

# STUDY OF STABLE MOTION PERCEPTION

Dissertation

zur Erlangung des Grades eines Doktors

der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät

und

der Medizinischen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt von

**Qinglin Li**

aus Xi'an, China

2017



Tag der mündlichen Prüfung:

07.12.2017

Dekan der Math.-Nat. Fakultät:

Prof. Dr. W. Rosenstiel

Dekan der Medizinischen Fakultät:

Prof. Dr. I. B. Autenrieth

1. Berichterstatter:

Prof. Dr. Nikos Logothetis/Georgios Keliris

2. Berichterstatter:

Prof. Dr. Uwe Ilg

Prüfungskommission:

Prof. Dr. Nikos Logothetis

Prof. Dr. Uwe Ilg

Prof. Dr. Georgios Keliris

Prof. Dr. Andreas Bartels

Prof. Dr. Annemie Van der Linden

Prof. Dr. Marleen Verhoye



**Erklärung / Declaration:**

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel:

„Study of stable motion perception“

selbständig verfasst, nur die angegebenen Quellen und Hilfsmittel benutzt und wo inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Ich bestätige, dass die falsche Abgabe einer Versicherung an Eides statt mit Freiheit zu drei Jahren oder mit Geldstrafe bestraft wird.

*I hereby declare that I have produced the work entitled "Study of stable motion perception" submitted for the award of a doctorate, on my own (without external help), have marked the sources and aids indicated and have marked passages included from other works whether verbatim or in content, as such. I swear upon oath that these statements are true and that I have not concealed anything. I am aware that making a false declaration under oath is punishable by a term of imprisonment of up to three years or by a fine.*

Tübingen, den 7.12.2017



Datum / Date

Unterschrift /Signature



*To my family,*

*for their endless love and support*





# CONTENT

<b>STUDY OF STABLE MOTION PERCEPTION .....</b>	<b>1</b>
<b>CONTENT .....</b>	<b>0</b>
<b>ABSTRACT .....</b>	<b>1</b>
<b>1. INTRODUCTION .....</b>	<b>2</b>
<b>1.1. Bias and context.....</b>	<b>3</b>
1.1.1. Subjective perception is biased .....	3
1.1.2. Context and individual differences .....	6
1.1.2.1. Modulation of contextual regularity and density .....	7
1.1.2.2. Beyond Global Precedence .....	7
<b>1.2. Motion Perception And Multi-Stability.....</b>	<b>8</b>
1.2.1. Motion integration and segregation .....	8
1.2.2. Multi-Stable Phenomenon and Perceptual Stabilization .....	10
<b>1.3. Neural correlates of post-stimulus motion perception .....</b>	<b>10</b>
1.3.1. Perceptual Dynamics .....	10
1.3.2. Persistent Neural Activity and Offset Response .....	11
<b>2. METHODS .....</b>	<b>13</b>
<b>2.1. Psychophysics and Bayesian modeling.....</b>	<b>13</b>
2.1.1. Subjects .....	13
2.1.2. Apparatus .....	13
2.1.3. Stimuli .....	13
2.1.4. Procedures .....	14
2.1.5. Psychophysics data analysis .....	15
2.1.6. Model Construction .....	15
<b>2.2. High-density laminar electrophysiological recording in monkey V1 .....</b>	<b>17</b>
2.2.1. Animals .....	17
2.2.2. Stimuli .....	17
2.2.3. Procedure .....	18
2.2.4. Apparatus .....	18
<b>3. RESULTS.....</b>	<b>19</b>
<b>3.1. Individual bias and perceptual stability.....</b>	<b>19</b>
<b>3.2. Offset response in awake monkey brain.....</b>	<b>21</b>
<b>4. DISCUSSION .....</b>	<b>24</b>
<b>4.1. Subjective motion perception .....</b>	<b>24</b>
<b>4.2. Dynamic laminar profile of visual motion offset response .....</b>	<b>26</b>
<b>REFERENCES.....</b>	<b>28</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>47</b>



## **ABSTRACT**

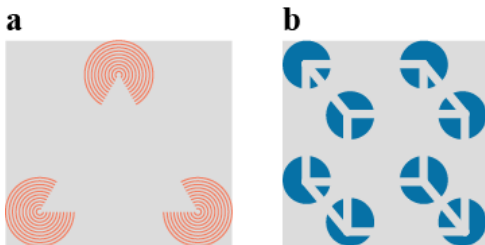
The goal of my PhD project was to combine approaches from human psychophysics, Bayesian modeling and animal electrophysiology to study the mechanisms underlying motion perception. The specific goals of the current project are to: **i)** investigate the relationship between perceptual bias and stabilization with contextual regularity and density in human observers, and develop novel Bayesian models of motion perception that can account for the data; **ii)** explore motion duration dependence of offset neural activity and its layer specificity in primary visual cortex (V1) of alert monkeys.

# 1. INTRODUCTION

*“Our perceptions cannot be mapped directly onto physical reality, and visual perception cannot be said to be a veridical representation of the external world—nor does it try to achieve such a goal. Rather, the goal of the visual system is to support behaviour that is evolutionarily advantageous, and it achieves this through statistically informed perceptual responses to sensory inputs based on previous behavioural outcomes.”*

*-(Soon et al., 2017)*

The world is ever changing, and perception is our primary contact with the outside world (Efron, 1969). The relationship between percepts and the objective environment has intrigued scientists and students for centuries. Sensory organs are thought to be closely related to the physical properties of the environment (Geisler, 2008); however, a long-standing question is whether cognitive system can directly measure such properties (Soon et al., 2017).



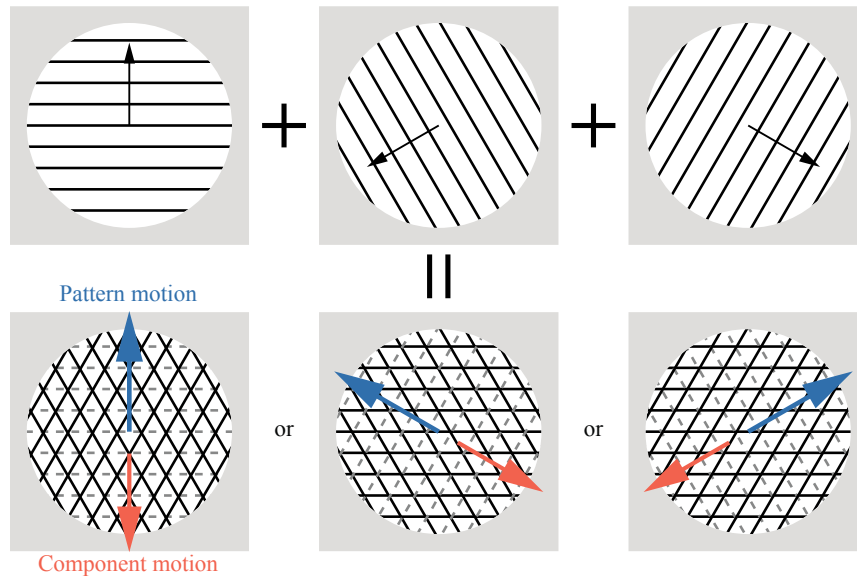
**Figure 2. Illustrations of subjective perception.** (a) Kanisza triangle (Kanizsa, 1955), adapted from (Zentall and Wasserman, 2012); (b) Subjective Necker cube, adapted from (Bradley and Petry, 1977). Both examples demonstrate the subject perception of nonexistent contours.

Sensory inputs are full of incompleteness, ambiguity and noise (Bach and Dolan, 2012; Carbon, 2014; Ernst and Bühlhoff, 2004; Jacobs, 2002; Kersten and Yuille, 2003; Knill and Pouget, 2004; Kornmeier et al., 2014; Rust and Stocker, 2010; Wei and Stocker, 2015). For instance, a retinal image is ambiguously created by the inverted projection of different images or objects onto the retina (Fig.1a) (Palmer, 1999; Pizlo, 2001). Local measurements by small receptive fields suffer from what is known as the aperture problem (Marr and Ullman, 1981; Wallach, 1935; Wuerger et al., 1996). The aperture problem states that direction-selective neurons cannot accurately estimate motion direction, and thus, are expected to respond with ambiguous and noisy one-dimensional motion estimates (Fig.1b). In addition, the brain is inherently noisy (Neri, 2010; Rolls, 2016) as neural responses vary while a stimulus remains constant. Such noise inducing trial-to-trial variability was observed across all levels (Calvin and Stevens, 1968; Faisal et al., 2008; Mainen and Sejnowski, 1995; Rolls, 2016; Sadaghiani et al., 2015; Scaglione et al., 2011; Shadlen and Newsome, 1998). Perception itself also varies on different occasions even stimulus is the same and salient (Supèr et al., 2001) depending on internal brain states (Platt and Glimcher, 1999; Shadlen and Kim, 1999; Thompson and Schall, 1999). Last but not least, what we perceive may not actually physically exist (see Fig.2).

Although the link between perception and physical property seems loose (Carbon, 2014; Soon et al., 2017), it is still hard to believe that perception is not a copy of the environment because the personal feeling of what we perceive is always coherent, seamless and robust. This begs the question - how can we obtain an unambiguous interpretation of the world and reach stable perception? Our brain takes advantage of information beyond sensory inputs from very different sources (Ernst and Bühlhoff, 2004). We do not ‘see’ the world with our eyes, but with our brains. According to Helmholtz, perception is the ‘best guess’ by our visual system, the product of unconscious inference based on the sensory input at hand and prior experience (Helmholtz, 1867). Constructivists proposed that perception is an active and constructive process (Gregory, 1980, 1997; Neisser, 2014); such ideas received quantitative support from Bayesian frameworks (Bach and Dolan, 2012; Kersten and Yuille, 2003; Knill and Pouget, 2004; Knill and Richards, 1996; Pouget et al., 2000; Rao et al., 2002; Weiss et al., 2002). Current conceptualizations of predictive coding state that

sensory inputs are not passively absorbed, but rather our brain actively predicts inputs across a cortical hierarchy (Jehee et al., 2006; Kok and de Lange, 2015; Spratling, 2008; Summerfield et al., 2006).

The question about whether the properties of sensory inputs can be measured perceptually is essentially ill-posed. The function of perception is not to measure, but to make sense of the environment (Malcolm et al., 2016), as guided by specific goals to ensure survival. Such goals may be developed through evolution (Goldstein, 2009), or from learning procedures (Chalupa and Werner, 2004). We should first ask why we perceive, instead of what or how we



**Figure 3.** *Cartoon version of tri-stable plaid, adapted from (Andrews and Schluppeck, 2000).* The stimulus consists of three superimposed square-wave gratings, separated from one another by an angle of 120 degrees, which moved at constant speed. At any given time, two of the components appear to group together into a moving ‘plaid’ pattern, which moves in the direction of the intersection of velocity constraint lines imposed by each individual component. At the same time, the remaining component appears to drift in the opposite direction. However, which of the three components group together appears to spontaneously shift over time.

perceive.

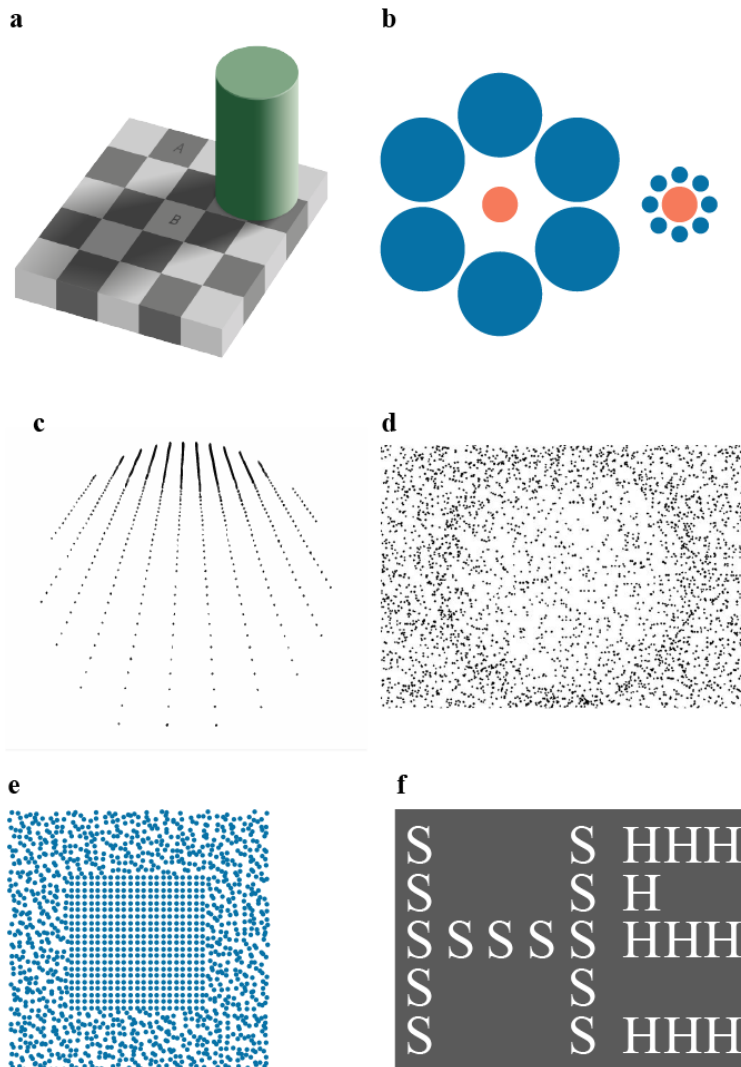
## 1.1. BIAS AND CONTEXT

### 1.1.1. SUBJECTIVE PERCEPTION IS BIASED

Perception is inherently a biased procedure influenced by information from various sources, which leads to perceptual errors and illusions (Brosch et al., 2010; Girshick et al., 2011; Graf et al., 2004; Haselton et al., 2015; Pessoa and Adolphs, 2010). An interesting historical phenomenon is that perception researchers have been more intrigued by biased perceptions than veridical or robust ones, whereby perceptual errors were described before basic perceptual processing was fully understood (Haselton et al., 2015). It was believed that perceptual bias provides a unique window to the underlying mechanisms of perception (Zoccolan et al., 2015).

Our perception can be biased by prior experiences, as perceptual abilities have a strong reliance on previous knowledge in front of ambiguity (Carbon, 2014; Knill and Pouget, 2004; Körding and Wolpert, 2004; Seriès and Seitz, 2013). Ample priors have been well described, for instance, the low-speed prior (see Fig.1b. Stocker and Simoncelli, 2006; Weiss et al., 2002); the light-from-above prior (see Fig.1c. Sun and Perona, 1998), and cardinal-orientation prior, also known as the oblique effect (see Fig.1d. Appelle, 1972; Girshick et al., 2011; Tomassini et al., 2010). Such priors

are thought to be learned from the properties of the environment (Girshick et al., 2011; Lewis et al., 2012; Ramachandran, 1988; Styles, 2005). According to the Bayesian approach, bias is represented in the form of a prior distribution, which is affected by the uncertainty of the stimulus (de Gardelle et al., 2010; Girshick et al., 2011; Heekeren et al., 2004; Morales et al., 2015; Putzeys et al., 2012; Tomassini et al., 2010; Watson et al., 2016; Webb et al., 2010; Wei and Stocker, 2015). The mediation of prior knowledge on perceptual bias is thought to be related to attention



**Figure 4. Illusions of contextual modulation.** (a) Adelson squares, also known as the checker shadow illusion (Adelson E. H.,1995). Areas A and B have identical brightness, but B is perceived as darker than A under the shadow context; (b) Ebbinghaus illusion, also known as Titchener Circles (Ebbinghaus, 1905). The size of the two orange circles is the same, but the right one appears larger than the left; (c) Perceiving depth and slant from the density dots pattern, adapted from (Braunstein and Payne, 1969); (d) Perceiving surface from density dots pattern, adapted from (Singer et al., 2008); (e) Pop-out effect of regularity, adapted from (Ouhana et al., 2013). (f) Navon's letters, adapted from (Navon, 1977).

and expectations. In general, the priority of a stimulus representation may be selectively affected by attention, and the interpretation of sensory information may be constrained by expectations (Schwarz et al., 2016a; Summerfield and Egnér, 2009; Sussman et al., 2016a). In electrophysiology studies, modulations of prior experience were found to reduce neural activity in sensory areas (Alink et al., 2010; Gong et al., 2017; Kok et al., 2012; Summerfield et al., 2008; Todorovic et al., 2011), suggesting that perceptual bias is mediated by the prior experiences of sensory representations. Interestingly, prior experiences are not constant or hard-wired (Kok et al., 2013; Seriès and Seitz, 2013), but interactively and dynamically adapted by the changing environment with high flexibility (Adams et al., 2004; Chalk et al., 2010; Roach et al., 2017; Sotiropoulos et al., 2011). For example, one recent study reported that the interaction of multiple priors affects perceptual stability with bi-stable stimuli. While compatible priors facilitate perceptual stabilization, conflict priors lead to enhanced rivalry (Zhang et al., 2017).

Cognition has limited computational capacity (Duncan, 1980; Fisher and L., 1982; Lavie, 1995; Motter, 1993; Palmer et al., 2011; Prinzmetal and Banks, 1983; Schneider and Shiffrin, 1977; Shiffrin and Gardner, 1972;

Spoehr and Lehmkuhle, 1982; Tsotsos, 1997, 1990; Yantis, 2001). Sensory inputs compete for neural representations within a mutually inhibitory network, and such competition is biased by attention towards the most relevant information, which is also known as the theory of biased competition (Desimone, 1998; Desimone and Duncan, 1995; Reynolds et al., 1999). Neurophysiological data confirmed that neural responses are biased in favor of attended objects, features or locations (Martinez-Trujillo and Treue, 2004; Rensink, 2000; Reynolds and Chelazzi, 2004; Reynolds et al., 1999; Yantis, 2001). Attentional influences on biased competition are implemented not only by top-down modulation with specific goals (Serences and Yantis, 2006; Sussman et al., 2016a) but also by bottom-up signals concerning the physical properties. For instance, it has been shown that neural activity is biased toward salient stimuli, with increasing neural activity (Reynolds and Chelazzi, 2004; Reynolds and Desimone, 2003; Serences et al., 2004). Recent studies on affective neuroscience have proposed a theory called the arousal-biased competition theory. According to this theory, a stimulus containing emotional cues increases arousal with enhanced levels of glutamate and norepinephrine in the locus coeruleus (LC), which consequently biases perception toward emotionally relevant stimuli (Cisler and Koster, 2010; Mather and Harley, 2016; Mather and Sutherland, 2011a; Mather et al., 2016; Sussman et al., 2016a, 2016b, 2017).

Prior knowledge and experience set specific expectations infer biased perceptual decisions (Schwarz et al., 2016b; Seriès and Seitz, 2013; Stocker and Simoncelli, 2006; Zelano et al., 2011). Accumulating evidence reveals that the processing of sensory inputs begins even before physical contact with the stimulus, suggesting a preparatory bias based on predictive sensory representations (McMains et al., 2007; Mesulam, 2008; Neisser, 2014; Sylvester et al., 2009). Researchers have shown that an expected stimulus is associated with higher perceptual sensitivity and shorter response latency (Polat and Sagi, 1994; Summerfield and Egner, 2009). The underlying mechanisms are recently thought to be associated with the state of pre-stimulus brain connectivity (Leske et al., 2015; Sadaghiani et al., 2015). Predictive coding theories state that sensory inputs are not passively processed (Jehee et al., 2006; Kok and de Lange, 2015; Spratling, 2008). A pre-stimulus predictive template is actively formed against the observed sensory information (Freeman, 1981; Friston, 2003; Kok et al., 2017; Mumford, 1992; Peelen et al., 2009; Summerfield et al., 2006; Zelano et al., 2011). Although the assumptions of predictive coding and biased competition are distinct, it is interesting to note the computational principles are equivalent (Spratling, 2008). Our perception can be biased not only by prior experiences but also by contextual modulation (Soon et al., 2017). Substantial evidence has revealed that perceptual bias is related to contextual information (Carandini and Heeger, 2011; Kapadia et al., 1995; Kohn, 2007; Levitt and Lund, 1997; Li, 1999; Louie et al., 2013; Ohshiro et al., 2011; Schwartz et al., 2007). Details about contextual modulation concerning subjective perception are discussed in the next section.

Bias is well-known as the default procedure in perception (Tomassini et al., 2010), which affects the direction and content of what we perceive (Bocanegra et al., 2012). An important question here is whether perceptual bias is a flaw of the visual system? Intuitively, bias presents a challenge because accuracy and precision are expected via evolution. A recent line of research revealed that we are not arbitrarily biased, and that bias is essentially a design feature of our cognitive system with evolutionary advantages (Haselton et al., 2015; Soon et al., 2017). For instance, it has been found that human perception is highly dependent on specific statistical properties of the natural environment (Geisler, 2008), the quantitative estimation of the cardinal-orientation prior the human observer matches to the statistic distribution of cardinal contours in natural environment (Girshick et al., 2011), indicating that perceptual bias is naturally adapted by taking advantage of stable properties of the environment (Ramachandran, 1988).

Perceptual advantages of bias have already been reported. For instance, *Bias facilitates orientation discrimination*. In orientation discrimination tasks, the discriminability of human subjects is best at biased cardinal directions (Appelle, 1972; Girshick et al., 2011; Tomassini et al., 2010), and with less variance (Girshick et al., 2011). Bias also resolves motion ambiguity; for example, Andrews and colleagues (Andrews and Schluppeck, 2000) designed a complex stimulus containing three superimposed moving gratings (see Fig.3). The gratings moved in different

directions with 120° difference, and one of them was always either vertical or horizontal. The perception of such stimulus by brief presentation was always a coherent pattern of two gratings moving in one direction, and the third grating moving in the opposite direction independently. Such stimulus represented an ambiguous situation since the combination of each two gratings were equally plausible. However, human subjects preferentially integrated two oblique gratings moving along cardinal axes, suggesting that motion ambiguity can be resolved by the cardinal directions to which they were predominantly biased. *Bias guides learning procedure.* Hutchinson and colleagues revealed that human memory systems are biased toward new information with enhanced encoding activity, suggesting that the learning procedure is guided by prior experience towards new aspects of the environment (Hutchinson et al., 2016). *Bias prioritizes threatening information.* In affective neuroscience, bias is often described as perceptual prioritization (Sussman et al., 2016a). It has been found that threatening stimuli have a higher processing priority with shorter reaction times and greater accuracy than non-threatening stimuli (Abdel Rahman, 2011; Anderson, 2005; Correll et al., 2011; Fox et al., 2002; Jusyte and Schönberg, 2014; Trawalter et al., 2008; Zeelenberg et al., 2006). These have been thought to be related with enhanced neural activity in fear-related brain areas (Bernier et al., 2017; Bishop, 2009; Etkin et al., 2004; Larson et al., 2005). This biased processing facilitates immediate reaction in the face of danger, which is important for survival and reproduction (Brosch et al., 2010; Haselton et al., 2015; New et al., 2007; Soon et al., 2017; Sussman et al., 2016a).

There is no doubt that our visual system is shaped by evolution (Goldstein, 2009). Inspired by some other cognitive phenomenon, such as central tendency (perceptual judgement tends toward the mean of the stimulus set; Vierordt, 1868 ), perceptual narrowing (perceptual discrimination narrows with experience during development, Hollingworth, 1910), and the pre-constancy period (infants younger than 5 months are sensitive to variant information, such sensitivity is gradually lost during development; Yang et al., 2015), we assume that bias itself may act as a functional filter, which actively coordinates the visual system with the environment promoting survival. Less important or unnecessary sensory inputs are filtered out from perception, leading cognitive systems to focus on the most relevant information.

#### 1.1.2.CONTEXT AND INDIVIDUAL DIFFERENCES

Visual information is rarely found in isolation (Bar, 2004; Krause and Pack, 2014). Our perception depends not only on the stimulus itself but also other stimuli which are presented simultaneously. Identical stimuli can be perceived differently in distinct contexts (Coen-Cagli et al., 2015). In general, the approach of contextual modulation deals with the interactions between the target stimulus and environment to explain how the target can be attributed from its context. The effects of contextual modulation have been described by studies on visual illusions and other phenomena (Carandini and Heeger, 2011; Kapadia et al., 1995; Kohn, 2007; Levitt and Lund, 1997; Li, 1999; Louie et al., 2013; Ohshiro et al., 2011; Schwartz et al., 2007). See some famous illustrations in Fig.4.

In recent years, the population-receptive-field (pRF) model provided a powerful method for detailing the precise population properties of cortical areas (Amano et al., 2009a; DeSimone et al., 2015; Dumoulin and Wandell, 2008; Victor et al., 1994). Using the pRF retinotopic mapping technique, recent studies have revealed that subjective perception is associated with various cortical properties, such as pRF size (Alvarez et al., 2015; Harvey and Dumoulin, 2011; Moutsiana et al., 2016; Schwarzkopf et al., 2014; Song et al., 2013a); cortical thickness and cortical surface area (Bergmann et al., 2016a, 2016b; Genç et al., 2015; Harvey et al., 2013; Schwarzkopf et al., 2011; Song et al., 2015); cortical magnification factor (Harvey and Dumoulin, 2011) and intra-cortical connectivity (Song et al., 2013b). Inter-subject variances of perceptual bias and the magnitude of contextual modulation are related to such functional architectures. For instance, it is already known that individual differences of the retinotopic surface of early visual areas are larger than that of cortical size (Dekaban and Sadowsky, 1978; Dougherty et al., 2003). Using an orientation



discrimination paradigm with center-surround contextual illusion, Song and colleagues found that subjects with larger surface area of primary visual cortex (V1) were less affected by contextual modulation, but more sensitive to tiny orientation differences (Song et al., 2013a). Similar results were found in object size judgement tasks (Delboeuf illusion), where identical stimuli were perceived as small with larger pRFs in V1 (Moutsiana et al., 2016).

#### 1.1.2.1. Modulation of contextual regularity and density

Evidence shows that humans exploit the statistical properties of the natural environment (Biederman, 1972; Biederman et al., 1973; Davenport and Potter, 2004; Geisler, 2008; Girshick et al., 2011; Joubert et al., 2007; Palmer, 1975; Ramachandran, 1988; Stansbury et al., 2013). Different from other statistical features, it is easy to notice objects or elements that are repeatedly or periodically distributed across the visual environment, a phenomenon known as regularity (Cecchetto and Lawson, 2017; Lin et al., 2006; Ouhana et al., 2013), see Fig.4e. Our perception of regular patterns is invariant to different viewing-angles or luminance conditions (Chetverikov, 2000). The human visual system possesses internal templates for regular patterns, indicating that regularity is a coded feature in human vision (Morgan et al., 2012; Ouhana et al., 2013). Attention is spontaneously biased toward regularities (Yu and Zhao, 2015; Zhao et al., 2013). Perceptually, regularly spaced elements tends to be grouped together as a coherent texture (Saarela et al., 2010), which has been claimed as a novel Gestalt grouping principle (van den Berg et al., 2011).

Additionally, percepts may be modulated by the surrounding density (see Fig4.c and d). It has been found that contextual density is associated with perceptual depth and slant (Cutting and Millard, 1984; Todd and Thaler, 2010), and that it affects surface integration and segmentation (Durgin, 2001; Durgin and Hammer, 2001). Researchers suggested that density is involved in early visual processing (MacKay, 1973) via an independent channel (Sun et al., 2016). However, the relationship between regularity and density is still unknown.

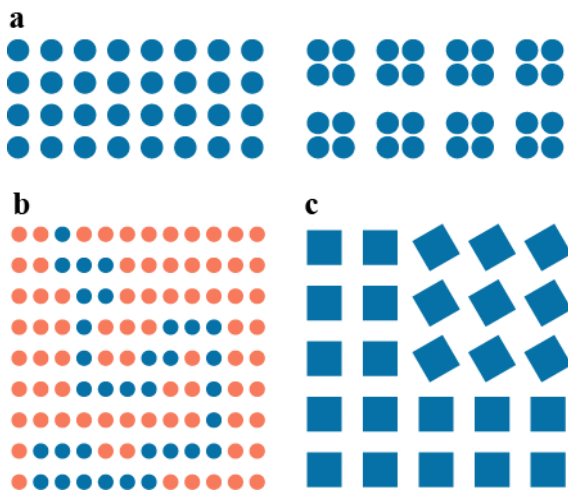
#### 1.1.2.2. Beyond Global Precedence

*“‘Gestalt.’ It is famously hard to translate the term into English—but also the German native speakers among the authors of this overview paper are not quite sure what exactly the term means in German. Probably the best—but incomplete—translation is still “configuration.” This translation misses the holistic aspect that was essential for the Gestalt psychologists’ early writings and that survives in the famous dictum: ‘the whole is something else than the sum of its parts’ [Koffka, 1935].”*

*-(Jäkel et al., 2016)*

Gestalt approaches elegantly describe the basic principles of perceptual organization (see Fig.5 for examples of Gestalt laws). Different from the assumption of associationism, which affirms that perception is a result of associating component sensations, Gestalt theories emphasize that global processing is different from the sum of the local analysis (Koffka, 1935; Wertheimer, 1912, 1923).

The principle of global precedence assumes that global configuration is prioritized over local feature analysis (Boff et al., 1986; Herzog and Clarke, 2014; Navon, 1977; Rider et al., 2016; Seriès and Seitz, 2013). In line with this hypothesis, researchers revealed that global stimulus structure, not local contextual manipulation, predicts perception and behavioural performance (Livne and Sagi, 2007; Manassi et al., 2012; Sayim et al., 2010). Global precedence theory does not state that local processing starts after the termination of global analysis, but that global features are aligned more closely to perception (see Fig.4f, Navon, 1981). Similarly, the reverse hierarchy theory assumes that local features are unconsciously processed by lower brain areas, and can only be consciously retrieved on demand after



**Figure 5. Gestalt principles of perceptual organization, adapted from (Wagemans et al., 2012).** (a) Proximity: objects close to each other appear to be in the same group; (b) Continuity: perceptual tendency to create a continuous object; (c) Common fate: objects with a common orientation will be perceptually grouped together.

global information has been perceived by higher brain areas (Hochstein and Ahissar, 2002). Recent studies have confirmed that conscious perception is dominated by global information (Campana et al., 2016; Haberman and Whitney, 2011; Parkes et al., 2001). Indeed, Gestalt theories not only assume that global processing has higher priority but also that local processing is constrained by global configuration. Wertheimer claimed that “the whole determines the appearance of its parts”. These ideas have been confirmed by prevalent predictive coding approaches:

*“The mapping from low- to high-level representation (e.g. from acoustic to word-level) is computed using the reverse mapping, from high- to low-level representation.”*

- (Chater and Manning, 2006)

Global perception is not formed by accumulating information from lower levels (Stokes et al., 2014). Instead, our brain actively predicts the lower level information with potential assumptions

(Bar, 2007; Chater and Manning, 2006; Friston, 2009; Friston et al., 2011).

## 1.2. MOTION PERCEPTION AND MULTI-STABILITY

### 1.2.1. MOTION INTEGRATION AND SEGREGATION

The visual cortex of primates contains a number of hierarchically and recurrently connected areas with retinotopic organization achieved by spatially distributed receptive fields (RFs) (Van Essen and Maunsell, 1983; Felleman and Van Essen, 1991; Van Essen et al., 1992). These RFs have variable sizes both within an area (increasing from the center of the visual field to the periphery) as well as across the visual hierarchy starting from very small RFs in primary visual cortex (V1) and gradually shifting to larger ones in higher areas (Freeman and Simoncelli, 2011). It has been speculated that this structure decomposes the visual field into high-resolution feature selective maps that are later integrated into global entities leading to our uniform perception of the world (Braddick, 1993; Maunsell and Newsome, 1987).

Plaid patterns - stimuli composed of two superimposed drifting gratings - have been very popular for the study of motion integration because of their ability to be integrated into a single coherent moving pattern or be segmented into the two underlying grating components, moving transparently over each other (Adelson and Movshon, 1982). Local measurements with small receptive fields induce an ambiguous and noisy one-dimensional motion estimation. These problems can be solved by the selective integration and pooling over time and space to reconstruct a global two-dimensional pattern (Adelson and Movshon, 1982; Braddick, 1993; Smith et al., 2005). Typically, the model for motion analysis consists of two stages. The first stage is the local motion detection of single gratings (Adelson and Bergen, 1985; Gizzi et al., 1990; Limb and Murphy, 1975; Movshon et al., 1985) in V1, with the second one integrating local motion signals from the previous stage to establish the direction of the entire pattern believed to exist in MT (middle temporal area, Movshon et al., 1985) and MST (medial superior temporal area, Khawaja et al., 2009).

Researchers found that a variety of attributes of the component gratings, such as, speed (Derrington and Suero, 1991; Kooi et al., 1992), direction (Burke and Wenderoth, 1993) and contrast (Stone et al., 1990), affect pattern

motion perception. Moreover, pattern motion response is delayed relative to component motion selectivity by approximately 100 ms (Born et al., 2010; Pack et al., 2001; Smith et al., 2005, 2010). These results suggest that pattern motion estimation is performed in higher brain areas, but not in V1, which supports the two-stage model of motion perception. However, an alternative view claims that pattern motion is highly dependent upon the local plaid features, instead of the combination of its motion components (Delicato and Derrington, 2005). Indeed, electrophysiological studies have already found neurons with pattern motion selectivity in early brain areas, such as, the V1 of macaques (Guo et al., 2004), marmosets (Tinsley et al., 2003) and mice (Juavinett and Callaway, 2015). Additionally, it has been proposed that simple bottom-up mechanisms, which do not consider stimulus structures, are insufficient for motion integration, rather this integration is reliant upon complex perceptual organization (Bruno and Bertamini, 2014).

Plaid patterns can also be perceived transparently as two component gratings sliding across each other. Motion transparency with plaid stimuli is simply taken as the failure of integration (Nishida, 2011). Alternatively, the mechanism of directional repulsion, which overestimates the directional difference between two motion signals, may play an important role in the processing of motion transparency (Marshak and Sekuler, 1979). However, mechanisms of surface segmentation were claimed to be involved in the determination of whether motion information is integrated or segmented (Stoner and Albright, 1996; Trueswell and Hayhoe, 1993). Generally, pattern direction can be directly derived from the locations where the two superimposed gratings intersect ('blob' or 'corner'), and the human visual system is more sensitive to the intersections of visual inputs among other features (Bergen and Julesz, 1983; Julesz, 1981). The luminance of intersection affects figure-ground segmentation of plaid patterns, which helps the dissociation of plaid into separate bars (Stoner et al., 1990). It is well-accepted that motion transparency depends crucially on spatiotemporal manipulations using random-dots (Braddick, 1997; Vidnyánszky et al., 2002). Two sets of randomly located dots moving in different directions would induce a robust transparent motion perception depending on the directional difference, and locally-paired dots could interrupt such transparent perception into a directionless flicker when the paired dots are moving towards their partners (Qian et al., 1994) Nevertheless, if the paired dots are not targeted to each other and move in different directions, for example, moving orthogonally to each other, then the resulting perception is coherent (Curran and Braddick, 2000; Matthews et al., 2000).

Previous studies have revealed that the visual system may have a preference for pattern motions (Hedges et al., 2011). Human subjects always perceive coherent motion after transparent motion adaptation (integrated motion aftereffect), unless two component motion patterns are distinctively different from each other (Alais et al., 2005; Curran et al., 2007; Verstraten et al., 1994, 1999). Bidirectional motion signals on each position are integrated locally to produce local unidirectional aftereffects, such effects are then integrated over space in higher brain areas (Curran and Benton, 2006; Lee and Lu, 2012; López-Moliner et al., 2004; Vidnyánszky et al., 2002). Researchers have indicated that precise local information is retained during the motion integration procedure (Lee and Lu, 2012; Scarfè and Johnston, 2011), which is important for motion feature representations (Braddick, 1993). For example, it has been found that the integrated motion aftereffects are weak for non-adapted locations with multi-aperture stimuli (Lee and Lu, 2012).

Despite the extensive use of plaid stimuli over the past decades in this field (Movshon et al., 1985), the underlying mechanism of motion coherence and transparency of such stimuli remains unclear. Importantly, it has been shown that the luminance of intersections within a plaid pattern could bias our perception, and it would be more chance to perceive coherent motion if the intersections are too bright or too dark (Stoner et al., 1990; Thiele and Stoner, 2003). But, Noest and van den Berg have already investigated how local manipulations of plaid stimuli influence coherent and transparent perception (Noest and van den Berg, 1993).

### 1.2.2. MULTI-STABLE PHENOMENON AND PERCEPTUAL STABILIZATION

At each moment, our visual system is challenged by an immense amount of information to process. However, only a small fraction of the inputs reaches our awareness. When our brain fails to reach a single interpretation of these inputs from the environment, perceptual multi-stability occurs. The issue of multi-stable phenomenon has intrigued scientific minds for decades. This phenomenon has been found in vision (Leopold and Logothetis, 1999), audition (Repp, 2007), olfaction (Repp, 2007; Zhou and Chen, 2009) and speech (Warren and Gregory, 1958). Crucial features are similar within and across modalities (Schwartz et al., 2012). For the visual modality, a number of ambiguous visual patterns have been described, such as, the Necker cube and binocular rivalry. Ambiguous perception can be evoked by static visual stimuli, and also by motion patterns, such as, moving plaids. Multi-stable stimuli provide unique insight into visual processing, as changes in perception are decoupled from changes in the stimulus. Thus, understanding multi-stable perception helps us to understand visual perception in general. Investigators suggest that both top-down and bottom-up processes are involved in multi-stable perception (García-Pérez, 1989). Others argue that high-level processing is not necessary, happening automatically in the form of low-level competition between the stimulus features (Akman et al., 2009; Wilson et al., 2000). Further, the changes in stimulus features can bias perception in one or another direction (Klink et al., 2012). The percentage of neurons modulating their activity in response to perceptual changes increases hierarchically along the visual information pathway. At early stages of visual processing (V1/V2), only 20% of neurons were found to exhibit perceptual modulations (Keliris et al., 2010; Leopold and Logothetis, 1999). Moreover, our group has demonstrated that this finding is also consistent with the level of perceptual modulation of the local field potentials (Keliris, et al., 2010). Neuronal activity related to perception increases to 40% in areas V4 and MT (Leopold and Logothetis, 1999) and reach 90% in the inferior-temporal cortex (IT) (Sheinberg and Logothetis, 1997).

Previous research has found strong evidence for perceptual stabilization mechanisms in the visual system; such as, the reorganization of sensory information during intermittent viewing (Leopold et al., 2002), top-down modulation of beta-band synchronization (Kloosterman et al., 2015), feedforward inhibition (Bollimunta and Ditterich, 2012), arousal (de Gee et al., 2014; Mather and Sutherland, 2011a), memory (Wimmer and Shohamy, 2012), and the interaction of multiple priors (Zhang et al., 2017). For instance, perception of a successive stimulus could be impaired if post-stimulus processing was not already finished for the first stimulus (Keysers and Perrett, 2002). The rate of perceptual switching can be stabilized by periodically removing the stimulus during binocular rivalry (Leopold et al., 2002). Research has proposed an offset transient suppression mechanism to explain the observation that ongoing processing is suppressed by offset transient neural activity (de Graaf et al., 2017). Another recent study indicated that the interaction of multiple priors affects perceptual stability with bi-stable stimuli: compatible priors facilitate stabilization, while conflict priors lead to enhanced rivalry (Zhang et al., 2017). It is well-known that motion perception could be influenced by context, such as, cross-orientation suppression (Heeger et al., 1997), surround suppression, opponent-direction suppression (Qian et al., 1994), and induced motion (Duncker, 1929). A recent study reported that coherent and transparent perception can be biased by the surrounding motion signals (Takemura et al., 2011), indicating the interactive modulation of nearby signals may potentially contribute to stabilizing neuronal representations.

### 1.3. NEURAL CORRELATES OF POST-STIMULUS MOTION PERCEPTION

#### 1.3.1. PERCEPTUAL DYNAMICS

The primate visual system is highly dynamic. Ongoing neural activity fluctuates even when sensory inputs are constant (Sadaghiani et al., 2015). Recent studies have proposed that perception is a continuous, rather than a discrete event, and that perceptual function is periodically modulated by specific brain rhythms (Ainsworth et al., 2012; Fries, 2015). In order to coherently perceive the environment as stable in space and time, ongoing processing is required to balance the

internal representations and to accumulate sensory evidence (Murphy et al., 2015). The perceptual system roughly and rapidly estimates initial sensory inputs, and then gradually refines this estimation (Born et al., 2010). Internal representations are continuously updated in response to perceptual modulation (Ma et al., 2014; Nakamura and Colby, 2002; Tajadura-Jiménez et al., 2016).

Nevertheless, systematic perceptual errors were observed from human subjects when a stimulus was presented briefly (Lorenceanu et al., 1993), indicating that the initial estimation is not veridical. During binocular rivalry, onset rivalry may be strongly biased, and its underlying mechanism was recently found to differ from sustained rivalry by stochastic processing (Attarha and Moore, 2015; Carter and Cavanagh, 2007; Stanley et al., 2011). Similar temporal evolution has been confirmed by eye-movement studies in the motion domain. For instance, substantial evidence has shown that initial eye-movements do not reflect the direction of pattern motion, rather the components (Born et al., 2006; Masson and Castet, 2002; Masson et al.; Pack and Born, 2001). The ongoing processing is also associated with the inherent noise of the neural system (Neri, 2010; Rolls, 2016). Thus, the observed trial-to-trial variability is thought to be related to the dynamic state of brain activity (Arieli et al., 1996; Platt and Glimcher, 1999; Shadlen and Kim, 1999; Thompson and Schall, 1999).

### 1.3.2.PERSISTENT NEURAL ACTIVITY AND OFFSET RESPONSE

An increasing number of studies have confirmed that stable perception is not only affected by ongoing visual processing but also by the pre-stimulus brain state (Jehee et al., 2006; Jensen et al., 2012; Kayser et al., 2016; Kok and de Lange, 2015; Leske et al., 2015; Lou et al., 2014; Sadaghiani et al., 2015; Spratling, 2008), which is influenced by perceptual history (Brascamp et al., 2008; Jellema and Perrett, 2003; Leopold et al., 2005; St John-Saaltink et al., 2016) and expectations (Freeman, 1981; Friston, 2003; Mumford, 1992; Peelen et al., 2009; Summerfield et al., 2006; Zelano et al., 2011).

Relevant neural processing continues even when the stimulus is removed from presentation; such post-stimulus processing with persistent neural activity has been identified in various brain areas (Akaike, 1977; Andersen et al., 1987; Funahashi et al., 1989; Fuster and Alexander, 1971; Levick and Zacks, 1970; Mendoza-Halliday et al., 2014; Meyer et al., 2007; Miller et al., 1996a, 1996b; Schiller, 1968; Tark and Curtis, 2009). Just this year researchers have drawn great attention to the relationship between working memory and neural persistence (Bolkan et al., 2017; Guo et al., 2017; Schmitt et al., 2017; van Kerkoerle et al., 2017), revealing the maintained neural activity in thalamocortical circuit in the absence of a stimulus plays an important role in working memory (Bray, 2017). Indeed, it has been proposed that post-stimulus neural activity is also associated with conscious access (Babiloni et al., 2006; Dehaene et al., 2006; Pun et al., 2012; Sergent and Dehaene, 2004), aftereffect (Huang et al., 2008; Li et al., 2017), visible persistence (Coltheart, 1980a, 1980b; Duysens et al., 1985; Di Lollo, 1980; Sperling, 1960), and can bias competing neural presentations to reach a stable percept (Keysers and Perrett, 2002).

Whether previous visual information can be maintained during post-stimulus processing in early brain areas (Pasternak and Greenlee, 2005) is currently a controversial topic. It has been suggested that visual information can be stored in working memory only at higher brain areas (Mendoza-Halliday et al., 2014). For example, sustained neural activity during the post-stimulus period was observed in MST and lateral prefrontal cortex (LPFC), but not in MT, suggesting that only higher brain areas are involved in working memory (Mendoza-Halliday et al., 2014). On the other hand, post-stimulus neural modulation was found to be strongly correlated with previous stimulus information in V1 (Super et al., 2001). One recent study has revealed that neural modulation of working memory was stronger in the superficial and deep layers of V1 (van Kerkoerle et al., 2017). These studies revealed that V1 contributes not only to sensory coding but also to post-stimulus memory activity.

Neurons from different brain areas and species can respond to the termination of stimulation with transient spiking activity (Bair et al., 2002; Duysens et al., 1996; Hartley et al., 2011; He, 2002; McLelland et al., 2009; Qin et al., 2007). Surprisingly, being the first state of post-stimulus neural activity, little attention have been given to the visual offset responses since the first study in 1927 (Adrian and Matthews, 1927). A handful of studies on anesthetized animals has demonstrated offset responses to static stimuli in V1, which are orientation-tuned and more accurate than onset responses with shorter latency and less variability (Bair et al., 2002; McLelland et al., 2010) suggesting that offset response is informative. However, to our knowledge no research has reported visual motion offset responses in behaving animals. The signature of motion offset responses in the alert brain is still unknown along with whether visual motion information is maintained, with layer specificity, during offset responses in V1.

## 2. METHODS

### 2.1. PSYCHOPHYSICS AND BAYESIAN MODELING

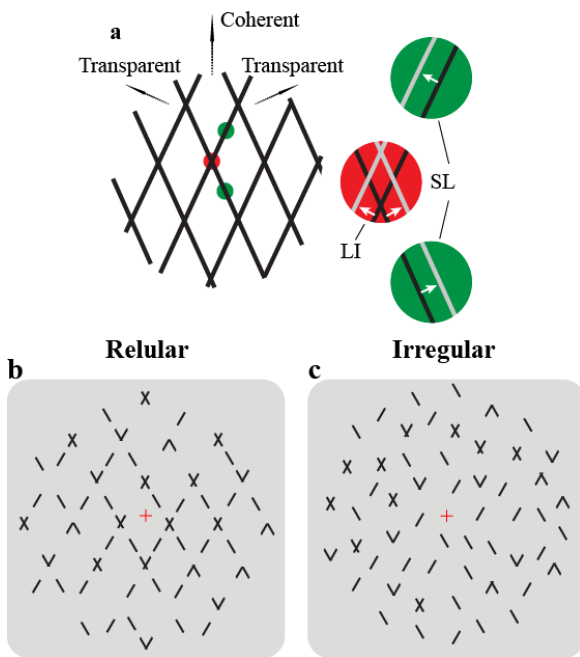
#### 2.1.1. SUBJECTS

Five subjects (4 females) participated in all of the experiments, four of whom were naïve to the aims of the study. All of them had either normal or corrected-to-normal vision. The study was approved by the ethical committee of the University of Tuebingen. Before participating in the experiment, written consent was obtained from each subject.

#### 2.1.2. APPARATUS

The experiments were performed in a dimly lit room. The stimuli were programmed using MatLab Psychophysics Toolbox (Brainard, 1997) and presented on a 17-inch CRT monitor with a resolution of  $1280 \times 1024$  and a refresh rate of 100 Hz. The monitor was gamma corrected with a mean luminance of  $15.6 \text{ cd / m}^2$ . The distance from the eyes of subject to the monitor was 43 cm. Subject responses were collected using a custom 2-button response box (see Procedures). The eye movements were monitored continuously using an infrared eye tracker (iView X™ Hi-speed, SMI).

#### 2.1.3. STIMULI

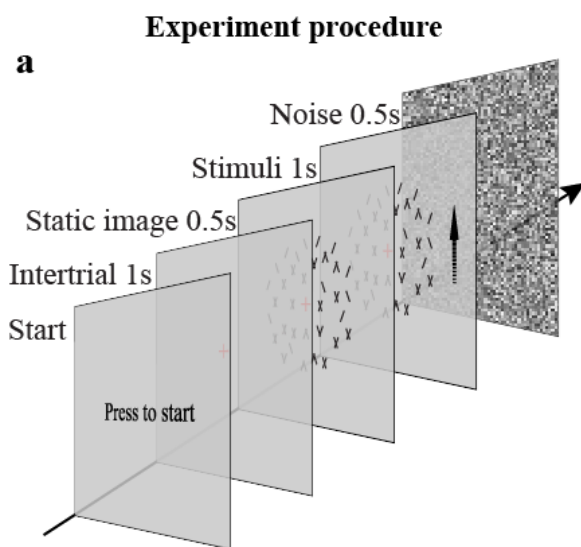


**Figure 6. Stimuli.** (a) Illustration of moving stimuli. The line-plaid (composed of two overlaying drifting line gratings) can be distinguished as two different local inputs, namely LI (red) and SL (green). The line locations at time  $t$  and  $t + \Delta t$  were plotted in black and gray, respectively, for each aperture; (b and c) Cartoon versions of regular and irregular stimuli. Note that in this cartoon all possible apertures are shown, while our experiments randomly selected a number as defined by the density.

The novel pseudo-line-plaid stimuli in this study were designed to mimic the standard upward-moving line-plaid by decomposing pattern into two different types of apertures (see Fig.6.a-c): separated lines (SL) and line intersections (LI). The mimicked plaid consisted of two identical superimposed asymmetric rectangular-wave gratings (Hupé and Rubin, 2003; Moreno-Bote et al., 2010; Takahashi, 2004) with a directional difference of  $120^\circ$  ( $\pm 60$  with respect to vertical). The spatial frequency of each grating was 1 cycle per degree with a speed of 2 degrees per second. In order to minimize the luminance effect of the intersection for plaid stimuli (Stoner et al., 1990; Thiele and Stoner, 2003), the width of the lines was set equal to one pixel. Thus, the gratings intersect at a point in the screen that amounts to approximately one pixel ( $0.038^\circ$ ). The color of the lines was black and the background was gray. The apertures of each type were identical, and their phases were locked according to the underlying classic plaid pattern. The spatiotemporal properties of SL and LI were identical over the full presentation cycle. The diameter of each aperture was around  $0.2^\circ$  of the viewing-angle, and they were non-overlapping. The potential locations for SL or LI were calculated beforehand as a location-grid, which followed the spatial organization of the real plaid stimulus (similar as showed in Fig.

6b). According to the location-grid, the vertical and horizontal distances between adjacent apertures was  $0.5^\circ$  and  $0.28^\circ$  of the viewing-angle, respectively. These locations were ‘labeled’ as SL and LI, which meant that the given aperture could be readout by its location. For example, LI can only be located on continuous vertical lines where the two gratings of the plaid stimulus intersect. The diameter of the whole stimulus pattern was  $23^\circ$  of the view-angle. In total, we used 720 potential aperture-locations. A rhombus-shaped mask was applied to each aperture, thus, no terminators (Pack et al., 2003) could be seen. The length of its vertical diagonal was  $0.2^\circ$  of the view-angle, and the vertical interior angle was  $60^\circ$ . In order to avoid the ‘flashing’ experience (lines from all apertures appearing on and off at the same time), we jittered the location of each mask vertically (up or down) with a random distance (maximum  $0.025^\circ$  of the view-angle) to avoid overlapping and adjusted the timing accordingly. A red fixation cross ( $0.2^\circ$  of the view-angle) was displayed at the center of the stimulus. Apertures were not located within the circular area ( $2^\circ$  of the visual-angle diameter) where the fixation cross was centered. Our stimuli were similar, but also different, in a number of ways to previously used multi-aperture stimuli (Amano et al., 2009b). For example, a) we used the line instead Gabor apertures, b) the aperture locations were selected according to the underlying plaid pattern, c) the number of apertures was manipulated (see below), and d) the proportion of different types was used to parametrically change perception.

For each experiment, the total number of apertures was chosen as one of three possible density conditions: low (L), medium (M), and high (H), with 180, 340 and 680 apertures, respectively. In addition, we parametrically manipulated the ratio between LI and SL along 11 homogeneously spaced proportions: 0% to 100%. In Experiment 2, the type of aperture did not depend on the structure of the underlying plaid pattern, that is, SL and LI could be located in any of on the locations on the underlying grid.



**Figure 7. Experiment procedure.** Subjects had to press a key on the response box to start a trial. After that, a red fixation cross was shown on the center of the monitor for 1 second. The luminance of the background is the same as the mean luminance of the following trial to exclude the influences of luminance changing. A static image in the following trial was presented for 0.5 seconds to avoid transitional eye movement. After that, the stimulus was shown for 1 second, and subjects had to report their perception of each condition.

#### 2.1.4. PROCEDURES

The same experimental pipeline was applied in Experiments 1 and 2 (see Fig.7). Subjects were instructed to press a key on the response box to start a trial. After that, a red fixation cross was shown on the center of the monitor for 1 second. Before the start of the trial, the background luminance was adjusted slightly to match the mean luminance, depending on the density condition, and to exhibit a homogeneous mean luminance across conditions and trials. First, a static image was presented for 0.5 seconds to control for transitional eye movements. Then, the stimulus started moving for 1 second, and subjects had to report their perception (either coherent or transparent) during this period by pressing one of the two keys. They were instructed to do so as fast as possible, and according to their first impression. In order to avoid potential adaptation effects, each trial was followed with a short (0.5 second) full-field Gaussian noise pattern with mean luminance equal to the average of all trials. Each psychometric point represents 30 measurements for each subject. All conditions were presented in a pseudo-randomized fashion with all conditions being presented once before a second repetition with a different order. The positions

for SL and LI were randomly selected from the location-grid for each repetition.



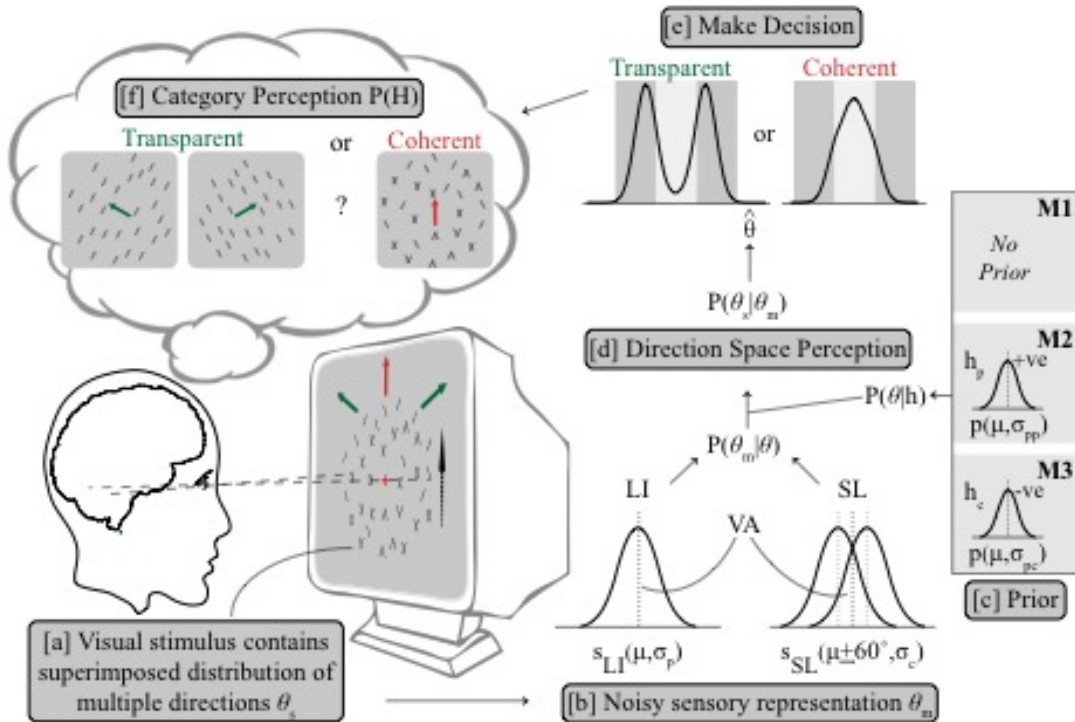
At the beginning of each block (330 trials), a standard 9-point eye-tracking calibration was performed. A new block would not begin if the subject's eye movements were not successfully calibrated. Subjects could freely take a break for up to 15 minutes after each block. For training, subjects performed 4 blocks of 15 trials before each experiment, and were instructed to fixate their eyes on the center of the monitor, using a chin-rest to prevent head movement.

### 2.1.5. PSYCHOPHYSICS DATA ANALYSIS

For each experiment, the ratio of coherence was treated as a measurement of bi-stability for each condition across repetitions for each subject. This was calculated by dividing the number of coherent response times by the total number of valid trials for that condition.

Data were fit using a cumulative Logistic psychometric function. A goodness-of-fit test was completed for each psychometric function using the Nelder-Mead-Simplex method to find the maximum likelihood. Residual deviances from the fits were used to quantify the goodness,

$$pDev = D_1 / D_0 \quad (1)$$



where  $D_1$  is deviance values from simulations that were larger than that of real data, and  $D_0$  is deviance values from all simulations. A better fit would indicate a greater value of  $pDev$  ( $pDev$  for all functions in this study: range: 0.12~1; mean: 0.79).

The raw eye movement data were interpolated with the method of the nearest neighbor after removing the blinks. The gaze positions in x- and y- dimensions were smoothed with a 200 ms running average window. At each time point, the absolute position of fixation was calculated. The trials with eye movements larger than  $2^\circ$  were excluded from further analysis.

### 2.1.6. MODEL CONSTRUCTION

Modeling transparent motion perception presents the challenge of separating a distribution of unlabeled signals along a directional space. Generally, the unlabeled signals arise from one source or from multiple sources, thereby posing a

difficult computational problem, which has been studied with vowel sounds (Beck and Kastner, 2009; Morgan et al., 2012). Here, we used the approaches cited above for inspiration, and consider the computational problem to be solved in an explicit two-step hierarchical processing with an initial unity/separation estimation, and subsequent direction perception which was made as a conditional estimate (Adelson and Movshon, 1982; Ouhanna et al., 2013; Pack and Born, 2001).

We constructed four models of the Bayesian observer faced with the decision on whether a given trial contains transparently or coherently moving stimuli. An optimal Bayesian model would average the probability of both hypotheses in this case, coherence dominated by components  $H=hc$ , and transparency dominated by the plaid pattern  $H=hp$ , determining the average probability distribution. For a difficult categorical perceptual decision like ambiguous global motion, we followed previous work (Adelson and Movshon, 1982; Beck and Kastner, 2009; Sato et al., 2007) and proposed that optimality might be sacrificed for a quick and self-consistent decision. The visual stimulus contains a superimposed distribution of multiple component directions,  $\theta_s$ , from which a sensory measurement of the same direction  $\theta_m$ , is made; an estimate contaminated by noise. Given the task at hand, in which the alternatives,  $hc$  (components) and  $hp$  (pattern), are not mutually compatible, we impose an assumption that ambiguity resolution forces the system to commit to just one of the alternatives and its corresponding prior distribution ( $P(\theta|hc)$  or  $P(\theta|hp)$ ), illustrated in Figure 8 (Morgan et al., 2012). The four model variants are driven by the following assumptions: M1 assumes no additional hypothesis and estimates the maximum  $P(\theta_m)$  to categorize direction (no prior); M2 makes a categorical decision after multiplication with an excitatory prior ( $hp$ ) on trials where early estimates suggest integration, which supports motion integration; M3 makes a decision after multiplication with an inhibitory prior ( $hc$ ) is suggested by early noisy computations, which supports motion segregation. The general Bayesian formulation for the probability of the alternative categorical hypotheses,  $H$ , is given by equation (1) which combines the likelihoods and priors,

$$P(H|\theta_m) = P(\theta_m|H)P(H)/P(\theta_m) \quad (1)$$

Applying model averaging over this posterior distribution results in each model of the corresponding equations (2),

$$\int P(\theta_s|\theta_m) d\theta = 1 \quad (2.1)$$

$$P(\theta_s|\theta_m) = P(\theta_s|\theta_m, H = hc)P(H = hc|\theta_m) + P(\theta_s|\theta_m, H = hp)P(H = hp|\theta_m) \quad (2.2)$$

where the composite posterior is obtained by adding both probabilities. We simplify the equation (2) by proposing an initial fast binary variable computation  $\chi$  (1, 2), corresponding to hypotheses  $H=hc$  and  $H=hp$  respectively. In each case, one alternative is selected and the unselected term is set to a probability of zero. The calculation uses an early noisy estimate of the direction measurement,  $\theta_m$ , to compute a vector average (VA) and a maximum likelihood estimate (MLE), before assigning a decision value of  $\chi = 1$ , if the MLE is closer to the VA, or pattern direction than a threshold direction, which is the component direction,  $\theta_c$ , and  $\chi = 2$  if the MLE is closer to the threshold direction,  $\theta_c$ . The conditional inference is therefore computed according to either,

$$P(\theta|\theta_m, \chi = 1) = P(\theta_m|\theta)P(\theta|h_p)/P(\theta_m) \quad (3),$$

in the coherent case or,

$$P(\theta|\theta_m, \chi = 2) = P(\theta_m|\theta)P(\theta|h_c)/P(\theta_m) \quad (4),$$

in the case of the transparent choice. In both equations (3) and (4), the likelihood term  $P(\theta_m|\theta)$  contains Gaussian functions of two components and one pattern terms whose width captures the sensory noise. The respective prior terms  $P(\theta|hp)$  and  $P(\theta|hc)$  are also both Gaussian terms centered on the VA direction which either enhance (hp) or inhibit (hc) the pattern to support integration or segregation, respectively. Simulated trials are fitted against psychometric data to study the interaction of sensory motion representations and prior distributions.

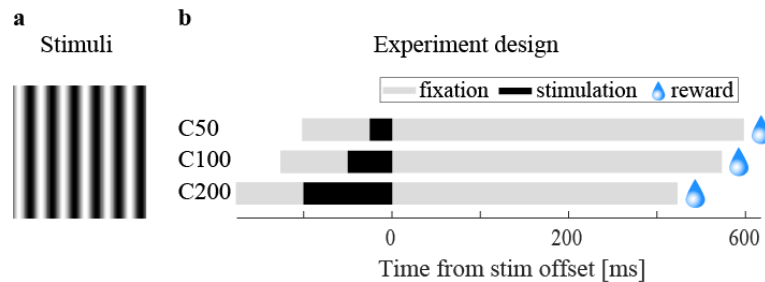
Each single simulated trial was run with a fixed set of input parameters,  $\theta_c = 60^\circ$  and  $\theta_p = 0^\circ$ , and the four model distribution parameters,  $\sigma_c$ ,  $\sigma_p$ ,  $\sigma_{pp}$  and  $\sigma_{pc}$ , were tested with the given ranges.

Once  $\chi$  has been determined, the posterior analytical function of Equation (4) was computed by combining Equations (3) and (5). RN draws are made across this posterior distribution with the results binned into a posterior discrete probability distribution  $PF(\theta)$ . An MAP estimation was used to obtain a final direction,  $\theta_f$ , and for the single trial, which simulated a forced-choice decision (i.e. transparent or coherent) based on the maximum direction (T:  $\theta_c/2 < |\theta_f|$  or C:  $\theta_c/2 > |\theta_f|$ ).

## 2.2.HIGH-DENSITY LAMINAR ELECTROPHYSIOLOGICAL RECORDING IN MONKEY V1

### 2.2.1.ANIMALS

Two adult awake monkeys (Macaque mulatta, Monkeys M1 and M2, 7-years-old, weighing 9 and 11 kg, respectively) participated in this study. All protocols and experiments were approved by the local authorities and the institutional representatives for animal protection. Experiments were performed according to the guidelines of the German Law for



**Figure 9. Stimuli and procedure.** (a) Stimuli. Full field drifting sinusoidal gratings were used in current study. A drifting grating moved into one of the eight directions for each trial, and its duration was manipulated according to experiment design. See text for details; (b) Experiment procedure. The duration of stimulation was manipulated in three conditions (C50, C100 and C200, with 50, 100 and 200 ms respectively). For each condition, monkeys were trained to keep fixating after stimulus offset until getting juice reward. The rewards were randomly delayed (125-250 ms).

Animal Protection, and in full compliance with the European Community guidelines for the care and use of laboratory animals (EUVS 86/609/EEC). Recording chambers (plastic, inner diameter 22 mm) for both monkeys were positioned over the left hemispheres for access to V1, as according to target stereotaxic coordinates. Chamber implantation was aided by high-resolution magnetic resonance anatomical imaging. The anatomical scan and recording chamber implantation was done while the animals were under general anesthesia (isoflurane).

### 2.2.2.STIMULI

Visual stimuli were full-field drifting sinusoidal gratings (see Fig.9a, spatial frequency: 1 cycles/degree, speed: 2 degree/second) of eight directions (0-315° at 45° increments), programmed using OpenGL-based stimulation toolbox. The duration of presenting drifting gratings was manipulated using a parametric design with three conditions (50, 100,

200 ms). All conditions were presented in a pseudo-randomized manner, and all conditions were presented once before a second repetition with a different order. Each condition was repeated at least 10 times during a session. The animals had to passively fixate the central fixation spot during each trial.

### 2.2.3.PROCEDURE

Each trial started with an indication sound, followed by a red fixation spot displayed on the center of the monitor against a gray background. After the pre-stimulus fixation period of 300 ms, a drifting grating was presented. Monkeys kept fixating when the grating turned off. The background during the post-stimulus period was with the same luminance as the pre-stimulus fixation period (see Fig.9b). Both monkeys were well-trained to maintain fixation during the whole trial. A trial would be aborted at any moment when the monkey failed to fixate or move the eye more than  $0.2^\circ$ . The total trial length was 1 s, followed by a randomized delay (range: 120 - 250 ms) in juice-reward.

### 2.2.4.APPARATUS

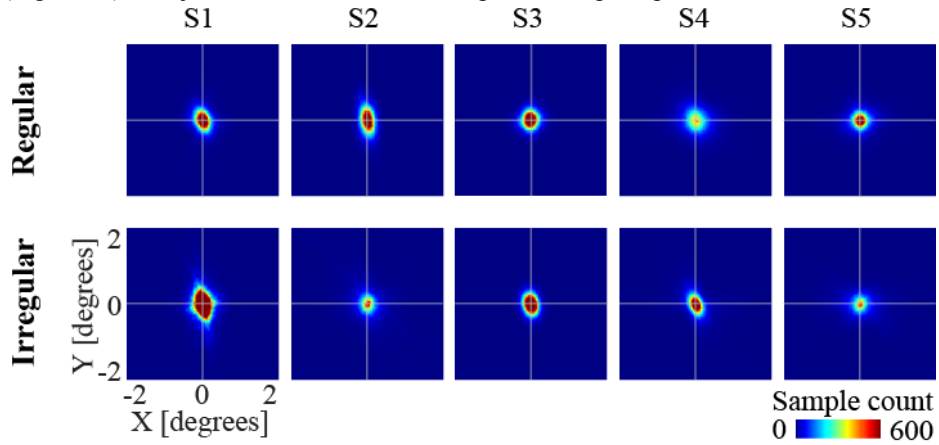
All stimuli were displayed on a LCD monitor (width: 60 cm, height: 34 cm) with a resolution of  $1920 \times 1080$  and a refresh rate of 100 Hz. The monitor was gamma corrected with a mean luminance of  $22.2 \text{ cd / m}^2$ . The distance from the eyes of monkeys to the monitor was 100 cm. All experiments were performed in a dimly lit room. The eye movements were monitored continuously online using an infrared eye tracker (IView X<sup>TM</sup> Hi-Speed Primate, SMI).

Recordings were performed using a 24-channel linear microelectrode (V-probes, Plexon) with an equal inter-contact spacing of  $100 \mu\text{m}$ , with contact impedance between 0.3 and  $0.5 \text{ M}\Omega$  at 1 kHz. The electrodes were guided into the brain manually using custom, adjustable micro-drives. Signals were amplified and recorded using Neuralynx Digital Lynx system (Neuralynx, USA). LFPs were recorded against a reference screw located at the posterior part of the skull. LFP was defined as band-pass filtered (.01–100 Hz) continuous signals recorded from each compartment of microelectrode, and were used as the basis of the analysis. The power of LFPs was calculated as the envelope amplitude of a Hilbert transformation from raw LFP signals. All data were processed using a custom MATLAB (The Mathworks Inc.) code.

### 3. RESULTS

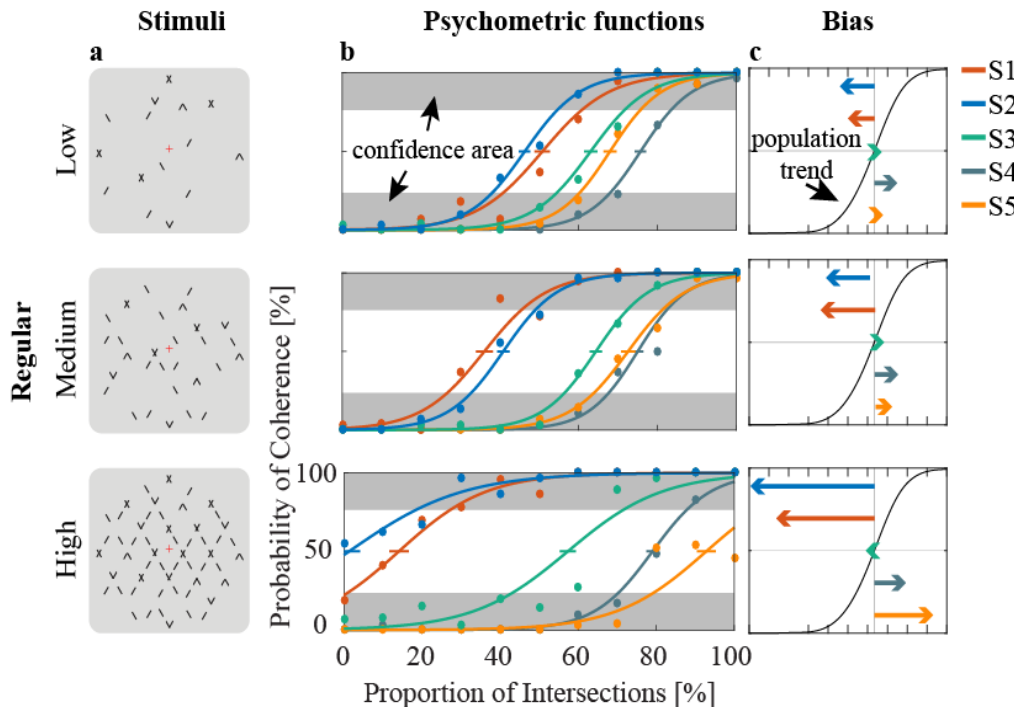
#### 3.1. INDIVIDUAL BIAS AND PERCEPTUAL STABILITY

Human psychophysics experiments were performed using novel bi-stable line-plaid stimuli, which we refer to as pseudo-plaids (Fig. 6b-c). Subjects were instructed to report their perception of either a coherent pattern moving



**Figure 10. Eye movement results.** Each subplot shows the averaged eye movement results of each subject from regular/irregular conditions.

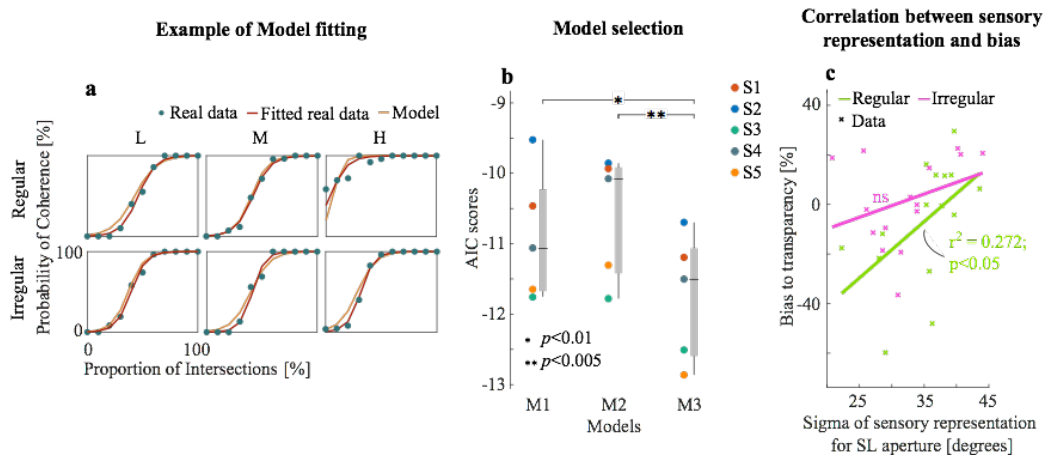
upward or two transparent surfaces sliding over each other (see Methods). Inspired by the geometric properties



typically used with moving line-plaids (Fig. 6a) (Adelson and Movshon, 1982; Pack and Born, 2001) and the nature of small receptive fields (RFs) in early visual areas, we developed a novel stimulus by decomposing the plaid into two types of apertures: separated lines (SL) and line intersections (LI). In this way, the stimuli could mimic two basic inputs that the visual system could experience locally: 1D- or 2D-motion (green/red apertures respectively in

**Figure 11. Estimation of bias and stability.** (a) Psychometric functions of each subject for Low (L), Medium (M) and High (H) density conditions in Experiment 1. The confidence area (in gray) was defined so that the probability of coherent or transparent perception was higher than 75%; (b) The direction and amount of bias for each subject corresponding to the conditions of (a). Error bars represent the standard error of the mean as estimated by 400 times resampling each psychometric function. (c) Bias estimation: in each plot, the direction of the arrow indicates the bias and the length of its denoted strength and bias direction.

Fig. 6a). We performed two experiments with the only difference being the aperture positioning. In Experiment 1 (regular, R) the structure of the mimicked plaid was maintained (Fig. 6b), whereas in Experiment 2 (irregular, I), the structures were spatially jittered (Fig. 6c). This manipulation retained the geometric plaid properties only for R stimuli,



**Figure 13.** (a) Example model fitting results from a representative subject. Empirical and model simulated psychometric functions were plotted for each experiment and conditions. (b) Comparison of AIC scores (see text). Individual subjects (left) and corresponding box plot (mean  $\pm$  SEM, right). Significance of statistical comparisons was marked with respect to M3. (c) Linear correlation between the sensory representation of SL aperture and the amount of bias to transparent perception. Data points were collapsed from all density conditions of regular/irregular experiments.

and thus, the underlying structure had a higher chance of being detected with increasing aperture density. The level of contextual organization was higher in Experiment 1 (R) as compared to Experiment 2 (I). For each experiment, the ratio between SL and LI was selected from a homogeneously spaced set of eleven ratios ranging from 0 (all SL) - 100% (all LI), and the total number of apertures was chosen from a set of three density conditions: Low (L), Medium (M), and High (H); with  $n=180, 340$  and  $680$ , respectively). Eye-movements were recorded during the experiments with a video eye-tracker and all subjects could consistently fixate within a circular window with radius  $0.4$  degrees of the visual angle (Fig. 10). For each subject, we first estimated the bias towards one of the two possible percepts (transparent or coherent) for the low-density condition by calculating the difference between the 50% coherence threshold taken from its fitted psychometric function and the same threshold was calculated from the population trend (Fig. 11). Then, we used linear mixed models with subjects as random factor to explore the relationship between bias and density for both regular and irregular conditions. The results displayed a significant linear fit between bias and density for regular ( $p < 0.005$ ), but not for irregular conditions ( $p > 0.05$ ; Fig. 12a).

To more explicitly estimate the stability of the two percepts for each condition, a perceptual stability index (PSI, Fig. 12b) was calculated for each subject. First, we defined the confidence area for stability as the probability of coherent or transparent perception (i.e. either side of the psychometric curve) which exceeded a threshold of 75%. Then, the PSI was calculated as the fraction of fitted data-points within the confidence area corresponding to the dominant percept and the rest of the points. As with the bias data, a linear regression analysis was performed on the PSI. The results showed a significant linear fit between PSI and density only for the regular ( $p < 0.05$ ), but not irregular conditions ( $p > 0.05$ ; Fig. 12b). To control for possible effects of threshold differences, we performed this analysis again with a threshold of 90%. Similar results were obtained, a significant linear correlation of PSI and density only for the regular condition ( $p < 0.05$ , results not shown).

To study the relative contribution of prior experience and the sensory representation of motion direction, we modeled the underlying motion perception task using a Bayesian framework (Sato et al., 2007; Stocker and Simoncelli, 2007). To this end, we used models of increasing complexity (no prior, a transparent prior or a coherent prior, conditional use of both priors). In the simplest model architecture (M1, no prior), the maximum likelihood was

estimated and categorized depending on whether it was closer to the coherent or transparent direction. For models M2 and M3, an initial noisy direction estimate was used to determine whether to apply an excitatory (M2) or an inhibitory (M3) prior, each of which required a single additional width parameter centered on the average direction. These would have an effect of biasing perception either towards the coherent (M2) or transparent (M3) direction. The basic assumptions and architecture of these models were inspired by a previous model for ambiguous perception (Meso et al., 2016a). Motion direction was represented as a multiplicative combination of Gaussian probability density functions representing the LI and SL aperture direction and variance (see Methods). The set of models, M1-M3 were tasked with a forced-choice decision on whether each simulated trial corresponded to a transparent or coherent direction over a number of conditions replicating Experiments 1 and 2.

An example of model-fitting results for representative subjects is shown in Fig. 13a. We then performed a model comparison based on the Akaike information criterion (AIC (Akaike, 1977)) to identify the optimal model architecture. AIC measurements are based off of the likelihood of fitting residuals to determine which model provides the best explanation for the data, giving a lower score for better fits, but penalizing models with more parameters. We calculated the AIC scores of M1-3 (Fig. 13b). A repeated one-way ANOVA was performed across all models. The results showed significant differences between models ( $F(2, 8) = 21.87, p < 0.001$ ). A post-hoc pairwise comparison showed that M3 was significantly better than M1 ( $p < 0.01$ ) and M2 ( $p < 0.005$ ), and we therefore considered M3 (transparent prior) as the most appropriate model. This could be interpreted as a general tendency of the visual system to separate motion components unless there is strong sensory evidence for integration into a single object (here provided by the line intersections (LI) apertures).

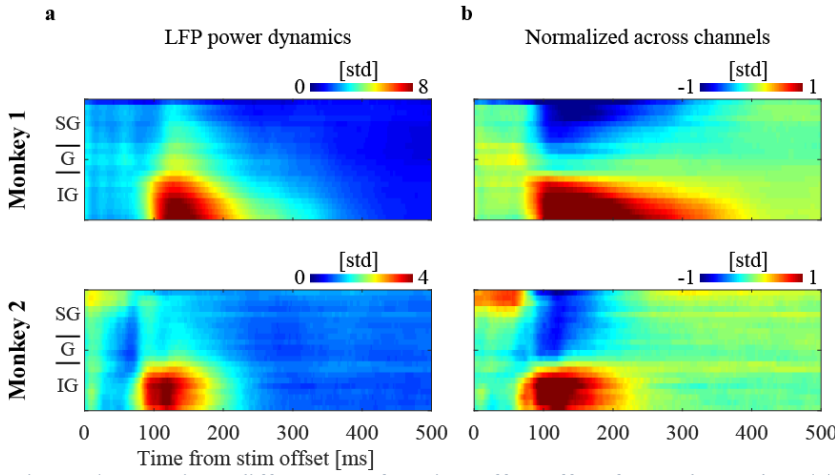
Further, we analyzed the relationship between the best model parameters of M3 and perceptual bias from empirical data to investigate the potential insights into the mechanisms of subjective biases. We found a significant linear correlation of bias and variability of the sensory representation (Gaussian sigma) for SL apertures ( $p < 0.05$ , Fig. 13c) only for the regular experiment, indicating that regularity influences the effectiveness of the sensory representation by increasing the standard deviation of its distribution.

### 3.2. OFFSET RESPONSE IN AWAKE MONKEY BRAIN

We used the following notations in the figures: n (not significant),  $p \geq .05$ ;  $.05 > * p > .005$ ;  $.005 \geq ** p > 1e-5$ ;  $1e-5 \geq *** p > 1e-10$ , and  $**** p \leq 1e-10$ .

We first calculated the amplitude of the LFP power and found that amplitudes increased transiently after stimulus offset for all layers and conditions (see Fig. 1c for example channels). The mean results across all sessions for each monkey (monkey 1:  $n = 9$ ; monkey 2:  $n = 8$ ) are denoted in Fig. 14a. The mean offset responses (averaged from 0-300ms post-stimulus) increased from superficial (supra-granular, SG; granular, G) to deep (infra-granular, IG) layers, where the mean from the superficial layers was significantly smaller than that of the deep layers (two-sample t-test; monkey 1:  $t(9) = -3.19, p < .05$ ; monkey 2:  $t(8) = -2.47, p < .05$ ; different time-window revealed similar results). In order to identify the differences between channels, we normalized the data across channels (see Fig. 14b). We can see clearly that superficial layers were relatively suppressed after stimulus offset, whilst deep layers were activated. Distinct responses were separated by the edge of layers 4 and 5.

In order to measure the offset response relative to the onset, we normalized the LFP power time-course at a trial-by-trial basis from the peak of onset transition to (see Fig. 15a, example results for condition C50). The magnitude of the offset response was found to be largest in the IG. Nevertheless, the magnitude of the offset response was smaller than for the



**Figure 14. Laminar differences of motion offset effect for each monkey.** (a) The mean LFP power time-course across sessions and conditions (see Supplementary Information for the results of each condition) and plotted as a function of time and cortical depth. SG, G and IG, supra-granular, granular and infra-granular layers, respectively; (b) Data from (a) was normalized across layers. Note that, consistent with both monkeys, distinct offset effects were found across layers: power amplitude decreased in supra-granular and granular layers (offset inhibition), but increased in the infra-granular layer (offset activation). Post-hoc pair-wise comparisons indicated the difference of ratios between SG and G was not significant ( $p > .05$ ), the ratio of IG was greater than G and IG ( $p = 9.56e-10$  for both). Similar results were found for monkey 2, but the main effect was significant ( $F(2, 45) = 57.61, p = 3.90e-13$ ). No significant difference was found between SG and G ( $p > .05$ ), and IG was greater than G and IG ( $p = 1.08e-09, p = 9.58e-10$ , respectively) for post-hoc comparisons.

We compared the variance amongst the channels in each layer (see Fig. 15b). The variance was calculated based on the LFP power time-course within a 300 ms time-window around the reference time point (100 ms before and 200 ms after) between channels of each layer (number of channels, SG: 8; G: 4 and IG: 8). The reference point was the median offset peak time for all channels. The variance values of each layer, collapsed across each time point and duration condition, were subjected to a one-way ANOVA for each monkey. For monkey 1, the results revealed a significant difference amongst layers ( $F(2, 750) = 147.57, p = 9.06e-55$ ). A post-hoc pair-wise analysis showed that the variance of G was significantly smaller than that of SG or G ( $p = 3.63e-08, p = 9.56e-10$  respectively). The variance of IG was larger than that of SG ( $p = 9.56e-10$ ). For monkey 2, the results revealed a significant difference amongst layers ( $F(2, 750) = 481.70, p = 2.81e-135$ ). A post-hoc pair-wise analysis revealed that the variance of G was significantly smaller than of SG or G ( $p = 3.33e-09, p = 9.56e-10$ ). The variance of IG was larger than that of SG ( $p = 9.56e-10$ ).

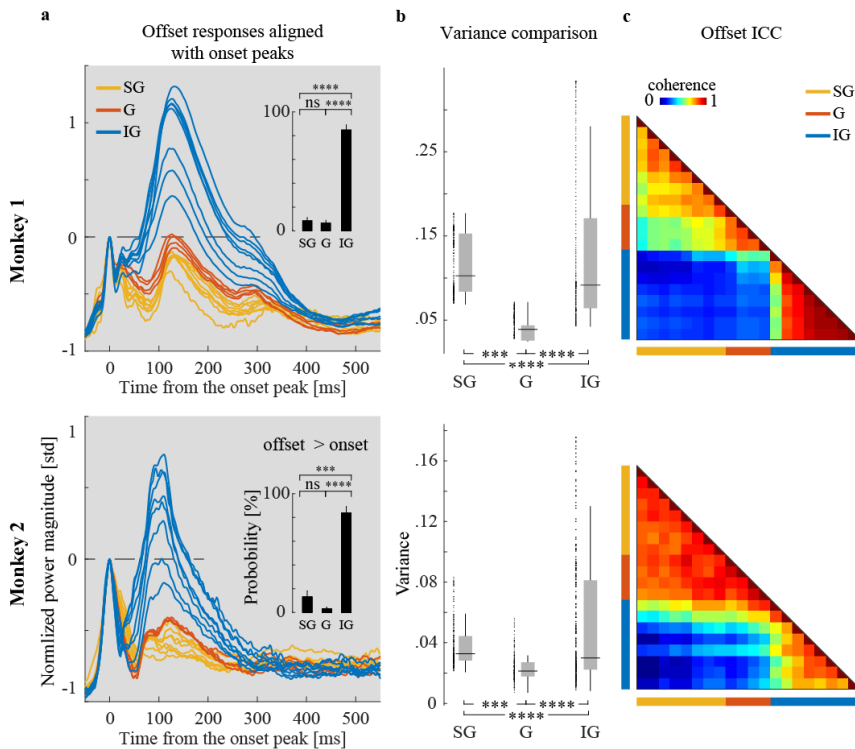
Next, we calculated the power coherence of the offset response (300 ms or a 200 ms time-window from the offset for monkeys 1 and 2, respectively) between channels. The coherence was estimated as the magnitude-squared coherence averaged by Welch's method. The mean inter-channel coherence for all sessions and duration conditions is shown in Fig. 15c, where we can see the abrupt decreasing of coherence from layer G to IG.

We analyzed the modulation of stimulation duration on offset response. The mean LFP dynamics after alignment of onset peaks across channels for each condition and monkey are displayed in Fig. 16a. The peak amplitude of the offset response within the time-window 50-250 ms after the offset of each trial was subjected to further statistical analyses. Note that, since we looked for the peak, using reasonably different time-windows would not change the results. We first collapsed the results from all duration conditions and compared the difference amongst layers with a one-way ANOVA (see Fig. 16b). Similar results were found for both monkeys: the main effect was significant (monkey 1:  $F(2, 46) = 118.67, p = 1.45e-13$ ; monkey 2:  $F(2, 46) = 57.10, p = 7.57e-14$ ), with post-hoc analyses revealing no significant differences between layers SG and G ( $p > .05$  for both monkeys), and IG was significantly larger than SG ( $p = 9.83e-10, p = 1.39e-09$ , for monkey 1 and 2, respectively) and G ( $p = 5.22e-09, p = 3.95e-07$ , for monkeys 1 and 2, respectively). We then performed a one-way repeated measures ANOVA for each layer. For monkey 1, the main effect

transient onset peak in SG and G for almost all trials. We computed the ratio of channels with offset peaks larger than onset peaks for all layers. The ratio for each layer across all duration conditions was plotted in the top-right corners of Fig. 15a. A one-way ANOVA was performed to compare the ratios across layers for each monkey. For monkey 1, the results revealed a significant main effect ( $F(2, 69) = 195.24, p = 3.90e-29$ ).

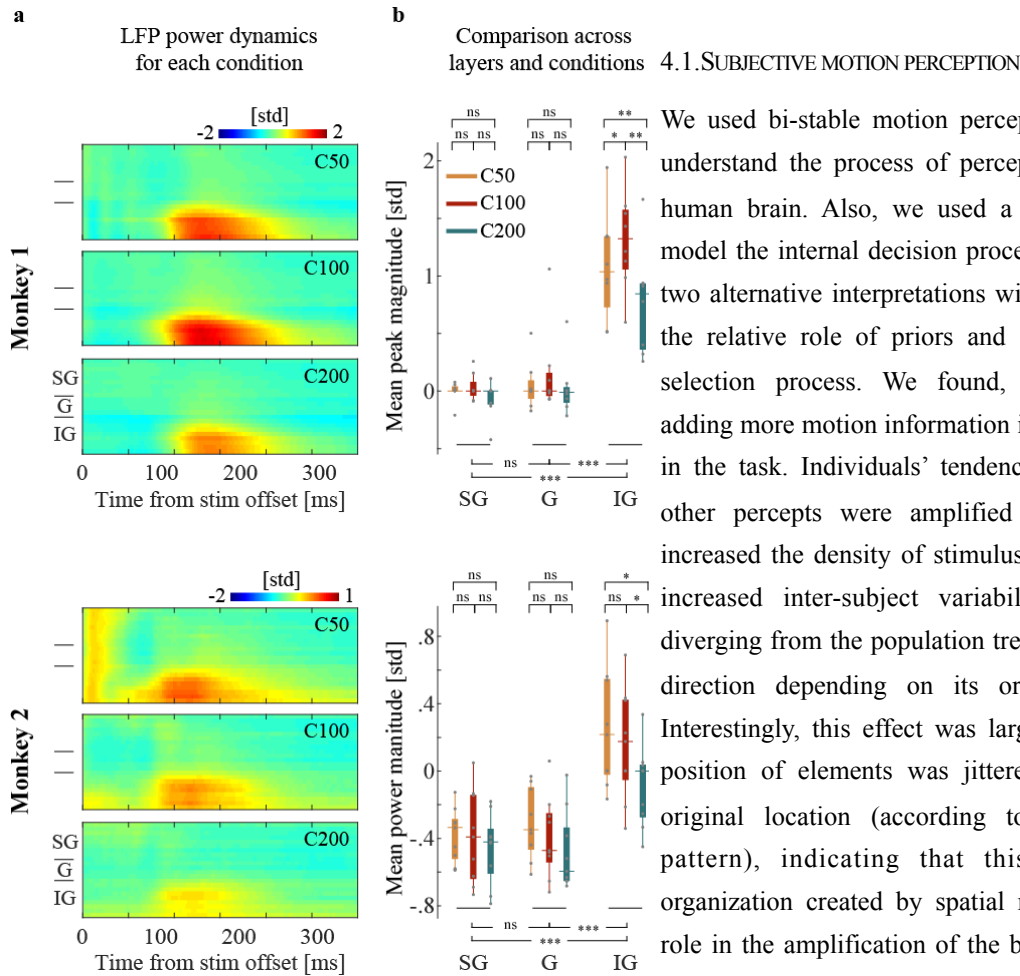


of the difference between peak amplitude was not significant for SG ( $F(2, 14) = 1.61, p > .05$ ) nor for G ( $F(2, 14) = 2.37, p > .05$ ), but significant for IG ( $F(2, 14) = 23.86, p = 3.08e-05$ ). A post-hoc analysis of layer IG showed that the mean peak amplitude for the duration condition of C200 was significantly less than that C50 and C100 (both  $p < .005$ ),



and C100 was larger than C50 ( $p < .05$ ). The results of monkey 2 showed no significant main effect for SG ( $F(2, 16) = 1.48, p > .05$ ). The main effect for G was significant ( $F(2, 16) = 5.69, p < .05$ ), but no significance was found from post-hoc comparisons (all  $p > .05$ ). A significant main effect for IG ( $F(2, 16) = 7.19, p = .005$ ) indicated that C200 was significantly smaller than C50 ( $p < .05$ ) and C100 ( $p < .05$ ), and no significant difference between C50 and C100 ( $p > .05$ ) was seen from post-hoc analysis.

**Figure 15. Relative offset responses and variance comparison for each monkey.** (a) Example relative offset responses for condition C50. Each trace signifies the mean LFP power time-course across trials and sessions, plotted as a function of time from the onset transition peak. Each trial was aligned from its onset peak after normalization. The mean probability of channels with offset peak larger than onset peak for each layer was collapsed across duration conditions and plotted in the top-right corners. Error-bar, SEM.; (b) Comparisons of within channel variances. The variance among channels in each layer was calculated based on the LFP power time-course within a 300 ms time-window around a reference time point (100 ms before, and 200 ms after the reference) between channels of each layer (SG: 8; G: 4 and IG: 8). The reference point was the median offset response peak time across all channels. The variance values of each layer were then collapsed across each time point and duration conditions for statistical comparison. Each data point (left) and box plot (right). (c) Mean offset inter-channel coherence (ICC) of LFPs was calculated between all pairs of channels (SG: 8; G: 4 and IG: 8; note that two channels between layers were not included) across all sessions and duration conditions during the offset period (0-300, 0-200 ms from stimulus offset for monkey 1 and 2 respectively). Each colored square indicates the value of pair wise coherence. The colored bars at left side and bottom edges indicate the channels from each layer. Coherence decreased abruptly between granular and infra-granular layers. This pattern consists with the ICC profile calculated by using common method (see text).



**Figure 16. Laminar differences and duration dependence of motion offset effect for each monkey.** (a) Temporal LFP power dynamics after stimulation offset. The mean LFP power time-course after stimulation offset was plotted as a function of time and cortical depth for each condition. Black horizontal bars indicate the edges of adjacent layers (SG, G and IG, supra-granular, granular and infra-granular layers, respectively); (b) Statistical comparisons of offset responses from the same data as (a) across layers and conditions with corresponding box plot (mean results from each session, gray dots on top).

modeled the sensory representations of 1D- and 2D-motion input signals as Gaussian processes with separate sigma parameters and, in addition, one of four different prior probability configurations. M3 (segregation prior) provided the best model suggesting that the visual system selectively applies inhibition within the direction space to help separate components. Importantly, it should be noted that M3 was the better model even in subjects that were biased towards coherent percepts. We conjectured that the brain, when faced with such tasks, applies a conditional implementation of separating priors on some critical trials and not an integrating one because integration might arise naturally from overlapping signal distributions (Mahani et al., 2005).

We used bi-stable motion perception as a tool to better understand the process of perceptual stabilization in the human brain. Also, we used a Bayesian framework to model the internal decision process, leading to one of the two alternative interpretations with the aim to understand the relative role of priors and sensory evidence in the selection process. We found, counter-intuitively, that adding more motion information increased response biases in the task. Individuals' tendencies to either one or the other percepts were amplified substantially when we increased the density of stimulus elements, leading to an increased inter-subject variability, with each subject diverging from the population trend with a magnitude and direction depending on its original bias (Fig. 12a). Interestingly, this effect was largely abolished when the position of elements was jittered with respect to their original location (according to the underlying plaid pattern), indicating that this form of contextual organization created by spatial regularity plays a major role in the amplification of the bias. As a measure of the effect of bias amplification, we computed a perceptual stability index and found that it linearly increased with element density (Fig. 12b).

To further understand the brain processes leading to this result, in particular with respect to prior information encoded in the brain of each participant, we built a battery of Bayesian models (M1-M3; see Methods) with the task to probabilistically select one of the two percepts on a trial-by-trial basis simulating the experiments. These

Bias stands at the core of signal detection theory (SDT) when applied to both living organisms and machines. In fact, Green and Swets, 1966, being the first to develop SDT approaches in psychophysics, directly criticized previously used methods for not being able to separate the sensitivity of subjects from their potential biases. In addition to the principle problem of detecting a signal within noise, our brains also face the problem of inherently ambiguous sensory inputs. Thus, to make veridical interpretations of the outside world, the brain needs to employ additional mechanisms, such as attention and prior experience (Desimone, 1998; Desimone and Duncan, 1995; Knill and Richards, 1996; Meso et al., 2016b; Rao et al., 2002). One theory suggested that objects simultaneously presented in the visual field compete and attention can bias the outcome of this competition (Desimone, 1998; Desimone and Duncan, 1995; Knill and Richards, 1996; Rao et al., 2002). Our results are consistent with the general framework of the biased competition hypothesis; however, attention does not seem to be the primary source of the observed biases as we did not manipulate attention across the different conditions and the subjects had to continuously perform the task of reporting their percepts so it should have remained constant. Moreover, individual bias directions were independent of the stimulus configuration (which was the same for all subjects) precluding bottom-up stimulus-driven attention effects. The subject specific results suggested a strong influence of prior experience or assumptions, and thus, we expected our modeling results to reveal that some subjects use a ‘coherence’ prior (M2) while others use a ‘component’ prior (M3). To our surprise, M3 (in comparison to M2; Fig. 12a) was a better model for all our subjects. This suggests that the selectivity of the visual system of each participant to the two motion signals (sensory input standard deviations) was more important for determining bias direction in comparison to the prior.

Furthermore, bias in our experiments was increased with stimulus element density. This was an unexpected finding, as previous studies have shown that increases in the density of random-dot-kinematograms (RDKs) result in decreased coherence thresholds (Barlow and Tripathy, 1997) or being unaffected (Eagle and Rogers, 1997; Talcott et al., 2000; Welchman and Harris, 2000). We note, however, that RDK experiments are closer to the foundations of SDT (i.e. detecting a signal within noise). We conjecture that in our scenario, competition between the two motion representations may be enhanced by density, resulting in the observed increase of bias towards a preferred representation, acting like a perceptual attractor. This is consistent with contrast-based motion signal increases, which result in stronger 2D motion attractors in a tri-stable motion stimulus (Meso et al., 2016b). In addition, research with RDKs has demonstrated that coherence thresholds in 5-6-year-olds were a) much higher and b) decreased with dot density in comparison to adults (Narasimhan and Giaschi, 2012). In our view, this indicates that sensory representations of this form, and perhaps the connectivity of the circuit, are also prone to changes by experience during development. This could explain the different directions of the biases in different subjects.

Interestingly, the bias-amplification and the increases in the perceptual stability index with density were largely abolished in the stimuli with jittered aperture positions. This is consistent with previous work demonstrating the importance of regularity (Barlow and Tripathy, 1997; Green and Swets, 1966), which appears to play a role in the selection of stable neural representations. Another interpretation is that the reduction of regularity eliminates, the correspondence of single stimulus elements to the underlying patterns, or ‘objects’, interfering with their spatial integration. This is consistent with studies that have demonstrated a precedence of global features in visual perception (Beck and Kastner, 2005; Phillips et al., 2015). Moreover, the influence of position jitter on bias indicates that the scale of integration cannot be completely local nor global as both these conditions should not elicit an effect. We, thus, suggest our results directly indicate that the motion integration mechanisms contributing to individual biases are of ‘meso-scale’, i.e. go beyond single-neuron receptive fields (RFs) in V1 to scales more typical for area V5/MT, but not the very large RFs found in size-invariant object selective areas like inferotemporal cortex (IT).

Previous research has found strong evidence for active perceptual stabilization mechanisms in the visual system, such as, the reorganization of sensory representations during intermittent viewing (Leopold et al., 2002), top-down

modulation of beta-band synchronization (Kloosterman et al., 2015), feedforward inhibition (Bollimunta and Ditterich, 2012), arousal (de Gee et al., 2014; Mather and Sutherland, 2011b), and memory (Wimmer and Shohamy, 2012). Our study suggests that bias may serve as an additional factor our brains use to stabilize our perception of the world.

#### 4.2. DYNAMIC LAMINAR PROFILE OF VISUAL MOTION OFFSET RESPONSE

In general, we found that, comparing to the fixation period, neural activity increased with a transition offset response for all V1 layers when a briefly presented moving stimulus turned off. Such persistent neural activity lasted around 300 ms from the offset of stimulation, and only the infra-granular (IG) layer was duration dependent. Similar results were observed in both monkeys. These results suggested that temporal dynamics are coded differently in V1 layers and highly dependent on the local network. In front of abrupt changes, temporal information is coded transiently in superficial layers, but additional duration information was maintained in deep layers. Superficial layers were relatively suppressed after stimulus removal, indicating that these layers may be more sensitive to ongoing sensory inputs. Importantly, we found that the transient offset response from the granular layer is highly consistent with the lowest inter-channel variability as compared to other layers, which is consistent with the latest observations that information encoded by neuronal populations in the granular layer is more accurate than in other layers (Ecker et al., 2010; Hansen et al., 2012).

Persistent neural activity after removal of stimulation has been found in various brain areas (Akaike, 1977; Andersen et al., 1987; Funahashi et al., 1989; Fuster and Alexander, 1971; Levick and Zacks, 1970; Mendoza-Halliday et al., 2014; Meyer et al., 2007; Miller et al., 1996a, 1996b; Schiller, 1968; Tark and Curtis, 2009), and the perception of a successive stimulus could be impaired if post-stimulus processing of the first one has not been finished (Keyser and Perrett, 2002). Recent research has reported that when the stimulus is abruptly turned off from one eye, the stimulus on the other eye would be invisible under binocular rivalry (de Graaf et al., 2017). The authors proposed a transient offset suppression mechanism, whereby ongoing processing is suppressed by offset transient neural activity. Our results may provide direct evidence for this hypothesis. While the robust offset response from the infra-granular layer may interrupt the processing of an ongoing stimulus, it has been known that pyramidal cells in V1 layer 6 project back to the dorsolateral geniculate nucleus (dLGN), from which V1 receives visual inputs (Bourassa and Deschênes, 1995; Jones E. G., 2007), and the suppression of inhibitory neurons in layer 6 throughout most of the superficial layers (Briggs and Callaway, 2001; Olsen et al., 2012). Thus, the offset transient neural activity of infra-granular layers may indicate active suppression processing.

Visual memory for briefly-presented stimuli can be maintained as iconic memory with rapid decay (Coltheart, 1980a, 1980b; Duysens et al., 1985; Di Lollo, 1980; Phillips and Routh, 1983; Sperling, 1960). Our results show that offset response lasted around 300 ms, and its amplitude was inversely dependent on the duration at the infra-granular layer. Consistent with previous research, our results suggest that visual motion information can be maintained in V1 (van Kerkoerle et al., 2017; Super et al., 2001), but only in the infra-granular layer. Additionally, since the monkeys were awake and behaving, our study provides direct evidence that the offset response is related to visual persistence. Visual persistence refers to a phenomenon that visual perception remains for a period of time when a briefly-presented stimulus turns off and the duration of visual persistence is inversely related to the stimulus duration and decays rapidly (Coltheart, 1980; Di Lollo, 1980; Sperling, 1960). One early study with anesthetized cats showed that the firing rates of some neurons from area 17 decreased with longer stimulation, while some others were independent of stimulus duration (Duysens et al., 1985). These observations are in line with our results indicating that only the infra-granular layer may be involved in this phenomenon.

It has been suggested that the offset response may be driven by the interaction between inhibition and disinhibition, that the offset response may originate from the released inhibition (Calford and Webster, 1981; Takahashi et al., 2004; Volkov and Galazjuk, 1991). Lateral suppression is a fundamental mechanism in visual system, where cross-orientation suppression of orientation-tuned neurons can be suppressed by cross-orientated stimuli (Morrone et al., 1982) and opponent direction suppression where different directions of motion suppress each other (Qian and Andersen, 1994). Thus, offset responses may reflect the released suppression after removal of oriented stimuli (McLelland et al., 2010), or an opponent motion signal. However, other researchers reported that the offset response is highly precise both in visual and auditory modalities (Bair et al., 2002; Qin et al., 2007), revealing that the offset response is mediated by an active mechanism (Qin et al., 2007). The hypothesis of suppression preceding offset response predicts increased neural activity with longer stimulation. However, our results exhibited an inverse correlation of durations for the infra-granular layer, and that the longest stimulation induced the weakest offset response in the infra-granular layer and a null effect for superficial layers. In addition, if the offset response was from the released suppression of stimulation, the amplitude of the offset response should be less or at most equal to the onset amplitude, but we found that the peak amplitude of the offset response was significantly larger than onset peak from infra-granular layer. Thus, our results revealed that the offset response in the infra-granular layer was actively “amplified”. Indeed, previous studies reported the spiking activity of the offset response was larger for longer stimulus durations (Duysens et al., 1985, 1996). Possible explanations could be that a) previous studies used a static stimulus, thereby different results from our study indicate that the offset response to moving stimuli may be potentially different; b) anesthetized animals were studied by those researchers, while we used alert monkeys; c) some of the previous studies took all post-stimulus neural activity into analysis, but not the transient offset response, it has been suggested that the transient offset response is different from the whole post-stimulus response.

In anesthetized animals, duration-dependent offset responses were found in the LGN with parvocellular cells (P cells), but not magnocellular cells (M cells) (McLelland et al., 2010). Additionally, some of the V1 neurons were not affected by the stimulus duration (Duysens et al., 1985). According to these results, researchers suggested that the duration-dependent offset responses may rely on cell types, and that the lack of duration dependence of the V1 neurons reflect the inputs from LGN M cells. However, V1 layer 4c receives inputs from both M and P LGN pathways (Blasdel and Lund, 1983; Hubel and Wiesel, 1972), and we found that there is no duration modulation in the granular layer with awake monkeys. Lastly, it had been shown that the offset response of complex cells is larger than simple cells in the anesthetized V1 (McLelland et al., 2010). However, we found that the amplitude of the offset response was highest only in the infra-granular layer, and it was already known that both layer 2/3 and 5 have a majority of complex cells (Orban, 1984). Therefore, the offset response in the awake brain may not reflect the downstream neural activity from LGN, but rather involves population neural activity under intra-cortical mechanisms regardless of cell type.

## REFERENCES

- Abdel Rahman, R. (2011). Facing good and evil: Early brain signatures of affective biographical knowledge in face recognition. *Emotion* *11*, 1397–1405.
- Adams, W.J., Graf, E.W., and Ernst, M.O. (2004). Experience can change the “light-from-above” prior. *Nat. Neurosci.* *7*, 1057–1058.
- Adelson, E.H., and Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* *2*, 284.
- Adelson, E.H., and Movshon, J.A. (1982). Phenomenal coherence of moving visual patterns. *Nature* *300*, 523–525.
- Adrian, E.D., and Matthews, R. (1927). The action of light on the eye: Part I. The discharge of impulses in the optic nerve and its relation to the electric changes in the retina. *J. Physiol.* *63*, 378–414.
- Ainsworth, M., Lee, S., Cunningham, M.O., Traub, R.D., Kopell, N.J., and Whittington, M.A. (2012). Rates and Rhythms: A Synergistic View of Frequency and Temporal Coding in Neuronal Networks. *Neuron* *75*, 572–583.
- Akaike, H. (1977). On entropy maximization principle. PR Krishnaiah (Ed.). In *Proc.Symp.Applications Statistics*, (North-Holland), pp. 27–47.
- Alais, D., Verstraten, F.A.J., and Burr, D.C. (2005). The motion aftereffect of transparent motion: two temporal channels account for perceived direction. *Vision Res.* *45*, 403–412.
- Alink, A., Schwiedrzik, C.M., Kohler, A., Singer, W., and Muckli, L. (2010). Stimulus Predictability Reduces Responses in Primary Visual Cortex. *J. Neurosci.* *30*.
- Alvarez, I., de Haas, B., Clark, C.A., Rees, G., and Schwarzkopf, D.S. (2015). Comparing different stimulus configurations for population receptive field mapping in human fMRI. *Front. Hum. Neurosci.* *9*, 96.
- Amano, K., Wandell, B.A., and Dumoulin, S.O. (2009a). Visual Field Maps, Population Receptive Field Sizes, and Visual Field Coverage in the Human MT+ Complex. *J. Neurophysiol.* *102*, 2704–2718.
- Amano, K., Edwards, M., Badcock, D.R., and Nishida, S. (2009b). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *J. Vis.* *9*, 4.1-25.
- Andersen, R.A., Essick, G.K., and Siegel, R.M. (1987). Neurons of area 7 activated by both visual stimuli and oculomotor behavior. *Exp. Brain Res.* *67*, 316–322.
- Anderson, A.K. (2005). Affective Influences on the Attentional Dynamics Supporting Awareness. *J. Exp. Psychol. Gen.* *134*, 258–281.
- Andrews, T.J., and Schluppeck, D. (2000). Ambiguity in the perception of moving stimuli is resolved in favour of the cardinal axes. *Vision Res.* *40*, 3485–3493.
- Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the “oblique effect” in man and animals. *Psychol. Bull.* *78*, 266–278.
- Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of Ongoing Activity: Explanation of the Large Variability in Evoked Cortical Responses. *Science* (80-. ). *273*, 1868–1871.
- Attarha, M., and Moore, C.M. (2015). Onset rivalry: factors that succeed and fail to bias selection. *Atten. Percept. Psychophys.* *77*, 520–535.

- Babiloni, C., Vecchio, F., Bultrini, A., Luca Romani, G., and Rossini, P.M. (2006). Pre- and poststimulus alpha rhythms are related to conscious visual perception: a high-resolution EEG study. *Cereb. Cortex* *16*, 1690–1700.
- Bach, D.R., and Dolan, R.J. (2012). Knowing how much you don't know: a neural organization of uncertainty estimates. *Nat. Rev. Neurosci.* *13*, 572.
- Bair, W., Cavanaugh, J.R., Smith, M. a, and Movshon, J.A. (2002). The timing of response onset and offset in macaque visual neurons. *J. Neurosci.* *22*, 3189–3205.
- Bar, M. (2004). Visual objects in context. *Nat. Rev. Neurosci.* *5*, 617–629.
- Bar, M. (2007). The proactive brain: using analogies and associations to generate predictions. *Trends Cogn. Sci.* *11*, 280–289.
- Barlow, H., and Tripathy, S.P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. *J. Neurosci.* *17*, 7954–7966.
- Beck, D.M., and Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nat. Neurosci.* *8*, 1110–1116.
- Beck, D.M., and Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Res.* *49*, 1154–1165.
- van den Berg, M., Kubovy, M., and Schirillo, J.A. (2011). Grouping by Regularity and the perception of illumination. *Vision Res.* *51*, 1360–1371.
- Bergen, J.R., and Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature* *303*, 696–698.
- Bergmann, J., Genç, E., Kohler, A., Singer, W., and Pearson, J. (2016a). Neural Anatomy of Primary Visual Cortex Limits Visual Working Memory. *Cereb. Cortex* *26*, 43–50.
- Bergmann, J., Genç, E., Kohler, A., Singer, W., and Pearson, J. (2016b). Smaller Primary Visual Cortex Is Associated with Stronger, but Less Precise Mental Imagery. *Cereb. Cortex* *26*, 3838–3850.
- Bernier, B.E., Lacagnina, A.F., Ayoub, A., Shue, F., Zelman, B. V., Krasne, F.B., and Drew, M.R. (2017). Dentate Gyrus Contributes to Retrieval as well as Encoding: Evidence from Context Fear Conditioning, Recall, and Extinction. *J. Neurosci.* *37*.
- Biederman, I. (1972). Perceiving real-world scenes. *Science* *177*, 77–80.
- Biederman, I., Glass, A.L., and Stacy, E.W. (1973). Searching for objects in real-world scenes. *J. Exp. Psychol.* *97*, 22–27.
- Bishop, S.J. (2009). Trait anxiety and impoverished prefrontal control of attention. *Nat. Neurosci.* *12*, 92–98.
- Blasdel, G.G., and Lund, J.S. (1983). Termination of afferent axons in macaque striate cortex. *J. Neurosci.* *3*, 1389–1413.
- Bocanegra, B.R., Huijding, J., and Zeelenberg, R. (2012). Beyond attentional bias: A perceptual bias in a dot-probe task. *Emotion* *12*, 1362–1366.
- Boff, K.R., Kaufman, L., and Thomas, J.P. (1986). *Handbook of Perception and Human Performance: Sensory processes and perception* (Wiley).
- Bolkan, S.S., Stujenske, J.M., Parnaudeau, S., Spellman, T., Rauffenbart, C., Abbas, A.I., Harris, A.Z., Gordon, J.A.,

- and Kellendonk, C. (2017). Thalamic projections sustain prefrontal activity during working memory maintenance. *Nat Neurosci* *XX*, XXX–XXX.
- Bollimunta, A., and Ditterich, J. (2012). Local computation of decision-relevant net sensory evidence in parietal cortex. *Cereb. Cortex* *22*, 903–917.
- Born, R.T., Pack, C.C., Ponce, C.R., and Yi, S. (2006). Temporal evolution of 2-dimensional direction signals used to guide eye movements. *J. Neurophysiol.* *95*, 284–300.
- Born, R.T., Tsui, J.M.G., and Pack, C.C. (2010). Temporal dynamics of motion integration. In *Dynamics of Visual Motion Processing: Neuronal, Behavioral, and Computational Approaches*, (Boston, MA: Springer US), pp. 37–54.
- Bourassa, J., and Deschênes, M. (1995). Corticothalamic projections from the primary visual cortex in rats: a single fiber study using biocytin as an anterograde tracer. *Neuroscience* *66*, 253–263.
- Braddick, O. (1993). Segmentation versus integration in visual motion processing. *Trends Neurosci.* *16*, 263–268.
- Braddick, O. (1997). Local and global representations of velocity: transparency, opponency, and global direction perception. *Perception* *26*, 995–1010.
- Bradley, D.R., and Petry, H.M. (1977). Organizational determinants of subjective contour: the subjective Necker cube. *Am. J. Psychol.* *90*, 253–262.
- Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat. Vis.* *10*, 433–436.
- Brascamp, J.W., Knapen, T.H.J., Kanai, R., Noest, A.J., van Ee, R., and van den Berg, A. V. (2008). Multi-timescale perceptual history resolves visual. *PLoS One* *3*, e1497.
- Braunstein, M.L., and Payne, J.W. (1969). Perspective and form ratio as determinants of relative slant judgments. *J. Exp. Psychol.* *81*, 584–590.
- Bray, N. (2017). Working memory: Persistence is key. *Nat. Rev. Neurosci.* *18*, 385–385.
- Briggs, F., and Callaway, E.M. (2001). Layer-specific input to distinct cell types in layer 6 of monkey primary visual cortex. *J. Neurosci.* *21*, 3600–3608.
- Brosch, T., Pourtois, G., and Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cogn. Emot.* *24*, 377–400.
- Bruno, N., and Bertamini, M. (2014). Perceptual organization and the aperture problem. In *Oxford Handbook of Perceptual Organization*, pp. 1–19.
- Burke, D., and Wenderoth, P. (1993). The effect of interactions between one-dimensional component gratings on two-dimensional motion perception. *Vision Res.* *33*, 343–350.
- Calford, M.B., and Webster, W.R. (1981). Auditory representation within principal division of cat medial geniculate body: an electrophysiology study. *J. Neurophysiol.* *45*, 1013–1028.
- Calvin, W.H., and Stevens, C.F. (1968). Synaptic noise and other sources of randomness in motoneuron interspike intervals. *J. Neurophysiol.* *31*, 574–587.
- Campana, F., Rebollo, I., Urai, A., Wyart, V., and Tallon-Baudry, C. (2016). Conscious Vision Proceeds from Global to Local Content in Goal-Directed Tasks and Spontaneous Vision. *J. Neurosci.* *36*, 5200–5213.
- Carandini, M., and Heeger, D.J. (2011). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* *13*, 51–62.



- Carbon, C.-C. (2014). Understanding human perception by human-made illusions. *Front. Hum. Neurosci.* *8*, 566.
- Carter, O., and Cavanagh, P. (2007). Onset rivalry: Brief presentation isolates an early independent phase of perceptual competition. *PLoS One* *2*, e343.
- Cecchetto, S., and Lawson, R. (2017). Regularity detection by haptics and vision. *J. Exp. Psychol. Hum. Percept. Perform.* *43*, 103–125.
- Chalk, M., Seitz, A.R., Series, P., S., T., and T., W. (2010). Rapidly learned stimulus expectations alter perception of motion. *J. Vis.* *10*, 2–2.
- Chalupa, L.M., and Werner, J.S. (2004). *The Visual Neurosciences* (MIT Press).
- Chater, N., and Manning, C.D. (2006). Probabilistic models of language processing and acquisition. *Trends Cogn. Sci.* *10*, 335–344.
- Chetverikov, D. (2000). Pattern regularity as a visual key. *Image Vis. Comput.* *18*, 975–985.
- Cisler, J.M., and Koster, E.H.W. (2010). Mechanisms of attentional biases towards threat in anxiety disorders: An integrative review. *Clin. Psychol. Rev.* *30*, 203–216.
- Coen-Cagli, R., Kohn, A., and Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. *Nat. Neurosci.* *18*, 1648–1655.
- Coltheart, M. (1980a). Iconic memory and visible persistence. *Percept. Psychophys.* *27*, 183–228.
- Coltheart, M. (1980b). The persistences of vision. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* *290*, 57–69.
- Correll, J., Wittenbrink, B., Park, B., Judd, C.M., and Goyle, A. (2011). Dangerous Enough: Moderating Racial Bias with Contextual Threat Cues. *J. Exp. Soc. Psychol.* *47*, 184–189.
- Curran, W., and Benton, C.P. (2006). Test stimulus characteristics determine the perceived speed of the dynamic motion aftereffect. *Vision Res.* *46*, 3284–3290.
- Curran, W., and Braddick, O.J. (2000). Speed and direction of locally-paired dot patterns. *Vision Res.* *40*, 2115–2124.
- Curran, W., Hibbard, P.B., and Johnston, A. (2007). The visual processing of motion-defined transparency. *Proc. R. Soc. B* *274*, 1049–1057.
- Cutting, J.E., and Millard, R.T. (1984). Three gradients and the perception of flat and curved surfaces. *J. Exp. Psychol. Gen.* *113*, 198–216.
- Davenport, J.L., and Potter, M.C. (2004). Scene Consistency in Object and Background Perception. *Psychol. Sci.* *15*, 559–564.
- Dehaene, S., Changeux, J.-P., Naccache, L., Sackur, J., and Sergent, C. (2006). Conscious, preconscious, and subliminal processing: a testable taxonomy. *Trends Cogn. Sci.* *10*, 204–211.
- Dekaban, A.S., and Sadowsky, D. (1978). Changes in brain weights during the span of human life: Relation of brain weights to body heights and body weights. *Ann. Neurol.* *4*, 345–356.
- Delicato, L.S., and Derrington, A.M. (2005). Coherent motion perception fails at low contrast. *Vision Res.* *45*, 2310–2320.
- Derrington, a, and Suero, M. (1991). Motion of complex patterns is computed from the perceived motions of their components. *Vision Res.* *31*, 139–149.

- Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 353, 1245–1255.
- Desimone, R., and Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annu. Rev. Neurosci.* 18, 193–222.
- DeSimone, K., Viviano, J.D., and Schneider, K.A. (2015). Population Receptive Field Estimation Reveals New Retinotopic Maps in Human Subcortex. *J. Neurosci.* 35.
- Dougherty, R.F., Koch, V.M., Brewer, A.A., Fischer, B., Modersitzki, J., and Wandell, B.A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *J. Vis.* 3, 1.
- Dumoulin, S.O., and Wandell, B.A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage* 39, 647–660.
- Duncan, J. (1980). The demonstration of capacity limitation. *Cogn. Psychol.* 12, 75–96.
- Duncker, K. (1929). Über induzierte Bewegung. *Psychol. Forsch.* 12, 180–259.
- Durgin, F.H. (2001). Texture contrast aftereffects are monocular; texture density aftereffects are binocular. *Vision Res.* 41, 2619–2630.
- Durgin, F.H., and Hammer, J.T. (2001). Visual aftereffects of sequential perception: dynamic adaptation to changes in texture density and contrast. *Vision Res.* 41, 2607–2617.
- Duysens, J., Orban, G.A., Cremieux, J., and Maes, H. (1985). Visual cortical correlates of visible persistence. *Vision Res.* 25, 171–178.
- Duysens, J., Schaafsma, S.J., and Orban, G.A. (1996). Cortical off response tuning for stimulus duration. *Vision Res.* 36, 3243–3251.
- Eagle, R.A., and Rogers, B.J. (1997). Effects of dot density, patch size and contrast on the upper spatial limit for direction discrimination in Random-dot Kinematograms. *Vision Res.* 37, 2091–2102.
- Ebbinghaus, H. (1905). *Grundzüge der psychologie* (Veit & Comp.).
- Ecker, A.S., Berens, P., Keliris, G.A., Bethge, M., Logothetis, N.K., and Tolias, A.S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science* 327, 584–587.
- Efron, R. (1969). *What is Perception?* (Springer Netherlands), pp. 137–173.
- Ernst, M.O., and Bühlhoff, H.H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169.
- Ernst, M.O., and Di Luca, M. (2012). Multisensory Perception: From Integration to Remapping. In *Sensory Cue Integration*, p.
- Van Essen, D.C., and Maunsell, J.H.R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends Neurosci.* 6, 370–375.
- Van Essen, D.C., Anderson, C.H., and Felleman, D.J. (1992). Information processing in the primate visual system: an integrated systems perspective. *Science* 255, 419–423.
- Etkin, A., Klemenhagen, K.C., Dudman, J.T., Rogan, M.T., Hen, R., Kandel, E.R., and Hirsch, J. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron* 44, 1043–1055.

- Faisal, A.A., Selen, L.P.J., and Wolpert, D.M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci.* *9*, 292–303.
- Felleman, D.J., and Van Essen, D.C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cereb. Cortex* *1*, 1–47.
- Fisher, D.L., and L., D. (1982). Limited-channel models of automatic detection: Capacity and scanning in visual search. *Psychol. Rev.* *89*, 662–692.
- Fox, E., Russo, R., and Dutton, K. (2002). Attentional Bias for Threat: Evidence for Delayed Disengagement from Emotional Faces. *Cogn. Emot.* *16*, 355–379.
- Freeman, W.J. (1981). A physiological hypothesis of perception. *Perspect. Biol. Med.* *24*, 561–592.
- Freeman, J., and Simoncelli, E.P. (2011). Metamers of the ventral stream. *Nat. Neurosci.* *14*, 1195–1201.
- Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron* *88*, 220–235.
- Friston, K. (2003). Learning and inference in the brain. *Neural Networks* *16*, 1325–1352.
- Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* *13*, 293–301.
- Friston, K., Mattout, J., and Kilner, J. (2011). Action understanding and active inference. *Biol. Cybern.* *104*, 137–160.
- Funahashi, S., Bruce, C.J., and Goldman-Rakic, P.S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* *61*, 331–349.
- Fuster, J.M., and Alexander, G.E. (1971). Neuron activity related to short-term memory. *Science* *173*, 652–654.
- García-Pérez, M. a (1989). Visual inhomogeneity and eye movements in multistable perception. *Percept. Psychophys.* *46*, 397–400.
- de Gardelle, V., Kouider, S., and Sackur, J. (2010). An oblique illusion modulated by visibility: non-monotonic sensory integration in orientation processing. *J. Vis.* *10*, 6.
- de Gee, J.W., Knapen, T., and Donner, T.H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *Proc. Natl. Acad. Sci. U. S. A.* *111*, E618-25.
- Geisler, W.S. (2008). Visual Perception and the Statistical Properties of Natural Scenes. *Annu. Rev. Psychol.* *59*, 167–192.
- Genç, E., Bergmann, J., Singer, W., and Kohler, A. (2015). Surface area of early visual cortex predicts individual speed of traveling waves during binocular rivalry. *Cereb. Cortex* *25*, 1499–1508.
- Girshick, A.R., Landy, M.S., and Simoncelli, E.P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci.* *14*, 926–932.
- Gizzi, M.S., Katz, E., Schumer, R. a, and Movshon, J. a (1990). Selectivity for orientation and direction of motion of single neurons in cat striate and extrastriate visual cortex. *J. Neurophysiol.* *63*, 1529–1543.
- Goldstein, E.B. (2009). *Sensation and Perception* (Cengage Learning).
- Gong, M., Jia, K., and Li, S. (2017). Perceptual competition promotes suppression of reward salience in behavioral selection and neural representation. *J. Neurosci.* *37*, 6242–6252.
- de Graaf, T.A., van Ee, R., Croonenberg, D., Klink, P.C., and Sack, A.T. (2017). Visual suppression at the offset of binocular rivalry. *J. Vis.* *17*, 2.
- Graf, E.W., Adams, W.J., and Lages, M. (2004). Prior depth information can bias motion perception. *J. Vis.* *4*, 427–433.

- Green, D.M., and Swets, J.A. (1966). Signal detection theory and psychophysics (Peninsula, Los Altos, Calif).
- Gregory, R.L. (1980). The Intelligent Eye (Weidenfeld & Nicolson).
- Gregory, R.L. (1997). Eye and Brain: The Psychology of Seeing (Princeton University Press).
- Guo, K., Benson, P.J., and Blakemore, C. (2004). Pattern motion is present in V1 of awake but not anaesthetized monkeys. *Eur. J. Neurosci.* *19*, 1055–1066.
- Guo, Z. V., Inagaki, H.K., Daie, K., Druckmann, S., Gerfen, C.R., and Svoboda, K. (2017). Maintenance of persistent activity in a frontal thalamocortical loop. *Nature* *545*, 181–186.
- Haberman, J., and Whitney, D. (2011). Efficient summary statistical representation when change localization fails. *Psychon. Bull. Rev.* *18*, 855–859.
- Hansen, B.J., Chelaru, M.I., and Dragoi, V. (2012). Correlated Variability in Laminar Cortical Circuits. *Neuron* *76*, 590–602.
- Hartley, D.E.H., Dahmen, J.C., King, A.J., and Schnupp, J.W.H. (2011). Binaural sensitivity changes between cortical on and off responses. *J. Neurophysiol.* *106*, 30–43.
- Harvey, B.M., and Dumoulin, S.O. (2011). The Relationship between Cortical Magnification Factor and Population Receptive Field Size in Human Visual Cortex: Constancies in Cortical Architecture. *J. Neurosci.* *31*, 13604–13612.
- Harvey, B.M., Klein, B.P., Petridou, N., and Dumoulin, S.O. (2013). Topographic Representation of Numerosity in the Human Parietal Cortex. *Science* (80-. ). *341*, 1123–1126.
- Haselton, M.G., Nettle, D., and Andrews, P.W. (2015). The Evolution of Cognitive Bias. In *The Handbook of Evolutionary Psychology*, (Hoboken, NJ, USA: John Wiley & Sons, Inc.), pp. 724–746.
- He, J. (2002). OFF responses in the auditory thalamus of the guinea pig. *J. Neurophysiol.* *88*, 2377–2386.
- Hedges, J.H., Stocker, A.A., and Simoncelli, E.P. (2011). Optimal inference explains the perceptual coherence of visual motion stimuli. *J. Vis.* *11*.
- Heeger, D.J., Carandini, M., and Movshon, J.A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *J. Neurosci.* *17*, 8621.
- Heekeren, H.R., Marrett, S., Bandettini, P.A., and Ungerleider, L.G. (2004). A general mechanism for perceptual decision-making in the human brain. *Nature* *431*, 859–862.
- Helmholtz, H. (1867). *Handbuch der physiologischen Optik* (Voss).
- Herzog, M.H., and Clarke, A.M. (2014). Why vision is not both hierarchical and feedforward. *Front. Comput. Neurosci.* *8*, 135.
- Hochstein, S., and Ahissar, M. (2002). View from the top: hierarchies and reverse hierarchies in the visual system. *Neuron* *36*, 791–804.
- Hollingworth, H.L. (1910). The Central Tendency of Judgment. *J. Philos. Psychol. Sci. Methods* *7*, 461.
- Huang, X., Levine, S., and Paradiso, M.A. (2008). Rebounding V1 activity and a new visual aftereffect. *J. Vis.* *8*, 25.1-10.
- Hubel, D.H., and Wiesel, T.N. (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *J. Comp. Neurol.* *146*, 421–450.

- Hupé, J.-M., and Rubin, N. (2003). The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids. *Vision Res.* *43*, 531–548.
- Hutchinson, J.B., Pak, S.S., and Turk-Browne, N.B. (2016). Biased Competition during Long-term Memory Formation. *J. Cogn. Neurosci.* *28*, 187–197.
- Jacobs, R.A. (2002). What determines visual cue reliability? *Trends Cogn. Sci.* *6*, 345–350.
- Jäkel, F., Singh, M., Wichmann, F.A., and Herzog, M.H. (2016). An overview of quantitative approaches in Gestalt perception. *Vision Res.* *126*, 3–8.
- Jehee, J.F.M., Rothkopf, C., Beck, J.M., and Ballard, D.H. (2006). Learning receptive fields using predictive feedback. *J. Physiol.* *100*, 125–132.
- Jellema, T., and Perrett, D.I. (2003). Perceptual History Influences Neural Responses to Face and Body Postures. *J. Cogn. Neurosci.* *15*, 961–971.
- Jensen, O., Bonnefond, M., and VanRullen, R. (2012). An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn. Sci.* *16*, 200–205.
- Jones E. G. (2007). *The Thalamus* (Cambridge University Press).
- Joubert, O.R., Rousselet, G.A., Fize, D., and Fabre-Thorpe, M. (2007). Processing scene context: Fast categorization and object interference. *Vision Res.* *47*, 3286–3297.
- Juavinett, A.L., and Callaway, E.M. (2015). Pattern and Component Motion Responses in Mouse Visual Cortical Areas. *Curr. Biol.* *25*, 1759–1764.
- Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature* *290*, 91–97.
- Jusyte, A., and Schönenberg, M. (2014). Subliminal cues bias perception of facial affect in patients with social phobia: evidence for enhanced unconscious threat processing. *Front. Hum. Neurosci.* *8*, 580.
- Kanizsa, G. (1955). *Margini quasi-percettivi in campi con stimolazione omogenea*.
- Kapadia, M.K., Ito, M., Gilbert, C.D., and Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron* *15*, 843–856.
- Kayser, S.J., McNair, S.W., and Kayser, C. (2016). Prestimulus influences on auditory perception from sensory representations and decision processes. *Proc. Natl. Acad. Sci. U. S. A.* *113*, 4842–4847.
- Keliris, G.A., Logothetis, N.K., and Tolias, A.S. (2010). The Role of the Primary Visual Cortex in Perceptual Suppression of Salient Visual Stimuli. *J. Neurosci.* *30*, 12353–12365.
- van Kerkoerle, T., Self, M.W., and Roelfsema, P.R. (2017). Layer-specificity in the effects of attention and working memory on activity in primary visual cortex. *Nat. Commun.* *8*, 13804.
- Kersten, D., and Yuille, A. (2003). Bayesian models of object perception. *Curr. Opin. Neurobiol.* *13*, 150–158.
- Keysers, C., and Perrett, D.I. (2002). Visual masking and RSVP reveal neural competition. *Trends Cogn. Sci.* *6*, 120–125.
- Khawaja, F.A., Tsui, J.M.G., and Pack, C.C. (2009). Pattern motion selectivity of spiking outputs and local field potentials in macaque visual cortex. *J. Neurosci.* *29*, 13702–13709.
- Klink, P.C., van Wezel, R.J.A., and van Ee, R. (2012). United we sense, divided we fail: context-driven perception of

- ambiguous visual stimuli. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 932–941.
- Kloosterman, N.A., Meindertsma, T., Hillebrand, A., van Dijk, B.W., Lamme, V.A.F., and Donner, T.H. (2015). Top-down modulation in human visual cortex predicts the stability of a perceptual illusion. *J. Neurophysiol.* 113, 1063–1076.
- Knill, D.C., and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719.
- Knill, D.C., and Richards, W. (1996). *Perception as Bayesian Inference* (Cambridge University Press).
- Koffka, K. (1935). *Principles of Gestalt Psychology* (Routledge & Kegan Paul).
- Kohn, A. (2007). Visual Adaptation: Physiology, Mechanisms, and Functional Benefits. *J. Neurophysiol.* 97, 3155–3164.
- Kok, P., and de Lange, F.P. (2015). Predictive Coding in Sensory Cortex. In *An Introduction to Model-Based Cognitive Neuroscience*, (New York, NY: Springer New York), pp. 221–244.
- Kok, P., Jehee, J.F.M., and de Lange, F.P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron* 75, 265–270.
- Kok, P., Brouwer, G.J., van Gerven, M.A.J., and de Lange, F.P. (2013). Prior Expectations Bias Sensory Representations in Visual Cortex. *J. Neurosci.* 33.
- Kok, P., Mostert, P., and Lange, F.P. De (2017). Prior expectations induce pre-stimulus sensory templates. *bioRxiv* 1–26.
- Kooi, F.L., De Valois, K.K., Grosf, D.H., and De Valois, R.L. (1992). Properties of the recombination of one-dimensional motion signals into a pattern motion signal. *Percept. Psychophys.* 52, 415–424.
- Körding, K.P., and Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature* 427, 244–247.
- Kornmeier, J., W?rner, R., Riedel, A., Bach, M., and Tebartz van Elst, L. (2014). A Different View on the Checkerboard? Alterations in Early and Late Visually Evoked EEG Potentials in Asperger Observers. *PLoS One* 9, e90993.
- Krause, M.R., and Pack, C.C. (2014). Contextual modulation and stimulus selectivity in extrastriate cortex. *Vision Res.* 104, 36–46.
- Larson, C.L., Ruffalo, D., Nietert, J.Y., and Davidson, R.J. (2005). Stability of emotion-modulated startle during short and long picture presentation. *Psychophysiology* 42, 604–610.
- Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 451–468.
- Lee, A.L.F., and Lu, H. (2012). Two forms of aftereffects induced by transparent motion reveal multilevel adaptation. *J. Vis.* 12, 3.
- Leopold, D. a., and Logothetis, N.K. (1999). Multistable phenomena: Changing views in perception. *Trends Cogn. Sci.* 3, 254–264.
- Leopold, D.A., Wilke, M., Maier, A., and Logothetis, N.K. (2002). Stable perception of visually ambiguous patterns. *Nat. Neurosci.* 5, 605–609.
- Leopold, D.A., Rhodes, G., Müller, K.-M., and Jeffery, L. (2005). The dynamics of visual adaptation to faces. *Proc. R.*

Soc. London B Biol. Sci. 272, 897–904.

Leske, S., Ruhnau, P., Frey, J., Lithari, C., Müller, N., Hartmann, T., and Weisz, N. (2015). Prestimulus network integration of auditory cortex predisposes near-threshold perception independently of local excitability. *Cereb. Cortex* 25, 4898–4907.

Levick, W.R., and Zacks, J.L. (1970). Responses of cat retinal ganglion cells to brief flashes of light. *J. Physiol.* 206, 677–700.

Levitt, J.B., and Lund, J.S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature* 387, 73–76.

Lewis, E., Lloyd, D.M., and Farrell, M.J. (2012). The role of the environment in eliciting phantom-like sensations in non-amputees. *Front. Psychol.* 3, 600.

Li, Z. (1999). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10530–10535.

Li, H., Liu, X., Andolina, I.M., Li, X., Lu, Y., Spillmann, L., and Wang, W. (2017). Asymmetries of Dark and Bright Negative Afterimages Are Paralleled by Subcortical ON and OFF Poststimulus Responses. *J. Neurosci.* 37.

Limb, J.O., and Murphy, J.A. (1975). Estimating the Velocity of Moving Images in Television Signals. *Comput. Graph. Image Process.* 4, 311–327.

Lin, W.C., Hays, J., Wu, C., Kwatra, V., and Liu, Y. (2006). Quantitative evaluation of near regular texture synthesis algorithms. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (IEEE), pp. 427–434.

Livne, T., and Sagi, D. (2007). Configuration influence on crowding. *J. Vis.* 7, 4.

Di Lollo, V. (1980). Temporal integration in visual memory. *J. Exp. Psychol. Gen.* 109, 75–97.

López-Moliner, J., Smeets, J.B.J., and Brenner, E. (2004). Components of motion perception revealed: two different after-effects from a single moving object. *Vision Res.* 44, 2545–2549.

Lorenceau, J., Shiffrar, M., Wells, N., and Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Res.* 33, 1207–1217.

Lou, B., Li, Y., Philiastides, M.G., and Sajda, P. (2014). Prestimulus alpha power predicts fidelity of sensory encoding in perceptual decision making. *Neuroimage* 87, 242–251.

Louie, K., Khaw, M.W., and Glimcher, P.W. (2013). Normalization is a general neural mechanism for context-dependent decision making. *Proc. Natl. Acad. Sci.* 110, 6139–6144.

Ma, W.J., Husain, M., and Bays, P.M. (2014). Changing concepts of working memory. *Nat. Neurosci.* 17, 347–356.

MacKay, D. (1973). Lateral interaction between neural channels sensitive to texture density? *Nature* 245, 159–161.

Mahani, A.S., Carlsson, A.E., and Wessel, R. (2005). Motion repulsion arises from stimulus statistics when analyzed with a clustering algorithm. *Biol. Cybern.* 92, 288–291.

Mainen, Z.F., and Sejnowski, T.J. (1995). Reliability of spike timing in neocortical neurons. *Science* 268, 1503–1506.

Malcolm, G.L., Groen, I.I.A., and Baker, C.I. (2016). Making Sense of Real-World Scenes. *Trends Cogn. Sci.* 20, 843–856.

- Manassi, M., Sayim, B., and Herzog, M.H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *J. Vis.* *12*, 13–13.
- Marr, D., and Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proc. R. Soc. London. Ser. B, Biol. Sci.* *211*, 151–180.
- Marshak, W., and Sekuler, R. (1979). Mutual repulsion between moving visual targets. *Science* *205*, 1399–1401.
- Martinez-Trujillo, J.C., and Treue, S. (2004). Feature-Based Attention Increases the Selectivity of Population Responses in Primate Visual Cortex. *Curr. Biol.* *14*, 744–751.
- Masson, G.S., and Castet, E. (2002). Parallel motion processing for the initiation of short-latency ocular following in humans. *J. Neurosci.* *22*, 5149–5163.
- Masson, G.S., Rybarczyk, Y., Castet, E., and Mestre, D.R. Temporal dynamics of motion integration for the initiation of tracking eye movements at ultra-short latencies. *Vis. Neurosci.* *17*, 753–767.
- Mather, M., and Harley, C.W. (2016). The Locus Coeruleus: Essential for Maintaining Cognitive Function and the Aging Brain. *Trends Cogn. Sci.* *20*, 214–226.
- Mather, M., and Sutherland, M.R. (2011a). Arousal-Biased Competition in Perception and Memory. *Perspect. Psychol. Sci.* *6*, 114–133.
- Mather, M., and Sutherland, M.R. (2011b). Arousal-Biased Competition in Perception and Memory. *Perspect. Psychol. Sci.* *6*, 114–133.
- Mather, M., Clewett, D., Sakaki, M., and Harley, C.W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behav. Brain Sci.* *39*, e200.
- Matthews, N., Geesaman, B.J., and Qian, N. (2000). The dependence of motion repulsion and rivalry on the distance between moving elements. *Vision Res.* *40*, 2025–2036.
- Maunsell, J.H., and Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. *Annu. Rev. Neurosci.* *10*, 363–401.
- McLelland, D., Ahmed, B., and Bair, W. (2009). Responses to static visual images in macaque lateral geniculate nucleus: implications for adaptation, negative afterimages, and visual fading. *J. Neurosci.* *29*, 8996–9001.
- McLelland, D., Baker, P.M., Ahmed, B., and Bair, W. (2010). Neuronal responses during and after the presentation of static visual stimuli in macaque primary visual cortex. *J. Neurosci.* *30*, 12619–12631.
- McMains, S.A., Fehd, H.M., Emmanouil, T.-A., and Kastner, S. (2007). Mechanisms of Feature- and Space-Based Attention: Response Modulation and Baseline Increases. *J. Neurophysiol.* *98*, 2110–2121.
- Mendoza-Halliday, D., Torres, S., and Martinez-Trujillo, J.C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. *Nat. Neurosci.* *17*, 1255–1262.
- Meso, A.I., Haruhana, K., Masson, G.S., and Gardner, J.L. (2016a). Repulsion of perceived visual motion direction as an emergent property of deciding to unify or segregate sources. (San Diego, CA), p.
- Meso, A.I., Rankin, J., Faugeras, O., Kornprobst, P., and Masson, G.S. (2016b). The relative contribution of noise and adaptation to competition during tri-stable motion perception. *J. Vis.* *16*, 6.
- Mesulam, M. (2008). Representation, inference, and transcendent encoding in neurocognitive networks of the human brain. *Ann. Neurol.* *64*, 367–378.



- Meyer, T., Qi, X.-L., and Constantinidis, C. (2007). Persistent Discharges in the Prefrontal Cortex of Monkeys Naive to Working Memory Tasks. *Cereb. Cortex* *17*, i70–i76.
- Miller, E.K., Erickson, C.A., and Desimone, R. (1996a). Neural mechanisms of visual working memory in prefrontal cortex of the macaque. *J. Neurosci.* *16*, 5154–5167.
- Miller, E.K., Erickson, C.A., and Desimone, R. (1996b). Neural Mechanisms of Visual Working Memory in Prefrontal Cortex of the Macaque. *J. Neurosci.* *16*.
- Morales, J., Solovey, G., Maniscalco, B., Rahnev, D., de Lange, F.P., and Lau, H. (2015). Low attention impairs optimal incorporation of prior knowledge in perceptual decisions. *Atten. Percept. Psychophys.* *77*, 2021–2036.
- Moreno-Bote, R., Shpiro, A., Rinzel, J., and Rubin, N. (2010). Alternation rate in perceptual bistability is maximal at and symmetric around equi-dominance. *J. Vis.* *10*, 1.
- Morgan, M.J., Mareschal, I., Chubb, C., and Solomon, J.A. (2012). Perceived pattern regularity computed as a summary statistic: implications for camouflage. *Proc. R. Soc. B Biol. Sci.* *279*, 2754–2760.
- Morrone, M.C., Burr, D.C., and Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proc. R. Soc. London. Ser. B, Biol. Sci.* *216*, 335–354.
- Motter, B.C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J. Neurophysiol.* *70*, 909–919.
- Moutsiana, C., de Haas, B., Papageorgiou, A., van Dijk, J.A., Balraj, A., Greenwood, J.A., and Schwarzkopf, D.S. (2016). Cortical idiosyncrasies predict the perception of object size. *Nat. Commun.* *7*, 12110.
- Movshon, J.A., Adelson, E.H., Gizzi, M.S., and Newsome, W.T. (1985). The analysis of moving visual patterns. *Pontif. Acad. Sci. Scr. Varia* *54*, 117–151.
- Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biol. Cybern.* *66*, 241–251.
- Murphy, P.R., Robertson, I.H., Harty, S., and O’Connell, R.G. (2015). Neural evidence accumulation persists after choice to inform metacognitive judgments. *Elife* *4*.
- Nakamura, K., and Colby, C.L. (2002). Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proc. Natl. Acad. Sci.* *99*, 4026–4031.
- Narasimhan, S., and Giaschi, D. (2012). The effect of dot speed and density on the development of global motion perception. *Vision Res.* *62*, 102–107.
- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cogn. Psychol.* *9*, 353–383.
- Navon, D. (1981). The forest revisited: More on global precedence. *Psychol. Res.* *43*, 1–32.
- Neisser, U. (2014). *Cognitive Psychology: Classic Edition* (Taylor & Francis).
- Neri, P. (2010). How inherently noisy is human sensory processing? *Psychon. Bull. Rev.* *17*, 802–808.
- New, J., Cosmides, L., and Tooby, J. (2007). Category-specific attention for animals reflects ancestral priorities, not expertise. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 16598–16603.
- Nishida, S. (2011). Advancement of motion psychophysics: review 2001-2010. *J. Vis.* *11*, 11.

- Noest, A.J., and van den Berg, A. V (1993). The role of early mechanisms in motion transparency and coherence. *Spat. Vis.* 7, 125–147.
- Ohshiro, T., Angelaki, D.E., and DeAngelis, G.C. (2011). A normalization model of multisensory integration. *Nat. Neurosci.* 14, 775–782.
- Olsen, S.R., Bortone, D.S., Adesnik, H., and Scanziani, M. (2012). Gain control by layer six in cortical circuits of vision. *Nature* 483, 47–52.
- Orban, G.A. (1984). *Neuronal Operations in the Visual Cortex* (Berlin, Heidelberg: Springer Berlin Heidelberg).
- Ouhanna, M., Bell, J., Solomon, J.A., and Kingdom, F.A.A. (2013). Aftereffect of perceived regularity. *J. Vis.* 13, 18–18.
- Pack, C.C., and Born, R.T. (2001). Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. *Nature* 409, 1040–1042.
- Pack, C., Berezovskii, V., and Born, R. (2001). Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. *Nature* 414, 905–908.
- Pack, C.C., Livingstone, M.S., Duffy, K.R., and Born, R.T. (2003). End-Stopping and the Aperture Problem. *Neuron* 39, 671–680.
- Palmer, Stephen E. (1975). The effects of contextual scenes on the identification of objects. *Mem. Cognit.* 3, 519–526.
- Palmer, S.E. (1999). *Vision science : photons to phenomenology* (MIT Press).
- Palmer, E.M., Fencsik, D.E., Flusberg, S.J., Horowitz, T.S., and Wolfe, J.M. (2011). Signal detection evidence for limited capacity in visual search. *Atten. Percept. Psychophys.* 73, 2413–2424.
- Parkes, L., Lund, J., Angelucci, A., Solomon, J.A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* 4, 739–744.
- Pasternak, T., and Greenlee, M.W. (2005). Working memory in primate sensory systems. *Nat. Rev. Neurosci.* 6, 97–107.
- Peelen, M. V., Fei-Fei, L., and Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature* 460, 94–97.
- Pei, Y.C., Hsiao, S.S., and Bensmaia, S.J. (2008). The tactile integration of local motion cues is analogous to its visual counterpart. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8130–8135.
- Pessoa, L., and Adolphs, R. (2010). Emotion processing and the amygdala: from a “low road” to “many roads” of evaluating biological significance. *Nat. Rev. Neurosci.* 11, 773–783.
- Phillips, W.A., and Routh, D.A. (1983). Short-Term Visual Memory [and Discussion]. *Philos. Trans. R. Soc. London B Biol. Sci.* 302.
- Phillips, W.A., Clark, A., and Silverstein, S.M. (2015). On the functions, mechanisms, and malfunctions of intracortical contextual modulation. *Neurosci. Biobehav. Rev.* 52, 1–20.
- Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision Res.* 41, 3145–3161.
- Platt, M.L., and Glimcher, P.W. (1999). Neural correlates of decision variables in parietal cortex. *Nature* 400, 233–238.
- Polat, U., and Sagi, D. (1994). Spatial interactions in human vision: from near to far via experience-dependent cascades of connections. *Proc. Natl. Acad. Sci. U. S. A.* 91, 1206–1209.

- Pouget, a, Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nat. Rev. Neurosci.* *1*, 125–132.
- Prinzmetal, W., and Banks, W.P. (1983). Perceptual capacity limits in visual detection and search. *Bull. Psychon. Soc.* *21*, 263–266.
- Pun, C., Emrich, S.M., Wilson, K.E., Stergiopoulos, E., and Ferber, S. (2012). In and out of consciousness: sustained electrophysiological activity reflects individual differences in perceptual awareness. *Psychon. Bull. Rev.* *19*, 429–435.
- Putzeys, T., Bethge, M., Wichmann, F., Wagemans, J., and Goris, R. (2012). A new perceptual bias reveals suboptimal population decoding of sensory responses. *PLoS Comput. Biol.* *8*, e1002453.
- Qian, N., and Andersen, R.A. (1994). Transparent motion perception as detection of unbalanced motion signals. II. Physiology. *J. Neurosci.* *14*, 7367–7380.
- Qian, N., Andersen, R. a, and Adelson, E.H. (1994). Transparent motion perception as detection of unbalanced motion signals. III. Modeling. *J. Neurosci.* *14*, 7381–7392.
- Qin, L., Chimoto, S., Sakai, M., Wang, J., and Sato, Y. (2007). Comparison Between Offset and Onset Responses of Primary Auditory Cortex on–off Neurons in Awake Cats. *J. Neurophysiol.* *97*.
- Ramachandran, V.S. (1988). Perception of shape from shading. *Nature* *331*, 163–166.
- Rao, R.P.N., Olshausen, B.A., and Lewicki, M.S. (2002). Probabilistic models of the brain : perception and neural function (Cambridge, Mass., London: The Mit Press).
- Rensink, R.A. (2000). Seeing, sensing, and scrutinizing. *Vision Res.* *40*, 1469–1487.
- Repp, B.H. (2007). Hearing a melody in different ways: Multistability of metrical interpretation, reflected in rate limits of sensorimotor synchronization. *Cognition* *102*, 434–454.
- Reynolds, J.H., and Chelazzi, L. (2004). ATTENTIONAL MODULATION OF VISUAL PROCESSING. *Annu. Rev. Neurosci.* *27*, 611–647.
- Reynolds, J.H., and Desimone, R. (2003). Interacting roles of attention and visual salience in V4. *Neuron* *37*, 853–863.
- Reynolds, J.H., Chelazzi, L., and Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *J. Neurosci.* *19*, 1736–1753.
- Rider, A.T., Nishida, S., and Johnston, A. (2016). Multiple-stage ambiguity in motion perception reveals global computation of local motion directions. *J. Vis.* *16*, 1–11.
- Roach, N.W., McGraw, P. V, Whitaker, D.J., and Heron, J. (2017). Generalization of prior information for rapid Bayesian time estimation. *Proc. Natl. Acad. Sci. U. S. A.* *114*, 412–417.
- Rolls, E.T. (2016). *Cerebral Cortex: Principles of Operation* (Oxford University Press).
- Rust, N.C., and Stocker, A.A. (2010). Ambiguity and invariance: two fundamental challenges for visual processing. *Curr. Opin. Neurobiol.* *20*, 382–388.
- Saarela, T.P., Westheimer, G., and Herzog, M.H. (2010). The effect of spacing regularity on visual crowding. *J. Vis.* *10*, 17–17.
- Sadaghiani, S., Poline, J.-B., Kleinschmidt, A., and D’Esposito, M. (2015). Ongoing dynamics in large-scale functional connectivity predict perception. *Proc. Natl. Acad. Sci.* *112*, 201420687.

- Sato, Y., Toyoizumi, T., and Aihara, K. (2007). Bayesian Inference Explains Perception of Unity and Ventriloquism Aftereffect: Identification of Common Sources of Audiovisual Stimuli. *Neural Comput.* *19*, 3335–3355.
- Sayim, B., Westheimer, G., and Herzog, M.H. (2010). Gestalt Factors Modulate Basic Spatial Vision. *Psychol. Sci.* *21*, 641–644.
- Scaglione, A., Moxon, K.A., Aguilar, J., and Foffani, G. (2011). Trial-to-trial variability in the responses of neurons carries information about stimulus location in the rat whisker thalamus. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 14956–14961.
- Scarfe, P., and Johnston, A. (2011). Global motion coherence can influence the representation of ambiguous local motion. *J. Vis.* *11*, 6.
- Schiller, P.H. (1968). Single unit analysis of backward visual masking and metacontrast in the cat lateral geniculate nucleus. *Vision Res.* *8*, 855–866.
- Schmitt, L.I., Wimmer, R.D., Nakajima, M., Happ, M., Mofakham, S., and Halassa, M.M. (2017). Thalamic amplification of cortical connectivity sustains attentional control. *Nature* *545*, 219–223.
- Schneider, W., and Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* *84*, 1–66.
- Schwartz, J.-L., Grimault, N., Hupé, J.-M., Moore, B.C.J., and Pressnitzer, D. (2012). Multistability in perception: binding sensory modalities, an overview. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* *367*, 896–905.
- Schwartz, O., Hsu, A., and Dayan, P. (2007). Space and time in visual context. *Nat. Rev. Neurosci.* *8*, 522–535.
- Schwarz, K.A., Pfister, R., and Büchel, C. (2016a). Rethinking Explicit Expectations: Connecting Placebos, Social Cognition, and Contextual Perception. *Trends Cogn. Sci.* *20*, 469–480.
- Schwarz, K.A., Pfister, R., and Büchel, C. (2016b). Rethinking Explicit Expectations: Connecting Placebos, Social Cognition, and Contextual Perception. *Trends Cogn. Sci.* *20*, 469–480.
- Schwarzkopf, D.S., Song, C., and Rees, G. (2011). The surface area of human V1 predicts the subjective experience of object size. *Nat. Neurosci.* *14*, 28–30.
- Schwarzkopf, D.S., Anderson, E.J., de Haas, B., White, S.J., and Rees, G. (2014). Larger Extrastriate Population Receptive Fields in Autism Spectrum Disorders. *J. Neurosci.* *34*, 2713–2724.
- Serences, J.T., and Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends Cogn. Sci.* *10*, 38–45.
- Serences, J.T., Yantis, S., Culbertson, A., and Awh, E. (2004). Preparatory Activity in Visual Cortex Indexes Distractor Suppression During Covert Spatial Orienting. *J. Neurophysiol.* *92*, 3538–3545.
- Sergent, C., and Dehaene, S. (2004). Neural processes underlying conscious perception: experimental findings and a global neuronal workspace framework. *J. Physiol. Paris* *98*, 374–384.
- Seriès, P., and Seitz, A.R. (2013). Learning what to expect (in visual perception). *Front. Hum. Neurosci.* *7*, 668.
- Shadlen, M.N., and Kim, J.-N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* *2*, 176–185.
- Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* *18*, 3870–3896.
- Sheinberg, D.L., and Logothetis, N.K. (1997). The role of temporal cortical areas in perceptual organization. *Proc. Natl.*

- Acad. Sci. *94*, 3408–3413.
- Shiffrin, R.M., and Gardner, G.T. (1972). Visual processing capacity and attentional control. *J. Exp. Psychol.* *93*, 72–82.
- Singer, J.M., Sheinberg, D.L., C., B.D., Y., C.A., and K., N. (2008). A method for the real-time rendering of formless dot field structure-from-motion stimuli. *J. Vis.* *8*, 8.
- Smith, M.A., Majaj, N.J., and Movshon, J.A. (2005). Dynamics of motion signaling by neurons in macaque area MT. *Nat. Neurosci.* *8*, 220–228.
- Smith, M.A., Majaj, N., and Movshon, J.A. (2010). Dynamics of pattern motion computation. In *Dynamics of Visual Motion Processing: Neuronal, Behavioral, and Computational Approaches*, (Boston, MA: Springer US), pp. 55–72.
- Song, C., Schwarzkopf, D.S., and Rees, G. (2013a). Variability in visual cortex size reflects tradeoff between local orientation sensitivity and global orientation modulation. *Nat. Commun.* *4*, 2201.
- Song, C., Schwarzkopf, D.S., Lutti, A., Li, B., Kanai, R., and Rees, G. (2013b). Effective connectivity within human primary visual cortex predicts interindividual diversity in illusory perception. *J. Neurosci.* *33*, 18781–18791.
- Song, C., Schwarzkopf, D.S., Kanai, R., and Rees, G. (2015). Neural population tuning links visual cortical anatomy to human visual perception. *Neuron* *85*, 641–656.
- Soon, C.S., Dubey, R., Ananyev, E., and Hsieh, P.-J. (2017). Approaches to Understanding Visual Illusions. pp. 221–233.
- Sotiropoulos, G., Seitz, A.R., Seriès, P., and Shams, L. (2011). Changing expectations about speed alters perceived motion direction. *Curr. Biol.* *21*, R883-4.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr. Gen. Appl.* *74*, 1–29.
- Spoehr, K.T., and Lehmkuhle, S.W. (1982). *Visual Information Processing* (W.H. Freeman).
- Spratling, M.W. (2008). Reconciling predictive coding and biased competition models of cortical function. *Front. Comput. Neurosci.* *2*, 4.
- St John-Saaltink, E., Kok, P., Lau, H.C., and de Lange, F.P. (2016). Serial Dependence in Perceptual Decisions Is Reflected in Activity Patterns in Primary Visual Cortex. *J. Neurosci.* *36*, 6186–6192.
- Stanley, J., Forte, J.D., Cavanagh, P., and Carter, O. (2011). Onset rivalry: the initial dominance phase is independent of ongoing perceptual alternations. *Front. Hum. Neurosci.* *5*, 140.
- Stansbury, D.E., Naselaris, T., and Gallant, J.L. (2013). Natural scene statistics account for the representation of scene categories in human visual cortex. *Neuron* *79*, 1025–1034.
- Stocker, A.A., and Simoncelli, E.P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* *9*, 578–585.
- Stocker, A.A., and Simoncelli, E.P. (2007). A Bayesian Model of Conditioned Perception. *Adv. Neural Inf. Process. Syst.* *2007*, 1409–1416.
- Stokes, D., Matthen, M., and Biggs, S. (2014). *Perception and Its Modalities* (Oxford University Press).
- Stone, L.S., Watson, a B., and Mulligan, J.B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Res.* *30*, 1049–1067.
- Stoner, G.R., and Albright, T.D. (1996). The interpretation of visual motion: evidence for surface segmentation

- mechanisms. *Vision Res.* *36*, 1291–1310.
- Stoner, G.R., Albright, T.D., and Ramachandran, V.S. (1990). Transparency and coherence in human motion perception. *Nature* *344*, 153–155.
- Styles, E.A. (2005). *Attention, Perception and Memory: An Integrated Introduction* (Psychology Press).
- Summerfield, C., and Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends Cogn. Sci.* *13*, 403–409.
- Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., and Hirsch, J. (2006). Predictive Codes for Forthcoming Perception in the Frontal Cortex. *Science* (80- ). *314*.
- Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.-M., and Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nat. Neurosci.* *11*, 1004–1006.
- Sun, J., and Perona, P. (1998). Where is the sun? *Nat. Neurosci.* *1*, 183–184.
- Sun, H.-C., Baker, C.L., Kingdom, F.A.A., T., V., and L, B.C. (2016). Simultaneous density contrast is bidirectional. *J. Vis.* *16*, 4.
- Super, H., Spekreijse, H., and Lamme, V.A. (2001). A Neural Correlate of Working Memory in the Monkey Primary Visual Cortex. *Science* (80- ). *293*, 120–124.
- Supèr, H., Spekreijse, H., and Lamme, V. a (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nat. Neurosci.* *4*, 304–310.
- Sussman, T.J., Jin, J., and Mohanty, A. (2016a). Top-down and bottom-up factors in threat-related perception and attention in anxiety. *Biol. Psychol.* *121*, 160–172.
- Sussman, T.J., Szekely, A., Hajcak, G., and Mohanty, A. (2016b). It's all in the anticipation: How perception of threat is enhanced in anxiety. *Emotion* *16*, 320–327.
- Sussman, T.J., Weinberg, A., Szekely, A., Hajcak, G., and Mohanty, A. (2017). Here Comes Trouble: Prestimulus Brain Activity Predicts Enhanced Perception of Threat. *Cereb. Cortex* *27*, 2695–2707.
- Sylvester, C.M., Shulman, G.L., Jack, A.I., and Corbetta, M. (2009). Anticipatory and Stimulus-Evoked Blood Oxygenation Level-Dependent Modulations Related to Spatial Attention Reflect a Common Additive Signal. *J. Neurosci.* *29*, 10671–10682.
- Tajadura-Jiménez, A., Marquardt, T., Swapp, D., Kitagawa, N., and Bianchi-Berthouze, N. (2016). Action Sounds Modulate Arm Reaching Movements. *Front. Psychol.* *7*, 1391.
- Takahashi, N. (2004). Effect of spatial configuration of motion signals on motion integration across space. *Swiss J. Psychol.* *63*, 173–182.
- Takahashi, H., Nakao, M., and Kaga, K. (2004). Cortical mapping of auditory-evoked offset responses in rats. *Neuroreport* *15*, 1565–1569.
- Takemura, H., Tajima, S., and Murakami, I. (2011). Whether dots moving in two directions appear coherent or transparent depends on directional biases induced by surrounding motion. *J. Vis.* *11*, 17.
- Talcott, J.B., Hansen, P.C., Assoku, E.L., and Stein, J.F. (2000). Visual motion sensitivity in dyslexia: evidence for temporal and energy integration deficits. *Neuropsychologia* *38*, 935–943.
- Tark, K.J., and Curtis, C.E. (2009). Persistent neural activity in the human frontal cortex when maintaining space that is off the map. *Nat Neurosci* *12*, 1463–1468.

- Thiele, A., and Stoner, G. (2003). Neuronal synchrony does not correlate with motion coherence in cortical area MT. *Nature* 421, 366–370.
- Thompson, K.G., and Schall, J.D. (1999). The detection of visual signals by macaque frontal eye field during masking. *Nat. Neurosci.* 2, 283–288.
- Tinsley, C.J., Webb, B.S., Barraclough, N.E., Vincent, C.J., Parker, A., and Derrington, A.M. (2003). The Nature of V1 Neural Responses to 2D Moving Patterns Depends on Receptive-Field Structure in the Marmoset Monkey. *J. Neurophysiol.* 90, 930–937.
- Todd, J.T., and Thaler, L. (2010). The perception of 3D shape from texture based on directional width gradients. *J. Vis.* 10, 17.
- Todorovic, A., van Ede, F., Maris, E., and de Lange, F.P. (2011). Prior Expectation Mediates Neural Adaptation to Repeated Sounds in the Auditory Cortex: An MEG Study. *J. Neurosci.* 31.
- Tomassini, A., Morgan, M.J., and Solomon, J.A. (2010). Orientation uncertainty reduces perceived obliquity. *Vision Res.* 50, 541–547.
- Trawalter, S., Todd, A.R., Baird, A.A., and Richeson, J.A. (2008). Attending to Threat: Race-based Patterns of Selective Attention. *J. Exp. Soc. Psychol.* 44, 1322–1327.
- Trueswell, J.C., and Hayhoe, M.M. (1993). Surface segmentation mechanisms and motion perception. *Vision Res.* 33, 313–328.
- Tsotsos, J.K. (1990). Analyzing vision at the complexity level. *Behav. Brain Sci.* 13, 423–445.
- Tsotsos, J.K. (1997). Limited Capacity of Any Realizable Perceptual System Is a Sufficient Reason for Attentive Behavior. *Conscious. Cogn.* 6, 429–436.
- Verstraten, F.A., Fredericksen, R.E., and van de Grind, W.A. (1994). Movement aftereffect of bi-vectorial transparent motion. *Vision Res.* 34, 349–358.
- Verstraten, F.A., van der Smagt, M.J., Fredericksen, R.E., and van de Grind, W.A. (1999). Integration after adaptation to transparent motion: static and dynamic test patterns result in different aftereffect directions. *Vision Res.* 39, 803–810.
- Victor, J.D., Purpura, K., Katz, E., and Mao, B. (1994). Population encoding of spatial frequency, orientation, and color in macaque V1. *J. Neurophysiol.* 72, 2151–2166.
- Vidnyánszky, Z., Blaser, E., and Papathomas, T. V. (2002). Motion integration during motion aftereffects. *Trends Cogn. Sci.* 6, 157–161.
- Vierordt, K. (1868). *Der Zeitsinn nach Versuchen.*
- Volkov, I.O., and Galazjuk, A. V (1991). Formation of spike response to sound tones in cat auditory cortex neurons: interaction of excitatory and inhibitory effects. *Neuroscience* 43, 307–321.
- Wagemans, J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M., and von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychol. Bull.* 138, 1172–1217.
- Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychol. Forsch.* 20, 325–380.
- Warren, R.M., and Gregory, R.L. (1958). An Auditory Analogue of the Visual Reversible Figure. *Am. J. Psychol.* 71, 612.

- Watson, T.L., Otsuka, Y., and Clifford, C.W.G. (2016). Who are you expecting? Biases in face perception reveal prior expectations for sex and age. *J. Vis.* *16*, 5.
- Webb, B.S., Ledgeway, T., and McGraw, P. V (2010). Relating spatial and temporal orientation pooling to population decoding solutions in human vision. *Vision Res.* *50*, 2274–2283.
- Wei, X.-X., and Stocker, A.A. (2015). A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. *Nat. Neurosci.* *18*, 1509–1517.
- Weiss, Y., Simoncelli, E.P., and Adelson, E.H. (2002). Motion illusions as optimal percepts. *Nat. Neurosci.* *5*, 598–604.
- Welchman, A.E., and Harris, J.M. (2000). The effects of dot density and motion coherence on perceptual fading of a target in noise. *Spat. Vis.* *14*, 45–58.
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.* *61*, 161–265.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.* *4*, 301–350.
- Wimmer, G.E., and Shohamy, D. (2012). Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science* *338*, 270–273.
- Wuerger, S., Shapley, R., and Rubin, N. (1996). “On the visually perceived direction of motion” by Hans Wallach: 60 years later. *Perception* *25*, 1317–1367.
- Yang, J., Kanazawa, S., Yamaguchi, M.K., and Motoyoshi, I. (2015). Pre-constancy Vision in Infants. *Curr. Biol.* *25*, 3209–3212.
- Yantis, S. (2001). *Visual Perception: Essential Readings* (Psychology Press).
- Yu, R.Q., and Zhao, J. (2015). The persistence of the attentional bias to regularities in a changing environment. *Attention, Perception, Psychophys.* *77*, 2217–2228.
- Zeelenberg, R., Wagenmakers, E.J., and Rotteveel, M. (2006). The impact of emotion on perception: Bias or enhanced processing? *Psychol. Sci.* *17*, 287–291.
- Zelano, C., Mohanty, A., and Gottfried, J.A. (2011). Olfactory predictive codes and stimulus templates in piriform cortex. *Neuron* *72*, 178–187.
- Zentall, T.R., and Wasserman, E.A. (2012). *The Oxford Handbook of Comparative Cognition* (Oxford University Press, USA).
- Zhang, X., Xu, Q., Jiang, Y., and Wang, Y. (2017). The interaction of perceptual biases in bistable perception. *Nat. Publ. Gr.* *7*, 1–8.
- Zhao, J., Al-Aidroos, N., and Turk-Browne, N.B. (2013). Attention is spontaneously biased toward regularities. *Psychol. Sci.* *24*, 667–677.
- Zhou, W., and Chen, D. (2009). Binaral Rivalry between the Nostrils and in the Cortex. *Curr. Biol.* *19*, 1561–1565.
- Zoccolan, D., Cox, D.D., and Benucci, A. (2015). Editorial: What can simple brains teach us about how vision works. *Front. Neural Circuits* *9*, 51.



## ACKNOWLEDGEMENTS

I thank Prof. Nikos Logothetis who has supported and supervised my research at MPI. Without him, the projects and my PhD were impossible to be finished. I thank Prof. Horst Herbert for his kindness and support during my PhD, he helped me out every time when I was in trouble. I also thank Prof. Uwe Ilg, as one of my double-PhD supervisors, he was always supportive.

I thank Prof. Georgios Keliris for his direct supervising of my PhD study in all aspects, I appreciate the opportunity to work closely with him. I thank Dr. Andrew Meso who taught me psychophysics and modelling with great patience, he provided me valuable advices and suggestions during our collaboration. I thank Prof. Roland Fleming, it's unforgettable about what I learned from him when we met at Giessen and the BCCN chalk talk. I thank Prof. Stelios Smirnakis who opened and broadened my scope of monkey research when we were talking at MPI. I thank Dr. Masataka Watanabe for his help, suggestions and interesting discussions.

I thank my friend Dr. Frederico Azevedo for his encouragement and help like always. Many thanks to my friend Joachim Werner for always being there to help.

I also thank all my friends and colleagues: Conchy Moya, Adi, Deniz Ipek, Mirko Lindig, Thomas Steudel, Ricardo Neves, Ulrike Passlack, Amalia Papanikolau, Carsten Klein, Daniel Zaldivar, Vishal Kapoor, Silvia Van Keulen, Andriani Rina, Renee Hartig, Yibin Shao, Maosen Wang, Yi He, Juan Felipe, Hamed Bahmani, Michael Ortiz-Rios, Matthias Munk, Jennifer Smuda, Hannes Goehring, Marcel Henni, Silvia Slesiona-Kuenzel, Yusuke Murayama. I am sincerely grateful to all of you who has inspired, encouraged and helped me in ways large and small during my years in Tuebingen.