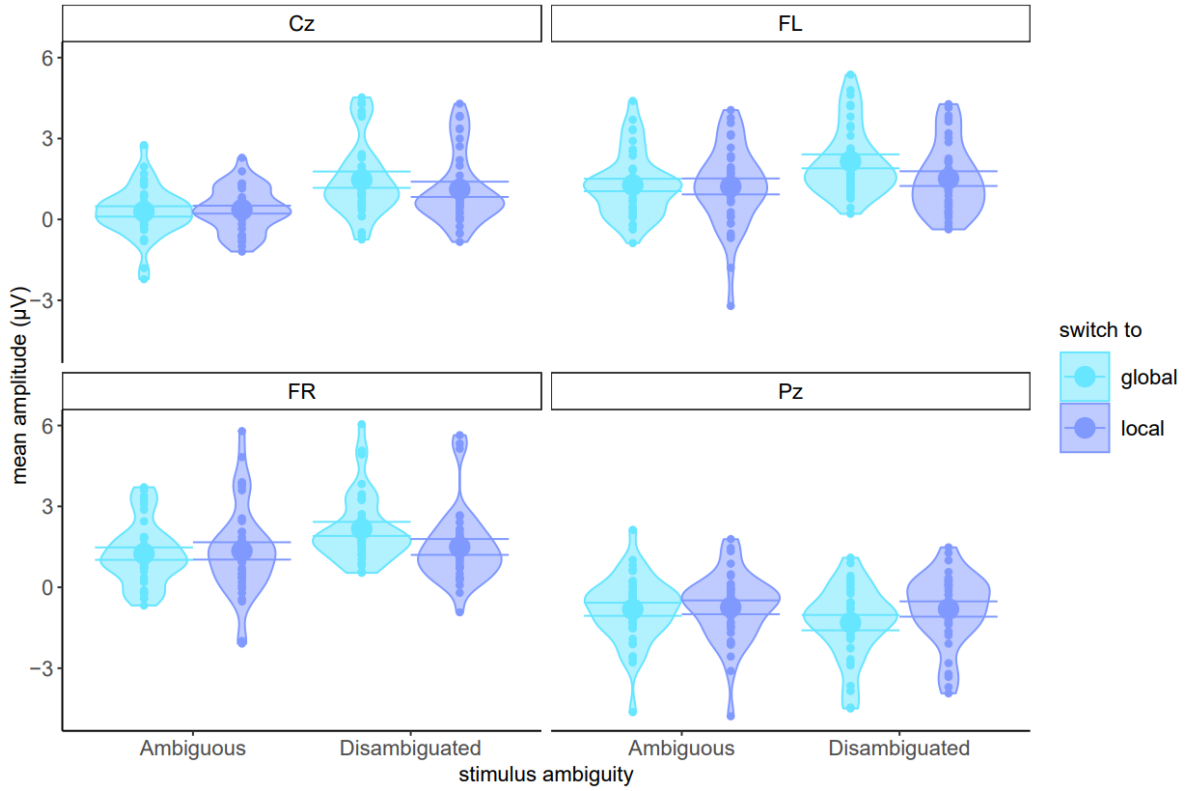


Figure 15. A. Individual mean amplitudes (in μV) in the time window between 150-250 ms post stimulus onset for all ambiguity and percept conditions, averaged across ROIs. Large points depict the mean per condition and error bars depict one standard error around the mean. Results of statistical tests are shown in grey for main effects and in black for interaction effects. **B.** Individual mean amplitudes (in μV) in the time window between 150-250 ms post stimulus onset for all ambiguity and percept conditions, for each ROI.

During the time window of the P200 effect, the main effect of ambiguity was highly significant ($F(1,19) = 38.38$, $p < 0.001$, with a small effect size (ges = 0.037)), with a higher amplitude for disambiguated trials compared to ambiguous trials (Table 3, Figures 13, 14 and 15). This effect is equivalent to the ambiguity/uncertainty reported by Kornmeier et al. (2016) and was also highly significant in the analysis of non-response trials in the first part of this chapter. The main effect of spatial ROI is highly significant as well ($F(1.66, 31.53) = 8.95$, $p < .001$, corrected for sphericity by means of Greenhouse-Geisser), with a large effect size (ges = 0.436), which indicates that the mean amplitude is significantly different across our 4 ROIs: the highest amplitude for FR and FL, followed by Cz and then Pz, which is in accordance with the results for non-response trials as well.

In the same time window, the main effect of percept was also significant: $F(1,19) = 6.09$, $p = 0.023$, with a very small effect size (ges = 0.003). More specifically, The mean amplitude was higher for trials during which participants experienced a perceptual switch towards the global percept ($M = 0.811$ µV, $SE = 0.105$ µV), compared to when they switched towards the local percept ($M = 0.689$ µV, $SE = 0.107$ µV).

Moreover, the interaction effect of ambiguity and percept was present as well: $F(1,19) = 19.75$, $p < .001$. Post-hoc tests indicate that the uncertainty effect is significant for local as well as global percepts ($p < .001$ for global, $p = .004$ for local). However, the difference in mean amplitude between trials with a switch to local compared to trials with a switch to global was significant only for disambiguated stimulus variants ($p < .001$).

P400

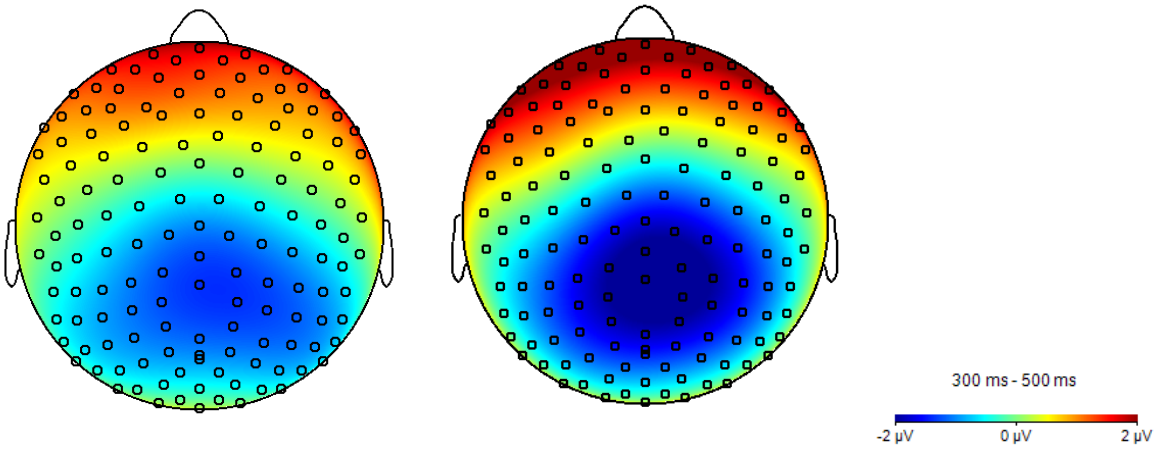


Figure 16. Topographic maps of difference waves for the P400 uncertainty effect in the time window of 300-500 ms post-stimulus-onset for trials during which the participant experienced a switch to the local (left) or global (right) percept. Voltages of disambiguated stimulus presentations were subtracted from ambiguous stimulus presentations.

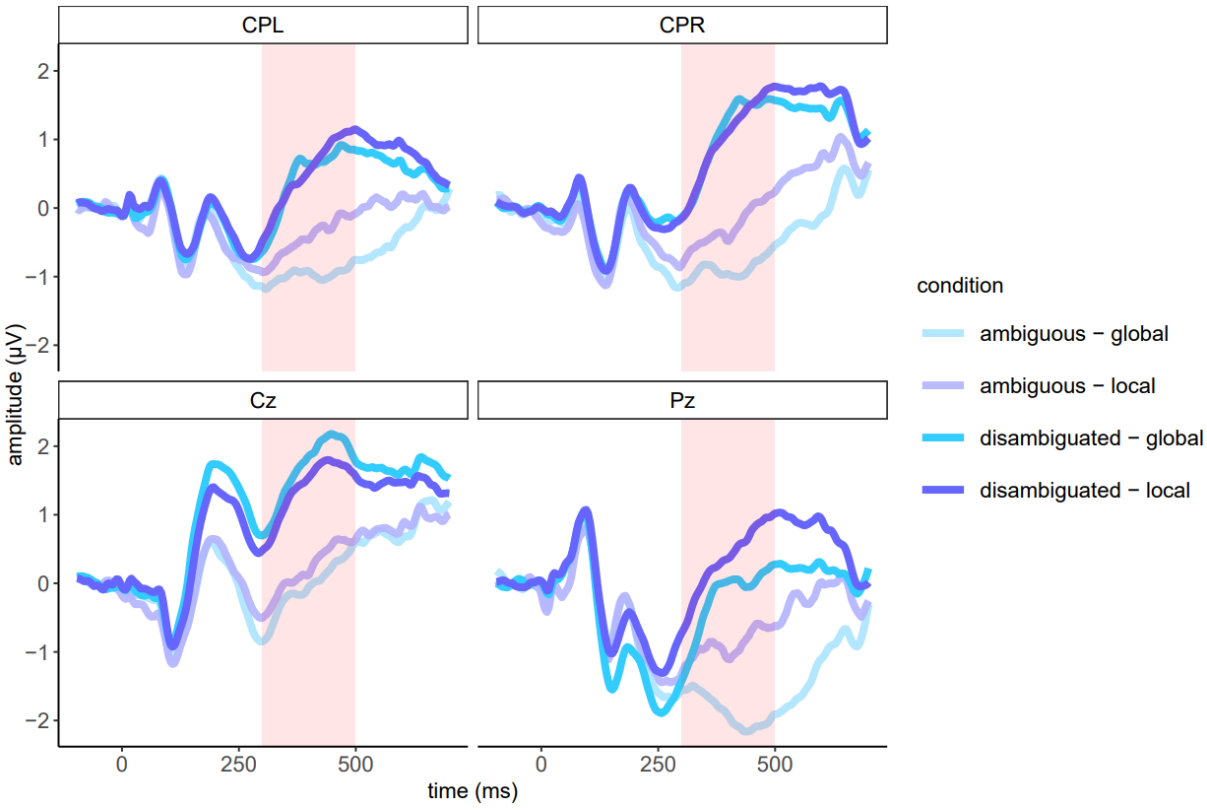
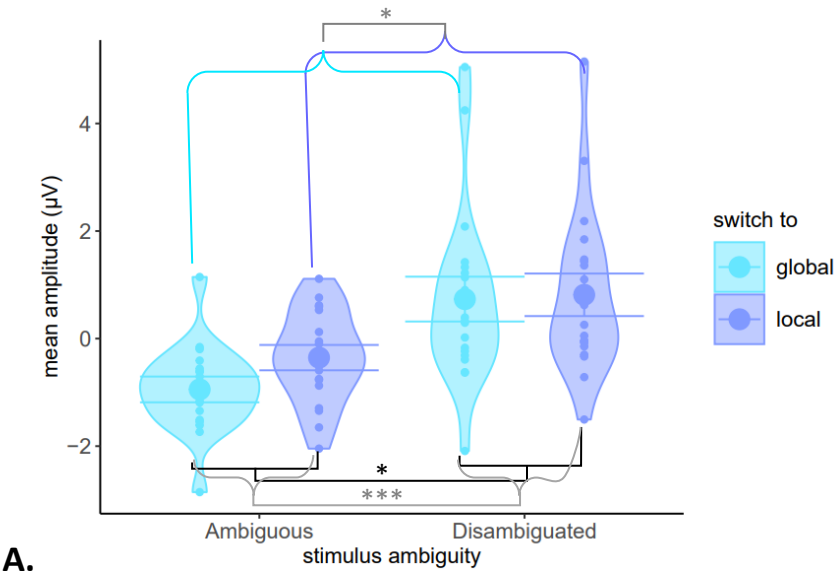


Figure 17. Event-Related Potential (ERP) traces for both percept and ambiguity conditions in 4 electrode ROIs: Cz, CPL, CPR, and Pz. Vertical areas in pink indicate the time window of interest for the P400 effect.

Table 4. Summary of the ANOVA in the time window of the P400 effect.

Effect	df	MSE	F	ges	p value
Ambiguity	1, 19	6.34	25.60 ***	.188	<.001
Switch to percept	1, 19	1.16	7.78 *	.013	.012
ROI	1.52, 28.84	5.98	9.88 **	.113	.001
Ambiguity x Switch to percept	1, 19	0.78	6.72 *	.007	.018
Ambiguity x ROI	1.73, 32.92	0.89	1.17	.003	.317
Switch to percept x ROI	1.61, 30.63	0.41	9.42 **	.009	.001
Ambiguity x Switch to percept x ROI	2.01, 38.22	0.23	0.38	<.001	.685



A.

B.

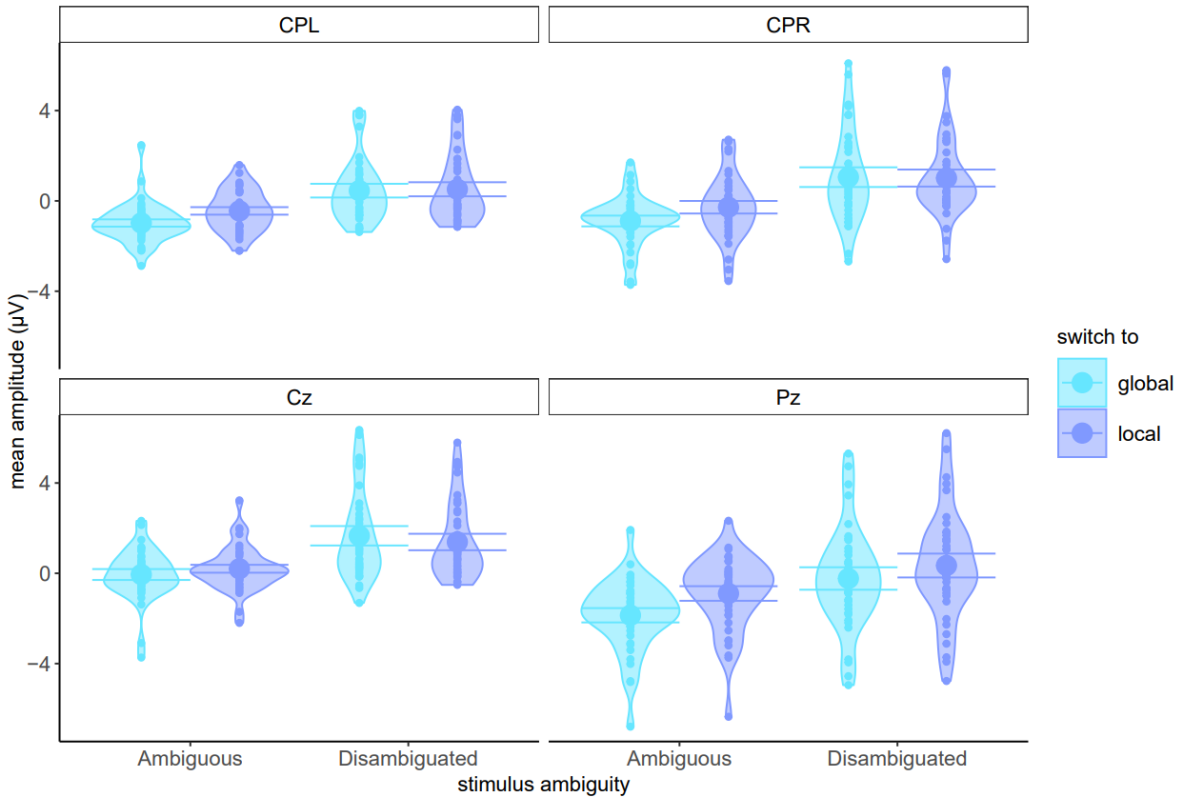


Figure 18. A. Individual mean amplitudes (in μV) in the time window between 300-500 ms post stimulus onset for all ambiguity and percept conditions, averaged across ROIs. Large points depict the mean per condition and error bars depict one standard error around the mean. Results of statistical tests are shown in grey for main effects and in black for interaction effects. **B.** Individual mean amplitudes (in μV) in the time window between 300-500 ms post stimulus onset for all ambiguity and percept conditions, for each ROI.

In the time window for the P400 effect, between 300 and 500 ms post-stimulus-onset, the difference between the mean amplitude when observers were presented with disambiguated stimuli (*M* = 0.774 µV, *SE* = 0.349 µV) compared to when they observed ambiguous stimuli (*M* = -0.650 µV, *SE* = 0.155 µV) was highly significant; $F(1,19) = 25.6$, $p < .001$, with a large effect size (ges = 0.188) (Table 4, Figures 16, 17 and 18). This effect is equivalent to the ambiguity/uncertainty reported by Kornmeier et al. (2016) and was also highly significant in the analysis of non-response trials in the first part of this chapter. The main effect of spatial ROI is significant as well ($F(1.52, 28.84) = 9.88$, $p = .001$, corrected for sphericity by means of Greenhouse-Geisser), with a medium effect size (ges = 0.113), which indicates that the maximum amplitude is significantly different across our 4 selected ROIs: Cz, Pz, CPL and CPR. More specifically, the mean amplitude for Cz was highest, followed by CPR, CPL and Pz, which is in accordance with the results for non-response trials as well.

In the same time window, the main effect of perceptual switch is also significant; $F(1,19) = 7.78$, $p = 0.012$, with a small effect size (ges = 0.013). The mean amplitude was between 300 and 500 ms post-stimulus-onset was higher for trials during which participants experienced a shift towards the local percept (*M* = 0.230 µV, *SE* = 0.242 µV) compared to when they experienced a switch towards the global percept (*M* = -0.106 µV, *SE* = 0.234 µV).

Importantly, the interaction effect of ambiguity and perceptual switch was significant as well; $F(1,19) = 6.72$, $p = 0.018$. More specifically, post-hoc tests show that the uncertainty effect was significant for trials with a switch to the local as well as to the global percept (all $p < .001$). However, the difference between mean amplitude for a perceptual switch towards the local (*M* = -0.354 µV, *SE* = 0.189 µV) vs. the global (*M* = -0.945 µV, *SE* = 0.177 µV) percept was significant for ambiguous stimuli ($p = .007$), but not for disambiguated stimuli.

3.2.3 Discussion

We found that the main effect of perceptual switch was significant in both time windows, for the P200 and the P400 effect. If we interpret the amplitude in terms of uncertainty associated to the percept (i.e., akin to Kornmeier et al., 2016), we can equate a lower amplitude in the ERP components to a higher level of uncertainty. This indicates that observers generally experience a higher uncertainty or ambiguity when perceiving the global percept compared to the local percept.

Post-hoc tests for the P200 component revealed that the local and global percepts of the ambiguous stimuli were associated with the most, but a relatively equal amount of uncertainty (or ERP amplitude), whereas the local percept of the disambiguated stimuli was associated to a lower level of uncertainty compared to both percepts of the ambiguous stimuli, but a significantly higher level of uncertainty compared to the global percept of disambiguated stimuli. This indicates that the global percept of a stimulus that does not evoke perceptual multistability is associated to the lowest level of uncertainty in the time window of the P200 effect, a finding that could be related to the fact that the percept associated to the disambiguated global stimulus entails visual information that can be represented more efficiently, grouping visual information into a lower amount of groups (i.e., 2) compared to the percept associated to the disambiguated local stimulus (i.e., 6). Alternatively, our disambiguated version of the global percept may have been made to elicit a more clear, global percept, whereas our disambiguated version of the local percept may elicit less, clear, local percepts. Future research may try to examine whether both versions of the disambiguated stimuli enable equally clear percepts.

From the significant interaction effect of ambiguity and perceptual switch in the time window of the P400 component, we learned that the mean amplitude for the local and global perceptual switch for disambiguated stimuli was higher in general, whereas the mean amplitude for ambiguous stimuli was lower, with the lowest amplitude for global perceptual switch trials. If we again adopt the interpretation of a lower amplitude in terms of an increase in uncertainty or ambiguity, this indicates that observers experience a higher uncertainty or ambiguity during a switch towards a local or global percept of the ambiguous stimulus, with the highest ambiguity and uncertainty for a switch towards the global percept of the

ambiguous stimulus. Nevertheless, the global percept is experienced with equal probability as the local percept and, as we saw in previous studies (Boeykens, Wagemans & Moors, 2019; Anstis & Kim, 2011), the global percept becomes dominant over time, seemingly spontaneously. It may seem contra-intuitive and unbeneficial from an evolutionary perspective for a percept that is associated to a higher level of uncertainty to be equiprobable in perception and to become more dominant over time. However, in the following, we will propose a potential account of how this may occur.

As we mentioned in the introduction, our visual system may prefer to evolve towards a global perceptual organization over time because the representation of the global percept is associated to a higher degree of efficiency in representing visual information. However, the process leading up to the global percept may initially be less efficient and take up more resources. The ambiguous information in the ambiguous blocks may cause observers to rely less on the ambiguous stimulus information and more on "higher-level" visual processes. Moreover, the ambiguous stimulus information may be more readily interpreted in correspondence with the local percept, whereas the stimulus information needs to be processed up to a higher level in order to be in correspondence with the global percept. Nevertheless, our visual system would continue to prioritize a global perceptual organization because of its benefits in efficiency in representation over time. Accordingly, the process leading up to the global percept would occur with more and more ease and take up less resources over time.

We may then interpret the uncertainty effect in terms of the ease of the process leading up to a perceptual representation, with a higher amplitude for a higher efficiency in processing. If this were the case, we would have to see a decrease in the uncertainty effect for the global percept over time. Unfortunately, the duration of this experiment was not long enough to replicate the behavioral effect of the 'shift to global' (Boeykens et al., 2019), which would be a good test-case for the increase in the dominance of the global percept and the potential associated increase in the ease of the process leading up to the global percept over time. Future research may also attempt to examine whether the 'shift to global' is associated with a decrease in the uncertainty effect for global percepts.

Kornmeier et al. (2020) decided to interpret the uncertainty effect in terms of ambiguity and uncertainty associated to the perception of certain stimuli because it increased for various types of ambiguous stimuli as well as stimuli which were low in visual acuity. However, the inefficiency in processing of and the necessity of the involvement of "higher-level" processes for the perception of ambiguous as well as low acuity stimuli is also higher compared to disambiguated and high acuity stimuli. Accordingly, if we frame the uncertainty effect in terms efficiency in processing, instead of the uncertainty associated to a perceptual decision, we may be able to align the results of our current study with previous research on the uncertainty effect.

Nevertheless, the P400 effect discussed here is rather late in timing to be associated to pre-perceptual processes and was interpreted in the context of a post-perceptual process by Kornmeier et al. (2016). More specifically, they suggested that the P400 effect reflects the uncertainty associated to the meta-perceptual evaluation of the previous perceptual decision. Therefore, the increase in the uncertainty effect for global perception of the motion stimulus may be due to post-perceptual processes. In this case, the post-perceptual processes may indeed be related to decision uncertainty, as suggested by Kornmeier et al. (2016; 2020). Alternatively, taking into account the tendency of our visual system to favour the global percept over time, the increase in the P400 effect for global perception may be due to an increase in top-down processes, perhaps silencing ambiguous stimulus information, in order to extend and facilitate a stable global percept. Further research should put these different speculative accounts to the test. As a metaperceptual evaluation of uncertainty associated to a previous perceptual decision should not necessarily result in an increase of the same percept for the subsequent stimulus presentation, analysing the correlation between the strength of the uncertainty effect within one trial and the maintenance of a percept in subsequent trials may provide some more insights into the underlying processes of the uncertainty effect.

3.3 General Conclusion

In the current study, we replicated the 'uncertainty' or 'ambiguity' effect which was previously reported by Kornmeier et al. (2016, 2020) for the perception of the 'local-global' ambiguous motion stimulus. The 'uncertainty' effect describes a higher ERP amplitude in the time window between 150 and 250 ms post stimulus-onset for frontocentral electrode locations (i.e., the P200 effect) and in the time window between 300 and 500 ms post stimulus-onset for parietocentral electrode locations (i.e., the P400 effect) for disambiguated compared to ambiguous stimuli.

The results of an analysis comparing the first and second half of the EEG experiment indicated that there was also a significant increase in ERP amplitude for the P200 and P400 components over the time course of the experiment. We hypothesized that if we would have replicated the finding of a 'shift to global' in the perception of the ambiguous 'local-global' stimulus over time, the perception of this stimulus would be associated to an increase in the amplitude of the P200 and P400 components over time, as the perception of this stimulus is associated to a decrease in ambiguity and uncertainty. However, we did not replicate the behavioral effect of the 'shift to global' over time and concluded that it may not occur over the shorter experiment duration of approximately 1.5 up to 2 h. Moreover, the interactions of the uncertainty effect with stimulus ambiguity for the P200 and P400 components did not reflect our hypothesis either. Therefore, we concluded that the general increase in ERP amplitude for the P200 and P400 components over time may be due to various other factors (e.g., fatigue) and cannot be interpreted with respect to stimulus ambiguity at this time.

In a subsequent analysis, we compared the trials for which participants experienced a switch towards the local percept to the trials for which participants experienced a switch towards the global percept. We found that the mean amplitude for a change in perception towards the local percept was significantly higher in both the time windows for the P200 and P400 components. The interaction between the uncertainty effect and percept in the time window of the P200 component indicated that the uncertainty effect was significant for trials during which participants experienced a switch towards the local percept, as well as trials during which participants experienced a switch toward the global percept. Moreover, a change in perception towards the local compared to the global percept was significant only for the

disambiguated stimulus variants. In contrast, although the 'uncertainty effect' was significant for both the local and global percept in the time window of the P400 component as well, the difference for a change in perception towards the local compared to the global percept was significant only for ambiguous stimulus presentations. If we interpret the amplitude in terms of uncertainty associated to the percept (i.e., akin to Kornmeier et al., 2016), we can equate a lower amplitude in the ERP components to a higher level of uncertainty.

The global percept of disambiguated stimuli was associated to the lowest level of uncertainty in the time window of the P200 effect, which could be related to the fact that the percept of the disambiguated global stimulus represents visual information that is more efficient, grouping information into a lower amount of groups (i.e., 2) compared to the percept of the disambiguated local stimulus (i.e., 6).

Results for the time window of the P400 effect indicated that a switch towards the global percept in ambiguous stimuli is associated to a higher level of uncertainty. Nevertheless, the global percept is experienced with equal probability as the local percept, which may be unbeneficial from an evolutionary perspective if it is in fact associated to a higher level of uncertainty.

We proceeded with an alternative interpretation of the uncertainty effect, framing it in terms of the efficiency associated to processes leading up to the perceptual representation, as the inefficiency in processing for the perception of ambiguous as well as low acuity stimuli is also higher compared to disambiguated and high acuity stimuli. Even though a global percept may represent visual information in a more efficient manner, the process leading up to this global percept may initially be less efficient and take up more resources. This process may eventually become more and more efficient over time, as our visual system continues to prioritize the global percept because of its higher efficiency in the representation of visual information. Nevertheless, as the P400 effect may be rather late in timing to be associated to pre-perceptual processes and was also interpreted as uncertainty in the context of a post-perceptual process by Kornmeier et al. (2016), we argued that the P400 effect may also be related to top-down processes that aim to extend and facilitate a stable subsequent percept. Future research should put these speculative accounts to the test.

References

Anstis, S., & Kim, J. (2011). Local versus global perception of ambiguous motion displays. *Journal of Vision, 11*(3): 13. https://doi.org/10.1167/11.3.13

Boeykens, C., Wagemans, J., & Moors, P. (2019). Does task relevance shape the 'shift to global' in ambiguous motion perception?. *Journal of Vision, 19*(1): 8. https://doi.org/10.1167/19.1.8

Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable Perception and the Role of the Frontoparietal Cortex in Perceptual Inference. *Annual Review of Psychology, 69*(1), 77–103. https://doi.org/10.1146/annurev-psych-010417-085944

Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, *141*, 247-257. https://doi.org/10.1016/j.visres.2017.01.008

Chater, N. (1996). Reconciling simplicity and likelihood principles in perceptual organization. *Psychological Review*, *103*(3), 566-581. https://doi.org/10.1037/0033-295X.103.3.566

de-Wit, L. H., Kubilius, J., Wagemans, J., & de Beeck, H. P. O. (2012). Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. *Journal of Vision*, *12*(11): 12. https://doi.org/10.1167/12.11.12

de-Wit, L., & Wagemans, J. (2015). Individual differences in local and global perceptual organization. In J. Wagemans (Ed.), *The oxford handbook of perceptual organization* (pp. 601-620). Oxford, UK: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199686858.013.028

Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision*, *8*(7): 2. https://doi.org/10.1167/8.7.2

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1521), 1211-1221. https://doi.org/10.1098/rstb.2008.0300

Gilbert, C. D. (1996). Plasticity in visual perception and physiology. *Current Opinion in Neurobiology, 6*(2), 269-274. https://doi.org/10.1016/S0959-4388(96)80083-3

Grassi, P. R., Zaretskaya, N., & Bartels, A. (2016). Parietal cortex mediates perceptual Gestalt grouping independent of stimulus size. *NeuroImage, 133*, 367-377. https://doi.org/10.1016/j.neuroimage.2016.03.008

Grassi, P. R., Zaretskaya, N., & Bartels, A. (2017). Scene segmentation in early visual cortex during suppression of ventral stream regions. *NeuroImage, 146*, 71-80. https://doi.org/10.1016/j.neuroimage.2016.11.024

Grasso, P. A., Gallina, J., & Bertini, C. (2020). Shaping the visual system: cortical and subcortical plasticity in the intact and the lesioned brain. *Neuropsychologia, 142*: 107464. https://doi.org/10.1016/j.neuropsychologia.2020.107464

Grzeczkowski, L., Clarke, A. M., Francis, G., Mast, F. W., & Herzog, M. H. (2017). About individual differences in vision. *Vision Research, 141*, 282-292. https://doi.org/10.1016/j.visres.2016.10.006

Hamaker, E. L. (2012). Why researchers should think" within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). The Guilford Press.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(5), 791-804. https://doi.org/10.1016/S0896-6273(02)01091-7

Jia, K., Zamboni, E., Kemper, V., Rua, C., Goncalves, N. R., Ng, A. K. T., ... & Kourtzi, Z. (2020). Recurrent processing drives perceptual plasticity. *Current Biology, 30*(21), 4177-4187. https://doi.org/10.1016/j.cub.2020.08.016

Joos, E., Giersch, A., Hecker, L., Schipp, J., Heinrich, S. P., Tebartz van Elst, L., & Kornmeier, J. (2020). Large EEG amplitude effects are highly similar across Necker cube, smiley, and abstract stimuli. *PloS One, 15*(5), e0232928. https://doi.org/10.1371/journal.pone.0232928

Karmarkar, U. R., & Dan, Y. (2006). Experience-dependent plasticity in adult visual cortex. *Neuron, 52*(4), 577-585. https://doi.org/10.1016/j.neuron.2006.11.001

Klink, P. C., Brascamp, J. W., Blake, R., & van Wezel, R. J. (2010). Experience-driven plasticity in binocular vision. *Current Biology, 20*(16), 1464-1469. https://doi.org/10.1016/j.cub.2010.06.057

Koenderink, J. (2019). Vision, an optical user interface. *Perception, 48*(7), 545-601. https://doi.org/10.1177/0301006619853758

Kornmeier, J., & Bach, M. (2012). Ambiguous Figures – What Happens in the Brain When Perception Changes But Not the Stimulus. *Frontiers in Human Neuroscience, 6*. https://doi.org/10.3389/fnhum.2012.00051

Kornmeier, J., Wörner, R., & Bach, M. (2016). Can I trust in what I see? EEG evidence for a cognitive evaluation of perceptual constructs. *Psychophysiology, 53*(10), 1507-1523. https://doi.org/10.1111/psyp.12702

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.

Miller, L. E., & Saygin, A. P. (2013). Individual differences in the perception of biological motion: links to social cognition and motor imagery. *Cognition, 128*(2), 140-148. https://doi.org/10.1016/j.cognition.2013.03.013

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement, 2*(4), 201-218. https://doi.org/10.1207/s15366359mea0204_1

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434. 10.1037/1082-989X.8.4.434

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). https://doi.org/10.1126/science.aac4716

Peirce, J. W. (2007). PsychoPy—psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8-13. 10.1016/j.jneumeth.2006.11.017

Phillips, W. L. (2019). Cross-cultural differences in visual perception of color, illusions, depth, and pictures. *Cross-Cultural Psychology: Contemporary Themes and Perspectives*, 287-308. https://doi.org/10.1002/9781119519348.ch13

Polich, J., & Kok, A. (1995). Cognitive and biological determinants of P300: an integrative review. *Biological Psychology, 41*(2), 103-146. https://doi.org/10.1016/0301-0511(95)05130-9

Sabeti, M., Boostani, R., & Rastgar, K. (2018). How mental fatigue affects the neural sources of P300 component?. *Journal of Integrative Neuroscience, 17*(1), 93-111. https://doi.org/10.3233/JIN-170040

Takeuchi, T., Yoshimoto, S., Shimada, Y., Kochiyama, T., & Kondo, H. M. (2017). Individual differences in visual motion perception and neurotransmitter concentrations in the human brain. *Philosophical Transactions of the Royal Society B: Biological Sciences, 372*(1714), 20160111. https://doi.org/10.1098/rstb.2016.0111

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172-1217. https://doi.org/10.1037/a0029333

Zaretskaya, N., Anstis, S., & Bartels, A. (2013). Parietal Cortex Mediates Conscious Perception of Illusory Gestalt. *Journal of Neuroscience, 33*(2), 523–531. https://doi.org/10.1523/JNEUROSCI.2905-12.2013

Zaretskaya, N., & Bartels, A. (2015). Gestalt perception is associated with reduced parietal beta oscillations. *NeuroImage, 112*, 61–69. https://doi.org/10.1016/j.neuroimage.2015.02.049

4. CSF study:

Exploration of the potential relationship of the "low-level" measure of contrast sensitivity and the "mid-level" perception of the local-global ambiguous motion stimulus.

4.1 Introduction

Perceptual organization, the process of giving structure to the information provided by our eyes, may lead to different percepts. For instance, during perceptual organization of the motion quartet (MQ, Figure 1 in the General Introduction), you may associate the top right dot in the first frame to the top left dot in the second frame, giving rise to a horizontal apparent motion percept. Alternatively, you may associate the top right dot in the first frame to the bottom right dot in the second frame, giving rise to the vertical apparent motion percept.

Some studies have suggested that there is a causal relationship between "low-level" perceptual organization and "mid-level" visual representations (He, Kersten, & Fang, 2012; Wokke et al., 2013; Feigin, Baror, Bar, & Zaidel, 2017). One "low-level" characteristic determining perception of an ambiguous stimulus may be the sensitivity to differences in luminance levels of the stimulus. The spatiotemporal contrast sensitivity function (CSF; Kelly, 1979) is a measure of the individual detection threshold of, or sensitivity to, the luminance contrast of drifting sinusoidal gratings with varying spatiotemporal characteristics (i.e., Gabor stimuli). The CSF is seen as a measure of "low-level" visual processing as it is believed to be closely related to the neural responses in primary visual cortex (V1). This argument is based on the channel hypothesis, which states that our visual system has dedicated, independent channels, or motion detectors, which are selectively responsive to specific spatiotemporal frequencies in luminance contrast. These detectors are assumed to be formed by optimal population decoding of V1 neurons, and perceptual decisions in a contrast sensitivity task are

assumed to be based on the information they provide (Graham, 1989; Dosher & Lu, 1998). The differing sensitivity for differing spatiotemporal frequencies is often seen as a consequence of limited neural resources and the optimal allocation of them to different frequencies, in order to facilitate motion perception as much as possible (Silverman, Grosof, De Valois, & Elfar, 1989).

Nakayama (1985) found that motion detectors specifically selective to the presentation of continuously drifting sinusoidal gratings with certain spatiotemporal frequencies are equally selective to the same sinusoidal gratings displaced discontinuously by distances that are one quarter of the stimulus period. More specifically, luminance contrast sensitivity for a drifting grating stimulus with spatial frequency *Fs* (cycles/deg) and temporal frequency *Ft* (cycles/sec or Hz) will be identical for discontinuous displacement of this grating stimulus by spatial distance *Ds* (deg) = 1/(4**Fs*) over temporal distance *Dt* (sec) = 1/(4**Ft*). Based on this finding, he concluded that the so called "motion detectors" of our visual system are not tuned to perceive only continuous motion. Rather, they "are fed by [pairs of] spatial frequency tuned receptive fields spaced so that they are in quadrature phase" (Nakayama, 1985, pp. 637). This structure was also employed in a modified version of the Reichardt motion detection model (van Santen & Sperling, 1984) and ensures the most optimal coding of visual motion information with the smallest number of motion detectors, maximizing the efficiency of a fixed number of neurons.

Nakayama (1985) goes on to show that sensitivity to drifting grating stimuli is comparable not only to discontinuously displaced grating stimuli at threshold, but also to apparent motion stimuli above detection threshold. (Gepshtein and Kubovy, 2007) continued this work and found that regimes of apparent motion in supra-threshold ambiguous stimuli can be predicted based on spatiotemporal contrast sensitivity at threshold levels. They started with considering the discrepancy between Korte's third law of apparent motion (1915), or space-time coupling, and later findings suggesting that space-time trade-off provides optimal conditions for apparent motion perception (i.e., maximal sensitivity). More specifically, space-time coupling means that, in order to achieve optimal perceived apparent motion, an increase in spatial distance between the two elements that are connected in apparent motion should be accompanied by an increase in temporal distance. Space-time trade-off means that an increase in spatial distance between the two elements that are connected in apparent motion

should be accompanied by an decrease in temporal distance, refuting Korte's law. Space-time tradeoff was found for several kinds of stimuli. Gepshtein and Kubovy (2007) employed dot motion lattices and showed that these seemingly contradicting findings can be reconciled by the optimization of speed estimation in a system with limited resources. Accordingly, whether space-time coupling or space-time trade-off occurs depends on the speed of motion, where coupling occurs for high speeds and trade-off for low speeds.

Considering the findings of Nakayama (1985) on the analogy of motion perception of grating stimuli at detection threshold and apparent motion perception above threshold, Gepshtein and Kubovy (2007) reasoned that both characteristics of motion perception at and above threshold might be the result of the optimal allocation of limited neural resources in our visual system. Therefore, the sensitivities to apparent motion perception may be related to the contrast sensitivity function. For comparison of the contrast sensitivity function (Kelly, 1979) to the ambiguous motion lattices, spatiotemporal frequencies were transformed into spatiotemporal distances by the quadrature rule of Nakayama (1985), where frequency *F* corresponds to a distance *D* of 1/(4**F*). As spatiotemporal properties leading to optimal motion perception (i.e., maximal sensitivity) all expend the maximal amount of resources, Gepshtein and Kubovy (2007) predicted that equally suboptimal spatiotemporal properties leading to equal sensitivities to motion perception (i.e., isosensitivity) should all expend an equally suboptimal amount of resources. More specifically, they found that a maximum probability of perceiving one apparent motion in a dot motion lattice with certain spatiotemporal properties is analogous to a maximum sensitivity point for equivalent spatiotemporal frequencies at detection threshold. Similarly, equivalence in probability of perceiving two different apparent motions in a dot motion lattice with certain spatiotemporal properties is analogous to equisensitivity points for equivalent spatiotemporal frequencies at detection threshold.

Zharikova, Gepshtein, and van Leeuwen (2017) found that perceptual organization of the 'local-global' ambiguous motion stimulus does not follow the principle of grouping by proximity. The local-global ambiguous motion stimulus enables both a relatively local interpretation of six dot pairs rotating and a relatively global interpretation of two hexagons pulsating. The perceptual experience of this stimulus is binary, implying that an observer usually has the ability to experience these two different percepts, but cannot experience both

percepts at the same time. The point of subjective equality (PSE) is the point at which two percepts are equally likely to be perceived by a given observer. If perception of the 'local-global' stimulus is ultimately determined by proximity, the distances between the elements of the local and global percept should be equal at the PSE between the two percepts. In contrast, the distances between the elements of the global percept were larger than the distances between the elements of the local percept at the PSE.

Similar to the drifting grating stimulus and apparent motion in dot lattices, spatial and temporal frequency are important stimulus properties of the 'local-global' ambiguous motion stimulus as well. According to Zharikova et al. (2017), the frequency of this stimulus is determined by how far apart in space and time visual information is presented. More specifically, the distance between adjacent dots determines spatial frequency, whereas the time between two successive frames determines temporal frequency. In line with Gepshtein and Kubovy (2007), Zharikova et al. (2017) reasoned that the CSF may be predictive for the perception in the 'local-global' stimulus. They stated that the likelihood of an observer experiencing perceptual grouping over a particular spatiotemporal distance would be related to the contrast sensitivity of this observer for this particular spatiotemporal distance. Therefore, as perceptual grouping is equally likely for the local and global percept at the PSE, spatiotemporal contrast sensitivity should be equal for the distance between the elements of the local percept (DsL, or w in their study) and the distance between the elements of the global percept (DsG, or b in their study) at the PSE (Figure 1).

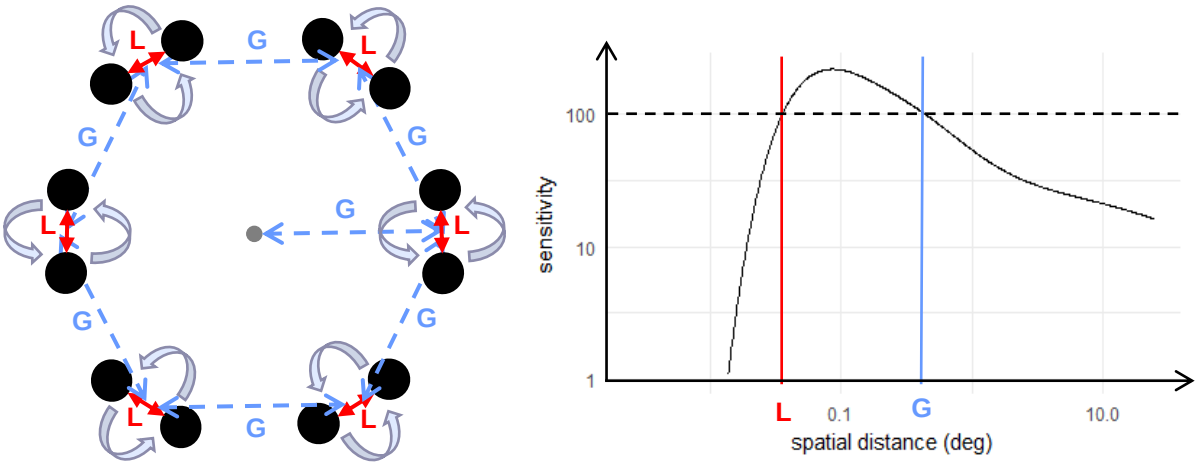


Figure 1. Left: Ambiguous motion stimulus with indicated apparent motion paths (in light blue arrows), distances between the elements of the local percept (L, in red), and distances between the elements of the global percept (G, in blue). An animated version of this stimulus can be found via the link https://osf.io/k46sj/?view_only=12317c5452064b92bec1c36df860d19c. **Right**: Example of contrast sensitivity function for temporal distance of 170 ms. The dashed line indicates equisensitivity for spatial distances L and G, corresponding to distances in the ambiguous motion stimulus at the point of subjective equality between the local and the global percept.

They measured individual contrast sensitivity in three observers for one temporal frequency (*Ft* = 1.47 Hz, corresponding to *Dt* = 170 ms) and multiple spatial frequencies (*Fs* = 0.1, 0.3, 0.7, 1.22, 2.45, 4.13, 6.28, 10.47, and 15.00 c/deg) and mapped this to these observers' percepts of the 'local-global' ambiguous motion stimulus, which they measured for the same temporal distance (*Dt* = 170 ms) and five different spatial distances between the elements of the local percept (*DsL* = 0.15, 0.25, 0.38, 0.64 and 1.28 deg). For comparison of the CSF to the 'local-global' stimulus, spatiotemporal frequencies were transformed into spatiotemporal distances by the quadrature rule here as well. They found correspondences between these two measurements and concluded that perception in the 'local-global' ambiguous motion stimulus can be predicted by individual contrast sensitivity as measured by the CSF. This suggests that measures of sensitivity to visual input are correlated for "low-level" and "mid-level" (i.e., perceptual organization into a "local" or "global" percept) measures within the classical visual hierarchy and may implicate a causal relationship between the two.

As the study of Zharikova et al. (2017) was carried out on a limited number of participants and the measures of correspondence between individual CSF and perception of the 'local-global' stimulus were rather limited as well, we sought to further examine this correspondence in more observers and stimulus conditions. We addressed the question whether the "low level" characteristic of spatiotemporal luminance contrast sensitivity could indeed be correlated to the perception of the 'local-global' ambiguous motion stimulus in two different empirical studies. We will outline both studies here, but will not provide as much detail on the methods and results as in the other empirical chapters, as the essential aspects of this chapter will be illustrated in the subsequent General Discussion. More specifically, during the implementation of these studies, we came to realize that there is a critical and fundamental problem in the conceptual interpretation of some of the elements in the theory relating contrast sensitivity to apparent motion perception. We will clarify this problem in the order that it became evident to us.

4.2 General Methods

Apparatus

Stimuli were presented on a CRT monitor (Sony GDN F520) with a resolution of 1024 x 768 pixels and a diagonal of 50 cm. A Bits# Stimulus Processor (Cambridge Research Systems) was coupled to this monitor in order to be able to display smaller differences in contrast for the CSF measurement. Participants were instructed to place their chin on a chin rest which was positioned at 57 cm from the monitor.

Stimuli

'Local-global' ambiguous motion stimulus

The 'local-global' ambiguous motion stimulus (Figure 1) consists of 6 pairs of dots with a radius of 0.2 deg presented on a uniform grey background (luminance 17.9 cd/m^2). The temporal distance between two successive frames is *Dt*. On each successive frame, the dots change slightly in position (i.e., a rotation around the pair-center of π/15 rad per frame) and when multiple frames are presented successively, apparent motion can be perceived. The apparent

motion can be ascribed to two different percepts: The local percept consists of six groups organized in dot pairs each rotating around themselves; the global percept consists of two groups organized in shrinking and growing hexagons of six dots. The spatial distance between the elements of the local percept is *DsL*, whereas the spatial distance between the elements of the global percept is *DsG*.

Drifting sinusoidal luminance contrast grating

The CSF was measured by means of horizontally drifting sinusoidal luminance contrast gratings (i.e., Gabor stimuli) with specific temporal and spatial frequencies, which were each presented for 500 ms on a uniform grey background (luminance 17.9 cd/m^2).

Procedure

Measurement of point of subjective equality (PSE) between 'local' and 'global' percept of the 'local-global' ambiguous motion stimulus

An individual between-dot distance DsG at the PSE between the 'local' and 'global' percept of the 'local-global' ambiguous motion stimulus was measured for each combination of spatial within-dot-pair distance DsL and temporal distance Dt by means of an adaptive staircase method (Psi, Kontsevich, & Tyler, 1999). Three Psi staircases of 50 trials were presented interleaved in randomized order in order to avoid sequence effects, resulting in 150 trials separated by a short break every 50 trials. Participants were instructed to indicate their 'local' vs. 'global' percept by means of a manual keyboard response, which was counterbalanced across participants. Based on the perceptual responses, the adaptive staircase determined the spatial between-dot distance DsG with which the 'local-global' stimulus was presented on each trial.

Measurement of (spatiotemporal) contrast sensitivity function (CSF)

Trials were presented in a random order with a short break after every 50 trials. Each trial started with the presentation of a central fixation point, followed by the presentation of the grating stimulus for 500 ms. Participants were instructed to indicate whether the grating stimulus was presented at the left or right side of the fixation point by means of a keypress, which was counterbalanced across participants. Similar to Zharikova et al. (2017), luminance

contrast for each these spatiotemporally different gratings was varied by a staircase procedure with a 1-up-3-down rule in the range of [0.004, 1] in order to determine the observer's CSF.

As the estimation of an observer's sensitivity from an adaptive psychophysical up-down staircase converges towards the end of the staircase, we can reliably estimate sensitivity by means of the last trials. We averaged the last 10 trials of each Psi staircase in order to obtain our estimates of sensitivity. Individual contrast sensitivity data based on the presented grating stimuli were used to fit the model of the full spatiotemporal CSF in order to gain a measure of the individual CSF across the full range of spatiotemporal frequencies. From the individual CSF, the equisensitivity points for distances DsL were estimated, corresponding to estimates of distances DsG at the PSE in the ambiguous motion stimulus. More specifically, out of this CSF, we extracted the range/slice where *Ft* is equivalent to *Dt* in the 'local-global' stimulus and derived the sensitivity for *DsL*. We then used this sensitivity to find the equisensitivity point, another distance *Ds* for which this sensitivity holds, which would then constitute the estimation of *DsG* from the CSF.

4.3 Experiment 1

In a first experiment, we measured individual contrast sensitivity and percepts of the 'local-global' ambiguous motion stimulus in 17 observers. Moreover, we examined whether both measures would remain stable over a time course of 5 days during which we manipulated "higher-level" relevance of percepts of the 'local-global' stimulus in an auxiliary task, equivalent to the research question and manipulations in Chapter 2, 'Relevance study', of this PhD thesis.

Stimuli

For the 'local-global' stimuli, the temporal distance Dt between successive frames was 33 ms and the distance between the elements of the local percept, DsL was held constant at an edge-to-edge distance of 0.39 deg. The distance between the elements of the global percept, DsG, was varied according to the adaptive procedure in the estimation of the PSE. During the execution of the task during the five successive days, DsG was held constant at the individual measure of the PSE. In order to facilitate the relevance manipulation, we applied a color scheme to the dots of the 'local-global' stimulus so that one dot in each pair was white

(luminance 76.8 cd/m^2), whereas the other was black (luminance 0.2 cd/m^2). For more details on the 'local-global' stimulus characteristics, we refer to Chapter 2 in this PhD thesis, 'Relevance study'. Variations of grating stimuli presented during the estimation of the CSF consisted of combinations of spatial frequencies *Fs* (0.1, 0.3, 1, 3.0, 9.0 c/deg) with temporal frequency *Ft* equal to 15 Hz. Other details of the stimuli are described in the General Methods.

Procedure

During the first and last day, we measured every observer's CSF and PSE between the 'local' and 'global' percept of the 'local-global' ambiguous motion stimulus. Individual CSF were modeled based on the responses to different grating stimuli presented at different contrast levels as described in the General Methods. Subsequently, out of the modeled spatiotemporal CSF, we extracted the range/slice where Ft is equivalent to Dt = 33 ms in the 'local-global' stimulus and derived the sensitivity for DsL. We then used this sensitivity to find the equisensitivity point, another distance Ds for which this sensitivity holds, which would then constitute the estimation of DsG from the CSF. The estimation of the individual PSE was carried out according to the General Methods. During the task, participants were presented with a 'local-global' stimulus that had a distance DsG equal to their estimated PSE.

Between the first and last day, observers were assigned to different relevance conditions, which would define the task they performed in 5 short sessions over the 5 successive days: 6 observers were assigned to the "local" condition, 6 observers were assigned to the "global" condition, and 5 observers were assigned to the "exposure" condition. In the "local" condition, observers performed a 'local task' with the 'local-global' stimulus, for which the local percept was relevant and the global percept irrelevant. In the "global" condition, observers performed a 'global task', for which the global percept was relevant and the local percept irrelevant. In the "exposure" condition, neither percept was relevant to perform the task. For more details on the manipulation of "higher-level" relevance in an auxiliary task, we refer to Chapter 2 in this PhD thesis, 'Relevance study'.

Results & Discussion

Our results on the perception of the 'local-global' stimulus indicated that there was a 'shift to global', implying that the 'global' percept became more dominant over time, irrespective of

the condition of task relevance. For a further discussion of this finding, we refer to Chapter 2, 'Relevance study', where we carried out a similar experiment and found similar results.

In addition, there seemed to be no clear relationship between individual perception of the 'local-global' stimulus and predictions of this perception from individual CSF in our data. This suggests that we did not find evidence for the conclusion of Zharikova et al. (2017), stating that individual CSF predicts perception of this stimulus.

However, during the fitting of the data based on the presented grating stimuli to the full spatiotemporal model of the CSF, we found that the model had not converged towards a reliable estimation of all parameters. Moreover, the parameter values were highly different from the theoretically expected values, and the fitted CSF did not correspond well to our observed data points. Therefore, we decided that we would carry out a second experiment to measure the CSF and the perception of the 'local-global' stimulus at precisely the same temporal distances in order to minimize the potential unknowns and errors in the fitting of the CSF. Moreover, as perception can be highly variable across spatiotemporal characteristics and we restricted our measurement of perception of the 'local-global' stimulus to only one combination of temporal distance Dt and spatial distance DsL, we decided to broaden the range of spatiotemporal characteristics studied in the second experiment. Lastly, in order to focus on an optimal measurement of the 'local-global' stimulus in correspondence to the CSF, we decided to remove the relevance manipulation and change the color scheme of the dots so that they would all have the same luminance values.

4.4 Experiment 2

In a second experiment, we aimed to minimize the uncertainties and difficulties with the fitting of the CSF and decided to measure the CSF for multiple temporal frequencies and the perception of the 'local-global' stimulus for temporal distances between frames that were identical to those temporal frequencies. Accordingly, we had psychophysical measurements at both identical temporal distances and the fitting of the CSF was therefore necessary only across spatial distances.

We again measured both individual perception of the 'local-global' ambiguous motion stimulus and contrast sensitivity, but this time across a wider range of spatiotemporal

frequencies, in five observers. Accordingly, we aimed to characterize perception of the ambiguous motion stimulus in interaction with spatiotemporal stimulus factors, and to explore whether the potential correspondence of individual CSF to ambiguous motion perception holds across this wider range of spatiotemporal frequencies.

Stimuli

Variations of 'local-global' ambiguous motion stimuli consisted of combinations of spatial within-dot-pair distances (0.14, 0.32, 0.68, and 1.40 deg measured from the edges of the dots) and temporal distances (166, 100, 83, and 33 ms between successive frames). In contrast to the previous experiment, all dots in the 'local-global' stimulus were now colored white (luminance 76.8 cd/m^2). Variations of grating stimuli consisted of combinations of spatial frequencies (0.1, 0.3, 1, 3.0, 9.0 c/deg) and temporal frequencies (1.5, 2.5, 3, 7.5 Hz, corresponding to temporal distances of 166 ms, 100 ms, 83 ms, 33 ms by the quadrature rule, which were also employed in the 'local-global' stimulus). Other details of the stimuli are described in the General Methods.

Procedure

For every combination of spatiotemporal distances in the 'local-global' stimulus, a PSE was measured by means of the adaptive procedure in the General Methods, resulting in 16 runs of 150 trials. Grating stimuli were presented in every combination of spatiotemporal frequencies. For each of these combinations, 30 trials were presented, resulting in a total of 600 trials. For every *measured* temporal frequency *Ft* (equivalent to *Dt* in the 'local-global' stimulus), we modeled the CSF and derived the sensitivity for *DsL*. We then used this sensitivity to find the equisensitivity point, another distance *Ds* for which this sensitivity holds, which would then constitute the estimation of *DsG* from the CSF. Other details of the procedure are described in the General Methods.

Results & Discussion

The results of perception of the 'local-global' stimulus across different spatiotemporal distances indicate that, overall, observers tend to perceive the ambiguous motion stimulus as more 'global' when there is a smaller temporal distance between frames and a larger spatial distance between the dots within a dot pair (i.e., local spatial distance DsL) (Figure 2). The

dependence on temporal distance between frames indicates that the global percept is strengthened for more continuous apparent motion. The finding of strengthened global perception for larger local spatial distance seems plausible since a larger distance between the elements of the local percept should result in a larger distance between the elements of the global percept, if there is indeed a violation of grouping by proximity (Zharikova et al., 2017).

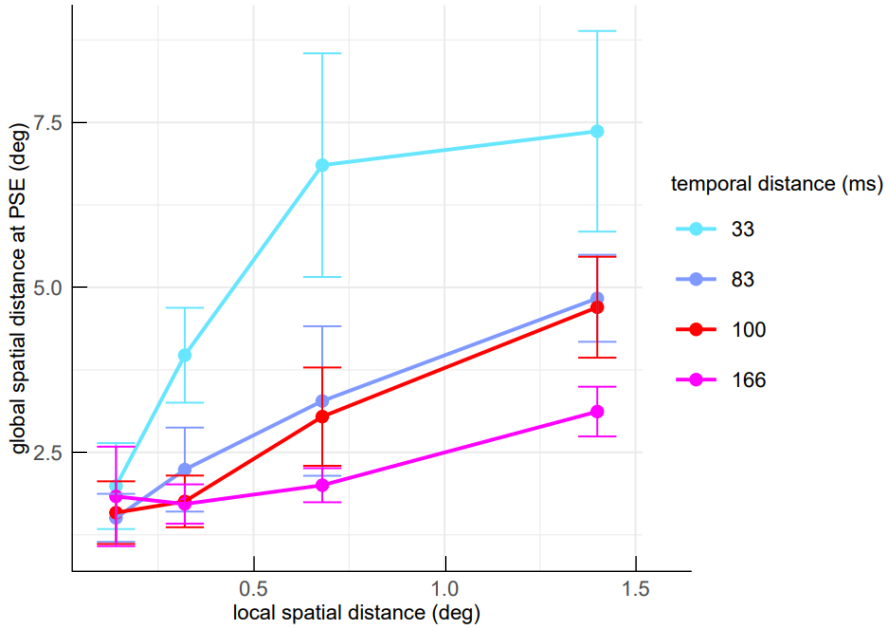


Figure 2. Mean spatial distances between the elements of the global percept at the point of subjective equality for different spatial distances between the elements of the local percept and different temporal distances between frames in the 'local-global' motion stimulus. Error bars indicate one standard error around the mean.

Another important conclusion is the fact that distances between the elements of the global percept, DsG, are systematically larger than distances between the elements of the local percept, DsL, at the PSE. Therefore, we replicate the finding of Zharikova et al. (2017), stating that grouping between the elements in the 'local-global' stimulus does not follow the Gestalt principle of grouping by proximity, as the distances between the elements of either percept are not identical at the PSE (i.e., the measure of transition ratio, corresponding to DsG/DsL at PSE in our measures, is larger than 1).

In figure 3, we displayed the measured contrast sensitivities for different spatiotemporal frequencies in all five individuals. The model of the CSF was fitted to these measured data, and from this fit we obtained predictions for DsG at the PSE. Importantly, the correlation

between the *estimated* DsG from individual CSF and individually *measured* DsG at the PSE in the 'local-global' stimulus across the measured spatiotemporal distances was merely 0.188. As can be seen in Figure 4, the measured DsG at the PSE are also systematically larger than the estimated DsG at the PSE. Our results indicate that, in contrast to the conclusions of Zharikova et al. (2017), we cannot predict perception of the 'local-global' ambiguous motion stimulus from measures of perception at threshold in the individual contrast sensitivity function.

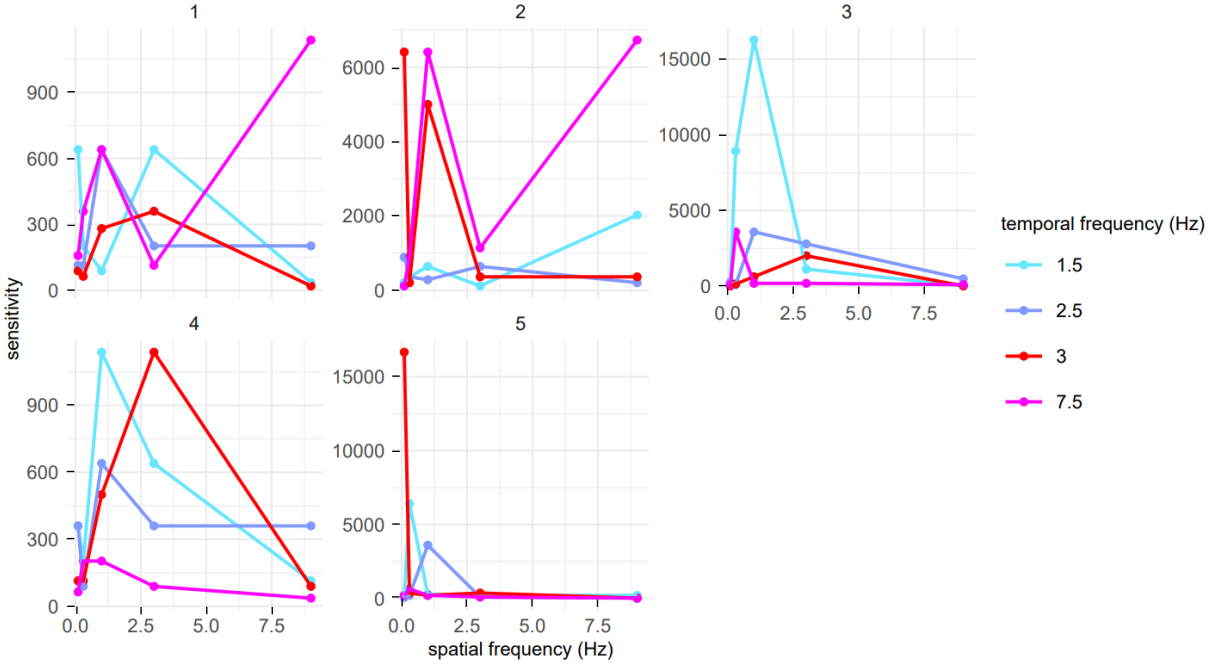


Figure 3. measured contrast sensitivities for different spatiotemporal frequencies in all five observers.

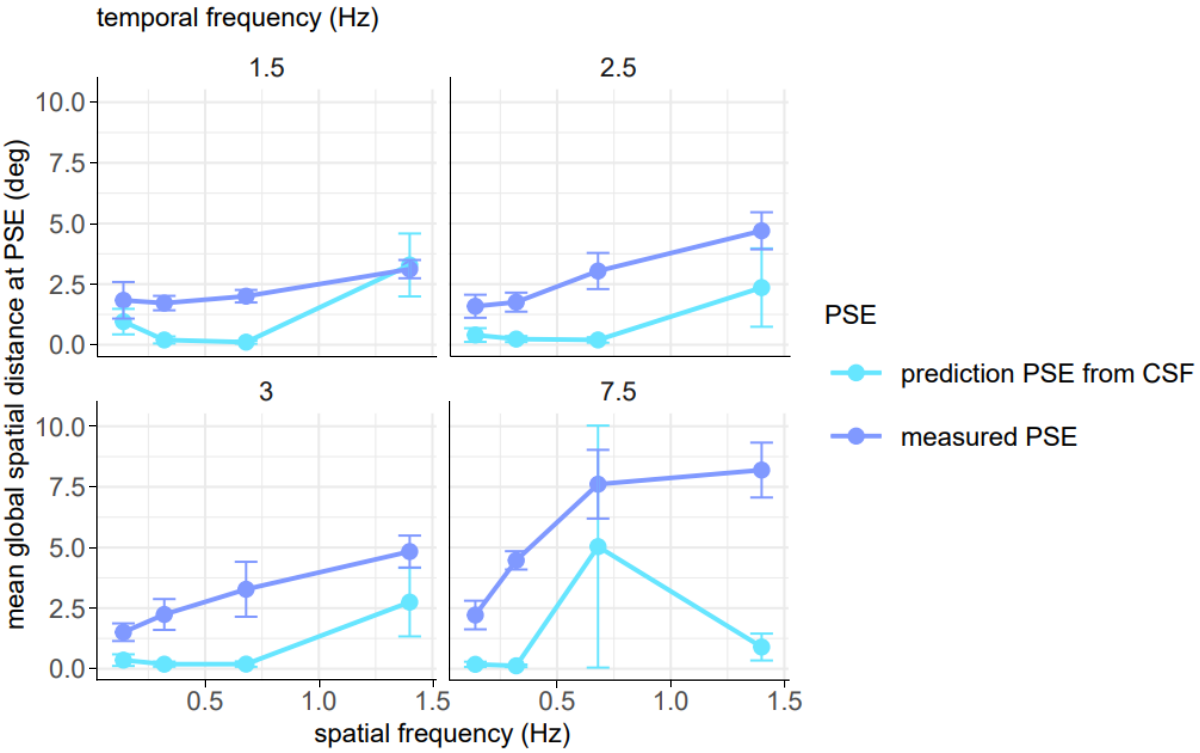


Figure 4. Means of estimations and measurements of global spatial distance at the PSE. Error bars depict one standard error around the mean.

One important observation complicating this study was the position of the CSF across spatial distances. For the majority of individuals, sensitivity was on the descending end of the CSF for DsL, resulting in an equisensitivity point corresponding to DsG on the ascending end of the CSF (Figure 5). In other words, the equisensitivity point corresponding to the sensitivity for the distance between the elements of the local percept (DsL) in the CSF, constituting the estimate for DsG, was often at a smaller distance than for DsL. If the relationship to perception in the ambiguous stimulus holds, this indicates that the distance between the elements of the global percept, DsG, should also be *smaller* than the distance between the elements of the local percept, DsL, at the PSE. In other words, the transition ratio (i.e., Ds/Ds at the PSE), as employed in the study of Zharikova et al. (2017), should be smaller than 1. However, the measured DsG at the PSE was larger than DsL for all observers in our study, indicating that the transition ratios were systematically larger than 1.

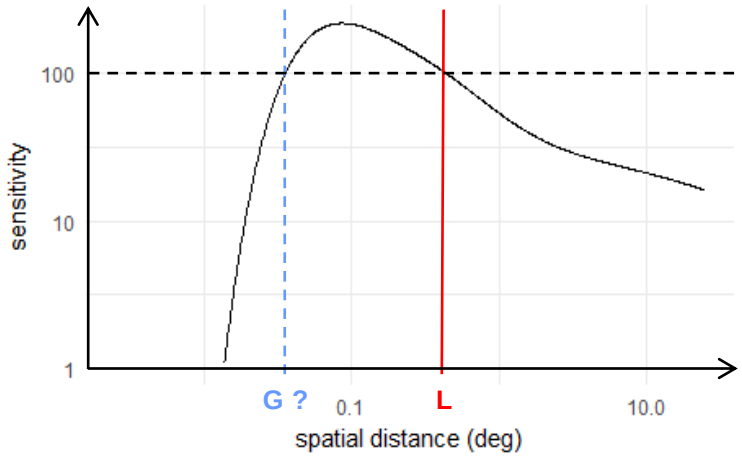


Figure 5. Hypothetical depiction of the problem encountered concerning the position of the CSF across spatial distances. When The CSF is at its descending end for spatial distance = L (DsL), the estimation of the equisensitivity point corresponding to distance G (DsG) will be at the ascending end of the CSF and will therefore be smaller, instead of larger than L (DsL).

Similarly to our results on the perception of the 'local-global' stimulus, Zharikova et al. (2017) also found that DsG was larger than DsL at the PSE in *most* of their participants, a finding supporting their conclusion of the defiance of grouping by proximity. They also reported that, in contrast to this, some of the individuals in their study had transition ratios predicted from CSF which were smaller than 1. However, they reported that these individuals had transition ratios measured during perception of the 'local-global' stimulus that were smaller than 1 as well. In principle, this would be logical if the measured transition ratios are indeed similar to the estimated transition ratios from CSF. However, this contrasts their earlier findings stating that transition ratios were generally larger than 1. Moreover, importantly, when the transition ratio is smaller than 1, the dots of different dot pairs start to overlap and the perception of this stimulus starts having a different appearance.

There is no way for the transition ratio to be smaller than 1 without changing the possible percepts phenomenologically. More specifically, as the distance between the elements of the global percept becomes smaller than the distance between the elements of the local percept, the global percept remains intact and becomes stronger, but the local percept is close to impossible to experience. We refer to the demonstration videos in the following link to try this out for yourself:

https://osf.io/8a6kn/?view_only=e688cc53fdd141b481f5906b8589b845.

The result is a new type of stimulus for which there is no distinction between a local vs. global percept (i.e., the possibility of a local percept is eliminated). Therefore, we believe that phenomenologically meaningful measured transition ratios cannot be smaller than 1.

4.5 General Discussion

Similar to Zharikova et al. (2017), we found that perception of the 'local-global' stimulus indeed violates grouping by proximity, with distances between the elements of the global percept, DsG, systematically larger than distances between the elements of the local percept, DsL, at the point of subjective equality (PSE). However, in contrast to Zharikova et al. (2017), we did not find a correspondence between the perception of the 'local-global' stimulus and predictions from individual luminance contrast sensitivity. Indeed, the correlation between individual perception of the 'local-global' stimulus and predictions from individually fitted CSF in our study was merely 0.188.

There may be multiple reasons for the discrepancy of our findings with the conclusions from the study of Zharikova et al. (2017). Firstly, the CSF is very difficult to measure and perception of ambiguous stimuli can be highly variable, adding a substantial amount of noise to all measurements. Moreover, even though the psychophysical measurements are based on a large number of trials per observer, the number of participants in the study of Zharikova et al. (2017) and our second experiment was rather low (i.e., three and five, resp.), leaving open the possibility that a study with a bigger sample size would provide more robust findings.

However, given this large inconsistency between our findings and the conclusions of Zharikova et al. (2017), we decided to examine the literature forming the basis of the theoretical predictions more thoroughly. The conclusions we derived based on this second investigation of the literature completely changes the perspective on this research question. However, the matter of measurement of luminance contrast sensitivity and predictions derived from this measurement is highly complex. In the following section, we aim to guide the reader through our analysis as didactically as possible.

Nakayama (1985)

As mentioned in the introduction section of this chapter, Nakayama (1985) found that luminance contrast sensitivity for a drifting grating stimulus with spatial frequency *Fs* (cycles/deg) and temporal frequency *Ft* (cycles/sec or Hz) will be identical for a discontinuous displacement of this grating stimulus by spatial distance *Ds* (deg) = 1/(4**Fs*) over temporal distance *Dt* (sec) = 1/(4**Ft*) (i.e., the quadrature rule, Figure 6). Here, it is important to keep in mind that the spatial frequency of a grating stimulus is a *within-frame* (i.e., static) characteristic, whereas the temporal frequency is an *across-frame* (i.e., motion) characteristic, and that the spatial and temporal distance for apparent motion are both *across-frame* characteristics. That is, the spatial frequency of a grating stimulus is measurable within one presentation frame of the stimulus, whereas its temporal frequency, spatial distance and temporal distance are measurable only across two (or more) successive presentation frames (Figure 6).

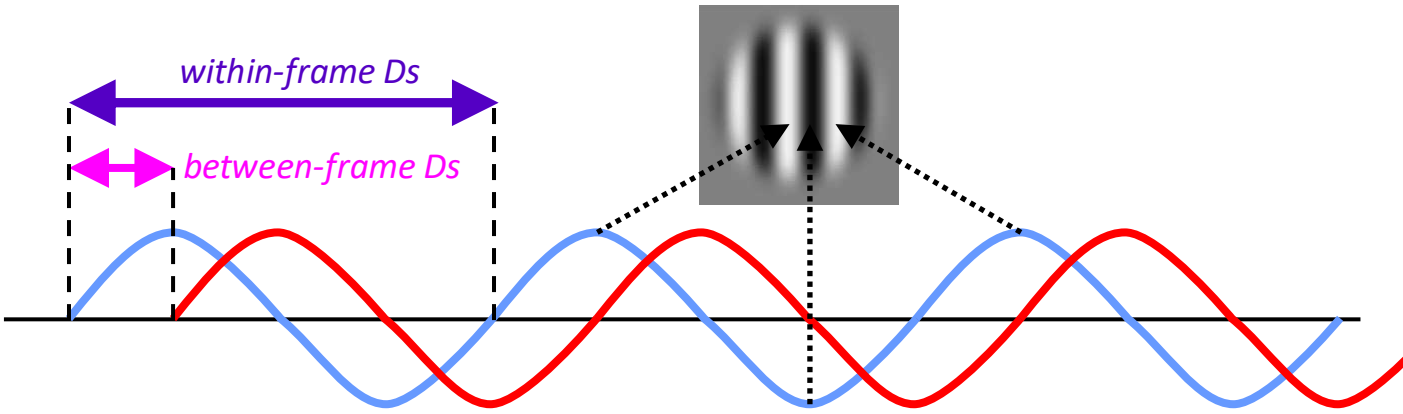


Figure 6. Sinusoidal waveform making up a luminance contrast grating stimulus. Apparent motion between a first frame (in blue) and a second frame (in red).

Gepshtein & Kubovy (2007)

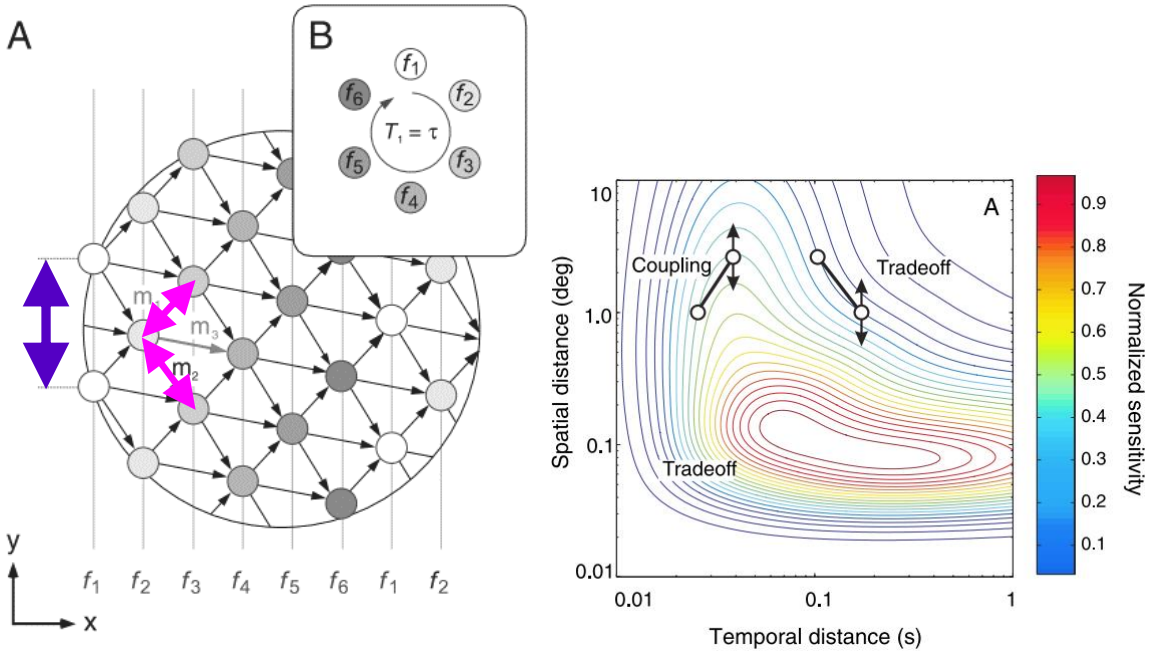


Figure 7. Left: Stimuli employed by Gepshtein and Kubovy (2007) in greyscale, superimposed by the equivalent distances to Nakayama (1985) and our study in color. Within-frame distance in purple, between-frame distance in pink. **Right:** The theoretical depiction of the spatiotempotal contrast sensitivity function (CSF) in distances by Gepshtein and Kubovy (2007).

Nakayama (1985) stated that the quadrature rule could be applied, not only to apparent motion of grating stimuli, but also to apparent motion of stimuli above threshold levels. As pointed out in the introduction, Gepshtein and Kubovy (2007) built on this conclusion and stated that equivalence in (or maximum) probability of perceiving two different (or one) apparent motions in a dot motion lattice with certain spatiotemporal properties is analogous to equisensitivity (or maximum sensitivity) points for equivalent spatiotemporal frequencies at detection threshold. They employed dot lattice stimuli with different spatiotemporal distances *between* successive frames and found that regimes of apparent motion in supra-threshold ambiguous dot lattice stimuli could indeed be predicted based on spatiotemporal contrast sensitivity at threshold levels (Figure 7). Based on the quadrature rule, they found similarities in the probability of (or sensitivity for) apparent motion perception along *between-frames* spatial distances *Ds* over temporal distances *Dt* and theoretical sensitivity for luminance contrast in grating stimuli with spatiotemporal frequencies corresponding to *between-frames* distances *Ds* and *Dt*.

Importantly, however, they clarified that the correspondence between sensitivity at threshold levels and sensitivity for apparent motion of supra-threshold stimuli in their "normative theory" was only defined by a *qualitative* agreement: "A normative theory predicts that estimates of the sensitivity of the visual system depend on the task and the stimuli used in obtaining the estimates. […] On this view, there would be little reason to expect a *quantitative* agreement between the equivalence contours obtained using different tasks and different stimuli" (Gepshtein & Kubovy, 2007, p. 12). This implies that there was little reason to expect a quantitative agreement between individual CSF and perception of the 'local-global' ambiguous motion stimulus, as studied by Zharikova et al. (2017), to begin with.

Zharikova, Gepshtein & van Leeuwen (2017)

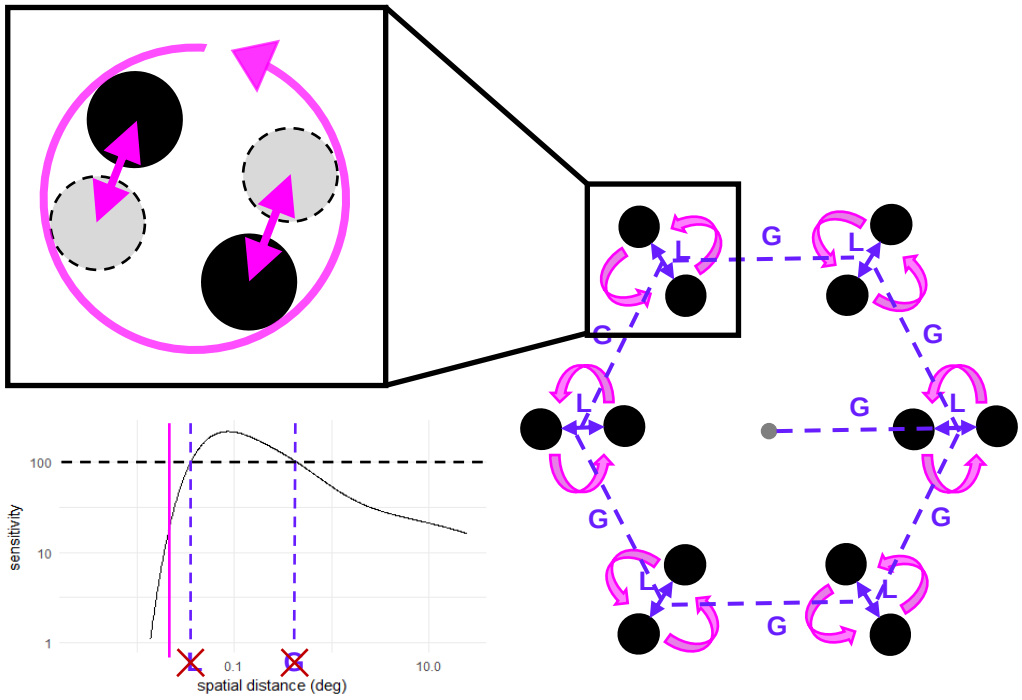


Figure 8. 'Local-global' ambiguous motion stimulus. Within-frame distance in purple, between-frame distance in pink. When comparing this Figure to previous Figure 1, it becomes clear that it is in fact not meaningful to translate the spatial distances in the 'local-global' stimulus (depicted in red and blue in the previous figure, both depicted in purple in the current figure) to spatial distances in the CSF as we did in the previous figure. Rather, the one distance that can be meaningfully translated to the spatial distances in the CSF is the between-frames distance, depicted here in pink, which is identical for the local and global percept.

As described in the introduction, Zharikova et al. (2017) followed the reasoning of Gepshtein and Kubovy (2007), but interpreted the relationship between perception at threshold levels to perception at supra-threshold levels in a more determined manner, implying quantitative predictions instead of merely qualitative ones. As mentioned above, a quantitative agreement between both measures may not be very likely. However, we failed to find even a qualitative agreement between both measures in our study, indicating that there may be more fundamental reasons for the discrepancy of our findings with those by Zharikova et al. (2017).

As perceptual grouping is equally likely for the local and global percept at the PSE, Zharikova et al. (2017) predicted that spatiotemporal contrast sensitivity may be equal for the distance between the elements of the local percept (*DsL*, or w in their study, with "w" for "within pairs") and the distance between the elements of the global percept (*DsG*, or b in their study, with "b" for "between hexagons") at the PSE (Figure 8). However, critically, in contrast to Nakayama (1985), Zharikova et al. (2017) interpreted the spatial distance Ds as the distance between elements *within* one presentation frame (i.e., the distance between the elements in the local percept as *DsL* and the distance between the elements of the global percept as *DsG*), instead of elements *between* successive frames. Moreover, they considered their interpretation of *Ds* as analogous to the measure of *Ds* corresponding to the predictions of Nakayama (1985) on sensitivity. In order to avoid confusion, we will refer to the interpretation of *Ds* in the work of Zharikova et al. (2017) as *within-frame Ds*, whereas we will refer to the other interpretation of *Ds* as *between-frame Ds*.

Given that Zharikova et al. (2017) interpreted the theory of Nakayama (1985) and Gepshtein and Kubovy (2007) as referring to *within-frame Ds* rather than *between-frame* distance *Ds*, the predictions of perception of the 'local-global' stimulus based on contrast sensitivity in the study of Zharikova et al. (2017) as well as in both of our experiments are meaningless. The correct interpretation of *between-frame* distance *Ds* would have been the distance connecting both ends of the rotation around the pair-center of π/15 rad per frame, which is identical for both the 'local' and 'global' percept. The elements of the 'local' and 'global' percept are grouped differently in perceptual organization, but are identical between successive frames, implying that the correct *between-frame* distance between the elements of the local percept, *DsL,* was identical to the correct *between-frame* distance between the elements of the global

percept, *DsG*. This results in only one point in the CSF and eliminates any meaningful comparison between the 'local' and 'global' percept based on contrast sensitivity (Figure 8).

Back to Nakayama (1985) and Gepshtein & Kubovy (2007)

Based on our analysis in the previous section, we can conclude that the focus on *within-frame* distance *Ds* was incorrect. However, the *within-frame* distance is surely of importance for perception at threshold levels and perception of apparent motion as well. Therefore, we asked whether within-frame distance Ds was included in the theory and analysis of Nakayama (1985) and Gepshtein and Kubovy (2007).

An important detail of the quadrature rule is the fact that Nakayama (1985) states that the sensitivity for a drifting grating will be the same as the sensitivity for apparent motion with a displacement along spatiotemporal *between-frame* distances *Ds* and *Dt* of the *same* grating stimulus that was employed in the drifting situation, which has the characteristic of a spatial frequency equal to *Fs* (corresponding to *within-frame Ds*). This aspect is crucial, as (static) grating stimuli of many different *within-frame* spatial frequencies *Fs* can be moved along *between-frame* spatial distance *Ds* and temporal distance *Dt* to create apparent motion. However, the sensitivity to these different apparent motions with different *within-frame* spatial frequencies will *only* be identical to the sensitivity for a continuously drifting grating stimulus (with *within-frame* spatial frequency *Fs* and temporal frequency *Ft*) when the stimulus used to create the apparent motion has a *within-frame* spatial frequency equal to the *Fs* of the continuously drifting grating stimulus.

Based on the quadrature rule, where frequency *F* corresponds to a distance *D* of 1/(4**F*), Gepshtein and Kubovy (2007) transformed the contrast sensitivity function (CSF, Kelly, 1979) along different spatiotemporal frequencies into a CSF-equivalent along different spatiotemporal distances (Figure 7, Right). However, importantly, the authors did not take into account that every point in this CSF-equivalent was still dependent on the fact that the (*within-frame*) spatial frequency of the stimulus (which was employed to induce apparent motion) had to be identical to its equivalent based on the (*between-frame*) spatial distance of this point in the CSF-equivalent. This may lead a reader of this CSF-equivalent graph to believe that sensitivity for apparent motion of *any* stimulus with *any* (*within-frame*) spatial frequency (or *within-frame* distance) characteristics can be predicted purely based on the *between-*

frame aspects of the apparent motion (i.e., the spatial and temporal distance between successive frames). On the contrary, as Nakayama (1985) pointed out and as is evident from the practice of measurement of the static spatial CSF, the within-frame spatial frequency properties are important for any measurement of visual sensitivity.

Perhaps Gepshtein and Kubovy (2007) did not take the *within-frame* distance into account because they related motion perception in threshold as well as supra-threshold levels to optimization of speed perception, rendering static properties less important. However, Gepshtein and Kubovy published an earlier paper in which they examined and emphasized precisely the differential influence of *within-frame* and *between-frame* distances in apparent motion (Gepshtein & Kubovy, 2000). Moreover, in the discussion of the later paper, they underline the importance of *within-frame* spatial frequency: "an optimal visual system should change the distribution of its sensitivity across the parameters of stimulation so that its estimate of stimulus frequency content becomes more reliable than its estimate of stimulus location" (Gepshtein & Kubovy, 2007, p. 12).

We conclude that, even though there is definitely some kind of relationship between spatiotemporal contrast sensitivity and apparent motion perception of ambiguous stimuli, this relationship is far from straightforward to characterize. Future research should reconsider the specific spatial characteristics of sensitivity to motion perception at threshold, and the theory of Gepshtein and Kubovy (2007) should be expanded to include predictions for *within-frame* spatial distances. Even so, given the wide variety of additional processes that can be associated to these different tasks (e.g., different demands in attention, different response strategies, etc.), expecting the relationship between these different levels of processing to be quantitatively similar may be asked too much. In fact, a qualitative correspondence between the two is much more likely.

References

Dosher, B. A., & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences*, *95*(23), 13988-13993. https://doi.org/10.1073/pnas.95.23.13988

Feigin, H., Baror, S., Bar, M., & Zaidel, A. (2017). Behaviorally relevant prior experience biases subsequent perception. *Journal of Vision*, *17*(10): 493. https://doi.org/10.1167/17.10.493

Gepshtein, S., & Kubovy, M. (2000). The emergence of visual objects in space–time. *Proceedings of the National Academy of Sciences*, *97*(14), 8186-8191. https://doi.org/10.1073/pnas.97.14.8186

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision*, *7*(8): 9. https://doi.org/10.1167/7.8.9

Graham, N. V. S. (1989). *Visual Pattern Analyzers*. Oxford University Press.

He, D., Kersten, D., & Fang, F. (2012). Opposite modulation of high-and low-level visual aftereffects by perceptual grouping. *Current Biology*, *22*(11), 1040-1045. https://doi.org/10.1016/j.cub.2012.04.026

Kelly, D. H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. *Josa*, *69*(10), 1340-1349. https://doi.org/10.1364/JOSA.69.001340

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729-2737. https://doi.org/10.1016/S0042-6989(98)00285-5

Nakayama, K. (1985). Biological image motion processing: a review. *Vision Research*, *25*(5), 625-660. https://doi.org/10.1016/0042-6989(85)90171-3

Silverman, M. S., Grosof, D. H., De Valois, R. L., & Elfar, S. D. (1989). Spatial-frequency organization in primate striate cortex. *Proceedings of the National Academy of Sciences*, *86*(2), 711-715. https://doi.org/10.1073/pnas.86.2.711

Van Santen, J. P., & Sperling, G. (1984). Temporal covariance model of human motion perception. *JOSA A*, *1*(5), 451-473. https://doi.org/10.1364/JOSAA.1.000451

Wokke, M. E., Vandenbroucke, A. R., Scholte, H. S., & Lamme, V. A. (2013). Confuse your illusion: feedback to early visual cortex contributes to perceptual

completion. *Psychological Science*, *24*(1), 63-71. https://doi.org/10.1177/0956797612449175

Zharikova, A., Gepshtein, S., & van Leeuwen, C. (2017). Paradoxical perception of object identity in visual motion. *Vision Research*, *136*, 1-14. https://doi.org/10.1016/j.visres.2017.04.008

5. MQ study:

Perception of the ambiguous motion quartet: A stimulus-observer interaction approach

This study was published in Journal of Vision in 2021

Reference

Boeykens, C., Wagemans, J., & Moors, P. (2021). Perception of the ambiguous motion quartet: a stimulus-observer interaction approach. *Journal of Vision*, *21*(13). https://doi.org/10.1167/19.1.8

Abstract

Visual perception is the result of a highly complex process depending on both stimulus and observer characteristics and, importantly, their interactions. Generating robust theories and making precise predictions in light of this complexity can be challenging, and the interaction of stimulus- and observer-related effects is often neglected or understated. In the current study, we examined inter- and intra-individual differences and the effects of a wide range of three stimulus characteristics (i.e., spatial distance, temporal distance, and spatial location). Our results indicate that not all individuals show the same group average stimulus-driven effects on the perception of a motion quartet and that these effects are not always equal across the entire stimulus range. Moreover, we observed that there are clear individual differences in spontaneous perceptual dynamics and that these can be overridden by some but not all stimulus manipulations. We conclude that considering different stimulus manipulations, different observers, and their interactions can provide a more nuanced and informative view on the processes governing visual perception. This study examines the effect of spatial distance, spatiotemporal distance, spatial location, and individual differences on the perception of the ambiguous motion quartet.

5.1 Introduction

5.1.1 (Meta-)theoretical context

The normative approach for scientific experiments is to start from a strong theory in which a number of independent variables (e.g., stimulus size), predicted to influence a dependent variable (e.g., some measure of the percept), are specified in a conceptually coherent framework (Fig. 1 "hourglass"). Subsequently, a critical design is chosen in order to test a hypothesis and advance the predefined theory by generating data that allow to confirm or reject the hypothesis. This critical test typically focuses on a narrow range of manipulations of the independent variables. Accordingly, the hypothesis is tested as rigorously as possible and conclusions can be generalized as broadly as the theory allows (Meehl, 1990; Debrouwere, 2020). This approach is optimal when you are dealing with a strong theory, meaning that the process studied is already well known and/or precise predictions can be made. However, there is considerable debate about the robustness of established knowledge of several cognitive functions (Greenwald, Pratkanis, Leippe, & Baumgardner, 1986; Open Science Collaboration, 2015; Fisher, Medaglia, & Jeronimus, 2018; Davis-Stober, & Regenwetter, 2019; Debrouwere, 2020). Processes studied in behavioral science are influenced by multiple interacting factors, making it extremely difficult to generate robust theories and make precise predictions (Open Science Collaboration, 2015; Debrouwere, 2020).

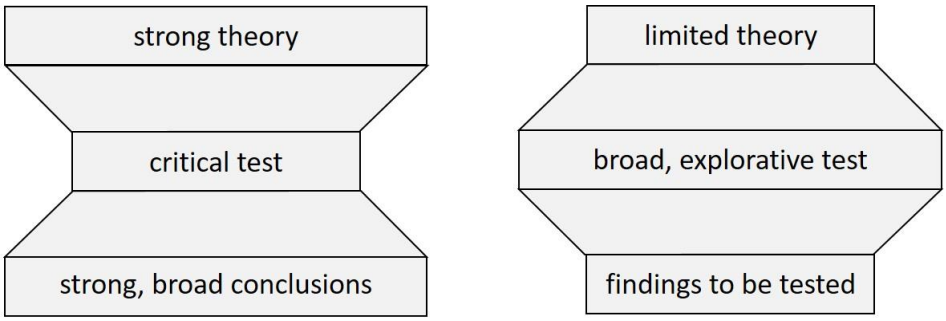


Figure 1. The "hourglass" vs. the "reverse-hourglass" approach. Left: The standard "hourglass" approach, in which we start from a strong, broad theory, test precisely, and have strong, generalizable conclusions. Right: The "reverse-hourglass" approach, in which we start from a limited theory, test broadly, and have exploratory findings, which should be followed-up by further research for replication, extension, and refinement

Although vision science is known for relatively strong established knowledge and theory, many of the broader processes studied in vision science can be similarly challenging for generating robust theories and making precise predictions. Visual perception is the result of a highly complex process depending on both stimulus and observer characteristics and importantly, their interactions. There are ample examples of studies in the literature reporting on the influence of stimulus or observer properties on perception. In addition to a wide variety of studies in which different stimuli are linked to different percepts in the average observer, research has shown that considerable individual variability exists when perceiving the same visual stimulus as well. A good illustration of individual variability is provided by perceptual multistability, where the same visual stimulus elicits different percepts (usually two or three) when viewed multiple times or for an extended period of time (e.g., a few minutes). Accordingly, there is intra-individual variability in the perception of the same stimulus. Inter-individual differences in these perceptual dynamics have been well-documented as well, and observers show reliable signatures in their "switching behavior" (e.g., the parameters of an observer's dominance duration distribution) (Brascamp, Becker, & Hambrick, 2018; Dieter, Sy, & Blake, 2017; Gallagher & Arnold, 2014).

Even well-known phenomena have been studied almost exclusively with an emphasis on either stimulus-related, or observer-related effects in isolation. Indeed, keeping either the observer or the stimulus constant (or treating it as such) while manipulating the other can be a great tool to study perception. However, studying their interaction is crucial to further advance our understanding of visual perception. We may find that some interaction effects due to variability in stimuli and observers are not negligible. Yet, in order to focus on either stimulus or observer characteristics, the influence of observer or stimulus characteristics, respectively, and their interplay is often neglected or understated in research where actually both are (or should be) independent variables (Mollon, Bosten, Peterzell, & Webster, 2017). One may assume that the variability of the dependent measure (e.g., variability of percepts in different observers or variability in a stimulus set) is normally distributed, leading to the notion that the average describes the individual (stimulus or observer), while this is potentially not the case (Wijnants, Cox, Hasselman, Bosman, 2013; Charest, & Kriegeskorte, 2015; Fisher, Medaglia, Jeronimus, 2018; Davis-Stober, & Regenwetter, 2019). For instance, stimulus-related effects of the average observer may not be informative for understanding any one

observer in particular (Fisher, Medaglia, Jeronimus, 2018). Alternatively, a sample may be taken from a very homogeneous group (e.g., all similar observers at one point in time or only one particular stimulus-type), causing the variability in the sample to be normally distributed, but compromising generalizability and making it potentially unrepresentative of the complexity of the processes that are studied. For instance, findings about the perception of well-known stimuli may not be replicated when the stimulus is only slightly different (Wijnants, Cox, Hasselman, Bosman, 2013). Similarly, as visual perception is a dynamic process, seemingly robust stimulus effects for one observer may change over time as well (Friston, & Kiebel, 2009; Koenderink, 2019; Molenaar, 2004; Hamaker, 2012).

Perhaps some of our theories are not yet capturing the essence of the broader studied process and perhaps we prevent further research progress when we keep measuring our dependent variables at particular isolated points in a non-uniform space that is influenced by multiple independent variables in ways that we are currently unable to predict. Think of trying to study a complex mechanism M in a group of individuals. We know that M has fluctuations due to internal factors and is influenced by external factors. Imagine that we measure the influence of a specific range of one external factor (e.g., 10-20 units) on M at a given point in time. At another point in time, we measure fluctuations in M due to internal factors. From these isolated studies, how should we predict M in the future? Some (curvi-)linear relationship between the external factor and M may exist within the scope of the studied units of the external factor (i.e., 10-20 units), but the effect of the external factor and its interactive effects may be very different outside this specific studied range. Even if we know what to expect on average within the studied scope of the external factor, internal factors or other external factors may still override the effect of the external factor at another point in time. Similarly, we may have an idea of the fluctuations of M due to internal factors, but surely this does not allow us to make precise predictions of M when the effect of an external factor is very strong. Moreover, internal factors and their interactive effects with external factors may depend upon which different individuals we study and may differ considerably from one individual to another. Therefore, even though we considered multiple influencing factors in isolation, it might still be close to impossible to replicate findings or predict M precisely in the future. Alternatively, we could attempt to study the influence of a wider range of external factors, whilst simultaneously considering internal factors in multiple individuals. Accordingly, we can

get a broader view of the complex mechanisms determining M. Importantly, we do not encourage drawing strong conclusions at the end of such an exploratory approach. These findings, since they are not the result of a critical test for a specific hypothesis, should be put to the test further on.

Likewise, when we study the effects of different stimulus and observer properties in isolation, and we only consider a narrow range of the stimulus and observer space, we may be missing valuable information. This may lead to theories lacking in explanatory or predictive value for particular observers and a wider range of stimuli. When our theories and knowledge of visual phenomena cannot account for this variability, difficulties in replication arise and the robustness of theories is compromised. In preparation of the current study, we considered multiple research questions concerning the perception of a multistable stimulus dependent on stimulus and observer properties, with corresponding potential operationalizations and designs according to the "hourglass" approach. However, we were confronted with its limitations in this context as mentioned above and concluded that it would be difficult to provide a critical test associated with a strong theory. Therefore, we found it worthwhile to consider a more exploratory approach, which may provide us with information on stimulus-observer effects that were not studied in combination before. This information may, in turn, shape our theories to better account for even main effects of simple stimuli. It implies that we will attempt to study a complex system as holistically as we can (instead of studying the effects of its components in isolation) in order to understand the complex system as a whole, as well as the effects of its components. Instead of adopting the standard "hourglass" approach (in which we start from a strong, broad theory, test precisely, and have strong, generalizable conclusions), we will start from a limited theory, test broadly, and have exploratory findings, which should be followed-up by further research for replication, extension, and refinement (Fig. 1, "reverse-hourglass").

5.1.2 Empirical context

For the current study, we considered observer- and stimulus-related effects on the perception of an ambiguous motion quartet stimulus (Fig. 2). A motion quartet (MQ) is an ambiguous apparent motion stimulus which consists of two frames of two dots (Gengerelli, 1948). In the first frame, one dot is displayed in the upper left corner (or the upper right) relative to center and the other dot is displayed in the lower right (or the lower left) corner relative to center. In the second display, one dot is displayed in the upper right corner (or the upper left) and the other dot is displayed in the lower left (or the lower right) corner. When the two frames are presented successively, or with a sufficiently short time delay in between, observers can see apparent motion between the dots in the separate frames. The influence of several stimulus characteristics on the perception of the MQ has been studied since long and we will briefly review some important findings. Although individual differences have almost always been reported, the focus remained in obtaining a general description of the perception of the MQ, assuming an absence of any noteworthy influence of observer characteristics. For instance, the temporal distance between the two frames may elicit a difference in perception of the motion quartet (Ramachandran & Anstis, 1983). Previous studies have also documented a general vertical bias, indicating vertical apparent motion even when distances between dots across frames are equally large in both directions (Gengerelli, 1948; Chaudhuri & Glaser, 1991, Genç et al., 2011). Further studies have concluded that the vertical bias in the perception of the MQ may be due to the efficiency of motion processing within one hemifield and hemisphere compared to across hemifields and hemispheres (Chaudhuri & Glaser, 1991; Genç et al., 2011).

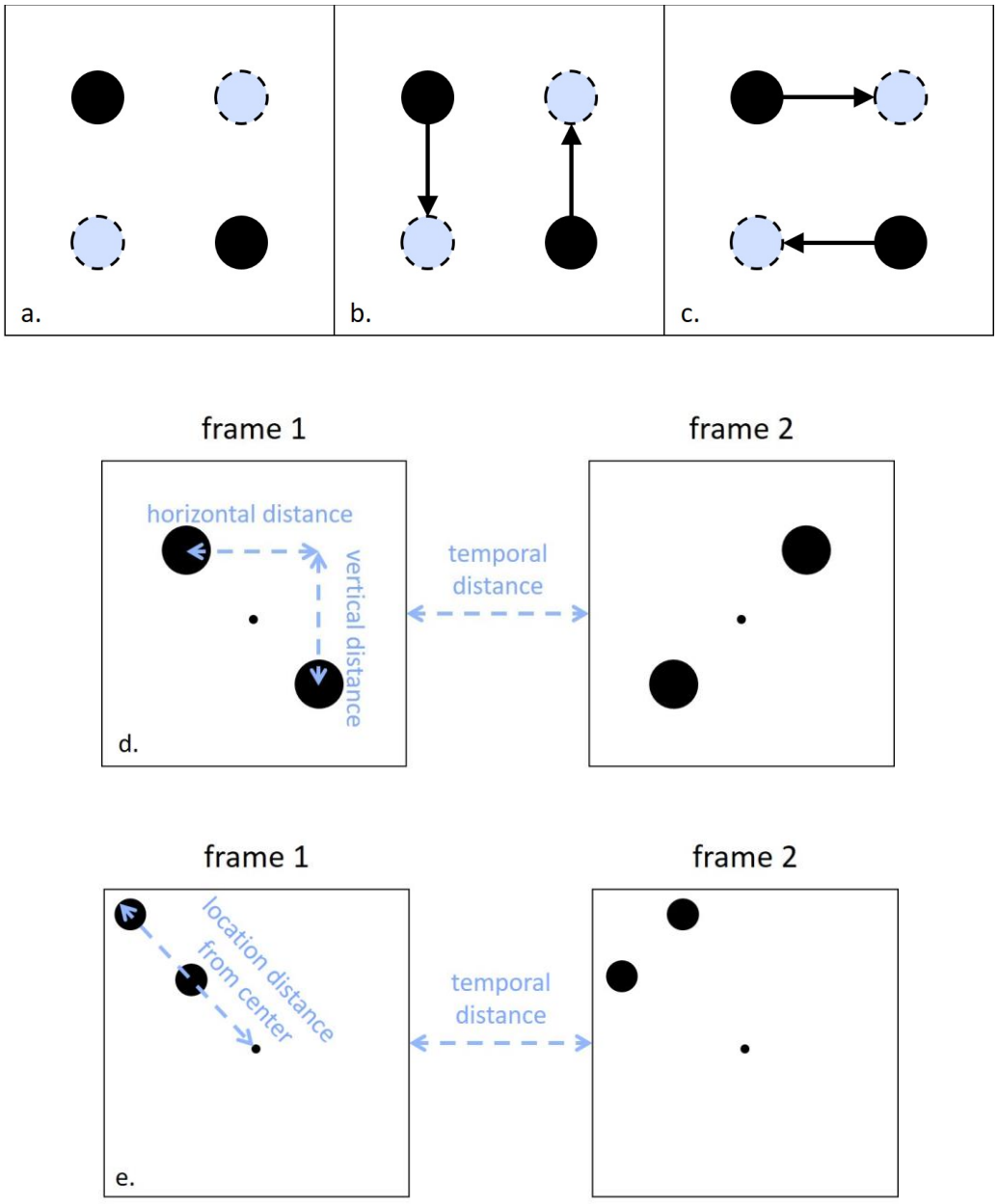


Figure 2. Ambiguous motion quartet stimuli. a: The ambiguous motion quartet (MQ) is depicted with the first frame in black and the subsequent frame in dashed lines. An animated demonstration can be found here: https://osf.io/2yktz/?view_only=ef697c56758547858ca4d02c63b62013. In an alternating presentation of these two frames, apparent motion can be perceived. **b:** Vertical motion percept of the ambiguous MQ. **c:** Horizontal motion percept of the ambiguous MQ. **d:** Ambiguous MQ in Spatial Scale and Spatiotemporal Scale experiment. **e:** Ambiguous MQ in Spatial Location experiment.

We were inspired by the studies of Mark Wexler and colleagues, in which perception of variations of different ambiguous motion stimuli was studied in different observers across multiple time scales (Wexler, Duyck, & Mamassian, 2015; Wexler, 2018). In the latter study, perception of the motion quartet was studied over time for stimulus presentations in different orientations. They found that observers exhibit particular patterns of percepts for these

ambiguous stimuli. Interestingly, these patterns are robustly idiosyncratic and can evolve over time. A fitted model showed that the results were in accordance with the idea that perception is governed by continuously evolving internal brain states, which can be described as a random walk, due to external factors and internal neural dynamics. They concluded that "Overall, these results exhibit the multidimensional complexity of internal states underlying the perception of even simple stimuli" (Wexler, 2018, p. 1).

The study of Wexler (2018) shed light onto the complex interplay of observer and stimulus characteristics in multistable perception, but its model provided only a description of the interaction of internal dynamics and the variation of one main stimulus characteristic. It is plausible that perceptual dynamics may be differentially affected by different external factors. If we strive for a robust, precise and generalizable description and prediction of multistable perception, further exploration is necessary. In order to further determine how internal neural dynamics and external factors shape the perception of the MQ, we sought to explore the stimulus-observer space more broadly. This exploratory approach may provide us with more information on effects that were not studied in combination before, which may, in turn, shape our theories to better account for even main effects of simple stimuli. In the current study, starting from the findings in the classic literature about the perception of the MQ (Gengerelli, 1948; Ramachandran & Anstis, 1983; Chaudhuri & Glaser, 1991), we examined inter- and intra-individual differences and the effects of a wide range of three stimulus characteristics (i.e., spatial distance, temporal distance, and spatial location), as a case study in adopting the "reverse-hourglass" approach.

In addition, we introduce recurrence quantification analysis as a tool to quantify the characteristics of the temporal dynamics different observers exhibit in response to different stimuli. A recurrence analysis looks into the number and duration of recurrences of a specific state (in our case, percept) in a dynamical system over time (Wallot, 2017). It is a nonlinear method of data analysis by means of which we aim to explore sequential effects of percepts in all observers.

5.2 Spatial Scale Experiment

In the first experiment, we examined individual differences in the perception of the MQ across different horizontal and vertical spatial distances (Fig. 2d). The previously reported vertical bias (Gengerelli, 1948; Chaudhuri & Glaser, 1991, Genç et al., 2011) has been demonstrated mainly by the prominence of vertical apparent motion in the average observer for a MQ with an equally large vertical and horizontal distance. Employing a similar MQ, but with a variable orientation, Wexler (2018) found more differences across observers, including a vertical or horizontal bias, as well as a clockwise or counterclockwise bias. Similarly, we examined the perception of a MQ with a variable spatial scale. By exploring perception of the MQ across different observers and a wider variety of spatial distances, we aim to more extensively characterize conditions under which the previously reported vertical bias is present and thereby understand the mechanisms underlying the perception of the MQ more fully.

5.2.1 Methods

Participants

29 university students (1 male, mean age = 18.9, *SD* = 2.23) participated in this study in return for course credit. All participants had normal or corrected-to-normal vision, provided informed consent, and were naïve with respect to the goals of the study. The study was approved by the Social and Societal Ethics Committee of KU Leuven (SMEC, approval code G-2017 05 834).

Apparatus

All stimuli were displayed on a CRT monitor (Sony GDM F520) with a refresh rate of 60 Hz, a diagonal length of 50 cm and a resolution of 1024 x 768 pixels. Participants were positioned in a dark room at a viewing distance of 57 cm from the monitor by means of a chin rest. Stimulus presentation and response registration was controlled by software programmed in Python using the PsychoPy library (Peirce, 2007).

Stimuli

All stimuli consisted of two white dots with a diameter of 0.78 deg (luminance 76.8 cd/m^2) and a central white fixation point with a diameter of 0.08 deg displayed on a grey background (luminance 17.9 cd/m^2). One stimulus presentation consisted of two successive stimulus displays with a duration of 320 ms each (Fig. 2d). There was no blank interval between these two displays and therefore the stimulus-onset-asynchrony (i.e., SOA) within one stimulus presentation is identical to stimulus display duration. In the first display, the dots were positioned in the top right and the bottom left relative to fixation. In the second display, the dots were positioned in the bottom right and top left relative to fixation (Fig. 2). Horizontal and vertical distances between the centers of the two dots varied in discrete steps of 0.3 deg between 0.6 deg to 5.4 deg.

Procedure and design

The horizontal and vertical distances between dots displayed across the two displays varied on each trial and were selected from 289 (17 x 17) unique combinations of these distances in a randomized order. This resulted in 289 trials separated in 3 short runs. Observers were instructed to report the perceived direction of apparent motion after stimulus presentation on each trial. When they saw apparent motion between the top right and top left dot, and between the bottom left and bottom right dot, they indicated that they saw "horizontal"[1] apparent motion. When they saw apparent motion between the top right and bottom right dot, and between the bottom left and top left dot, they indicated that they saw "vertical" apparent motion. During the inter-trial-interval, the fixation point was displayed for 1000 ms.

Data-analysis

All data analysis was conducted in R 3.6.0. Essentially, the data consist of a single measurement of an observer's percept for each unique stimulus (determined by the horizontal and vertical distance between the dots). Proportions of vertical vs. horizontal

[1] In the literature, perception of this apparent motion stimulus is often described as horizontal or vertical motion. However, two different kinds of percepts are possible under such a description. That is, some observers report to see two dots moving independently horizontally or vertically, while others perceive the dots as belonging to a single object rotating clockwise or counterclockwise.

perceptual reports were derived from these data, providing a coarse representation of perceptual bias. Our goal was to quantify an observer's tendency to perceive the stimuli either more in the horizontal or vertical direction, in relation to horizontal and vertical distance. In order to do this more precisely, we fitted a generalized linear model with a logistic link function to the data of the perceptual responses of each observer. Intuitively, this boils down to estimating a multidimensional psychometric function. We included linear and quadratic effects of both vertical and horizontal distance to assess how the perceptual responses changed in function of vertical and horizontal distance. Based on the estimated coefficients of this model, we derived a curve of subjective equality across all vertical and horizontal distances (i.e., the estimated transition curve between horizontal and vertical percepts, made up from the estimated point of subjective equality, or PSE, along horizontal and vertical distances). Furthermore, the model coefficients were used to quantify the size of the transition zone between horizontal and vertical percepts (i.e., the inverse steepness of the psychometric function) and change/stability of the aspect ratio between horizontal and vertical distance at the PSE along the curve of subjective equality (i.e., the significance of quadratic effects).

5.2.2 Results and Discussion

Visual inspection and proportions of perceptual reports over different horizontal and vertical distances show that there are substantial individual differences in perceptual biases of the ambiguous MQ (Fig. 3 and 4, Supplementary Table 2 and Fig. A for all participants). Perceptual bias here is defined as a tendency to perceive apparent motion between dots in one direction when the physical distance between these dots in this direction is not the shortest option available (i.e., not grouping by proximity). A large subset of the individuals shows a vertical bias, which was also reported in previous studies (Gengerelli, 1948; Chaudhuri & Glaser, 1991). However, unexpectedly, based on the existing literature, an almost equally large subset of individuals seems to have no clear bias at all, approximating grouping by proximity. Moreover, a smaller number of individuals even shows a horizontal bias, which is in accordance with the findings for a different MQ in the study of Wexler (2018), but which had not been pointed out in other studies before. These results indicate that individual differences in perception of the MQ might be more important than previously assumed. In principle, this variability in

perceptual bias extends far enough to yield subgroups differing categorically in perceptual bias (i.e., vertical, none, horizontal). However, individual variability in perception of the MQ might be best described on a continuous scale. Either way, further exploration of this variability will provide new insights that would be hidden when only considering a ubiquitous vertical bias. Moreover, we did not vary the order of both frames of the MQ stimulus, although there might be even more individual variability when this stimulus factor is also taken into account (Wexler, 2018). There are various possible explanations for the observed individual variability in perceptual bias, ranging from individual differences in "low-level" processing of visual input to differences in voluntary control of perception and perceptual reports. However, the influence of voluntary control is most probably limited, as Ramachandran and Anstis (1983) found that participants could not change their perception of a MQ by will for stimuli with an SOA shorter than 465 ms, and Wexler (2018) found that the effect of voluntary control on perception of a MQ was significant only for SOA longer than or equal to 400 ms.

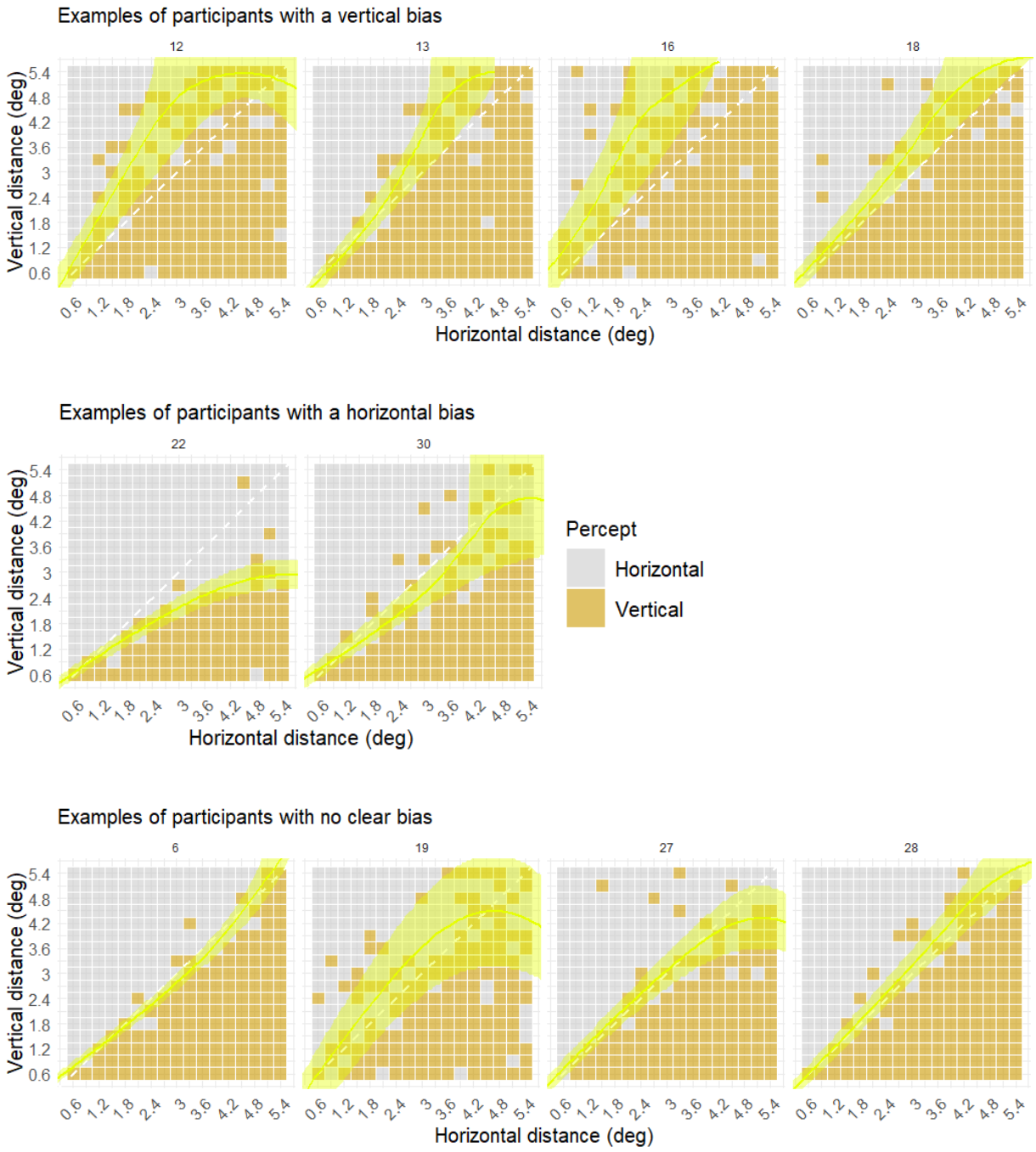


Figure 3. Percepts of ambiguous MQ across spatial scale (Spatial Scale experiment). Each plot represents all perceptual reports for one participant across all different vertical and horizontal distances. Each colored square in a plot represents one perceptual report, which may be 'Horizontal' or 'Vertical'. The yellow overlay represents the "transition zone" between horizontal and vertical percepts across spatial distances, derived from a fitted generalized linear model. From top to bottom: Examples of participants with a vertical bias, a horizontal bias, and no clear bias. The small number on top of each plot indicates the participant number. Plots for all participants can be found in Supplementary Figure A.
https://osf.io/tahqw/?view_only=f0b5de9a282140fa973df0763529b720

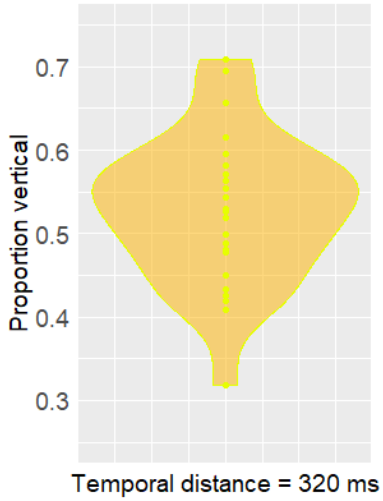


Figure 4. Density plot of proportions of vertical percepts per participant (Spatial Scale experiment). Each dot inside the density plot represents one participant. There are more participants with a vertical bias, as can be derived from the relatively heavy top part of the density plot. However, there appear to be substantial individual differences, ranging from a vertical to a horizontal bias.

As described in the section 'Data-analysis', we fitted a generalized linear model and estimated transition curves from horizontal to vertical percepts across different vertical and horizontal distances. Visual inspection of these transition curves, as well as the individual variability in (non-)significance of the quadratic effects in the model, imply that there are individual differences in the stability of perceptual bias, or the aspect ratio (AR) at their PSE, over spatial scale (Fig. 5, Supplementary Table 1). Whereas some individuals remain more or less constant in the extent of their perceptual bias (i.e., rather linear transition curves and nonsignificant quadratic effects), others seem to increase or decrease in perceptual bias over horizontal and/or vertical distance (i.e., rather nonlinear transition curves and significant quadratic effects). This interaction with spatial scale has not been studied or emphasized before, yet our findings indicate that spatial scale might be an important factor to take into account when discussing the strength of the perceptual bias.

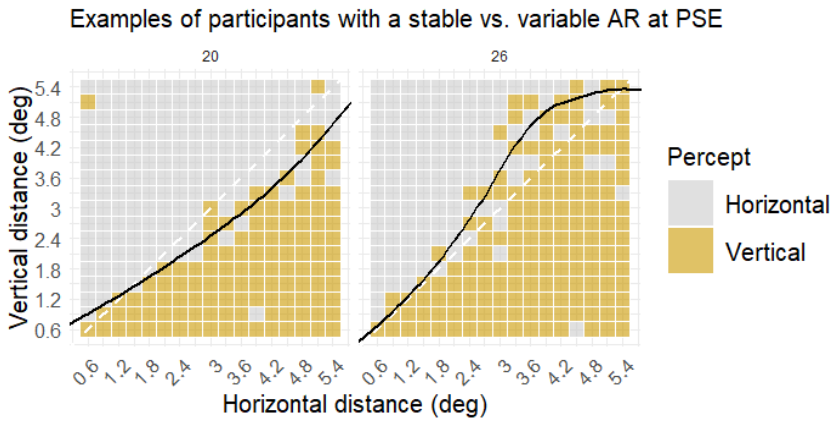


Figure 5. Percepts of two exemplary participants with a stable vs. variable AR at PSE (Spatial Scale experiment). The black line in the plot represents the curve of subjective equality (i.e., PSE across different spatial distances), which was derived from the fitted generalized linear model. Left: A black line that remains relatively linear across spatial distances indicates that the AR at PSE is stable across spatial distances. Right: A black line that is relatively curvilinear across spatial distances indicates that the AR at PSE is variable across spatial distances.

The coefficients for vertical and horizontal distance in the fitted model also provided a measure for the size of the transition zone between both percepts across spatial distances. More specifically, one may interpret the fitted transition from horizontal to vertical percepts across spatial distances as a 3-dimensional psychometric function, where the change from one perceptual state to another occurs not only across one stimulus dimension (such as vertical distance), but across two stimulus dimensions (i.e., horizontal and vertical distance). A higher term estimate and steeper slope in both dimensions of this fitted psychometric function corresponds to a smaller transition zone, or less variability in percepts around the transition curve from horizontal to vertical percepts.

Here, we can see considerable individual differences as well (Fig. 7). Figure 6 provides examples of two observers with a small (i.e., observer nr. 6) vs. large (i.e., observer nr. 10) transition zone, corresponding to two observers at either end of the scatter plot in Figure 7. The size of the transition zone may be interpreted in several ways, which are not mutually exclusive. One possibility is that, for a larger transition zone, there is a larger effect of internal noise, which can manifest itself in perception in several ways across the visual processing hierarchy (i.e., the stability or instability of the dynamic systems starting from processing visual input and leading to the organization of the percept, the decision process and the generation of a response). In any case, the size of the transition zone has mostly been

neglected in previous research, although individual differences in the size of the transition zone can be a highly valuable addition to what is known about the effect of perceptual bias on the perception of the MQ. It may be worthwhile to determine, for instance, in which manner the perception of the MQ by an individual with a very small transition zone in combination with a strong vertical bias is driven "low-level" and "higher-level" factors. In some studies, a strong influence of perceptual bias is interpreted as demonstrating a stronger top-down influence on perception, compared to a more stimulus-driven influence of grouping by proximity on perception (Nikolaev, Gepshtein, & van Leeuwen, 2016). Yet, a strong perceptual bias might also be the result of a "low-level" distortion of perception (Carter & Cavanagh, 2007). Likewise, as mentioned before, the size of the transition zone may be the result of various factors along the visual-cortical hierarchy.

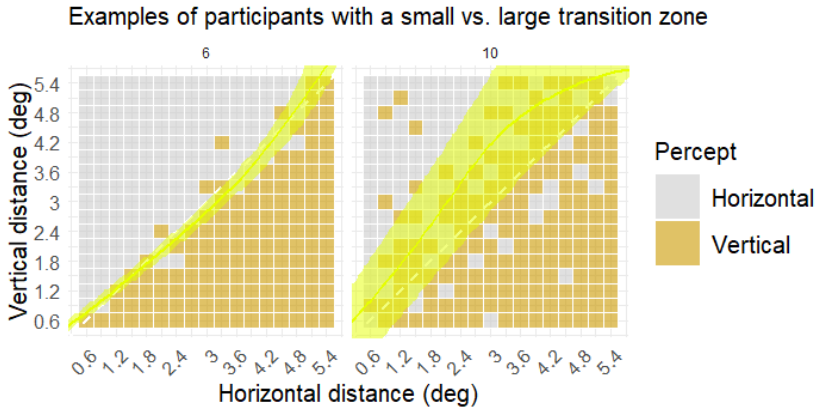


Figure 6. Percepts of two exemplary participants with a small (i.e., observer nr. 6) vs. large (i.e., observer nr. 10) transition zone (Spatial Scale experiment). The yellow overlay in the plots represents the size of the transition zone (i.e., the range of spatial distances for which the participant has highly variable perceptual reports), which was derived from the fitted generalized linear model. Left: A small transition zone indicates that the participant has an abrupt switch of perceptual reports across spatial distances, with little variability in perceptual reports on either side of the transition curve. Right: A large transition zone indicates that the participant has a gradual switch of perceptual reports across spatial distances, with high variability in perceptual reports on either side of the transition curve.

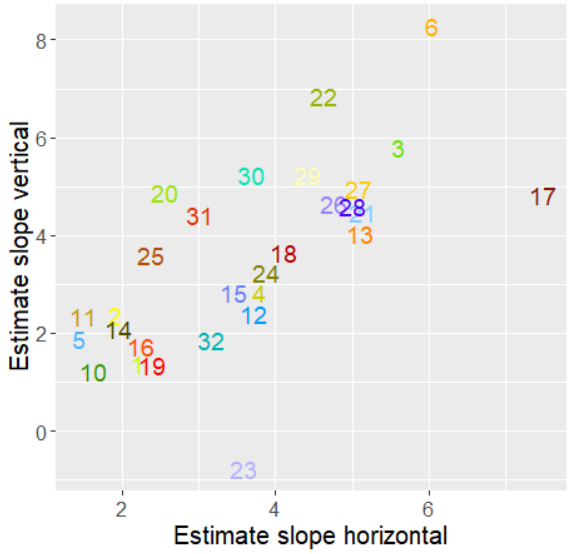


Figure 7. Estimates for the horizontal distance and vertical distance term in the generalized linear model for each participant (Spatial Scale experiment). Term estimates for each participant are indicated by the participant numbers in different colors. The higher the term estimate, the steeper the slope of the psychometric function, the smaller the transition zone. As can be seen in this plot, participant 6 and 10 (Fig. 6) are on either ends of the range.

5.3 Spatiotemporal Scale Experiment

In the second experiment, we examined individual differences in the perception of the MQ across different horizontal and vertical spatial distances, as well as different temporal distances (Fig. 2d). Previous research has reported a differential perception of the MQ across different temporal distances, or SOAs (Ramachandran & Anstis, 1983). More specifically, Ramachandran and Anstis (1983) found that participants experienced hysteresis, with percepts of the MQ being less variable and less ambiguous, for stimuli with an SOA shorter than 400 ms. Again, it may be worthwhile to specify under which circumstances the impact of SOA is present, and what it specifically entails, thereby contributing to a further understanding of the perception of the MQ.

5.3.1 Methods

Participants

22 university students (4 male, mean age = 18.95, *SD* = 1.84) participated in this study, which was otherwise identical to the Spatial Scale experiment with respect to participant-related

information. Two participants were excluded because their random responses indicated they did not report their percept.

Apparatus

Apparatus was identical to the Spatial Scale experiment.

Stimuli

The horizontal and vertical distances between the two dots in the stimuli varied in discrete steps of 0.4 deg between 0.6 deg to 5.4 deg. Stimuli were otherwise identical to the Spatial Scale experiment, except for an additional variability in temporal characteristics, with the duration of stimulus displays (i.e., SOA) ranging between 160 ms and 640 ms in steps of 160 ms. Both displays within one stimulus presentation had identical durations (Fig. 2d).

Procedure and design

The temporal, horizontal and vertical distances between dots displayed across the two displays varied on each trial and were selected from 676 unique combinations of these distances in a randomized order. This resulted in 676 trials, separated in 7 short runs. The procedure was otherwise identical to the Spatial Scale experiment.

Data-analysis

All data analysis was identical to the Spatial Scale experiment.

5.3.2 Results and discussion

Visual inspection of perceptual reports in all observers does not reveal clear differences in perceptual biases of the ambiguous MQ as a function of different temporal distances (Fig. 8, Supplementary Fig. B for all participants). However, a one-way repeated measures ANOVA on the proportions of perceptual reports (Supplementary Table 4) shows that the effect of temporal distance on perceptual bias is statistically significant ($F(3,57) = 3.165$, $p = .031$). More specifically, vertical perceptual bias is most prominent for shorter temporal distances (160 and 320 ms) and, on average, less clear for longer temporal distances (480 and 640 ms). Moreover, there are more individual differences in perceptual bias for longer temporal distances, as can

be seen in Figure 8. Nevertheless, for comparison, Figure 4 displays individual perceptual biases in the Spatial Scale experiment and shows that also for an SOA as short as 320 ms, there can still be a significant amount of individual variability in perceptual biases. Our results show that temporal distance can be of crucial importance in the manifestation of perceptual bias in some observers and may have impacted results of various previous studies using an SOA of 100 to 250 ms (Gengerelli, 1948; Chaudhuri & Glaser, 1991, Genç et al., 2011).

Given the fact that there is less individual variability for shorter temporal distances, it is possible that perception of the MQ relying more on "low-level" processes is more homogeneous across individuals than when stimulus presentation time is sufficiently long to allow more elaborate "mid-" and "higher-level" visual processing. However, shorter presentation times do not necessarily exclude "higher-level" processes. As mentioned in the Spatial Scale experiment, there are various possible explanations for the observed individual variability in perceptual bias. Factors that may be at the basis of individual differences in perception of the MQ can be, for instance, early processing of visual input, perceptual grouping, motion perception, perceptual decision making, and further research is needed to further disentangle these factors. Moreover, we cannot exclude the influence of voluntary control, or the possibility that participants willingly altered their percepts when the SOA was longer. However, the main source of variability is most probably not merely voluntary control, as we also found considerable variability in perceptual bias for stimuli with an SOA that is below the previously reported thresholds of 465 ms and 200-400 ms for voluntary control (Ramachandran & Anstis, 1983; Wexler, 2018).

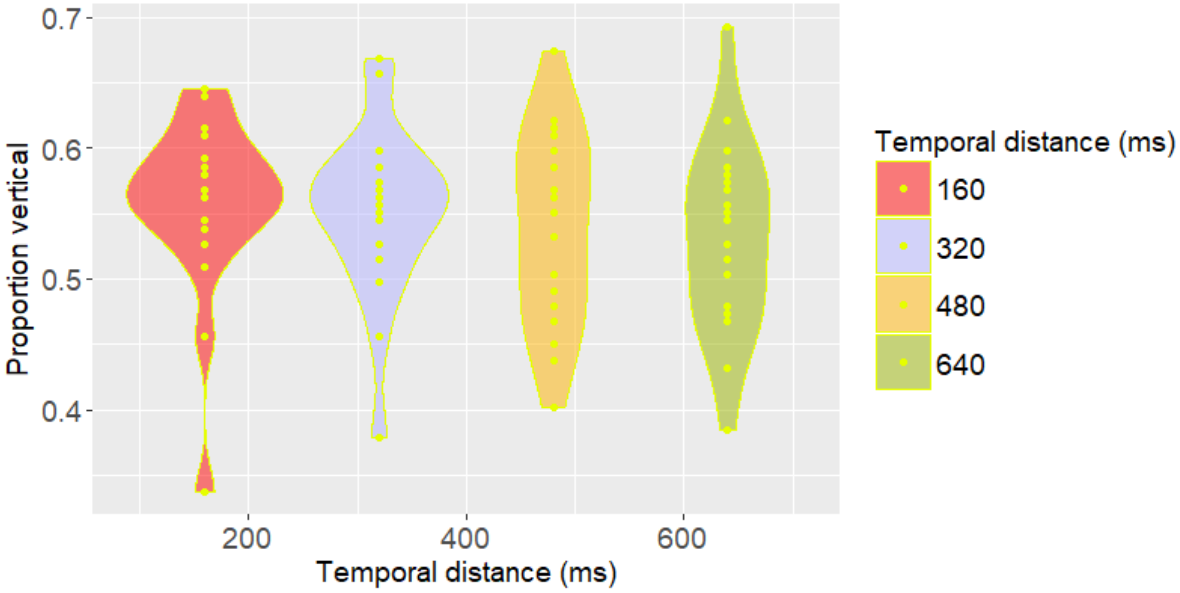


Figure 8. Density plot of proportions of vertical percepts per participant per temporal distance (Spatiotemporal Scale experiment). Each dot inside the density plot represents one participant. There are relatively less inter-individual differences and relatively more participants with a vertical bias for smaller temporal distances, as can be derived from the relatively heavy top part of the first two density plots. However, there appear to be substantial individual differences for all temporal distances, ranging from a vertical to a horizontal bias. Percept plots for all participants can be found in Supplementary Figure B.
https://osf.io/tahqw/?view_only=f0b5de9a282140fa973df0763529b720

Individual differences in variability of AR at PSE are more prominent for longer temporal distances. By means of the (non-)significance of quadratic effects in the fitted model we may derive that a subset of the tested individuals differs in the stability of perceptual bias over spatiotemporal distances (Fig. 9, Supplementary Table 3). This rather complicated effect can be described as follows. There is little or no significant change in AR at PSE across spatial distances for short temporal distances (160 ms) in any of the observers. However, for a subset of observers, change in AR at their PSE across spatial distances is prominent for longer temporal distances (320, 480 and 640 ms). For the other observers, whether the change in AR at PSE across spatial distances is small or large remains relatively stable across temporal distances. These results suggest that, aside from an increase in inter-individual variability for longer temporal distances, in some individuals, we see an increase in intra-individual variability for longer temporal distances as well.

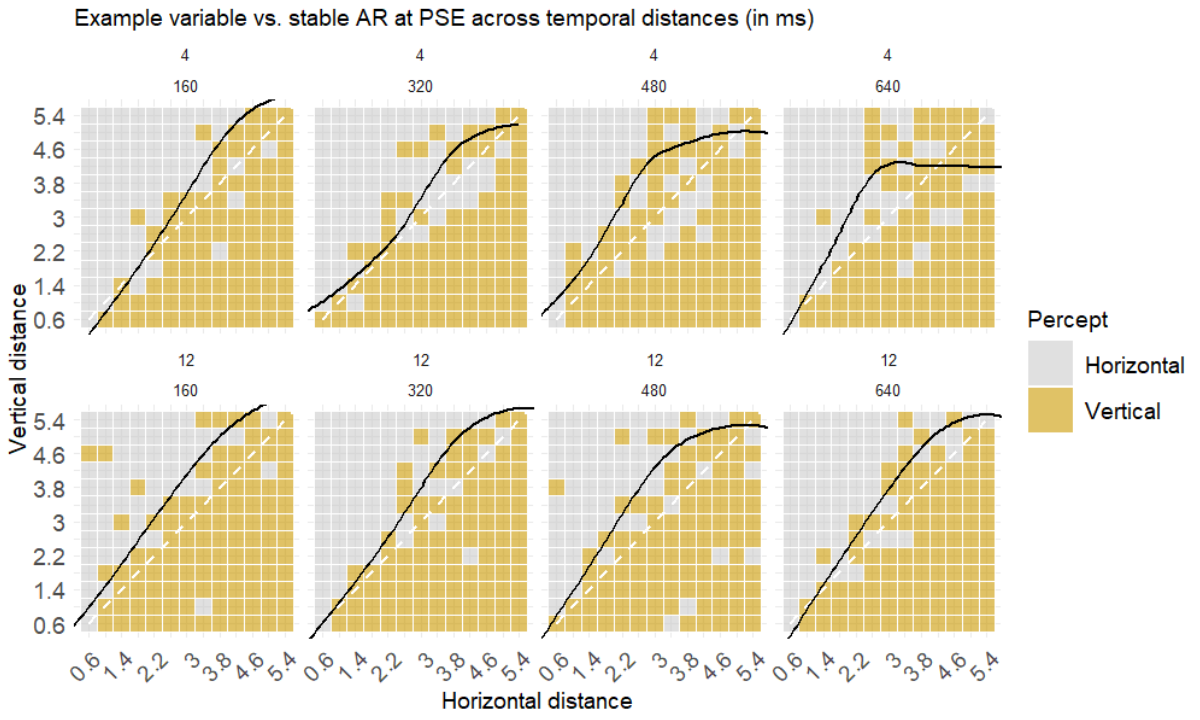


Figure 9. Percepts of two exemplary participants with no change vs. change in the stability of AR at PSE across temporal distances (Spatiotemporal Scale experiment). The black line in the plot represents the curve of subjective equality (i.e., PSE across different spatial distances) that was derived from the fitted generalized linear model. Top: The stability in AR at PSE across spatial distances changes over temporal distances. Bottom: The stability in AR at PSE across spatial distances remains relatively stable over temporal distances.

Lastly, individual differences in the size of the transition zone are more prominent for shorter temporal distances. A subset of the observers seems to have a small transition zone for a short temporal distance and larger transition zone for longer temporal distances, whereas other observers have a larger transition zone across all four temporal distances (Fig. 10). We fitted the same generalized linear model and estimated transition curves from horizontal to vertical percepts across different vertical and horizontal spatial distances for all different temporal differences. Further inspection of individual differences in the term estimates revealed that this subset of observers indeed have a higher term estimate for shorter temporal distances and a lower term estimate for longer temporal distances (Fig. 11), corresponding to a smaller and larger transition zone, respectively. The study of Ramachandran and Anstis (1983) reported that observers experience less variability and ambiguity in their perception of a MQ with AR = 1 for shorter temporal distances. One way to compare this variability in percepts of the MQ in their study to our results, is to look at variability of percepts around the diagonal, as the AR is closest to 1 here. Since the transition zone largely overlaps with the diagonal in

our results, the size of the transition zone can roughly reflect the ambiguity or variability of percepts around the diagonal. This comparison indicates that there is indeed an overall decrease in variability for shorter temporal distances, in accordance with the findings of Ramachandran and Anstis (1983). Interestingly, our results demonstrate that this overall decrease in variability, or overall decrease in the size of the transition zone, is driven only by a subset of the individuals with a smaller transition zone for shorter temporal distances (Figure 11).

Overall, these results suggest that there are not only more inter-individual differences in perceptual bias for longer temporal distances, there are more intra-individual differences in percepts along the transition curve for longer temporal distances as well. We may again speculate about which processes underlie an increase of the size of the transition zone for longer temporal distances, but further research should inquire into the specific roles of various stages of visual processing.

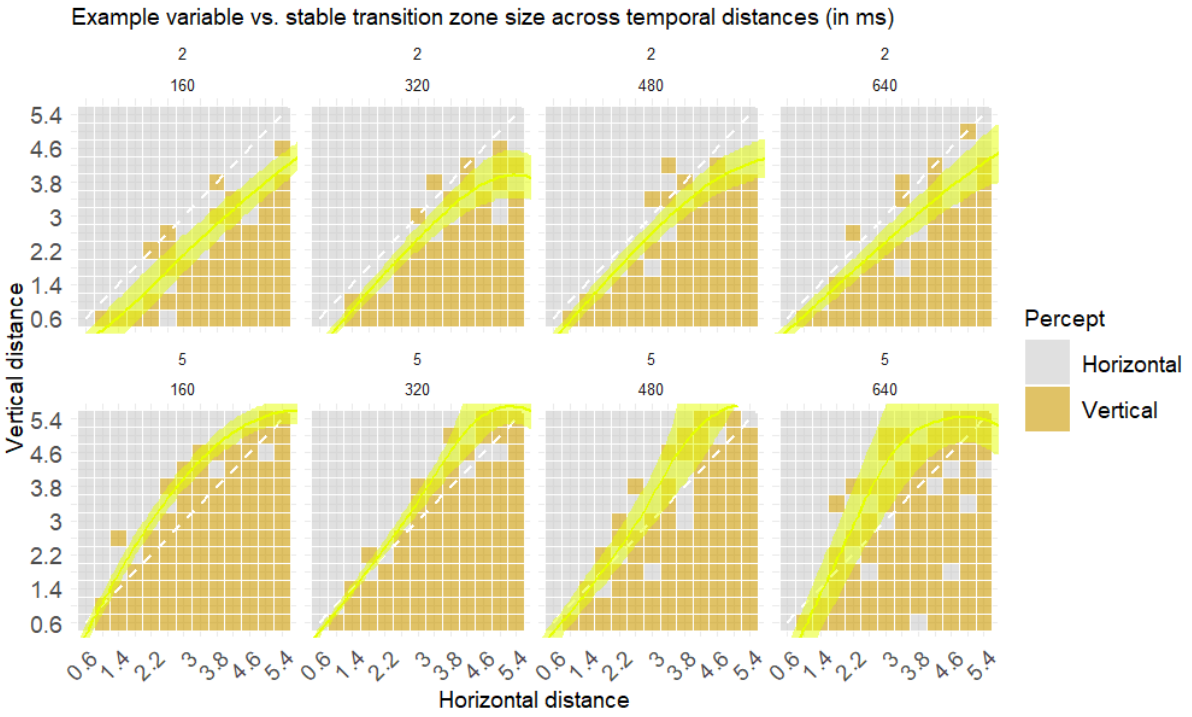


Figure 10. Percepts of two exemplary participants with no change vs. change in the size of the transition zone across temporal distances (Spatiotemporal Scale experiment). The yellow overlay in the plots represents the size of the transition zone (i.e., the range of spatial distances for which the participant has highly variable perceptual reports), which was derived from the fitted generalized linear model. Top: For this participant, the size of the transition zone remains relatively stable across temporal distances. Bottom: For this participant, the size of the transition zone becomes larger across temporal distances.

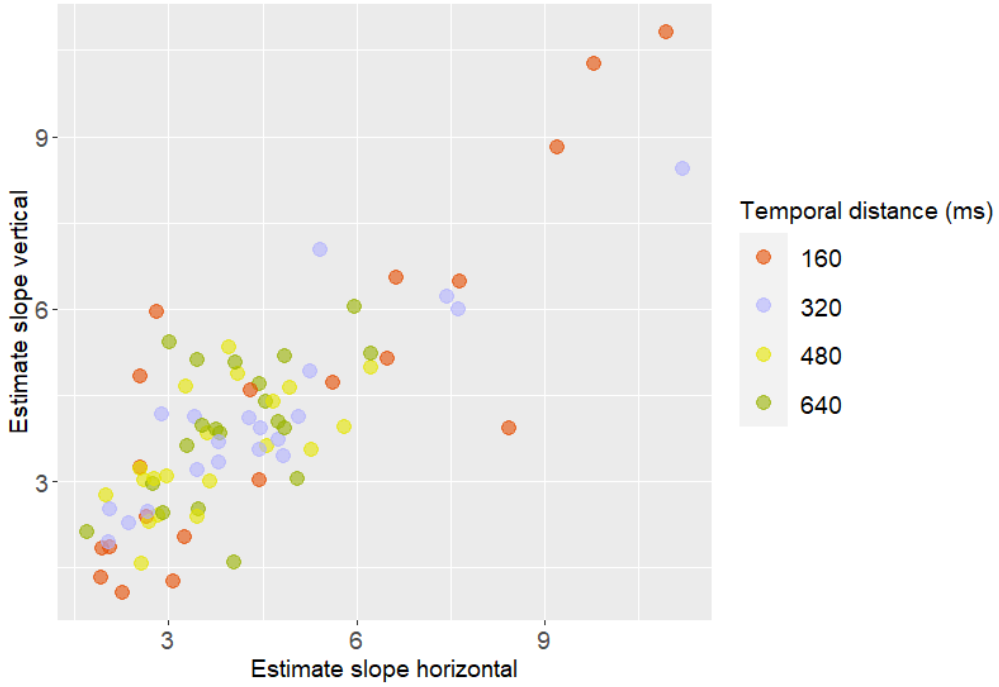


Figure 11. Estimates for the horizontal distance and vertical distance term in the generalized linear model for all participants per temporal distance (Spatiotemporal Scale experiment). Term estimates for each participant are indicated by dots in different colors per temporal distance. The higher the term estimate, the steeper the slope of the psychometric function, the smaller the transition zone. As can be seen in this plot, the smaller the temporal distance, the higher the term estimates.

5.4 Spatial Location Experiment

In the third and final experiment, we examined individual differences in the perception of the MQ across different stimulus locations (Fig. 2e). Analogous to the study of Wexler (2018) and our first two experiments, the manipulation of spatial location induces some variability in the stimulus that was studied before, which allows us to characterize the conditions under which previously reported perception of the MQ is established. In combination with studying differences across observers, this will enable us to better understand the perception of the MQ. Moreover, previous research has stated that a vertical bias in perception of the MQ may be due to the efficiency of motion processing within one visual hemifield (Chaudhuri & Glaser, 1991; Genç et al., 2011). By means of varying the location of the MQ across the two hemifields, we can inquire into the robustness of this finding across stimuli and observers.

5.4.1 Methods

Participants

36 university students (8 male, mean age = 18.8, *SD* = 1.27) participated in this study, which was otherwise identical to the previous experiments with respect to participant-related information.

Apparatus

All apparatus were identical to the previous experiments.

Stimuli

The stimuli were slightly different from those used in the previous two experiments. They consisted of two successive displays of two white dots with a diameter of 0.5 deg (luminance 76.8 cd/m^2) and a central fixation point with a diameter of 0.2 deg displayed on a grey background (luminance 17.9 cd/m^2). In the first display, the dots were positioned in the top right and the bottom left relative to the (not displayed) center between them. In the second display, the dots were positioned in the bottom right and top left relative to the (not displayed) center between them (Fig. 2e). The diagonal distance between the centers of the dots was kept constant at 4.24 deg, resulting in a MQ with an AR equal to 1. Across trials the position of the two dots varied on an invisible 11x11 grid, spanning -6.75 to 6.75 deg in horizontal and vertical directions away from the fixation point, which was positioned in the center of the display. The duration of each display (i.e., SOA in our experiment) was 500 ms.

Procedure and design

The location of the MQ varied on each trial and was selected in a random order. This resulted in 121 trials executed in one run. Observers were instructed to report the perceived direction of apparent motion after stimulus presentation on each trial, identical to the first two experiments. They were instructed to keep their eyes on the fixation point, in order to perceive the MQ without any eye movements towards its location, which was further facilitated by employing randomly chosen and equally probable stimulus locations around the

center of the display. An eye tracker was not employed in this experiment. During the inter-trial-interval, the fixation point was displayed for 500 ms.

Data-analysis

The data in this experiment also consist of a single measurement of an observer's percept for each unique stimulus (determined by the spatial location of the MQ). Proportions of vertical vs. horizontal perceptual reports were derived from these data. Since there was no clear structure in the data based on spatial location (i.e., vertical and horizontal displacement of the MQ from center), no model dependent on spatial location was fitted to the data. In order to assess the effect of spatial location in a more coarse manner, we compared proportions of percepts across 4 visual fields, or quadrants (top-left, top-right, bottom-left, bottom-right) in all participants by means of a test of equal proportions[2]. Since the data consist of a single measurement per spatial location, measurement error plays a substantial role. In order to minimize the influence of measurement error and ensure that we did not miss any meaningful structure in the variability of perceptual reports, we also smoothed the data and repeated the test of equal proportions. The smoothing was obtained by means of sliding a 3 by 3 window across each perceptual report and replacing this value with the average of the 9 perceptual reports in the window.

5.4.2 Results and discussion

There are clear individual differences in (the dynamics of) perceptual reports of a motion quartet with AR = 1 presented across different locations in the visual field (Fig. 12). Whereas most individuals are variable in their percepts over different spatial locations, a smaller subset of the observers appears to be highly stable. Surprisingly, variability in perceptual reports does not seem to be driven by the location of presentation and seems to occur irrespective of these stimulus manipulations. As can be seen by visual inspection of the examples in Figure 12 (see also Supplementary Fig. C for all participants and smoothed data), there seems to be no clear structure of percepts across spatial location. Results of the test of equal proportions also

[2] The test of equal proportions consists of a weighted sum of squared deviations between the observed proportions in each group and the overall proportion for all groups. The test statistic has an approximate χ^2 distribution.

indicated that there is no significant difference in proportions of percepts compared across the 4 quadrants for any of the participants, except participant no. 040 (χ_3 = 10.52, p = .015). However, the smoothed data did not yield a significant result for this participant, or any other (Supplementary Fig. C). Therefore, we may conclude that there is no clear, meaningful influence of spatial location on the perception of this MQ with AR = 1. Again, we cannot exclude the influence of voluntary control, as the SOA was 500 ms. However, the absence of any dependence of percepts on spatial location is most probably not merely due to voluntary control, since the percepts in the spatiotemporal scale experiment were still clearly defined by stimulus characteristics when the SOA was as long as 640 ms. Interestingly, in contrast to the current findings, various previous studies have reported an effect of spatial location on the perception of visual stimuli (Moutsiana et al., 2016; Schwarzkopf, 2019; Finlayson et al., 2020). This indicates that perception of different basic visual stimuli is differentially affected by location in the visual field, which would be worthwhile for further exploration. One of many possible distinctions here may be the difference between static and dynamic stimuli.

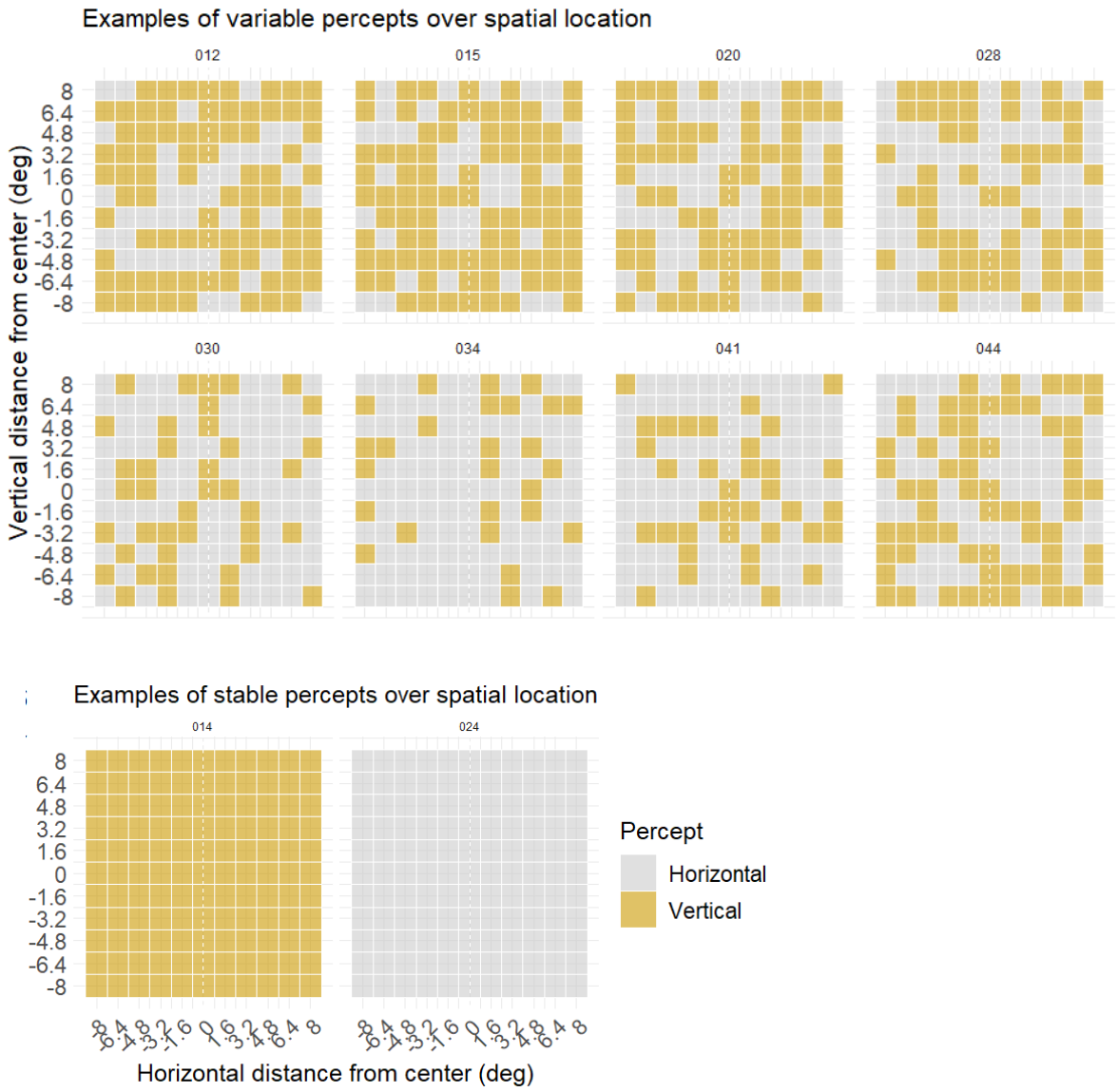


Figure 12. Percepts of ambiguous MQ across spatial location (Spatial Location experiment). Each different plot represents all perceptual reports for one participant across different vertical and horizontal distances. Each colored square in a plot represents one perceptual report, which may be 'Horizontal' or 'Vertical'. Top: Examples of participants with variable percepts across spatial location (or time). Bottom: examples of participants with a stable percept across spatial location (or time). Plots for all participants can be found in Supplementary Figure C.
https://osf.io/tahqw/?view_only=f0b5de9a282140fa973df0763529b720

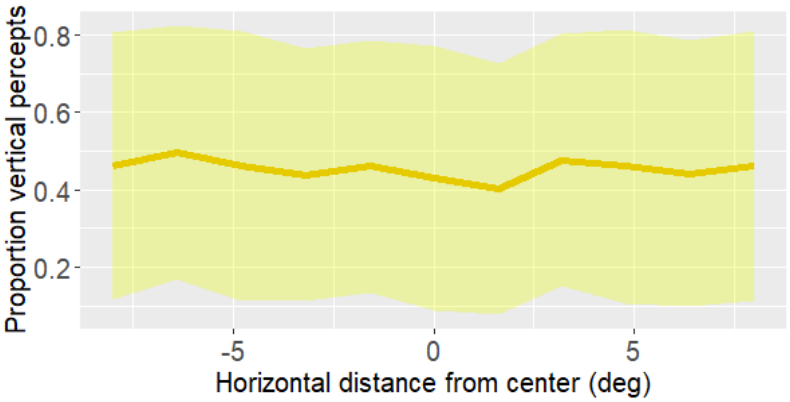


Figure 13. Average proportion of vertical percepts across horizontal distance from center (Spatial Location experiment). The thick line represents the average across participants, the yellow overlay represents the range of one standard deviation in both directions.

A remarkable finding here is that, on average, observers do not tend to report a vertical apparent motion percept more often when the motion quartet is presented in the vertical midline (i.e., horizontal distance of MQ from center is around zero) of the display (Fig. 13). This contrasts with previous findings which suggested that the vertical bias was due to the hemifield(s) and hemisphere(s) of presentation location (Chaudhuri & Glaser, 1991; Genç et al., 2011). Moreover, the data in this experiment indicate that there is no vertical bias at all (Fig. 13). There may be several reasons for these conflicting results. In principle, it is possible that the observers in the current study differed in their perceptual bias from the observers in previous studies by chance. However, since the current and the study of Genç et al. (2011) both contain data from approximately 35 participants, this would be rather unlikely. Since we also observed a slight vertical bias on average in the first two experiments, it is more plausible that the specific stimulus properties in our study were substantially different from previous studies. For instance, it is highly probable that the SOA of 500 ms had an impact on perceptual bias, since we know from the Spatiotemporal Scale experiment that longer SOA gives rise to more individual variability and on average, a less prominent vertical bias. However, in contrast to the Spatial Location experiment, the stimuli with longer temporal distances (i.e., 480 and 640 ms) in the Spatiotemporal Scale experiment still yielded a vertical bias *on average*. This suggests that factors other than SOA are causing these conflicting findings. This is an important consideration to be made, given that the MQ stimuli in the current study are not extensively different from the stimuli employed in previous research (other than SOA), and given that MQ stimuli employed in other studies often differ as much.

5.5 Sequential effects of percepts

The results of the last experiment showed substantial individual differences in variability of the perception of an identical motion quartet across space. However, perception did not seem to be influenced by stimulus location. As mentioned in the introduction, perception is the result of both stimulus and observer characteristics, and their interactions. So far, we have discussed the effects of stimulus manipulations (e.g., the overall effect of temporal distance on perceptual bias across all participants) and the observer-dependent interactive effects of these stimulus manipulations (e.g., the observer-dependent effect of temporal distance on the size of the transition zone). When in the spatial location experiment, variability in perception was not driven by changes in a stimulus characteristic (i.e., spatial location) and we did not see any observer-dependent interactive effects of this stimulus characteristic either, observer characteristics may be causing the variability.

Spontaneous dynamics in internal factors of the observer over time may be at the basis of the perceptual variability observed in the Spatial Location experiment. Moreover, if spontaneous dynamics are indeed a strong factor determining perception of the MQ, it is possible that they have played a role in the first two experiments assessing the influence of spatiotemporal characteristics as well. In order to get some more insight into the temporal characteristics of the perceptual dynamics over the time course of the experiment, we employed recurrence quantification analysis. For comparison, we performed this analysis on all three experiments. A recurrence analysis looks into the number and duration of recurrences of a specific state in a dynamical system over time (Wallot, 2017). It has been employed in eye-movement research (Anderson et al., 2013), where each state in the dynamical system of an observer's eye movements roughly corresponds to the same location of an eye movement. In the case of the current study, the dynamical system corresponds to an observer's multistable perception of the ambiguous stimulus, with the states corresponding to one of two possible percepts (vertical or horizontal apparent motion), recurrences of which correspond to an identical percept on particular trials. By means of the recurrence analysis, we aim to explore sequential effects of percepts in all observers.

5.5.1 Methods

We examined individual perceptual dynamics by means of a recurrence quantification analysis (RQA) and its corresponding recurrence plots. A recurrence plot in the context of the current study is a plot with recurrence points displaying, for each moment within the measured time period, the times at which the dependent variable of interest has the same value as at that moment. For instance, if you want to know whether the reported percept of an observer was identical at time point 1 and time point 5 (Figure 14, Right), you can look for a recurrence point at the intersection on the plot of time point 1 on the x-axis and time point 5 on the y-axis, or vice versa (i.e., the plot is symmetrical along the diagonal).

A RQA evaluates visualizations in a recurrence plot and examines the number and duration of recurrences, which may be quantified by different measures. We will consider three of these measures, recurrence rate (RR), determinism (DET), and laminarity (LAM). The formulas for these three recurrence measures are the following:

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} \boldsymbol{R}(i,j)$$

$$DET = \frac{\sum_{l=l_{min}}^{N} lP(l)}{\sum_{l=1}^{N} lP(l)}$$

$$LAM = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=1}^{N} vP(v)}$$

RR is a measure of the density of recurrence points in the recurrence plot. DET is the proportion of all recurrence points on diagonals made up from multiple recurrence points. LAM is the proportion of all recurrence points on the vertical (or horizontal) lines made up from multiple recurrence points. Accordingly, the recurrence plot provides a direct visualization of (single) repetitions of the same state (i.e., recurrence points, which are measured by RR and correspond to the re-occurrence of a specific percept), multiple consecutive repetitions of the same state (i.e., vertical or horizontal lines of recurrence points, which are measured by LAM and correspond to the observer being 'stuck' in one specific percept from one moment in time to the next), and repetitions of the same sequence of

different states (i.e., diagonal lines of recurrence points, which are measured by DET and correspond to the re-occurrence of a pattern of percepts).

All these visualizations and measures may also be visualized and measured in a more classic treatment of percept sequences, such as in Figure 14 (Left). More specifically, a recurrence point (single re-occurrence of the same state and the derivative measure RR) may be seen as every instance of a certain state, excluding its first occurrence. The vertical or horizontal lines of recurrence points in a recurrence plot (i.e., multiple consecutive repetitions of the same state and the derivative measure LAM) may be seen as multiple consecutive instances of the same state. And finally, the diagonal lines of recurrence points in a recurrence plot (i.e., repetitions of the same sequence of different states and the derivative measure DET) may be seen as multiple occurrences of the same sequence of different states. However, a percept sequence is more suitable to display the specific states in a sequence, whereas the recurrence plot is more suitable to display the recurrence relationships among states in this sequence in a straightforward manner. Moreover, quantitative measures of sequential analysis can be obtained directly from the recurrence plots, whereas data re-arrangement or supplementary analysis is necessary in order to obtain some of the equivalent measures from the classic percept sequence visualization.

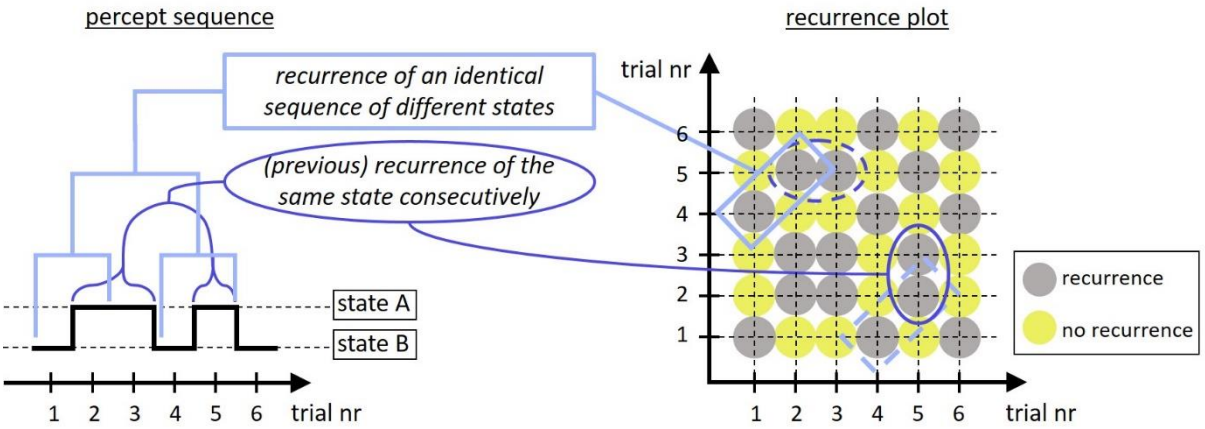


Figure 14. Example of a percept sequence and its corresponding recurrence plot. The light blue rectangles indicate the recurrence of an identical sequence of different states, as quantified by DET. The dark blue oval shapes indicate a (previous) recurrence of the same state consecutively, as quantified by LAM. A comparison of these two different types of sequential representations shows that a percept sequence is more suitable to display the specific states in a sequence, whereas the recurrence plot is more suitable to display the recurrence relationships among states in this sequence. In this hypothetical example, with 18 recurrence points in a recurrence plot of 36 points in total, including 8 points on diagonals of minimum 2 recurrence points (excluding the main diagonal) and 6 points on vertical lines of minimum 2 recurrence points, the recurrence measures can be calculated as follows:

$$RR = \frac{1}{36} * 18$$

$$DET = \frac{8}{18}$$

$$LAM = \frac{6}{18}$$

5.5.2 Results and discussion

The recurrence plots from the Spatial Location experiment (Fig. 15, Supplementary Fig. E for all participants) clearly show individual differences in perceptual dynamics over time. Whereas some individuals are "stuck" in the same percept throughout the entire experiment, other participants show slightly more variability in their percepts over time or have highly variable percepts over time. Even without looking at observer ID, we can derive that individuals that have a stable percept over spatial location (Fig. 10, bottom) will have recurrence plots that display a stable percept over time as well (Fig. 12, top), irrespective of which percept this was. However, without observer ID it is not possible to determine which individual plots of percepts over spatial location (Fig. 10) correspond to medium or high variability over time (Fig. 12), since these plots do not provide information about the time (i.e., trial number) at which there was a stimulus presentation on each spatial location. Since the individual differences we saw in percepts of the motion quartet across the visual field did not seem to be determined by changes in stimulus characteristics (i.e., location), we may assume that these were due to individual differences in dynamics of internal factors. Interestingly, in contrast to the Spatial Location experiment, individual differences in perceptual dynamics over time seem to be completely absent in the Spatial Scale and Spatiotemporal Scale experiment (Fig. 16, Supplementary Fig. F for all participants). All individuals in these experiments appear to be highly variable in their perceptual dynamics.

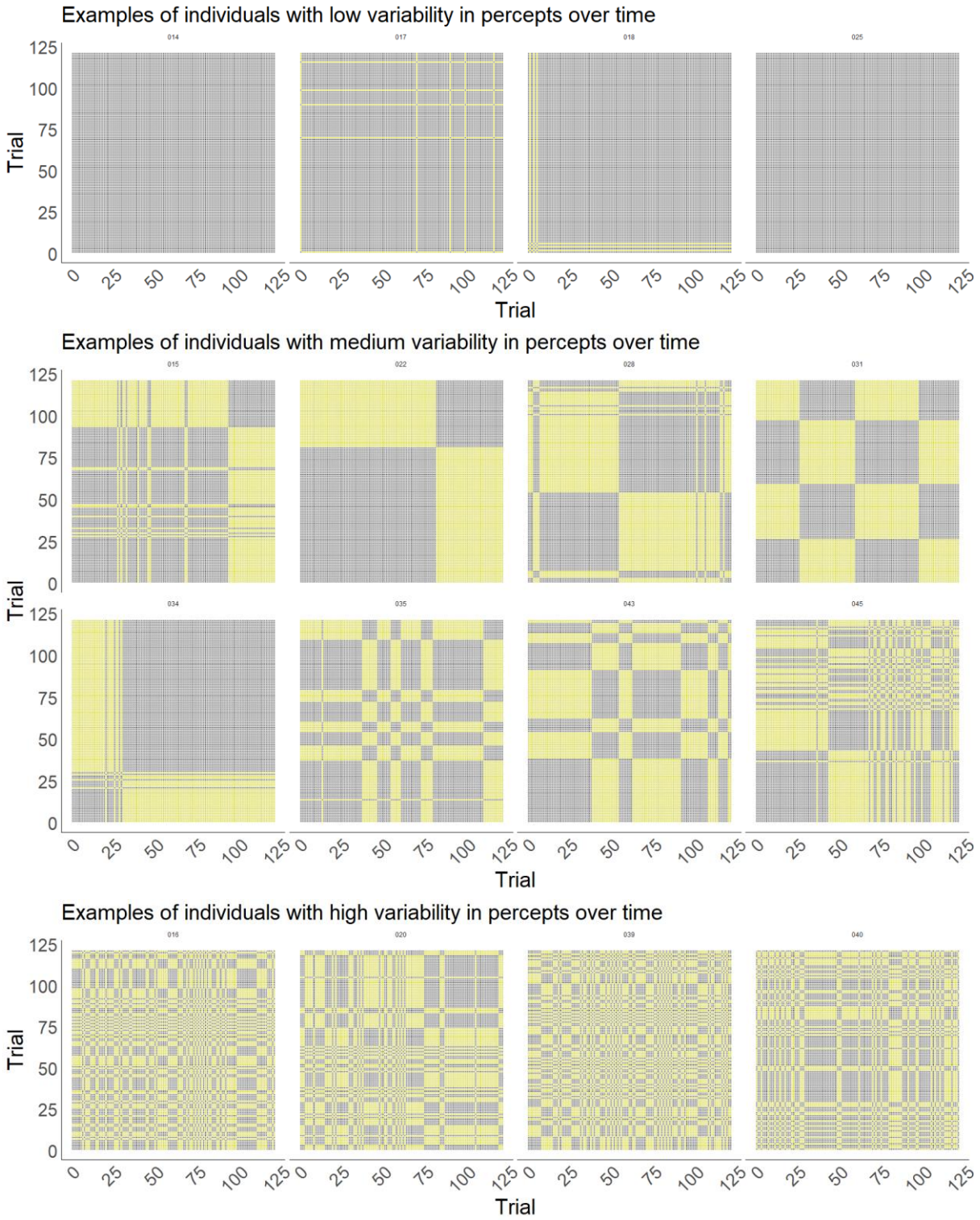


Figure 15. Examples of recurrence plots of the Spatial Location experiment. Each plot contains recurrence points displaying in gray, for each moment, the times at which the perceptual report is the same as the one at that moment. For instance, an identical reported percept of an observer at trial 1 and trial 50 will be displayed by a gray recurrence point at their intersection on the plot. From top to bottom: Examples of individuals with a stable percept over time, individuals with a medium variability and individuals with a high variability in percepts over time. Plots for all participants can be found in Supplementary Figure E.
https://osf.io/tahqw/?view_only=f0b5de9a282140fa973df0763529b720

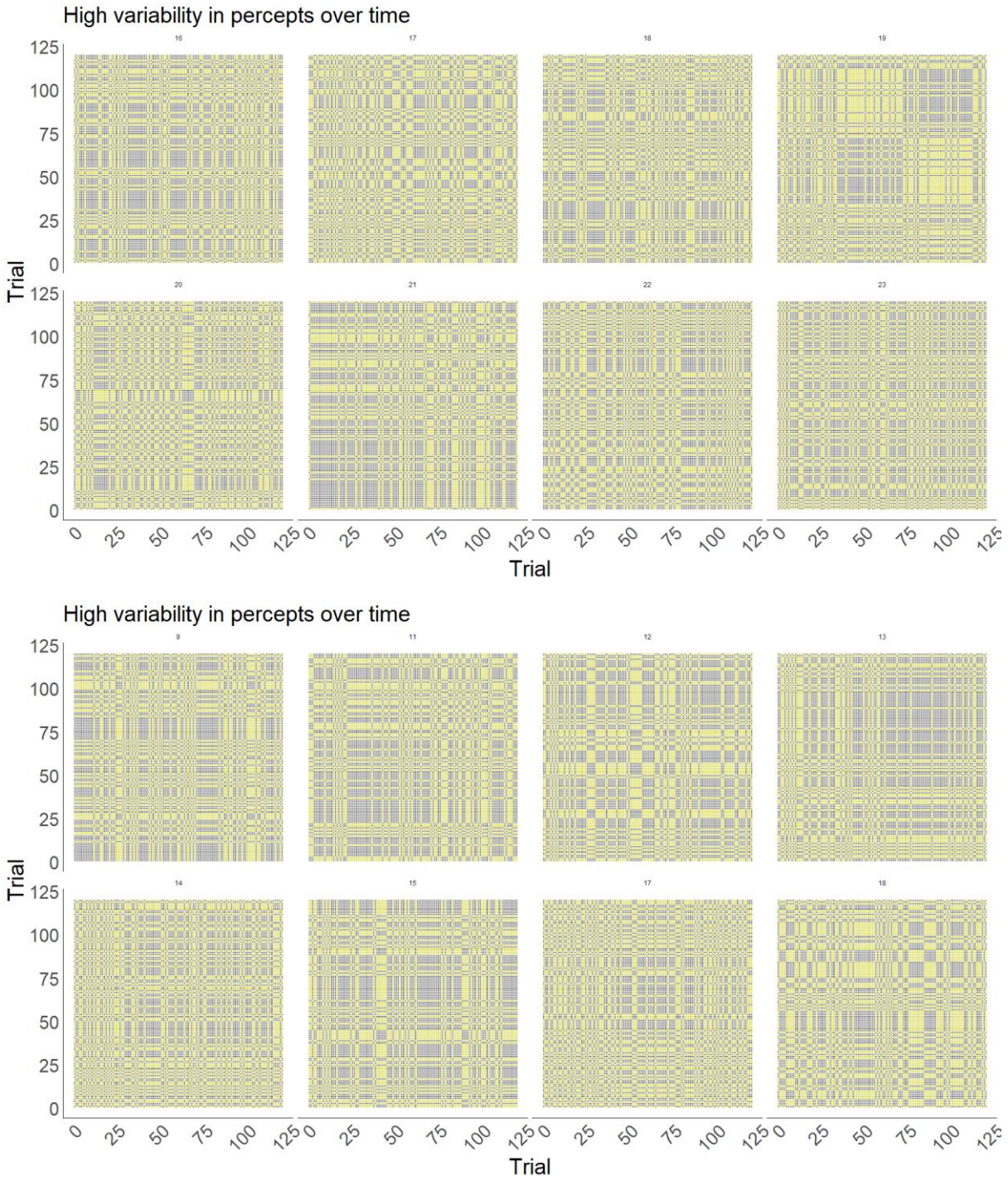


Figure 16. Examples of recurrence plots of the Spatial Scale and Spatiotemporal Scale experiments. Each plot contains recurrence points displaying in gray, for each moment, the times at which the perceptual report is the same as the one at that moment. Top: Examples of individuals from the Spatial Scale experiment, all with a high variability in percepts over time. Bottom: Examples of individuals from the Spatiotemporal Scale experiment, all with a high variability in percepts over time. Plots for all participants can be found in Supplementary Figure F.
https://osf.io/tahqw/?view_only=f0b5de9a282140fa973df0763529b720

Likewise, in the results of the RQA we can see a similar difference between the Spatial and Spatiotemporal Scale experiments and the Spatial Location experiment. In Figure 17, density plots of observers with different values for different measures of the RQA are displayed. Whereas observers have relatively evenly spread and highly variable RR values (i.e., measure of the density of recurrence points in the recurrence plot) in the Spatial Location experiment, they have more dense and homogeneously low RR values in the Spatial and Spatiotemporal Scale experiments. Similarly, whereas observers have more variable and higher DET values (i.e., proportion of all recurrence points on the diagonals, which correspond to the re-occurrence of a pattern) in the Spatial Location experiment, they have more homogeneously low DET values in the Spatial and Spatiotemporal Scale experiments. This implies that observers in the Spatial Location experiment are more likely to have the same percept multiple times (i.e., RR) and that they are more likely to have the same sequence of percepts multiple times (i.e., DET). Moreover, as the values of LAM (i.e., proportion of all recurrence points on the vertical lines, indicating identical values from one moment to the next) show, observers in the Spatial Location experiment are also more likely to have the same percept multiple times in a row, implying that observers tend to get "stuck" in one percept.

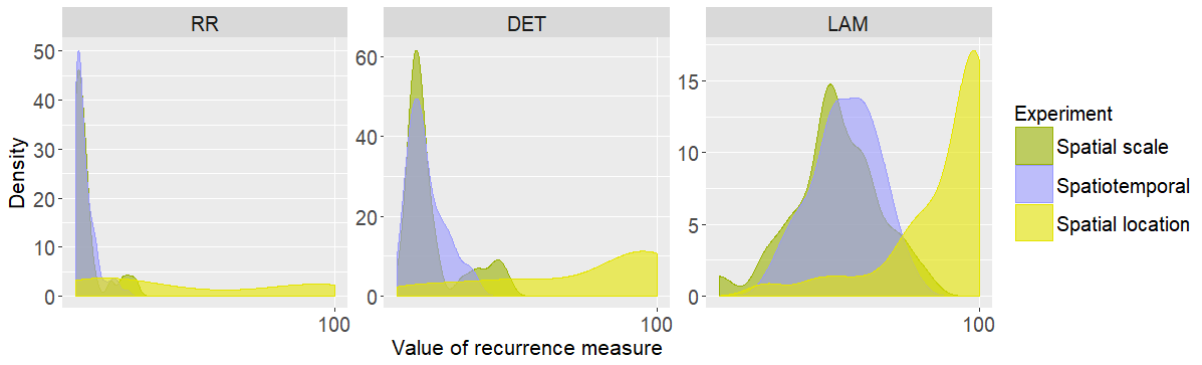


Figure 17. Density plots of RQA measures for all participants and all experiments. Each color in the density plots displays the values of the recurrence measure for all participants in the experiment corresponding to this color. Each density plot corresponds to a different RQA measure.

Yet, it is highly unlikely that all individuals in the Spatial Scale and Spatiotemporal Scale experiments were so homogeneous in the dynamics of their internal factors that these individual differences would be so small. Since spatial distance was altered on every trial and we saw a clear stimulus effect of spatial distance on perception in these experiments (i.e., percepts were clearly systematically driven by vertical and horizontal distances as can be seen

in Fig. 3), a more probable explanation would be that individual differences in dynamics of internal factors were overridden by a stronger stimulus-driven effect of spatial and perhaps temporal distance. Likewise, Leopold et al. (2002) found that individual differences in the perception of different ambiguous stimuli could be changed by manipulating the length of the inter-stimulus-interval. While the inter-individual differences in this study were left undiscussed, they can be observed in some of their figures. In the current study, the individual differences in dynamics of internal factors are most probably still present in the Spatial and Spatiotemporal Scale experiment, but they may have been cast to the background by a more dominant influence of stimulus manipulations. We cannot know for sure that temporal distances were a causal factor in the difference in recurrence plots for the Spatiotemporal Scale and Spatial Location experiment, as the recurrence plots from the Spatial Scale experiment did not seem to differ from the Spatiotemporal Scale experiment.

A consideration to be made is whether some individuals are more susceptible to stimulus differences overriding spontaneous internal dynamics than others. Accordingly, individual differences in the recurrence plots in the Spatial Location experiment may be caused by individual differences in the influence of overriding stimulus effects, rather than individual differences in spontaneous internal dynamics. Following this line of thought implies that observers with the highest variability in percepts over time would also be the most susceptible to overriding stimulus effects. However, even when considering only the data of participants with high variability in percepts over time, the absence of any structure in percepts across spatial location remains. Therefore, it is indeed plausible that the individual differences observed in the spatial location experiment are driven by spontaneous internal dynamics.

Why is spatial scale a more disruptive stimulus influence for spontaneous internal dynamics than spatial location? One important difference between the Spatial Scale experiment and the Spatial Location experiment is the temporal distance between the two frames of the MQ, or SOA. However, even when we consider only the temporal distance of 480 ms (differing from SOA of the Spatial Location experiment by only 20 ms) in the Spatiotemporal Scale experiment, we can see a clear structure of percepts dependent on spatial distance, whereas this is completely absent for the Spatial Location experiment. This indicates that the main stimulus properties of spatial scale vs. spatial location are most likely at the basis of differences in the RQA. It would be very worthwhile for future studies to further investigate why spatio(-

temporal) distances are more disruptive of spontaneous internal dynamics than spatial location.

We may draw a parallel between the RQA and other types of sequential analysis. Research on sequential effects, or history effects, has shown that there are many potential implications of the characteristics of previous stimuli and previous percepts on the perception of a stimulus in the current trial (e.g., priming, adaptation, hysteresis, perceptual stabilization) (Gepshtein & Kubovy, 2005; Pearson & Brascamp, 2008; Murphy et al., 2014).

The effect of the previous percepts on the current percept is what we have looked into in the recurrence analysis. In the Spatial and Spatiotemporal Scale experiments, we did not see any strong short- or longer-lasting effects of percepts in previous trials on the percept in the current trial. This can be derived from the patterns in the recurrence plots, as well as the RQA measures. More specifically, if there were a strong dependence on percepts in previous trials or an effect of perceptual stabilization, we would see a less variable pattern in the recurrence plots of these experiments and a less clear relationship of percepts to stimulus characteristics of the current trial, neither of which is the case. Moreover, as perceptual stabilization is usually seen for presentations of the same stimulus, the stimulus changes in each trial were probably also too large for any stabilization to occur.

However, in the Spatial Location experiment, there is a clear relationship between the previous percepts and the percept in the current trial for some observers. This can be derived from the low variability of the patterns in some of the recurrence plots, as well as the RQA measures. Perceptual stabilization might have happened, but only for the observers with a low variability in percepts across time. However, this would not explain why perceptual stabilization in these observers occurs irrespective of stimulus location and does not occur in a location-contingent or retinotopic manner, as has been suggested in these previous research on perceptual stabilization and memory (Chen & He, 2004; Pearson & Brascamp, 2008; Knapen et al., 2009; van Dam, 2010; Murphy et al., 2014).

We may speculate about whether a strong relationship between previous percepts and the percept in the current trial is causal, or whether there is a less direct relationship. For instance, if there is a causal relationship between the previous percepts and the current one, a participant who has ‘low variability over time’ will have a stronger effect of percept history

than an observer who has 'high variability over time', and will therefore be more likely to be "stuck" in one percept. On the other hand, if there is no causal relationship between previous percepts and the current one, a participant who has 'low variability over time' may have an inherent temporary bias towards one percept, unrelated to percept history, and will therefore be more likely to be "stuck" in one percept. Importantly, either way, dynamic processes within the visual system of the observer, and not differences between external stimuli, appear to be at the basis of patterns of percepts, and individual differences in these patterns, in the Spatial Location experiment.

The recurrence analysis did not look into any effects of previous stimuli on the perception of the current stimulus. For the Spatial and Spatiotemporal Scale experiments, the effect of the previous stimuli may have had an effect on the perception of the current MQ. However, no long lasting effects (i.e., building up across multiple trials and lasting for multiple trials) of previous stimuli can be present, since the order of stimuli is random and the eventual pattern of percepts is unmistakably formed by current trial stimulus characteristics. Therefore, we should only consider the effects of the stimuli in the previous trial. Moreover, as percepts are almost exclusively influenced by stimulus characteristics of the current trial (e.g., a clear effect of horizontal distance), there can be a strong effect of the previous trial only around the diagonal (AR approximately = 1), as we see more variability in percepts and more individual differences there. A problem would arise if these differences around the diagonal, which partially constitute the individual differences of bias (horizontal, vertical, no bias), change in AR around PSE (change, no change), and size of the transition zone (small or large), would be due to merely a coincidental difference in the stimuli that were presented before the trials around this diagonal.

Since the differences between observers are made up from multiple trials and are situated in a particular "region" of stimulus characteristics (e.g., a clustering of more horizontal percepts exactly around the diagonal is what constitutes an observer with a horizontal bias), it is quite unlikely that these differences are merely due to a coincidental difference in previous trials. However, as a proof of concept, we computed the median AR of the previous trials for all trials on the diagonal in the spatial scale experiment and found no diverging median AR for observers assigned to a particular perceptual bias, or for any observers in general. This indicates that the AR of the stimuli previous stimuli, or stimulus bias, was not a driving force

for differences in perceptual bias. Therefore, we believe it is unlikely that differences between observers are due to differences between previous (random sequences of) stimuli in these experiments.

In the Spatial Location experiment, there was no clear effect of stimulus characteristics in the current trial at all, which makes any potential effects of stimulus characteristics of previous trials on the percept in the current trial even less likely. Therefore, there is no indication for previous stimuli being causally related to the differences between observers in this experiment. A coincidental difference of previous stimuli cannot be consistent enough across multiple (29, 22, and 36) participants to influence a difference between the Spatial and Spatiotemporal experiment vs. the Spatial Location experiment. Since there is also no reason to believe that differences in previous stimuli caused any differences between observers in any of the experiments, we do not believe that effects of previous stimuli were an important factor in any of the findings we discussed.

5.6 Conclusions

In the current study, we aimed to illustrate the richness of information that can be uncovered when considering the interaction of various stimulus- and observer-related factors. We demonstrated that considering different stimulus manipulations, different observers and their interactions can provide a more nuanced and informative view on the processes governing visual perception. Firstly, we showed that not all individuals show the same group average stimulus-driven effects on the perception of a MQ. Secondly, we showed that stimulus-driven effects are not always equal across the entire stimulus range. Thirdly, we showed that there are clear individual differences in spontaneous perceptual dynamics and that these can be overridden by some (e.g. spatial scale), but not all (e.g. spatial location) stimulus manipulations. This study provided a demonstration of how the interaction of stimulus and observer properties should indeed not be understated (Mollon, Bosten, Peterzell, & Webster, 2017). Moreover, we clearly saw that findings based on aggregations (i.e., perception across observers or stimuli) often fail to reflect the constituent processes in the individuals of which they are thought to be indicative (i.e., perception for a particular individual or perception of a similar, but slightly different stimulus). As previously pointed out by Wexler (2018), visual perception of the MQ can be highly complex, depending on various stimulus and observer

characteristics and their interactions. When this can be demonstrated for simple, well-known stimuli such as the MQ, they will most probably be at least as important for less extensively studied phenomena as well. We were only able to shed more light onto this complexity thanks to the "reverse-hourglass" approach, and we would like to encourage more research to consider opting for a similar approach. The results of the current study may give rise to further, possibly more targeted, exploration, which will ultimately provide a foundation for a stronger, more comprehensive and more robust theory.

References

Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods*, *45*(3), 842-856. https://doi.org/10.3758/s13428-012-0299-5

Brascamp, J. W., Becker, M. W., & Hambrick, D. Z. (2018). Revisiting individual differences in the time course of binocular rivalry. *Journal of Vision*, *18*(7): 3. https://doi.org/10.1167/18.7.3

Carter, O., & Cavanagh, P. (2007). Onset rivalry: Brief presentation isolates an early independent phase of perceptual competition. *PLoS One*, *2*(4), e343. https://doi.org/10.1371/journal.pone.0000343

Charest, I., & Kriegeskorte, N. (2015). The brain of the beholder: honouring individual representational idiosyncrasies. *Language, Cognition and Neuroscience*, *30*(4), 367-379. https://doi.org/10.1080/23273798.2014.1002505

Chaudhuri, A., & Glaser, D. A. (1991). Metastable motion anisotropy. *Visual Neuroscience*, *7*(5), 397-407.

Chen, X., & He, S. (2004). Local factors determine the stabilization of monocular ambiguous and binocular rivalry stimuli. *Current Biology*, *14*(11), 1013-1017. https://doi.org/10.1016/j.cub.2004.05.042

Davis-Stober, C. P., & Regenwetter, M. (2019). The 'paradox' of converging evidence. *Psychological Review, 126*(6), 865-879. https://doi.org/10.1037/rev0000156

Debrouwere, S., & Rosseel, Y. (2021). The conceptual, cunning and conclusive experiment in psychology [published online ahead of print December 8, 2021]. *Perspectives on Psychological Science,* https://doi.org/10.1177/17456916211026947.

Dieter, K. C., Sy, J. L., & Blake, R. (2017). Individual differences in sensory eye dominance reflected in the dynamics of binocular rivalry. *Vision Research*, *141*, 40–50. https://doi.org/10.1016/j.visres.2016.09.014

Finlayson, N. J., Neacsu, V., & Schwarzkopf, D. S. (2020). Spatial Heterogeneity in Bistable Figure-Ground Perception. *i-Perception*, *11*(5), 2041669520961120. https://doi.org/10.1177/2041669520961120

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences of the U.S.A.*, *115*(27), E6106-E6115. https://doi.org/10.1073/pnas.1711978115

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1211-1221. https://doi.org/10.1098/rstb.2008.0300

Gallagher, R. M., & Arnold, D. H. (2014). Interpreting the temporal dynamics of perceptual rivalries. *Perception*, *43*(11), 1239 – 1248. https://doi.org/10.1068/p7648

Genç, E., Bergmann, J., Singer, W., & Kohler, A. (2011). Interhemispheric connections shape subjective experience of bistable motion. *Current Biology*, *21*(17), 1494-1499. https://doi.org/10.1016/j.cub.2011.08.003

Gengerelli, J. A. (1948). Apparent movement in relation to homonymous and heteronymous stimulation of the cerebral hemispheres. *Journal of Experimental Psychology*, *38*(5), 592-599. https://doi.org/10.1037/h0062438

Gepshtein, S., & Kubovy, M. (2005). Stability and change in perception: spatial organization in temporal context. *Experimental Brain Research*, *160*(4), 487-495. https://doi.org/10.1007/s00221-004-2038-3

Greenwald, A. G., Pratkanis, A. R., Leippe, M. R., & Baumgardner, M. H. (1986). Under what conditions does theory obstruct research progress?. *Psychological Review*, *93*(2), 216-229. https://doi.org/10.1037/0033-295X.93.2.216

Hamaker, E. L. (2012). Why researchers should think" within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of Research Methods for Studying Daily Life* (pp. 43–61). New York: The Guilford Press.

Knapen, T., Brascamp, J., Adams, W. J., & Graf, E. W. (2009). The spatial scale of perceptual memory in ambiguous figure perception. *Journal of Vision*, *9*(13): 16. https://doi.org/10.1167/9.13.16

Koenderink, J. (2019). Vision, an optical user interface. *Perception*, *48*(7), 545-601. https://doi.org/10.1177/0301006619853758

Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, *5*(6), 605-609. https://doi.org/10.1038/nn0602-851

Meehl, P. E. (1990). Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports*, *66*(1), 195-244. https://doi.org/10.2466/pr0.1990.66.1.195

Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice?. *Vision Research*, *141*, 4-15. https://doi.org/10.1016/j.visres.2017.11.001

Molenaar, P. C. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201-218. https://doi.org/10.1207/s15366359mea0204_1

Moutsiana, C., De Haas, B., Papageorgiou, A., Van Dijk, J. A., Balraj, A., Greenwood, J. A., & Schwarzkopf, D. S. (2016). Cortical idiosyncrasies predict the perception of object size. *Nature Communications*, *7*(1), 12110. https://doi.org/10.1038/ncomms12110

Murphy, A. P., Leopold, D. A., & Welchman, A. E. (2014). Perceptual memory drives learning of retinotopic biases for bistable stimuli. *Frontiers in Psychology*, *5*, 60. https://doi.org/10.3389/fpsyg.2014.00060

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. https://doi.org/10.1167/16.14.11

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods,* 162(1–2), 8–13, https://doi.org/10.1016/j.jneumeth.2006.11.017.

Pearson, J., & Brascamp, J. (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences*, *12*(9), 334-341. https://doi.org/10.1016/j.tics.2008.05.006

R Core Team. (2014). *R: A language and environment for statistical computing*. From http://www.R-project.org/

Ramachandran, V. S., & Anstis, S. M. (1983). Perceptual organization in moving patterns. *Nature*, *304*(5926), 529-531.

Schwarzkopf, D. S. (2019). Size perception biases are temporally stable and vary consistently between visual field meridians. *i-Perception*, *10*(5), 2041669519878722. https://doi.org/10.1177/2041669519878722

Wallot, S. (2017). Recurrence quantification analysis of processes and products of discourse: A tutorial in R. *Discourse Processes*, *54*(5-6), 382-405. https://doi.org/10.1080/0163853X.2017.1297921

Wexler, M. (2018). Multidimensional internal dynamics underlying the perception of motion. *Journal of Vision*, *18*(5): 7. https://doi.org/10.1167/18.5.7

Wexler, M., Duyck, M., & Mamassian, P. (2015). Persistent states in vision break universality and time invariance. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(48), 14990–14995. https://doi.org/10.1073/pnas.1508847112

Wijnants, M., Cox, R., Hasselman, F., Bosman, A., & Van Orden, G. (2013). Does sample rate introduce an artifact in spectral analysis of continuous processes?. *Frontiers in Physiology*, *3*, 495. https://doi.org/10.3389/fphys.2012.00495

6. General Discussion

In the following, we will proceed with a General Discussion of the work in this PhD thesis, starting with a summary and some general conclusions discussed according to the structure of the General Introduction in the section 'Summary and general conclusions'. This will be followed by a meta-theoretical reflection of the implications and caveats associated to the concepts "low-" and "high-level" in the section 'Level with me: a re-evaluation of the use of "low-" and "high-level" in visual perception'. Subsequently, some theoretical aspects that are related to the general conclusions will be discussed in the section 'Some (meta-)theoretical observations and speculations'. Finally, we will touch upon some of the shortcomings in the present studies and suggestions for future research in the section 'Looking back and looking forward'.

6.1 Summary and general conclusions

Visual perception and perceptual organization

In the introduction section of this PhD thesis, we defined perceptual organization in terms of Gestalt psychology, as well as in terms of the empiricist approach to visual perception. Throughout the empirical chapters in this PhD thesis, we saw that the processes governing perceptual organization indeed cannot be understood without considering insights from both theoretical accounts.

In the first chapter, 'Relevance study', we saw that the perceptual organization of the 'local-global' ambiguous motion stimulus is characterized by a 'shift to global', or dominance of the global percept over time, which can be explained according to the Gestalt principle of simplicity, indicating that the global percept is a better Gestalt compared to the local percept of the 'local-global' stimulus. However, as we saw in the second empirical chapter, 'EEG study', the 'local' percept was associated with activity in more posterior electrode positions, compared to the 'global' percept, as would be predicted from the hierarchical visual system put forward in the empiricist approach. We further discuss how Gestalt psychology compares to the empiricist approach in the section 'Some (meta-)theoretical observations and speculations'.

Motion perception & Apparent motion perception

All of the stimuli employed in this PhD thesis were defined by a form of apparent motion perception, a process that has already been thoroughly studied in the literature. We stressed the complexity involved in this seemingly straightforward process in the introduction, and throughout the empirical chapters found that motion perception can indeed be highly intricate, even in processes which were thought to be well-known.

In the third chapter, 'CSF study', we saw how a descriptive theory for sensitivity to luminance contrast in moving grating stimuli at threshold levels may be translated to motion stimuli above threshold levels. However, many distance-related parameters and types of information are involved in the description of motion and our visual system, as well as we as researchers, are faced with the complex task of interpreting these different parameters and types of information in a correct sense. Even though sensitivity for motion perception at threshold and apparent motion perception in ambiguous stimuli undoubtedly share common underlying mechanisms, we concluded that it is not trivial to establish any direct relationship between these two types of motion perception.

In the fourth chapter, 'MQ study', we studied in more depth the perception of the motion quartet, the typical example stimulus demonstrating the phenomenon of apparent motion. We found that its perception is influenced by a complex interaction of spatiotemporal distance between the elements of the stimulus and the observer's perceptual dynamics.

Multistable perception of ambiguous stimuli

In the introduction chapter of this PhD thesis, we provided an analysis of all the different ways in which multistable perception of ambiguous stimuli may manifest, and how this may be comparable to the perception of unambiguous stimuli, or perception in general. Here, we noted that there may be more than one form of ambiguity, and that ambiguity may even exist for stimuli that do not seem to be ambiguous based merely on observable behavior.

In the 'Relevance study' chapter, the 'local-global' stimulus was an ambiguous stimulus that can evoke inter- as well as intra-individual variability in percepts. We adapted the stimulus for each observer so that the stimulus would be most ambiguous to them at the start of the experiment, implying that we studied the stimulus at the PSE between the local and global

percept. Therefore, there might be additional processes associated to the competition between different perceptual representations that are not present during the perception of unambiguous stimuli. Towards the end of the experiment, observers perceived the 'local-global' stimulus as more global, indicating that the stimulus was no longer at the observer's PSE and bringing the perception of the stimulus closer to the scenario of the perception of unambiguous stimuli.

In the 'EEG study' chapter, we saw that the difference in perception of the ambiguous 'local-global' stimulus at the observer's PSE and the perception of the disambiguated version of this stimulus could be reflected in the ERP 'uncertainty effect' put forward by Kornmeier et al. (2016). As discussed, the 'uncertainty effect' could be a reflection of the additional processes associated to the competition between different perceptual representations. However, as the timing of this effect is rather late (200 ms and 400 ms post-onset) to be pre-perceptual, future research should further examine the precise nature of the processes associated to these components.

In the 'MQ study' chapter, we studied different kinds of ambiguous stimuli, ranging from stimuli that evoke intra-individual differences to stimuli that evoke inter-individual differences to relatively unambiguous stimuli. This approach allowed us to reach more comprehensive and nuanced conclusions about variability in multistable perception of apparent motion.

Variability in multistable perception: individual differences

As we argued in the introduction chapter, individual differences that are not the result of an a priori defined manipulation or group are often neglected in research. Nevertheless, we found multiple indications of the importance of individual differences in this PhD thesis.

In the first chapter, 'Relevance study', we indeed saw a 'shift to global' in the perception of the 'local-global' stimulus on average, but we clearly saw substantial individual differences as well. However, in the subsequent chapter, 'EEG study', the 'shift to global' was not found to be significant on average. We speculated that the individual differences in the perception of this same stimulus may be one of the reasons (aside from other factors such as the difference in the duration of the experiment) why this was the case.

In the third chapter, 'CSF study', we decided to examine individual differences in the general population in more depth, and aimed to relate individual differences in the "low-level" measure of luminance contrast sensitivity to individual differences in the perception of the 'local-global' stimulus. Although fundamental problems in the theory made a direct comparison impossible, the measured individual CSFs and PSEs in the perception of the 'local-global' stimulus as such did show a great amount of individual differences.

In the 'MQ study', individual differences in the perception of a MQ were examined as the main topic of the chapter. We found that there are far more individual differences in the perception of the MQ than previously assumed, and that these individual differences are also sometimes influenced by stimulus characteristics. We concluded that studying individual differences is fundamental to further advance our understanding of visual perception.

Origins of variability in perception

When considering the origins of variability in perception, we proposed multiple theoretical frameworks that had individual differences and variability in perception as a logical consequence. According to vision as unconscious inference (von Helmholtz, 1867), our processing resources are limited, which implies that we are not able to represent the world in a fully veridical manner and that we have to make inferences based on likeliness in our previous experiences. According to vision as a user interface (Hoffman, Singh, & Prakash, 2015; Koenderink, 2011; Koenderink, 2019), we do not need to represent the world in a veridical manner and we can rely on any other type of information to obtain perception that supports adaptive behavior. An important consequence of both theories is the fact that a limited amount of visual information and various other factors, such as behavioral relevance, influence perception.

In the first chapter, 'Relevance study', we saw that behavioral relevance of a certain percept in an auxiliary task did not significantly influence the perception of the 'local-global' stimulus. Therefore, we can conclude that the most perceived (i.e., according to the account of vision as unconscious inference based on likeliness) or most task relevant (i.e., according to the account of vision as a user interface) percept of the 'local-global' stimulus is not favored over time. However, the perception of this stimulus did change over time, resulting in a bias towards the global percept. The results of this experiment support the notion that our visual

system strives towards a better, simpler Gestalt. Accordingly, perception of the 'local-global' stimulus is not only determined by the Gestalt principle of simplicity, its perception is also influenced by experience with the stimulus. However, future research should determine precisely why this perceptual organization is favored over time.

In the 'EEG study', we studied the ERP 'uncertainty effect', describing a difference in voltage for the perception of the ambiguous 'local-global' motion stimulus (more uncertainty associated with its perception) compared to disambiguated versions of this stimulus (more certainty associated with its perception). When the perception of a stimulus is associated to a higher degree of uncertainty, it may also be less behaviorally useful (i.e., according to the account of vision as a user interface) compared to a percept that is associated to a lower degree of uncertainty. We found that a perceptual switch towards the global percept was associated to a higher uncertainty (i.e., lower amplitude is associated to a higher uncertainty according to the interpretation of Kornmeier et al (2016)), and may therefore be associated to a lower usefulness according to the user interface theory as well. However, given the fact that the global percept was perceived with an equal, and not a lower probability as the local percept, it would be unlikely that the user interface theory accounts for the findings of the 'EEG study'. The fact that the global percept was perceived with an equal probability as the local percept and the fact that our results indicated an equal dominance of both percepts over the course of the experiment as well (i.e., no significant change in the PSE between both percepts in the beginning compared to the end of the experiment), may be accounted for by the idea that likeliness determines perception in the account of vision as unconscious inference. Nevertheless, this conclusion would be in contrast with the findings and conclusions of the 'Relevance study'.

A theory stating that perception at threshold as well as above threshold levels is determined by an optimal allocation of limited resources (i.e., in accordance with the account of vision as unconscious inference based on likeliness) (Gepshtein & Kubovy, 2007), was the basis of the third chapter, 'CSF study'. However, because of the problems we experienced in this chapter, we are currently unable to draw any conclusions about the basis of our findings for the theory of vision according to unconscious inference.

A further meta-theoretical analysis of the phylogenetic and ontogenetic origins of variability in multistable perception according to Gestalt psychology compared to the vision as a process

of unconscious inference and the empiricist approach is discussed in the section 'Some (meta-) theoretical observations and speculations'.

Factors determining variability in multistable perception

In this thesis, we examined the influence of different factors on variability in visual perception in terms of different levels of the visual system, described in terms of their "position" in the visual system (i.e., "low-level", "mid-level", and "high-level").

In *Chapter 2, 'Relevance study',* we examined the role of mere exposure vs. experienced "high-level" relevance in an auxiliary task on the perception of the 'local-global' ambiguous motion stimulus over the time course of five successive days. We found that the influence of "high-level" relevance in an auxiliary task did not influence the perception of this stimulus over time. Rather, "low-level" mere exposure seemed to be sufficient to cause a seemingly spontaneous 'shift to global' in the perception of the 'local-global' stimulus over time.

Chapter 3, 'EEG study', measured electrophysiological correlates of perception of the ambiguous 'local-global' motion stimulus vs. disambiguated versions of this stimulus over time, and between a change in perception towards the local compared to the global percept. Here, we replicated the 'uncertainty' or 'ambiguity' effect which was previously reported by Kornmeier et al. (2016, 2020) for the perception of the 'local-global' ambiguous motion stimulus. We showed that the early ERP component P1 was not significantly different for any of the conditions of interest, indicating that the effects we obtained were most probably not related to any differences in "low-level" processing. The relatively early ERP component P200 (150 and 250 ms post stimulus-onset) was present over frontal electrode sites, whereas the later component P400 (300 and 500 ms post stimulus-onset) was present over centroparietal electrode sites. However, in the absence of any source localization, it is not possible for us to draw any certain conclusions about the "low-level" or "high-level" cortical localization of their underlying activity. However, we can conclude that the earlier P200 effect mostly seems to reflect differences in the 'uncertainty' or 'efficiency' associated to percepts disambiguated stimuli, whereas the P400 effect mostly seems to reflect differences in the 'uncertainty' or 'efficiency' associated to percepts ambiguous stimuli. We proposed an alternative interpretation of the 'uncertainty effect', framing it either in terms of the efficiency associated to processes leading up to the perceptual representation, or in terms of the presence of top-

down processes that aim to extend and facilitate a stable subsequent percept. Future research should put these speculative accounts to the test.

Chapter 4: 'CSF study' set out to examine the relationship between the "low-level" measure of individual luminance contrast sensitivity and the perception of a 'local-global' ambiguous motion stimulus which enables both a 'local' and 'global' apparent motion percept. However, because of the problems we encountered, we are currently unable to draw any conclusions concerning the role of the "low-level" measure of contrast sensitivity. Nevertheless, given the wide variety of additional processes that can be associated to the differing measurements of contrast sensitivity and above threshold perception of apparent motion, we concluded that expecting the relationship between these different levels of processing to be quantitatively similar may be less likely than the expectation of a qualitative correspondence.

In *Chapter 5, 'MQ study'*, we explored the influence of stimulus variability on individual differences in the perception of a motion quartet stimulus. Differences in stimuli may be characterized as the primary factor influencing bottom-up visual processing. We found that the effect of different spatiotemporal distances in the MQ stimuli clearly interacted with various forms of individual differences in the perception of the MQ, as well as observer-related perceptual dynamics.

The wide variety along different levels of the visual system in the studies carried out in this PhD thesis gave us a rich and comprehensive idea of the enormous complexity involved in visual perception. However, during the progress of this thesis, we encountered some caveats associated with the assumption of different levels in a hierarchy of the visual system and became aware that the terms low, mid and high level are sometimes used to mean different things throughout the literature. The following section in the present discussion chapter, 'level with me', aims to address the ambiguity inherent to the terms "low-" vs. "high-level" in the research of this PhD thesis and in vision science in general, and aims to illustrate where the distinction between different levels can be useful and where it is counterproductive for developing a better understanding of visual perception.

6.2 Level with me: A re-evaluation of the use of “low-“ and “high-level” in visual perception.

An abundance of “low-level” and “high-level” in the literature

The concepts “low-“, “mid-“, and “high-level” are ubiquitous in the literature on visual perception and used across various topics in vision science. Sometimes they are employed indirectly in order to embed the research of interest within a well-known framework of levels of processing and/or a cortical hierarchy. For instance, one could study Bayesian decoding of features in visual perception and frame differences between these features in terms of “low-“ and “high-level” (Ding et al., 2017). Other times, and sometimes concurrently, they are employed when explicitly assessing the effect of either “low-“ or “high-level” factors, which are subsequently operationalized in one way or another. For instance, one could study the effect of “low-“ and “high-level” influences on eye movements and proceed to operationalize what could be qualified as “low-“ vs. “high-level” (Schütt et al., 2019).

Examples of the employment of the terms “low-“ and “high-level” in the literature on visual perception in general range from studies on eye movements (e.g., Le Meur, Le Callet & Barba, 2007), to studies on cognitive functions (e.g., Geisler & Chou, 1995), to studies on the potential borders between perception and cognition (e.g., Firestone & Scholl, 2016), and so on. Studies on motion perception have also steadily drawn on the difference between “low-“ and “high-level”. Examples of these studies deal with topics on apparent motion (e.g., Braddick, 1980), biological motion (e.g., Mather, Radford & West, 1992), effects of language on motion perception (e.g., Meteyard, Bahrami & Vigliocco, 2007), computational approaches to motion perception (e.g., Stein & Hebert, 2009), psychopharmacology (e.g., Carter et al., 2004), and so on.

Research on multistable perception of ambiguous stimuli, the main topic of this PhD thesis, also draws heavily on the terms “low-“ and “high-level”. What is more, the debate on the potential involvement of different levels in multistable perception (Brascamp et al., 2018, Long & Toppino, 2004) has made the distinction between different levels in the visual system fundamentally important to the study of ambiguous stimuli. Research examining the EEG

neural correlates (Kornmeier, Hein, & Bach, 2009; Kornmeier, Pfäffle, & Bach, 2011), as well as the large volume of studies exploring fMRI neural correlates (Blake & Logothetis, 2002; Sterzer, Kleinschmidt, & Rees, 2009) of multistable perception have explicitly framed contributing different factors as "low-level" or "high-level".

An implicit assumption that we all agree on the meaning of "low-level" and "high-level"

There has been some discussion on the definitions of the borders between perception and cognition (Rolfs & Dambacher, 2016; Block, 2016; Firestone & Scholl, 2016), concepts which are undoubtedly intertwined with, but most certainly not identical to the concepts of "low-level" and "high-level". Yet, researchers rarely question whether the way we qualify properties and/or processes as "low-level" or "high-level" is understood in the same way by all different readers, even within one research field, and whether the different frameworks in which these terms are embedded are all equally clear about their definition. It is as though these concepts are so universally employed that their meaning should be deeply rooted in the language of visual perception. There is the implicit assumption that they are well-established, self-explanatory, and that we are all in agreement about their meaning and implications.

Indeed, if you are a vision scientist, you will probably assume that you have a fairly clear idea of what these terms imply to you, or even what they imply in general. And often, this assumption will be correct. However, language can be tricky. As the hypothesis of linguistic relativity states, the structure of our language, or the way we conceptualize the world, can affect our understanding of it and the way we think about it (Wolff & Holmes, 2011). Moreover, because of semantic relationships, the terms "low-" and "high-level" are easily associated to other concepts or attributes, and with the implicit assumption that these terms are self-explanatory comes the implicit assumption that we all make the same associations to these words. Indeed, on the rare occasions when an explicit definition of "low-" and "high-level" is given (this might be in student books), they are usually framed by means of the associations commonly made with them (e.g., the complexity of stimulus characteristics associated with them, the brain areas associated with them, the variety of the processes associated with them, the timing of the processes associated with them, etc.).

For instance, imagine that we have no idea of what a bike is, and we are scientists trying to figure out the structure and functioning of a bike. During the progress of our understanding, we decided to start employing the concepts “low-level” and “high-level” to refer to different characteristics of the bike. However, even though all scientists assume that other scientists will understand the implications of the concepts “low-level” and “high-level” when they use these concepts, there is no real consistency in the interpretations of these concepts among different scientists. Scientist A mostly refers to the concepts to describe the length (“short” vs. “long”) of the bike. Scientist B mostly refers to the concepts to describe the color (“red” vs. “blue”) of the bike. Scientist C mostly refers to the concepts to describe the weight (“light” vs. “heavy”) of the bike. Sometimes, the characteristics “short”, “red” and “light” co-occur in the same bike. Similarly, a bike is sometimes “long”, “blue” and “heavy” at the same time. This is the reason why all three kinds of adjectives got bundled up in the concepts “low-“ vs. “high-level” over time. However, importantly, the characteristics “short”, “red” and “light” do not always co-occur in the same bike, and neither do the characteristics “long”, “blue” and “heavy”. In fact, while some research and theories suggest that length, color and weight are associated to each other in some way, other research and theories suggest that they are not necessarily associated to each other at all. Imagine then, one of these scientists reporting findings of a new study about the bike, and describing these findings in concepts of “low-level” and “high-level”, without specifying whether he/she is referring to length, color, or weight. How should the other scientists know that their implicit assumptions of certain interpretations of these concepts are not what the scientist is reporting results about at all? Perhaps it may seem to the scientists as if they are all working on the same problem in “bike science”, when they are in fact studying different problems altogether? How should different scientists and different research domains work together to develop a common understanding and better theory about the bike? Irrespective of which interpretations of “low-level” or “high-level” are correct in the actual structure and functioning of the bike, no scientific progress is possible if scientists A, B, and C do not specify their *own* interpretations. Therefore, it may prove useful to summarize different interpretations of the concepts “low-level” and “high-level”, if only so scientist A, B and C can realize how their understanding of the bike may be limited by the confusion and conflicts in the possible implications of, for instance, the characteristics that are all associated to “low-level”: “short”, “red” and “light”.

Dissecting some possible meanings of “low-level” and “high-level” and what these imply

Thus, we may ask what these potential meanings and implications of the concepts "low-level” and “high-level” are. Defining concepts as general as these, which are assumed to be common knowledge and yet are employed in a variety of different contexts, can be a perilous endeavor. There are multiple interpretations in the literature, which are not always mutually exclusive. In the following, we will try to define the ways in which the terms “low-“ and “high-level” are commonly interpreted in the field of vision science, and often more specifically, in the field of multistable perception. Accordingly, we aim to illustrate that these potentially different interpretations may not always result in a common understanding of “low-“ and “high-level”. Importantly, bear in mind that the meanings of the terms “low-“ and “high-level” are not bound by or restricted to the definitions and implications mentioned here. We do not aim to provide a “correct” definition of the terms “low-“ and “high-level”. Rather, we would like to point out the various ways in which they are employed and implicitly defined in the literature. Accordingly, this allows us to become aware of the potential conflicts between different interpretations that may be adopted by anyone reading these concepts, and the risks associated to the use of these concepts without a clearly stated definition that is intended by the author using these concepts.

The terms “low-level” and “high-level” are adjectives that can describe different types of **nouns**. The implications and meaning of these adjectives, however, is dependent on their association to and interpretation by other **adjectives**.

More specifically, when a **noun** is seen as “low-level” or “high-level”, this noun may be **conceptual or neural**, and may be a **representation** (e.g., in awareness, in a computational network, etc.) **or process**. The combination of these options leads to the distinction of a **conceptual representation** (e.g., a local vs. global percept), a **conceptual process** (e.g., perceptual grouping by similarity), a **neural representation** (e.g., the neural representation of a face), or a **neural process** (e.g., firing of neurons). Importantly, the interpretations of “low-level” and “high-level” as a conceptual representation and a neural process are most prevalent in vision science, but the potential interpretations as a noun are not restricted to these

combinations, as there are conceptual theories of different levels of visual processing as well as hypotheses about different levels of neural representations.

When describing these different types of nouns, the meaning of the terms "low"- and "high-level" is implied by their association to other **adjectives**. We will make a non-exhaustive list of potential adjectives here.

Potential conflicts (i.e., with other interpretations) associated to the interpretations in each of the following scenarios are depicted in blue.

For many adjectives, a comparison is made to how Gestalt psychology would define the adjective associated to "low-" and "high-level" and this comparison is depicted in italics. It is important to note, however, that Gestalt psychologists themselves did not define their theory in terms of "low-" and "high-level".

Examples corresponding to the work in this PhD thesis in each of the following scenarios are depicted in frames.

1) A first adjective that is commonly related to the terms "low-" and "high-level" is their **spatial distance from sensory input in terms of cortical location of the underlying neural process or representation**.

   The interpretation of distance by cortical location has applications for the definitions by neural process and neural representation only. This follows from the characterization of a classic visual hierarchy, where different levels of processing and representation are embedded in different locations from lower-located, more posterior to higher-located, more anterior sites across the human cortex. This visual hierarchy follows the feed-forward, or bottom-up, flow of electrochemical information in the brain, starting in the retinae and proceeding through the lateral geniculate nuclei (LGN), the primary visual cortex (V1), V2, and so on (Figure 1). This information flow is then typically divided into a more temporal ventral stream, proceeding from V3 through region V4, and a more parietal dorsal stream, proceeding from V3 through region V5 or MT. The dorsal stream is associated to motion perception, as most neurons in MT/V5 respond strongly to items moving in a specific direction.

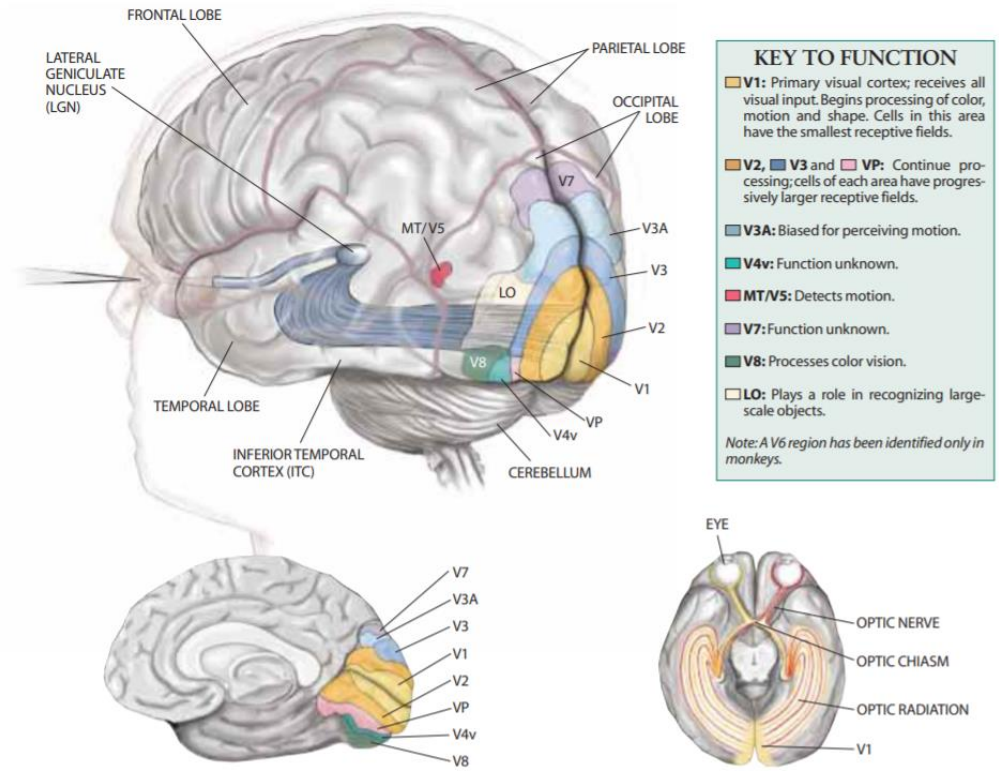


Figure 1. Cortical hierarchy of vision, adopted from Logothetis (1999)

Applied to the definition by neural representation, a “low-level” representation could be the representation of, for instance, the orientation of a line in primary visual cortex, whereas a “higher-level” representation could be, for instance, the representation of a face associated with the fusiform face area in temporal cortex.

In Chapter 2 of this PhD thesis (the ‘Relevance study’), a neural representation of single dot motion could be considered as “low-level”, whereas a neural representation of pulsation or rotation direction of a group of dots (a hexagon or multiple pairs) could be considered as “high-level”. Moreover, we may consider the EEG correlates of a percept to be a neural representation of this percept in the ‘EEG study’ chapter of this PhD thesis. Accordingly, the frontal topography of the ERP in the P400 time window during perception of ambiguous stimuli can be considered as “higher-level” compared to the central topography during the perception of disambiguated stimuli.

Applied to the definition by neural process, a “low-level” process could be the firing of single cells in primary visual cortex, whereas a “higher-level” process could be the pooling of activity across different neurons, thereby increasing the receptive field size.

An even "higher-level" process could be activity in associative areas or the coordination of large-scale activity across a network of different cortical areas.

In the 'Relevance study' chapter of this PhD thesis, motion perception, typically associated to neuronal activity in area MT, could be considered as a "low-level" neural process, whereas a "high-level" process could be the activity of the dorsal attention network during the perception of an ambiguous motion stimulus. Moreover, in the 'EEG study', the frontocentral topography of the processes giving rise to the P200 effect can be considered as "higher-level" compared to the parietal topography of the processes giving rise to the P400 effect.

In contrast to this approach of a classic cortical hierarchy that was mostly advocated by empiricists, Gestalt psychology did not believe in successive summation processes in a hierarchical system. Rather, they claimed that perception is the result of a flexible and dynamic process that leads to the most "Prägnant" perceptual organization of stimuli. In terms of their neural mechanisms, Köhler (1920/1938) referred to holistic electric fields across the whole cortex.

Even though a vast amount of neurophysiological evidence supports the bottom-up order of processing from "low-level" to "high-level" in a classic cortical hierarchy, there is currently also a consensus about the existence of equally important top-down processes, which contradict the order from a supposed "low-level" to a supposed "high-level" of cortical locations (Logothetis, 1999). Visual processing may therefore not merely be the result of an orderly bottom-up series of processes. Alternatively, it can be the result of a sequence of processing loops, as is proposed by a wide variety of theoretical accounts, such as the Reverse Hierarchy Theory Hochstein and Ahissar (2002) and models of predictive coding (Rauss & Pourtois, 2013). For instance, according to the Reverse Hierarchy Theory, visual processing occurs in two stages where a first bottom-up processing sweep to "higher-level" cortical locations is followed by a slower interaction of bottom-up and top-down processes between different cortical locations. One may ask what meaning the concepts "low-" and "high-level" still have in the face of these newer theoretical models of perception that do not necessarily impose a fixed preceding role on early sensory areas, and whether it may

not be equally useful to speak in terms of, for instance, posterior and anterior instead of “low-level” and “high-level”.

Moreover, the interpretations of “low-“ and “high-level” according to the adjective of cortical location can even be quite problematic on their own. For instance, consider the neural process of single neuron firing in frontal cortex. We would define the recording of a single neuron as “low-level” according to a general conception of different levels of analysis, or according to a different conception of cortical location by spatial scale. However, when you consider the adjective of cortical location here, it would be rather odd to categorize activity in frontal cortex as “low-level”.

2) A second adjective that is commonly related to the terms “low-“ and “high-level” is their **temporal distance from sensory input, or timing: early vs. late.**

This adjective often pertains to the timing and duration of a process and is not always applicable to representations. However, the timing of a representation may also be interpreted as how much time or processing (e.g., in a computational model or in the brain) it takes to reach this representation.

For instance, applied to the definition by conceptual representation, we can consider different interpretations. According to a classic visual hierarchy, a “low-level” representation could be a simple visual feature, such as a line, and is considered to be reached early, whereas a “higher-level” representation could be a more complex configuration of visual features, such as a face, and is considered to be reached later (e.g., in a computational model complying to a classic visual hierarchy).

In contrast, according to Gestalt psychology, the perception of the whole, or Gestalt, which can be a holistic configuration such as a face, is seen as taking precedence in our perceptual experience. Accordingly, this representation is reached earlier than the perception of the elements that are part of the whole, such as the lines, which are not automatically part of our perceptual experience.

In the work of this PhD thesis, the representation of moving dots in the 'local-global' stimulus can be early compared to the representation of local rotation or global pulsation direction to be reported for the task in the 'Relevance study'. *In contrast, the local or global percept could also be considered as taking precedence in our perceptual experience, without the need to analyze moving dots first.*

Applied to the definition by conceptual process, we can again consider different interpretations. According to a classic visual hierarchy, simple processes such as detecting the orientation of a line occur earlier than the processing of more complex configurations of visual features (e.g., in a computational model complying to a classic visual hierarchy).

In contrast to this, in Gestalt psychology, perceptual organization and phenomena such as 'grouping by proximity' are considered to be primary, whereas analysis of single visual features would be something that does not occur spontaneously and is certainly not considered as a process that needs to be completed before perceptual organization of a stimulus.

In the work of this PhD thesis, an early process could be the influence of contrast sensitivity on the perception of the 'local-global' stimulus, whereas a late process could be the influence of the task-relevance on the perception of the 'local-global' stimulus in the 'Relevance study'. *In accordance with the Gestaltist approach, "grouping by common fate" can be interpreted as taking precedence compared to any analysis of individually moving dots.*

Applied to the definition by neural representation, the representation of, for instance, the orientation of a line is reached early compared to the representation of a more complex configuration, such as a face.

In the work of this PhD thesis, the neural representation of dot motion in the 'local-global' stimulus can be early compared to the neural representation of rotation or pulsation direction to be reported for the task in the 'Relevance study'.

Applied to the definition by neural process, the firing of neurons sensitive to, for instance, a particular orientation occurs early compared to "higher-level" processes.

In the work of this PhD thesis, a "low-level" neural process could be the processing giving rise to an early event-related component in EEG measurements, such as the visual P1 component we observed in the 'EEG study', whereas a "higher-level" process could be the further processing giving rise to a later event-related component, such as the P400 component we observed in the 'EEG study'.

Interestingly, the Reverse Hierarchy Theory of Hochstein and Ahissar (2002) already advocated to dissociate the structural, anatomical aspects (cf. our adjective of cortical location for neural representations and processes) and functional process-related temporal aspects (cf. this adjective of timing) from the concepts of "low-" and "high-level". According to the theory, in a first and fast feed-forward processing sweep to "higher-level" cortical locations, perception of the "gist", or a categorization, of a stimulus is enabled quickly. Subsequently, a slower interaction of bottom-up and top-down processes between different cortical locations enables a more detailed representation of the stimulus.

3) A third distinction between "low-" and "high-level" pertains to their **similarity to the sensory input: specific vs. comprehensive/abstract**.

Applied to the definition by conceptual representation, a "low-level" representation of a visual feature, such as a line, is more specific and similar to the sensory input, whereas a "higher-level" representation, such as the emotion of a face, is less similar to the sensory input and constitutes a more comprehensive and complex representation of the visual information.

In the 'Relevance study' chapter of this PhD thesis, a specific, "low-level" representation could be a dot in the 'local-global' ambiguous motion stimulus, whereas a comprehensive, "higher-level" representation could be the rotation or pulsation direction that is task-relevant during the perception of this stimulus.

The conception of "higher-level" here is in accordance with what Gestalt psychology would define as a good Gestalt: the more comprehensive and efficient a representation is, the better the Gestalt. However, due to the focus on phenomenology in Gestalt psychology, only the local and global percept would be considered, whereas the representation of a dot that is not a part of our conscious experience would not be defined as "lower-level".

Applied to the definition by conceptual process, the processing of simple features, such as the orientation of a line, could be considered as "low-level" because it is more specific and similar to the sensory input, whereas processing more complex information, such as whether the face seems happy, could be considered as "high-level" because it requires a more comprehensive and complex form of processing.

In the 'Relevance study' chapter of this PhD thesis, a specific, "low-level" process that is similar to the sensory input would be dot motion perception in the 'local-global' ambiguous motion stimulus, whereas a comprehensive, "higher-level" process would be the detection of the rotation or pulsation direction that is task-relevant.

Processes implicated in Gestalt psychology, such as "perceptual grouping by common fate", are not framed with a distinction between "low-" and "high-level", but they can be considered as aiming for the most comprehensive process. Accordingly, dot motion perception can also be seen as less comprehensive than rotation or pulsation direction perception in the 'local-global' stimulus.

The same logic applies to the definition by neural representation and neural process, where "low-level" neurons are also more specific in their response profiles compared to neurons higher up the visual cortical hierarchy.

In the 'EEG study' chapter of this PhD thesis, the neural representation of the disambiguated stimuli could be considered to be more specific and similar to the sensory input, whereas the neural representation of the ambiguous stimuli may be more comprehensive and less similar to the sensory input. Moreover, a "low-level" process that is similar to the sensory input would be the detection of a visual stimulus giving rise to P1 component, whereas a "higher-level" process that is less similar to the sensory input and more comprehensive and complex would be the detection of perceptual uncertainty giving rise to the P200 and P400 components.

The following two adjectives are often linked to the concepts of "low-" and "high-level", but are less directly related to them.

4) **Consciousness**

The application of the adjective of consciousness to the different definitions of nouns we proposed for the previous adjectives may be less straightforward, as consciousness in itself is a very difficult concept to define. Nevertheless, the terms "low-level" and "high level" are often linked to accounts of consciousness and therefore, we think it is important to mention some of these associations here as well. Typical "low-level" functions are usually not associated with consciousness (e.g., processing of visual features in neurons with small receptive fields), whereas typical "high-level" cognitive functions are often associated with awareness (e.g., thinking, memory, reasoning, etc.).

Applied to the definition by a conceptual representation, a "low-level", unconscious representation would be the representation of a single visual feature (e.g., a line), which does not automatically enter our conscious experience, whereas we experience a "higher-level" representation, such as the emotion of a face, consciously.

In the 'Relevance study', the representation of a moving dot does not enter our conscious experience, whereas the local percept of rotation or the global percept of pulsation is experienced consciously. Indeed, ambiguous stimuli have often been employed to study the mechanisms of consciousness, as they enable the transition from one conscious percept to another (Logothetis, 1999).

In Gestalt psychology, the focus of analysis is mostly restricted to phenomenology. Therefore, it is in accordance with the notion that percepts are part of our conscious experience. However, this does not necessarily imply that the Gestaltists would qualify a percept as "higher-level" compared to the representations of visual features that do not reach awareness. Rather, they believe that our perceptual experience is dominant and that perception of single visual features occurs only after scrutiny.

Here we can again refer to the Reverse Hierarchy Theory (Hochstein & Ahissar, 2002), where the first and fast feedforward sweep, resulting in a conscious percept in the end, is dissociated from a slower, more detailed analysis of the visual input.

Applied to the definition by conceptual process, the processing of visual features does not typically enter our awareness, whereas we might assume that a "high-level" process, such as thinking about the emotions of a face, does. However, it is important to note that thinking about a line can also enter our awareness and that the processing of emotions itself does not necessarily enter our awareness.

Similarly, in the 'Relevance study', motion perception could be an example of a "low-level" process that occurs unconsciously, but based on this experiment we are unable to determine whether the "higher-level" process of deriving the rotation or pulsation direction requires awareness.

Applied to the definitions by neural representation and neural process, we could state that the "low-level" neural representation and process of, for instance, a line and the processing of its orientation in early visual cortex, does not typically enter awareness. However, whether a "low-" or "high-level" neural representation or process is associated to a conscious experience critically depends on the definition of "low-level" and "high-level" neural representations and processes that you adopt, with respect to the previous adjectives. For instance, a common conception of the neural basis of awareness states that conscious perception is associated with wide-spread neural activity over sensory as well as fronto-parietal brain areas, which is supported by long-distance functional connectivity and synchrony within a broad neural network (Sergent & Dehaene, 2004). This implies that, for instance, neural processes in primary visual

cortex can be "high-level" if they are part of wide-spread neural activity, according to the adjective of consciousness. However, these same neural processes will be considered as "low-level" according to the adjective of cortical location, leading to a conflict in interpretations.

5) **Local vs. Global**

Similar to the concept of consciousness, the concepts 'local' and 'global' are not clearly defined and multiple possible interpretations exist, depending on the definition of 'low-level' and 'high-level' that you adopt with respect to the previous adjectives. As discussed in the 'EEG study' chapter of this PhD thesis, according to classic empiricists, the terms can be applied to the characteristics of a hierarchical visual system, where 'local' information on different levels is 'pooled' and summarized into more 'global' representations "higher up" in the visual system. In this interpretation, the terms 'local' and 'global' may sometimes be intertwined with different levels in a hierarchical visual system (cf. the adjective of cortical location for neural representations and processes). Indeed, switches to more local and global perceptual organizations of ambiguous stimuli are associated with increases in activity at lower and higher stages (cf. the adjective of cortical location for neural representations) of the visual system, respectively (Fang, Kersten & Murray, 2008; de-Wit, Kubilius, Wagemans & Op de Beeck, 2012). However, even though there are some commonalities between these concepts, it is important to note that they are not necessarily identical.

According to Gestalt psychology, the concepts 'local' and 'global' are both descriptive of states or percepts in awareness (cf. the adjective of consciousness for conceptual representations) that are associated to "mid-level" processing and are therefore considered to be on the same level, not tied to different levels in a visual hierarchy. Nevertheless, the local percept is seen as more comprehensive (cf. the adjective of similarity to the sensory input for conceptual representations) compared to the global percept.

Your “low-“ or “high-level” might not be mine

As you can see, there is a considerable amount of confusion possible within one adjective or across different adjectives that are applicable to the same noun, and even for adjectives that are applied to different nouns. Even though some implications of this vast number of potential meanings are in mutual agreement, there will also often be a discrepancy between what is meant by the author and what is understood by the reader of the concepts “low-“ and “high-level”. Moreover, even within the meaning of one possible noun and adjective, the importance of where the boundary between these different levels is placed is often underestimated, as an arbitrary border can be placed very differently for different people, implicating yet another form of terminological ambiguity.

It is important to note that adopting a differentiation of different levels of representation and processing in visual perception certainly also has its benefits, irrespective of whether this differentiation can be applied similarly to every phenomenon or in every kind of theory. Any form of categorization is made for pragmatic reasons and makes communication about visual perception more convenient in many ways, as long as it is clear what is meant with these terms. Moreover, as these concepts are so deeply rooted in the language concerning visual perception, we believe it would not be feasible to banish these terms. Rather, we would like to advocate a more consistent and clearly delineated use of the concepts “low-level” and “high-level”. When reading a text that mentions these concepts, we might automatically assume their meaning, but when faced with explicit definitions of these concepts, we may realize that their meaning can be more variable. We should be aware that “your low- or high-level might not be mine”.

An important conclusion to be drawn is that the terms “low-level” and “high-level” do not have a strictly defined meaning when isolated from the context in which they are employed. The meaning they aim to convey should therefore be specified further instead of merely relying on the context in which they are employed. A potential way of further specifying this meaning can be according to the nouns and adjectives employed in this chapter.

6.3 Some (meta-) theoretical observations and speculations

Visual perception and perceptual organization

At the time when the Gestaltist and empiricist approaches to vision were most dominant, the available research tools and methods were not yet what they are today. The empiricists decided to focus only on what was directly measurable, resulting in a theory that was stimulus-driven. In Gestalt psychology, on the other hand, researchers believed that we should not ignore what we could not directly measure and decided that our phenomenological experience should guide our understanding of perception.

Our current methods allow for more and more complex processes to become observable, and the observable contents that were prioritized by empiricists are getting "closer" to the contents of phenomenology that were prioritized by the Gestaltists, and vice versa. We found that some implications of the classic empiricist approach, such as a strictly bottom-up hierarchical visual system, as well as some implications of the classic Gestaltist approach, such as the notion of a psychophysical isomorphism between perceptual organization and the brain events underlying it (Köhler, 1920/1938), have failed. On the other hand, some other implications of the classic empiricist approach, such as the importance of the cortical hierarchy in the bottom-up stream of visual processing, as well as some implications of the classic Gestaltist approach, such as the primacy of Gestalts and global perceptual organizations, have a well-established and still growing body of evidence.

Moreover, as our current research methods allow for a wider scope in both approaches, their *current* interpretations may not have to be considered as directly opposing anymore. Rather, their main focus has been different from the start, with observable elements for the empiricists and phenomenology for the Gestaltists, and can be complementary and integrated in a complete account of visual perception.

One example of a theory that may be considered as an attempt at reconciling both the observable and phenomenology is the Reverse Hierarchy Theory (Hochstein & Ahissar, 2002), which proposes that our instant conscious perception of the "gist" in a scene is the result of a fast bottom-up sweep of processing, and that further analysis of the details of a scene requires a slower interaction of top-down and bottom-up processes.

Another, more recent theory trying to reconcile observable entities and our conscious experience of them was put forward by Herzog et al. (Herzog, Kammer, & Scharnowski, 2016; Herzog, Drissi-Daoudi & Doerig, 2020; Doerig, Schurger & Herzog, 2021). In their theory, Herzog proposes that conscious perception is the result of fundamental, unconscious, discrete and idiosyncratic processing periods of about 400 ms long. More specifically, they argue that our perceptual experience may be continuous, but that this is in fact the result of sequential, discrete windows, called fundamental processing periods, during which processing of visual input occurs unconsciously. Within each of these fundamental processing periods, changes in the external stimulus may occur and be processed unconsciously, but they will only be noticeable in our perceptual experience if these changes are different between the beginning and end of this fundamental processing period. Accordingly, the fact that Gestalts are primary in our perceptual experience, as well as the fact that there is clear evidence for the neural processing of "low-level" stimulus characteristics before "higher-level" stimulus characteristics, is accounted for in the theory. Moreover, the idiosyncratic nature of these fundamental processing periods accounts for individual differences in perception.

This notion can be applicable to the example of apparent motion in this PhD thesis. For instance, consider apparent motion between two dots in two successive frames, where our perceptual experience is characterized by a change in position along the apparent motion path between the two dots, even though the position of the dot in the second frame can only be processed at the end of the apparent motion path. This phenomenon may be accounted for by the notion that our conscious percept is a construction of the summary of visual information that it has processed over a fundamental processing period, during which the apparent motion may have taken place.

Interestingly, the discussion of the distinction between observable (neural) processes and the phenomenon of conscious perception is at the core of multistable perception of ambiguous stimuli, as they allow for a change in conscious perception without a change in the stimulus, therefore resulting only in a change in (neural) processes that underlie the construction of a percept and its conscious experience.

Current dominant research in multistable perception is mostly concerned with the debate about the involvement of "high-level" processes, and the distinction between perception and cognition, but that may raise some questions to be addressed in future research: Why don't

we focus more on the distinction between the construction of a percept and the conscious experience of it? Is this question addressed by focusing on the distinction between different levels in a hierarchy or the distinction between perception and cognition? Is focusing on the distinction between the construction of a percept and the conscious experience of it not feasible because there is no way to know whether a percept is completely constructed if it is not consciously experienced? Does this mean that those two processes are the same?

Origins of variability in perception

Furthermore, we considered phylogenetic (i.e., all development throughout the evolution of a species) and ontogenetic (i.e., all development throughout the lifespan of an individual) development potentially leading up to variability in visual perception.

The focus of Gestalt psychology is mainly on descriptions of perceptual organization which can be predicted by stimulus characteristics (i.e., grouping principles) and which are relatively stable across individuals and over time or experience. The focus of the empiricist approach, and of accounts of perception as 'unconscious inference', on the other hand, is mainly on perception which can be predicted by experiences that an observer has with the external world. Accordingly, Gestaltists found that empiricists minimized the active role of the human observer too much, since perception as a result of experience with the world can be seen in terms of an observer passively taking in the external world without having an inherent determining role in perception. In contrast, it can be argued that the idea of 'grouping principles' that determine perception of a stimulus can be interpreted in terms of a passive observer as well, as the experience of an individual's visual system with the stimulus in the past has no effect on the interpretation of this stimulus.

In a way, we could argue that Gestalt psychology focused on the phylogenetic factors in determining perception, which are relatively stable over time and across individuals, whereas the empiricists focused on the ontogenetic factors in determining perception, which are more variable across and within observers. A focus on either phylogenetic and ontogenetic factors does not exclude the fact that both accounts may partially be reconciled and may both determine perception up to a certain extent.

6.4 Looking back and looking forward

With our reflections and the results of the 'MQ study' chapter in mind, it is important to consider the complexity of the possible interactions of the results of inter- and intra-individual differences in multistable perception in all empirical chapters in this PhD thesis and the specific stimuli we employed.

Aside from the main differences in the blocked design and duration of the experiment, another reason for the fact that we did not replicate the 'shift to global' from the 'Relevance study' in the 'EEG study' may be the color scheme employed to enable the 'global' task manipulation, which incidentally also provided global stimulus grouping cues. With the results of the second chapter, 'EEG study', in mind, it may be interesting to find a way to manipulate the relevance of either percept of the 'local-global' ambiguous motion stimulus without the need for a color scheme that changes the color of some but not all of the dots in the stimulus. It may be worthwhile to examine whether the 'shift to global' in the 'Relevance study' would still occur irrespective of the manipulation of task relevance, or whether this manipulation would have been more influential if this color scheme was not employed. We expect that it would still occur, as previous research studying the perception of the 'local-global' stimulus has also reported a 'shift to global' over time, and did not have the color scheme that we employed in the 'Relevance study'. However, the influence of auxiliary task relevance might be more present, as it cannot be counteracted by an additional influence of stimulus characteristics.

Vice versa, it would be worthwhile to examine the EEG neural correlates for the same experiment as the one we carried out in the 'EEG study' chapter, but with a relevance manipulation or mere exposure across five successive days, as in the 'Relevance study' chapter, between the first and last part of the EEG experiment. As described in the 'EEG study', we expect that the 'shift to global' would occur in this case, and the neural correlates of this change in perception may be worthy of further investigation.

As we failed to adequately study the relationship between individual contrast sensitivity and perception of the 'local-global' stimulus in the 'CSF study', future research may take our considerations into account and examine this relationship, as well as whether the 'shift to

global' in perception of the 'local-global' stimulus is accompanied by a change in the individual's CSF.

The 'local-global' stimuli employed in the first three empirical chapters are different from the MQ stimulus employed in the last empirical chapter. These stimuli differ in the sense that the percepts enabled by the MQ stimulus are identical in "globality", whereas the percepts enabled by the 'local-global' stimulus differ in "globality". As we saw in the 'EEG study', the influence of "globality" in the multistable perception of ambiguous motion stimuli is highly important. It would be interesting to examine how our findings in the first three chapters differ when a MQ stimulus, instead of the 'local-global' stimulus is employed. For the 'Relevance study', we expect that the influence of auxiliary task relevance would be stronger and significant, as the influence of the efficiency of Gestalt quality of simplicity would be equal for percepts that are equal in "globality". In the 'EEG study', we expect that the significant difference of the amplitude for perception of the local compared to the global percept of disambiguated stimuli in the time window of the P200 effect, as well as the significant difference of the amplitude for perception of the local compared to the global percept of ambiguous stimuli in the time window of the P400 effect, would disappear when a MQ stimulus is employed.

Similarly, it would be interesting to study the perception of the 'local-global' stimulus with a wider variety of stimulus characteristics and a more in-depth analysis of individual differences, as we did in the 'MQ study'. Perhaps the intra-individual 'shift to global' in the perception of the 'local-global' stimulus over time would differ depending on the spatial distances between the elements of the local and global percept, as this would also entail that the stimulus is not exclusively studied for characteristics that enable an individual PSE between the local and global percept, and would result in a variability of the ambiguity of this stimulus. Moreover, it may be worthwhile to examine the intra-individual variability in the perception of the 'local-global' stimulus over time, equivalent to the RQA and recurrence plots we employed in the 'MQ study'. Perhaps these findings might reflect an increase in the dominance of the global percept over time by means of an decrease in the intra- and inter-individual variability of perceptual dynamics.

References

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13-21. https://doi.org/10.1038/nrn701

Braddick, O. J. (1980). Low-level and high-level processes in apparent motion. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *290*(1038), 137-151. https://doi.org/10.1098/rstb.1980.0087

Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable perception and the role of the frontoparietal cortex in perceptual inference. *Annual Review of Psychology*, *69*, 77-103. https://doi.org/10.1146/annurev-psych-010417-085944

Block, N. (2016). Tweaking the concepts of perception and cognition. *Behavioral and Brain Sciences*, *39*, 21-22.

Carter, O. L., Pettigrew, J. D., Burr, D. C., Alais, D., Hasler, F., & Vollenweider, F. X. (2004). Psilocybin impairs high-level but not low-level motion perception. *Neuroreport*, *15*(12), 1947-1951.

de-Wit, L. H., Kubilius, J., Wagemans, J., & de Beeck, H. P. O. (2012). Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. *Journal of Vision*, *12*(11): 12. https://doi.org/10.1167/12.11.12

Ding, S., Cueva, C. J., Tsodyks, M., & Qian, N. (2017). Visual perception as retrospective Bayesian decoding from high-to low-level features. *Proceedings of the National Academy of Sciences*, *114*(43), E9115-E9124. https://doi.org/10.1073/pnas.1706906114

Doerig, A., Schurger, A., & Herzog, M. H. (2021). Hard criteria for empirical theories of consciousness. *Cognitive Neuroscience*, *12*(2), 41-62. https://doi.org/10.1080/17588928.2020.1772214

Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision*, *8*(7): 2. https://doi.org/10.1167/8.7.2

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences*, *39*. https://doi.org/10.1017/S0140525X15000965

Geisler, W. S., & Chou, K. L. (1995). Separation of low-level and high-level factors in complex tasks: visual search. *Psychological Review*, *102*(2), 356-378. https://doi.org/10.1037/0033-295X.102.2.356

Gepshtein, S., & Kubovy, M. (2007). The lawful perception of apparent motion. *Journal of Vision*, *7*(8): 9. https://doi.org/10.1167/7.8.9

Herzog, M. H., Kammer, T., & Scharnowski, F. (2016). Time slices: what is the duration of a percept?. *PLoS Biology*, *14*(4), e1002433. https://doi.org/10.1371/journal.pbio.1002433

Herzog, M. H., Drissi-Daoudi, L., & Doerig, A. (2020). All in good time: long-lasting postdictive effects reveal discrete perception. *Trends in Cognitive Sciences*, *24*(10), 826-837. https://doi.org/10.1016/j.tics.2020.07.001

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(5), 791-804. https://doi.org/10.1016/S0896-6273(02)01091-7

Hoffman, D. D., Singh, M., & Prakash, C. (2015). The interface theory of perception. *Psychonomic Bulletin & Review, 22*(6), 1480-1506. https://doi.org/10.3758/s13423-015-0890-8

Joos, E., Giersch, A., Hecker, L., Schipp, J., Heinrich, S. P., Tebartz van Elst, L., & Kornmeier, J. (2020). Large EEG amplitude effects are highly similar across Necker cube, smiley, and abstract stimuli. *PloS One, 15*(5), e0232928. https://doi.org/10.1371/journal.pone.0232928

Koenderink, J. (2011). Vision as a user interface. *Human Vision and Electronic Imaging XVI* (Vol. 7865), 786504-786513. https://doi.org/10.1117/12.881671

Koenderink, J. (2019). Vision, an optical user interface. *Perception, 48*(7), 545-601. https://doi.org/10.1177/0301006619853758

Köhler, W. (1920). *Die physischen Gestalten in Ruhe und im stationären Zustand. Eine natur-philosophische Untersuchung [The physical Gestalten at rest and in steady state].* Braunschweig, Germany: Vieweg und Sohn.

Köhler, W. (1938). Physical Gestalten. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 17–54). London, England: Routledge & Kegan Paul. (Original work published 1920)

Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: when bottom-up and top-down coincide. *Brain and Cognition*, *69*(1), 138-147. https://doi.org/10.1016/j.bandc.2008.06.005

Kornmeier, J., Pfäffle, M., & Bach, M. (2011). Necker cube: stimulus-related (low-level) and percept-related (high-level) EEG signatures early in occipital cortex. *Journal of Vision*, *11*(9): 12. https://doi.org/10.1167/11.9.12

Kornmeier, J., Wörner, R., & Bach, M. (2016). Can I trust in what I see? EEG evidence for a cognitive evaluation of perceptual constructs. *Psychophysiology, 53*(10), 1507-1523. https://doi.org/10.1111/psyp.12702

Le Meur, O., Le Callet, P., & Barba, D. (2007). Predicting visual fixations on video based on low-level visual features. *Vision Research*, *47*(19), 2483-2498. https://doi.org/10.1109/ICIP.2007.4379119

Logothetis, N. K. (1999). Vision: a window on consciousness. *Scientific American*, *281*(5), 68-75.

Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychological Bulletin*, *130*(5), 748-768. https://doi.org/10.1037/0033-2909.130.5.748

Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *249*(1325), 149-155. https://doi.org/10.1098/rspb.1992.0097

Meteyard, L., Bahrami, B., & Vigliocco, G. (2007). Motion detection and motion verbs: Language affects low-level visual perception. *Psychological Science*, *18*(11), 1007-1013. https://doi.org/10.1111/j.1467-9280.2007.02016.x

Rauss, K., & Pourtois, G. (2013). What is bottom-up and what is top-down in predictive coding?. *Frontiers in Psychology*, *4*, 276. https://doi.org/10.3389/fpsyg.2013.00276

Rolfs, M., & Dambacher, M. (2016). What draws the line between perception and cognition. *Behavioral and Brain Sciences*, *39*, 1-77.

Schütt, H. H., Rothkegel, L. O., Trukenbrod, H. A., Engbert, R., & Wichmann, F. A. (2019). Disentangling bottom-up versus top-down and low-level versus high-level influences on eye movements over time. *Journal of Vision*, *19*(3): 1. https://doi.org/10.1167/19.3.1

Sergent, C., & Dehaene, S. (2004). Neural processes underlying conscious perception: experimental findings and a global neuronal workspace framework. *Journal of Physiology-Paris*, *98*(4-6), 374-384. https://doi.org/10.1016/j.jphysparis.2005.09.006

Stein, A. N., & Hebert, M. (2009). Occlusion boundaries from motion: Low-level detection and mid-level reasoning. *International Journal of Computer Vision*, *82*(3), 325-357. https://doi.org/10.1007/s11263-008-0203-z

Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, *13*(7), 310-318. https://doi.org/10.1016/j.tics.2009.04.006

von Helmholtz, H. (1867). *Handbuch der physiologischen Optik: mit 213 in den Text eingedruckten Holzschnitten und 11 Tafeln (Vol. 9)*. Voss.

Wolff, P., & Holmes, K. J. (2011). Linguistic relativity. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 253-265. https://doi.org/10.1002/wcs.104