



THE UNIVERSITY *of* EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.

This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

ACQUISITION AND INFLUENCE OF EXPECTATIONS
ABOUT VISUAL SPEED

GRIGORIOS SOTIROPOULOS



Doctor of Philosophy

Institute for Adaptive and Neural Computation

School of Informatics

University of Edinburgh

2014

ABSTRACT

It has been long hypothesized that due to the inherent ambiguities of visual input and the limitations of the visual system, vision is a form of “unconscious inference” whereby the brain relies on assumptions (aka expectations) to interpret the external world. This hypothesis has been recently formalized into Bayesian models of perception (the “Bayesian brain”) that represent these expectations as prior probabilities. In this thesis, I focus on a particular kind of expectation that humans are thought to possess – that objects in the world tend to be still or move slowly – known as the “slow speed prior”. Through a combination of experimental and theoretical work, I investigate how the speed prior is acquired and how it impacts motion perception. The first part of my work consists of an experiment where subjects are exposed to simple “training” stimuli moving more often at high speeds than at low speeds. By subsequently testing the subjects with slow-moving stimuli of high uncertainty (low contrast), I find that their perception gradually changes in a manner consistent with the progressive acquisition of an expectation that favours progressively higher speeds. Thus subjects appear to gradually internalize the speed statistics of the stimulus ensemble over the duration of the experiment. I model these results using an existing Bayesian model of motion perception that incorporates a speed prior with a peak at zero, extending the model so that the mean gradually shifts away from zero. Although the first experiment presents evidence for the plasticity of the speed prior, the experimental paradigm and the constraints of the model limit the accuracy and precision in the reconstruction of observers’ priors. To address these limitations, I perform a different experiment where subjects compare the speed of moving gratings of different contrasts. The new paradigm allows more precise measurements of the contrast-dependent biases in perceived speed. Using a less constrained Bayesian model, I extract the priors of subjects and find considerable interindividual variability. Furthermore, noting that the Bayesian model cannot account for certain subtleties in the data, I combine the model with a non-Bayesian, physiologically motivated model of speed tuning of cortical neurons and show that the combination offers an improved description of the data. Using the paradigm of the second experiment, I then explore the role of visual experience on the form of the speed prior. By recruiting avid video gamers (who are routinely exposed to high speeds) and nongamers of both sexes, I study the differences in the prior among groups and find, surprisingly, that subjects’ speed priors depend more on gender than on gaming experience. In a final series of experiments similar to the first, I also test subjects on variations of the trained stimulus configuration – namely different orientations and motion directions. Subjects’ responses suggest that they are able to apply the changed prior to different orientations and, furthermore, that the changed prior persists for at least a week after the end of the experiment. These results provide further support for the plasticity of the speed prior but also suggest that the learned prior may be used only across similar stimulus configurations, whereas in sufficiently different configurations or contexts a “default” prior may be used instead.

ACKNOWLEDGEMENTS

I would like to express my gratitude towards my principal supervisor, Peggy Seriès, for her guidance, insight and invaluable feedback; for being attentive and involved in my work, demanding high quality but letting me pursue my research interests; and for being patient with my disorganization. Many thanks go to Aaron Seitz for his valuable guidance on the experimental part of my work and for a fruitful collaboration. I also thank Alan Stocker for providing me with additional details about his published Bayesian model (implemented in this thesis) and Dan Berbec for his contribution to the work of Chapter 5. Lastly, I thank Dagmara Panas for meticulously reading my thesis and for her numerous comments that significantly improved clarity and readability.

DECLARATION

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

Grigorios Sotiropoulos

CONTENTS

1	INTRODUCTION	1
1.1	Perception as Bayesian Inference	2
1.2	Overview of thesis	5
2	BACKGROUND AND RESEARCH MOTIVATION	6
2.1	Expectations in perception	6
2.1.1	Types of expectations	6
2.1.2	Impact of expectations	9
2.1.3	Expectations and environmental statistics	11
2.2	Outstanding questions and research motivation	15
2.2.1	Can long-term priors be changed?	15
2.2.2	How specific are priors?	17
2.2.3	Research focus: speed prior	21
3	SPEED PRIOR CHANGE AND ITS EFFECT ON PERCEIVED MOTION DIREC- TION	30
3.1	Introduction	30
3.2	Methods	31
3.2.1	Experiment	31
3.2.2	Bayesian model	34
3.3	Results	37
3.3.1	Experiment	37
3.3.2	Modelling	41
3.4	Discussion	44
4	SPEED PRIOR RECONSTRUCTION AND INTERACTIVE EFFECTS OF CON- TRAST	46
4.1	Introduction	46
4.2	Psychophysical experiment	48
4.2.1	Methods	48
4.2.2	Results	50
4.3	Modelling	51

4.3.1	Methods	51
4.3.2	Results	55
4.4	Discussion	60
5	SPEED PRIOR VARIABILITY AND VISUAL EXPERIENCE	65
5.1	Introduction	65
5.2	Psychophysical experiments	66
5.2.1	Methods	66
5.2.2	Results	68
5.3	Modelling	74
5.3.1	Methods	74
5.3.2	Results	74
5.4	Discussion	77
6	SPECIFICITY AND ROBUSTNESS OF THE LEARNED SPEED PRIOR	82
6.1	Introduction	82
6.2	General methods	82
6.3	Experiment 1	83
6.3.1	Methods	83
6.3.2	Results and discussion	84
6.4	Experiment 2	88
6.4.1	Methods	88
6.4.2	Results and discussion	89
6.5	Experiment 3	93
6.5.1	Methods	93
6.5.2	Results and discussion	94
6.6	General discussion	98
7	DISCUSSION	101
7.1	Limitations	101
7.1.1	Reconstruction of the changing prior	101
7.1.2	Bayesian models	102
7.2	Relation to other work	104
7.2.1	Motion direction perception and locus of the speed prior	104
7.2.2	Specificity of priors	108
7.2.3	Relation to other types of learning	110

7.3	Conclusions	111
8	APPENDIX	126
8.1	Supplementary Procedures of Chapter 4	126
8.1.1	Derivation of Equation 17	126
8.1.2	Model validation dataset	126
8.1.3	PSE Variability	127
8.1.4	Derivation of Equation 20	128
8.2	Supplementary Figures of Chapter 4	129
8.3	Supplementary Figures of Chapter 5	130

INTRODUCTION

Throughout its life, an animal is faced with uncertainty about the external environment. In the case of vision, the projection of a 3D physical stimulus onto the retina most often results in a loss of information about the true properties of the stimulus, i.e. there is inherent ambiguity in the information entering the brain. Thus it is often the case that multiple physical stimuli give rise to the same retinal image. For example, an object forming an elliptical pattern on the retina may indeed have an elliptical shape but it may also be a disc viewed with a slant (Figure 1). How does the visual system “choose” whether to see an ellipse or a slanted disc?

Furthermore, even in the absence of external ambiguity, internal neural noise or physical limitations of sensory organs, such as limitations in the optics of the eye or in retinal resolution, may also result in information loss and prevent the brain from detecting details in the visual scene that are necessary to disambiguate it.

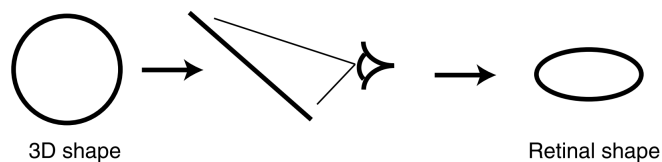


Figure 1: An elliptical pattern forming on the retina can be either an ellipse viewed upright or a circle viewed at an angle. (image adapted from Knill, 2007)

Confronted with these types of uncertainty, the brain must somehow make a guess about the external world; that is, it has to estimate the identity, location and other properties of the objects and processes that generate the sensory input. This estimation may rely on sensory cues (from the same or different modalities) but also on assumptions, or *expectations*, about the external world. Due to the fact that the brain uses cues and expectations in order to *infer* the percept, perception has been characterized as a form of “unconscious inference” – a hypothesis put forth by Helmholtz¹

¹ In contrast, forward inference refers to the computation of the 2D projection from the physical stimulus. In visual perception, this amounted to the measurement of the stimulus. However, for the measurement to be of any ecological value, there must be a one-to-one correspondence between the physical stimulus

as early as in 1867 (reprinted in von Helmholtz, 1962). According to the Helmholtzian view, vision is an instance of *inverse* inference, whereby the visual system estimates the true properties of a physical stimulus from its 2D projection on the retina, i.e. *inverts* the process of the mapping of the stimulus onto the retina, with the help of expectations about the properties of the stimulus.

1.1 PERCEPTION AS BAYESIAN INFERENCE

Helmholtz’s view of perception as unconscious inference has seen a resurgence in popularity in recent years in the form of the “Bayesian brain” hypothesis (Knill and Pouget, 2004; Ma et al., 2006). According to this hypothesis, the brain combines sensory input and prior expectations in a principled, probabilistically optimal² way that is akin to Bayesian inference (Kersten et al., 2004) – the method of inference that uses Bayes’ rule to update the probability distribution of a hypothesis based on accumulated evidence. In perception, the hypothesis is a particular property of the world (call it S) whereas the accumulated evidence is the sensory input (call it I). As an example, I could correspond to a matrix of light intensity (luminance) values at each point of the visual field. In Bayesian statistics, the conditional probability of the input given the state of the world $P(I|S)$ is called the *likelihood function*. Inference amounts to computing the *posterior probability* $P(S|I)$ (or simply the *posterior*) of a property of the world given the image I . According to Bayes’ rule, the posterior is given by

$$P(S|I) = \frac{P(I|S)P(S)}{P(I)} \quad (1)$$

where $P(S)$ is the *prior probability* of S (or simply the *prior*). $P(S)$ represents mathematically one’s *expectations* about what S might be prior to receiving any data. $P(I|S)$ and $P(S)$ represent the observer’s knowledge about the statistical structure of the world. This knowledge is called the *generative model* (Kersten et al., 2004; Vilares and Kording, 2011). $P(I)$ is the prior probability of the image; it does not depend on S and serves as a normalizing factor, ensuring that the posterior sums to unity. Bayesian

and its projection. Helmholtz realized that in most cases in vision, such a one-to-one mapping does not exist.

² Biological and computational constraints often prohibit the implementation of a truly optimal solution, in which case the brain will necessarily adopt a suboptimal strategy, such as an approximation. It has been argued that the inferential processes of the brain are in some cases suboptimal and that this suboptimality has been often misinterpreted as noise (Beck et al., 2012).

inference is the mathematically optimal way of probabilistically combining sensory input and expectations; an observer who uses Bayesian inference is thus called an *ideal observer*.

The posterior contains all the necessary information to make inferences about S by assigning a probability to each value of S . But what is the value that is actually perceived? Bayesian Decision Theory provides a framework for answering this question by defining a gain (or loss) function and choosing the percept (value of S) that maximizes the expected gain (or minimizes the expected loss) over all stimuli. If the goal of the animal is to have the fewest possible mismatches between perception and reality, the value of S that achieves this (call it S^*) is simply the most probable value, i.e. the mode of the posterior $P(S|I)$ and is called the *maximum a posteriori* (MAP) solution:

$$S^* = \arg \max_S P(S|I) \quad (2)$$

The MAP solution is optimal under the above criterion of least mismatches; however, it is not the only possibility. If, for example, the loss function is proportional to the squared difference of the inferred and actual percept $(S - S^*)^2$, which is a commonly used statistical measure, it can be shown that the optimal solution is not the mode but the *mean* of the posterior (which is generally different to the mode for skewed or multimodal distributions).

Taking either the mode or the mean of the posterior is a deterministic solution; for a given posterior, S^* is always the same. In perceptual experiments, however, there is very often trial-to-trial variability in subjects' responses: repeated presentations of the same stimulus do not lead to identical percepts every time. This variability has generally been modelled in two ways.

One way is to assume that perception and/or responses are corrupted by noise. The first type of noise, *perceptual* noise, could be in the image itself or in the neural activity of the visual system. Signal Detection Theory (SDT), a framework used to explain perceptual detection and discrimination tasks, assumes that the variability in subject responses comes from perceptual variability due to noise in the signal (i.e. the stimulus) and its sensory representations in the brain. Response noise, usually referred to as *decision* noise, is an additional source of noise between perception and

response - for example, neural noise in motor areas of the brain that introduces variability in reporting what was seen, even when the percept itself is noiseless (Mamassian et al., 2002). It has been argued that the variability observed in many perceptual experiments cannot be explained by SDT alone and instead requires the addition of decision noise (Mueller and Weidemann, 2008), and the concept of decision noise has been used in Bayesian models of perception to explain trial-to-trial variability (Weiss et al., 2002).

Another way to model trial-to-trial variability is to assume a stochastic rule for choosing S^* . The most popular approach is *probability matching* whereby S^* is simply a sample from the posterior. Thus across trials, the relative frequency of a particular percept is equal to its posterior probability. It can be shown that probability matching is not an optimal strategy under the standard loss criteria discussed above. However, the optimality of a decision rule based on the mode or the mean of the posterior rests on the assumption that the posterior is correct or that the environment is static; when either of these is not true, probability matching can be more useful because it increases exploratory behaviour and provides opportunity for learning (Mamassian et al., 2002; Wozny et al., 2010; Battaglia et al., 2011). Probability matching and more generally posterior sampling has also been proposed as a mechanism to explain multistable perception (Gershman et al., 2012), whereby an ambiguous image results in two (or more) interpretations that spontaneously alternate in time. For example, looking at Figure 2, right, we might see a vase at one time but two opposing faces at another time.

Two things should be noted regarding the strategies for choosing S^* . First, a MAP solution may be the only viable approach when the entire posterior cannot be well approximated, i.e. it may be possible to infer the most likely percept but not the probabilities for all possible percepts. Second, drawing from machine learning literature, some authors have suggested a smooth transition between MAP and probability matching by means of exponentiation of the posterior (Battaglia et al., 2011): raising the posterior to the power of $k > 1$ makes the posterior sharper. By increasing k the posterior approaches a delta function centred at the mode and thus sampling from that posterior is simply the MAP solution. Posterior sampling in the brain may then be a common strategy at various levels of perception and cognition and the brain

may adjust the value of the exponent k to switch from probability matching to MAP and everything in between (Moreno-Bote et al., 2011).

1.2 OVERVIEW OF THESIS

In this thesis I ask whether expectations about the speed of moving objects (assumed long-term and possibly hardwired) can be changed and, by using Bayesian models that incorporate this expectation in the form of a prior probability on speed, examine the temporal dynamics, stability and specificity of the changing prior. In chapter 2, I review the existing literature on the types, role and adaptability of priors in visual perception and then focus on previous theoretical and experimental work on the speed prior. In chapter 3, I describe a motion perception experiment that provides evidence for the plasticity of the speed prior and extend an existing Bayesian model, endowing it with a prior that changes to accommodate the speed statistics of the experimental environment. In chapter 4, I investigate the potential for change of the speed prior using a paradigm that allows a precise reconstruction of the prior from psychophysical data, uncovering a novel effect of contrast on perceived speed in the process. In chapter 5, I ask whether differences in everyday visual experience or gender can account for the differences in speed priors and the associated perceptual biases observed in humans. In chapter 6, I examine how generalizable or context-specific the speed prior is and whether it is robust in time. Finally, in chapter 7 I put my experimental and theoretical work in perspective and discuss its implications, its limitations and the prospect of unification of Bayesian models with physiologically grounded principles of processing of visual motion.

BACKGROUND AND RESEARCH MOTIVATION

2.1 EXPECTATIONS IN PERCEPTION

2.1.1 *Types of expectations*

Perceptual expectations refer to our prior beliefs about properties of the external world when interpreting sensory input. These beliefs might come from spatiotemporally confined cues that help disambiguate a scene. In Figure 2, the ambiguous “Rubin’s vase” shown in the middle panel is very likely to be interpreted as two faces if one’s gaze is closer to the left panel (or if the right panel is not attended) and as a vase if one’s gaze is closer to the right panel because of the spatial proximity of the ambiguous image to the unambiguous versions. The same effect would be observed with temporal proximity, such as when Rubin’s vase is presented shortly after the unambiguous image of a vase or two faces. Such expectations develop immediately and affect mostly the currently processed sensory input. This kind of expectations, induced by cues that are local in space or time and have immediate and short-term effects, have been recently dubbed “contextual expectations” (Serriès and Seitz, 2013). There are various types of contextual expectations. In a psychophysical experiment conducted by Sekuler and Ball (1977), expectations were in the form of explicitly given information regarding the possible directions of motion of a random-dot stimulus. The stimulus moved either upward or rightward and observers had to guess which trials contained a stimulus and which did not but were told which of the two directions would be presented if the stimulus did appear in a trial (the stimulus contrast was low so that its detection was challenging). In another psychophysical experiment (Haijiang et al., 2006), expectations were formed via Pavlovian conditioning, i.e. association between the sense of rotation of an object, as suggested by stereo and occlusion cues, and a signal (which could be a sound or the screen position of the stimulus). When the sense of rotation was later made ambiguous by removing the

stereo and occlusion cues, subjects' perception of sense depended on the presence of the signal.

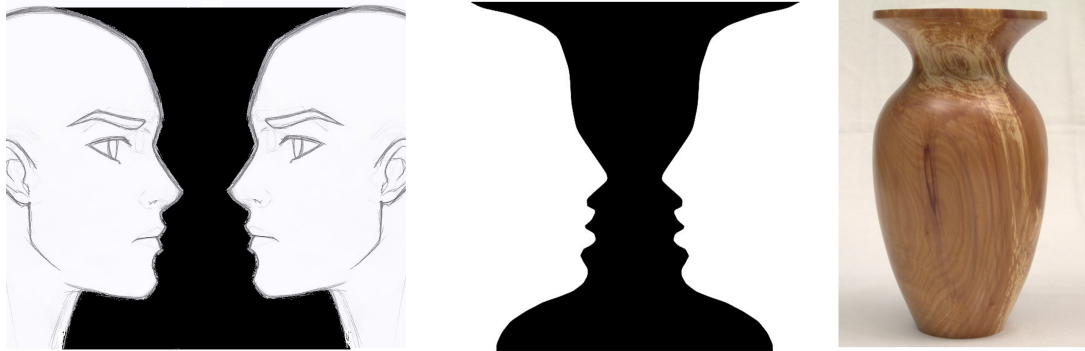


Figure 2: Rubin's vase. The black-and-white image in the middle can be seen as a black vase or as two opposing white faces. The unambiguous images on the left and right provide the context that primes the visual system to choose one of the interpretations for the ambiguous image.

Contextual expectations are not the only kind of expectations; another kind, conceptually more akin to Bayesian priors, are expectations based on general, or *prior*, knowledge about the world. [Serriès and Seitz \(2013\)](#) call these *structural expectations*. The expectation that shapes are commonly isotropic ([Knill, 1998, 2007](#)) is one such example: when humans see an elliptical pattern such as that of [Figure 1](#), they assume that it is a circle viewed at a slant rather than an ellipse because circles are isotropic. Another example of a structural expectation is that light comes from above (or perhaps from above-left; [Mamassian and Goutcher, 2001](#)). The light-from-above prior is used by humans ([Ramachandran, 1988; Mamassian and Landy, 2001; Adams et al., 2004](#)) as well as other animals ([Sun and Perona, 1998](#)) when inferring the properties of an object from its apparent shading. In shape-from-shading judgements, as well as in figure-ground separation tasks, another expectation also influences perception – that objects tend to be convex rather than concave ([Kanizsa and Gerbino, 1976; Adams and Mamassian, 2004; Peterson and Salvagio, 2008](#)). [Figure 3B](#) is interpreted by humans as a set of bumps instead of dents whereas [C](#) is typically interpreted as a set of black objects in white background (and not vice versa) because under this interpretation the objects are convex. A convexity prior seems to exist not just for objects within a scene (e.g. bumps on a surface) but also for the entire surface itself: subjects are better at local shape discrimination when the surface is globally convex rather than concave ([Langer and Bulthoff, 2001](#)). A related, recently reported expect-

ation is that depth (distance between figure and ground) is greater when the figure is convex rather than concave. This results in an increase in perceived depth in the case of convex figures given the same binocular disparity (Burge et al., 2010). Other examples of expectations are that objects tend to be viewed from above (Mamasian and Landy, 1998); that objects are at a distance of 2-4 m from ourselves (specific distance tendency, Gogel, 1965; Owens and Leibowitz, 1976); that objects in nearby radial directions are at the same distance from ourselves (equidistance tendency, Gogel, 1965); that cardinal (horizontal and vertical) orientations are more likely than oblique orientations (Tomassini et al., 2010; Girshick et al., 2011); and that people's gaze is directed towards us (Mareschal et al., 2013). The last example that I mention here is the expectation that has been the subject of my doctoral research - that objects in the world tend to move slowly or be still. In Bayesian models of motion perception, this is typically referred to as the *slow speed prior* (Weiss and Adelson, 1998; Weiss et al., 2002; Hürlimann et al., 2002; Stocker and Simoncelli, 2006; Montagnini et al., 2007; Bogadhi et al., 2011; Hedges et al., 2011).

The distinction between contextual and structural expectations is not always clear-cut. For example, when you see a dark, pistol-shaped object in the bathroom after you've taken off your glasses and your vision is blurred, you will likely see that object as a hairdryer (Figure 3C). The exact same shape seen in a workshop will be perceived as a drill. The context, bathroom sink vs workbench, helps disambiguate the object - a contextual expectation. However, this disambiguation relies on prior knowledge that hair dryers are more common in bathrooms and drills are more common in workshops; these are structural expectations. For the remainder of this thesis, structural expectations will be referred to simply as "expectations" whereas contextual expectations will be always qualified as such. In a Bayesian context, (structural) expectations will be referred to as "priors".

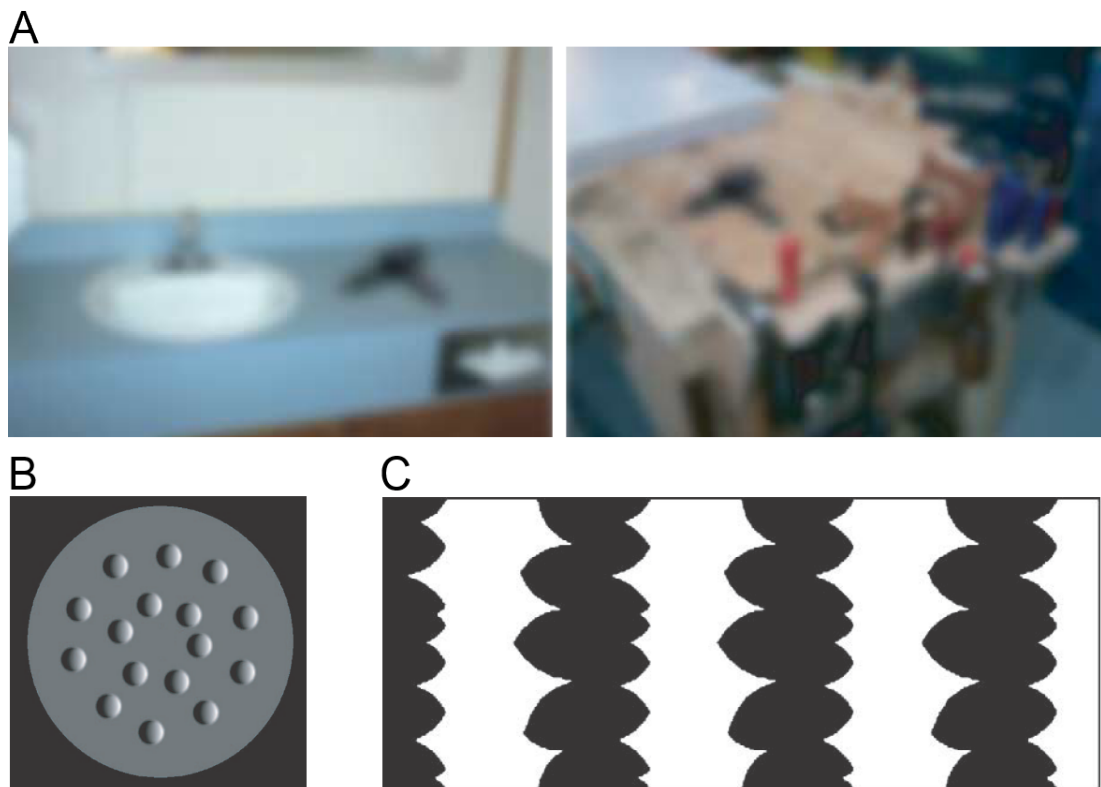


Figure 3: **A**, interplay between contextual and structural expectations. The black object on the left is typically perceived as a hair dryer because it has a pistol-like shape (structural expectation), it appears to be in a bathroom (contextual expectation) and we know that hair dryers are typically found in bathrooms (structural expectation). The identical-looking black object on the right is perceived as a drill since the context implies that the scene is a workshop (image adapted from Bar, 2004). **B**, convexity expectation for shape-from-shading. Shaded circles are most often seen as bumps illuminated from the right instead of dents illuminated from the left (image adapted from Adams, 2007). **C**, convexity expectation for figure-ground separation. Black regions are seen as convex objects in a white background more often than white regions being seen as concave objects in a black background (image adapted from Peterson and Salvagio, 2008).

2.1.2 Impact of expectations

Expectations help us infer the state of the environment (or aspects of interest of that state, such as particular objects) in the face of uncertainty. In the “bathroom sink vs workbench” example, expectations help disambiguate the dark object in the middle of each image in Figure 3. In some cases, multiple expectations must be employed to disambiguate a scene and perception is the result of the interaction of these expectations (Mamassian and Landy, 2001). Mamassian and Landy (1998) found that the perceived curvatures of a family of 3D-interpretable line drawings can be explained

under three types of expectations: that surfaces tend to be convex, that surface contours are aligned with the principal lines of curvature and that objects are viewed from above. Apart from aiding with object identification or shape judgements, expectations can impact perception in several other ways. For example, in the aforementioned study by Sekuler and Ball (1977), subjects were better and faster at detecting the presence of the stimulus at the expected directions. Similar improvements in detection performance and reaction times at expected motion directions were observed by Chalk et al. (2010) and improvements in luminance detection, brightness discrimination, orientation discrimination and form discrimination were observed by Downing (1988) at the spatial location where stimuli were expected to appear.

Sometimes, however, expectations can result in biased or illusory perception. In some cases, expectations about a particular measurable property influence the perceived magnitude of that property. For example, the aforementioned expectation of slow speeds is thought to result in underestimation of the speed of low-contrast stimuli (Stocker and Simoncelli, 2006), which explains, among other things, why drivers tend to drive faster in foggy conditions in real (Snowden et al., 1998) and simulated (Horswill and Plooy, 2008) environments. Expectations about one property can even influence the perception of another property: Wallach observed that a line moving behind a circular aperture, with its endpoints concealed such that its true direction cannot be recovered (Figure 4), always appears to move perpendicularly to its orientation. He hypothesized that the reason for this remarkably universal perception is an innate preference of the visual system for slow speeds and the fact that the perpendicular direction is the one corresponding to the slowest speed compatible with the retinal image sequence¹. Similarly, Weiss et al. (2002) showed that a Bayesian model incorporating a slow speed prior can explain, among other things, the difference in perceived direction between rhombus-shaped moving stimuli of different contrasts (see 2.2.3 for a detailed discussion).

¹ Wallach published his findings, originally in his doctoral thesis, in German, as a paper in the journal *Psychologische Forschung* in 1935. Although Wallach later translated a small part of his work in English, it was not until 60 years later that the entire paper was translated by Wuerger et al. (1996).

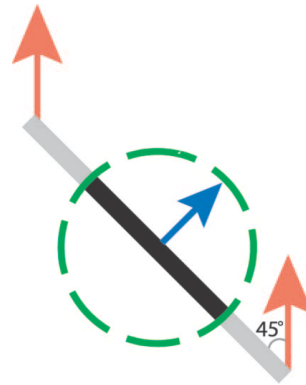


Figure 4: The aperture problem: the direction of motion of a line visible only through the green circle is ambiguous, as only the motion component normal to the line (blue arrow) is visible. True motion is upward (orange arrow) but the motion component parallel to the line is not visible. Humans invariably perceive the direction to be that of the normal component (image adapted from [Pei et al., 2008](#))

2.1.3 *Expectations and environmental statistics*

If our perceptual systems are to perform well, that is, if the interpretation of an ambiguous or noisy scene is to match the real world as closely as possible, our expectations will need to accurately reflect the structure of the world. In Bayesian terms, this means that our priors must closely approximate the statistics of the environment. Using the previous “bathroom sink vs workbench” example, we need to have a prior that favours the presence of drills in workshops. If that were not the case, we might interpret the object in the workshop as a hair dryer, which would lead to an incorrect inference more often than not because, statistically speaking, drills are more common in workshops than hair dryers are.

Our visual input - the images forming on the retina - although very varied, is only a small subset of the entire possible image space. Natural images are highly redundant, containing many statistical regularities that the visual system may exploit to make inferences about the world ([Simoncelli, 2003](#); [Geisler, 2008](#)). If expectations are to facilitate vision, they should approximate environmental statistics and, indeed, there is considerable evidence that in many cases they do. Even certain perceptual biases are thought to be byproducts of expectations that match natural scene statistics. [Tomassini et al. \(2010\)](#) found that the magnitude of the perceptual bias of subjects judging the orientation of elongated stimuli (Gabor patches) was proportional to

the uncertainty in orientation judgements. The bias was always towards the cardinal orientations (vertical and horizontal), indicating a tendency to expect cardinal orientations in the world. This expectation is in accord with the distribution of orientations in natural scenes (Coppola et al., 1998; Girshick et al., 2011). Girshick et al. (2011) formalized this expectation as a prior on orientation in a Bayesian model of perception and, using a previously developed method of extracting the prior from psychophysical data (Stocker and Simoncelli, 2006), found that subjects' priors matched the distribution of orientations in natural images to a considerable degree. Further evidence for a match between expectations and natural scene statistics comes from Yang and Purves (2003). The authors used laser range scanning to measure the distances of all unoccluded points in a variety of natural scenes. They found that the distribution of distances has a peak at about 3 m, which is in the middle of the range of 2-4 m found in studies on the specific distance tendency (see 2.1.1) by Gogel (1965) and Owens and Leibowitz (1976). Another example comes from Geisler and Perry (2009), who measured the pairwise statistics of edge elements from contours found in natural images (Figure 5). They found that the way humans group edge elements of occluded contours matches the performance of an ideal observer Bayesian model, i.e. a model with a prior reflecting the statistics of the natural image database. This suggests that perceptual mechanisms of contour grouping are closely related to natural image statistics. Finally, the influence of the convexity expectation in figure-ground separation (Kanizsa and Gerbino, 1976; Peterson and Salvagio, 2008) may also have a basis in natural scene statistics. Burge et al. (2010) found that subjects expect greater distances between convex figures and background than between concave ones and background. These expectations were in accord with the statistics of a collection of luminance and range images obtained from indoor and outdoor scenes.

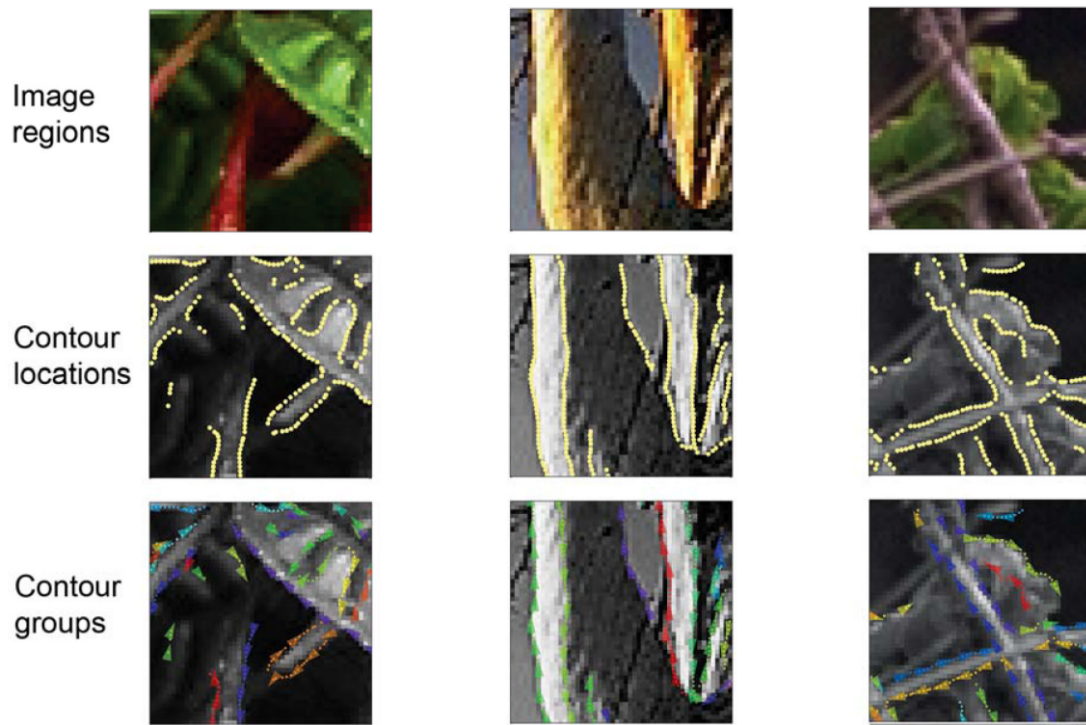


Figure 5: Contour locations and contour groups for small example patches from three different natural images. The contour locations and contrast polarities were detected by an automatic algorithm. The contour groups were obtained by hand segmentation. The direction of the arrows in the bottom images indicates the contrast polarity of the contour; the different colours represent different groups - note that because colours are randomly selected from a colour palette, different groups may have a similar colour (image and caption adapted from [Geisler and Perry, 2009](#)).

Research on the statistical structure of time-varying images is scarcer than research on static images. There are relatively few time-varying image databases, compared to static image databases and, moreover, inferring a distribution of object speeds from image sequences is complicated by the fact that retinal motion (the motion of patterns forming on the retina) does not correspond to object motion in a trivial way. One problem is that object speed can be inferred from retinal speed (the rate of movement of an object's projection on the retina) only if the distance of the object from the observer is known. Another problem is that retinal motion can be produced both by the independent motion of objects in the view field and by self-motion. In the case of self-motion, a further complication is that properties of motion signals are not independent of the position in the view field ([Zanker and Zeil, 2005](#)). For example, during self-motion such as walking or driving, radial motion flow in the periphery is greater than near the centre of view (foveal and parafoveal motion) and objects in the periphery appear to move centrifugally. The combination of object and self-

motion determine the statistics of retinal motion fields, i.e. the *optical flow*. A number of studies have looked at the statistics of optical flow fields by analyzing synthetic but realistic flow fields generated via simulated 3D camera motion through natural images that contain depth information. Calow and Lappe (2007) analyzed the statistics of flow fields in several retinal locations and found that the local distributions of speeds are to a good approximation log-normal. Interestingly, Zhang et al. (2013) showed that observers' peripheral vision had a strong centrifugal bias (static and even slow centripetally moving gratings presented briefly in the periphery appeared to move centrifugally). These biases were well accounted for by a Bayesian model with a speed prior that favoured centrifugal motion. Roth and Black (2005) found that, in synthetic data of optical flow generated using a range image database and camera motion simulation, the distribution of speeds peaks at zero (i.e. most points in the camera frame field have zero flow) and falls off with speed as a Laplace distribution, exhibiting significantly heavy tails. Dong and Atick (1995) adopted a simpler approach, analyzing a variety of video clips for certain simple statistical regularities and showing that the power spectrum of the clips was consistent with a variety of speed distributions $p(v)$ – in other words, their image analysis cannot fully constrain $p(v)$. However, they noted that a distribution that follows an inverse power law, i.e. is of the form

$$p(v) \propto \frac{1}{(v + v_0)^n} \quad (3)$$

(with v_0 and n as positive constants) leads to an analytical formula for the power spectrum function that matches well the measured power spectra of their natural image database. Equation 3 is a decreasing function of speed and thus a prior of this form would favour slow speeds.

Thus the work of Dong and Atick (1995), Roth and Black (2005) and Calow and Lappe (2007) suggests that a slow speed prior is consistent with natural scene statistics. As we saw in 2.1.1, Weiss et al. (2002) and Stocker and Simoncelli (2006) suggest that human observers possess such a prior. The model of Weiss et al. (2002) that incorporates a parametric slow speed prior (Gaussian centred at zero speed) is able to account for a wide variety of perceptual phenomena. Moreover, Stocker and Simoncelli (2006), by using a non-parametric approach to reverse-engineer the speed

prior from psychophysical data, found that observers' priors favoured slow speeds. Interestingly, the priors closely matched an inverse power law, exhibiting significantly heavier tails than a Gaussian. The two models are discussed in more detail in [2.2.3.2](#).

2.2 OUTSTANDING QUESTIONS AND RESEARCH MOTIVATION

While expectations have been studied for more than a century, the formalization of expectations as priors in Bayesian models of perception is relatively recent. Although by now there is a considerable body of theoretical work on Bayesian models that account for a wealth of perceptual data from human and non-human observers, there are several outstanding questions. In this section I lay out the questions that motivated my doctoral research, as well as any relevant theoretical and experimental work.

2.2.1 *Can long-term priors be changed?*

The modification of perceptual expectations can be achieved in a variety of ways. In [2.1.1](#) I gave two examples of how contextual expectations can be rapidly changed. Structural expectations on the other hand are presumably learned either through experience, by exposure to the statistics of the environment ([Simoncelli, 2003](#); [Geisler, 2008](#)), or are innate, hardwired in the brain by evolution. Certain structural expectations can be quickly learned from stimulus statistics in experimental setups. [Chalk et al. \(2010\)](#) presented subjects with random-dot stimuli moving in certain two directions more often than in others. Subjects were asked to report whether they saw motion and what the direction was. After minutes of exposure to the stimulus ensemble, subjects' responses indicated a bias towards the two frequent directions. Furthermore, in trials where subjects reported seeing motion when there wasn't any (a form of hallucination), the perceived motion was in the two frequent directions.

Although direction expectations are structural, there is no evidence, to my knowledge, to suggest that certain motion directions in natural environments are more common than others. Downward motion due to gravity and horizontal motion due to locomotion on flat ground could plausibly be more common than other directions in nature; however, the complications mentioned in [2.1.3](#) (briefly, the fact that retinal motion does not correspond to object motion in a simple way), the statistics of retinal

motion directions are likely very different from the statistics of motion in nature with respect to any fixed frame of reference. In absence of environmental regularities that may have shaped the prior to favour certain directions more than others, the brain may be more likely to adapt to the statistics of an experimental stimulus ensemble. But is it possible to change long-term priors that have been shaped according to the statistics of natural scenes, either during development or by evolution? [Adams et al. \(2004\)](#) answered this question affirmatively in the case of the light-from-above prior. They presented subjects with shaded hemispheres² at a variety of orientations and ask them to report whether they are bumps (convex) or dents (concave). Based on their responses, they inferred that the initial light prior of all subjects was roughly overhead. They subsequently exposed subjects to similar stimuli in a 1.5-hour session but this time giving them haptic (active touch) feedback on whether the hemispheres were bumps or dents. Crucially, however, the shading of the hemispheres and the haptic feedback were consistent with a light direction that deviated $\pm 30^\circ$ from overhead. When the subjects were then tested (without haptic feedback), their responses had changed: the 30° and -30° groups had shifted the mode of their priors by 9° and -13° on average, respectively. Thus, with just 1.5 hours of training, humans were able to update their light-from-above prior. Subsequent research by the same group confirmed the plasticity of this prior ([Adams et al., 2010](#)). Interestingly, the light-from-above prior does not seem to be mutable in chickens, as revealed by an experiment with chickens reared in an environment illuminated from below ([Hershberger, 1970](#)).

To my knowledge, the only other attempt at changing a long-term prior is the work of [Knill \(2007\)](#). The prior in question is that objects tend to be isotropic, so that, for example, circles are more common than ellipses in the world (see also [Figure 1](#)). This tendency for isotropy is assumed to reflect environmental statistics, although to date this has not been examined directly. Under the assumption of isotropy, the apparent aspect ratio (AR) of a surface serves as a cue to slant (referred to as *foreshortening* cue): if the surface is assumed to be circular, the deviation of AR from 1 is proportional to slant. In the experiments of [Knill \(2007\)](#), subjects were asked to judge the slant of elliptical surfaces presented on a screen. Subjects were provided with stereo glasses so that stereoscopic cues to depth were available to them. The experiment was set up in a way that the foreshortening cues conflicted with the stereoscopic cues:

² The stimuli in [Adams et al. \(2004\)](#) were similar to the ones in [Figure 6, B](#): the hemispheres there can be seen as bumps illuminated from the right or as dents illuminated from the left.

the stereoscopic cues were always in perfect agreement with the true magnitude of surface slant, which was indicated to the subjects at the end of each trial by haptic feedback, whereas the foreshortening cues were less correlated with true slant. Subjects completed 5-7 one-hour sessions (one per day) of exposure to two different stimulus ensembles. The first ensemble consisted of ellipses with ARs very close to 1 whereas the second consisted of ellipses of random ARs (and thus circles, which have an AR of 1, were no more common than ellipses with $AR \neq 1$). In the course of the experiment, subjects exposed to the second ensemble relied progressively less on the foreshortening cues and more on the stereoscopic cues, which implies that they started seeing elliptical patterns more often as ellipses of random ARs (and less often as slanted circles). This adaptation was not seen in subjects exposed to the first stimulus ensemble: the weights subjects gave to the two cues did not change across sessions. According to the author, this is indirect evidence that subjects exposed to a large proportion of irregular (anisotropic) shapes changed their assumptions about regularity and no longer assumed that surfaces are mostly circular. In the Bayesian model of Knill (2007), the prior is a mixture of a component where AR is equal to 1 (Dirac delta function) and a component where AR is normally distributed. Learning consisted of changing the mixing coefficient to favour the second component more. Interestingly, results were similar when haptic feedback was omitted but the disparity between the two cues was large (much larger than the disparities in the initial experiment). This suggests that when the conflict between foreshortening and stereoscopic cues is large enough, the visual system will gradually downweight the less reliable – that is, the foreshortening – cue (as stereoscopic cues are more reliable), even in the absence of feedback. Overall, the work of Knill (2007) suggests that subjects are able to adapt to the AR statistics of the world, either via explicit (with feedback) or implicit learning.

2.2.2 *How specific are priors?*

Although the research is limited, we saw that some priors appear to be mutable by means of “statistical learning” - exposure to unambiguous (or disambiguated with feedback) stimuli with distributions different from the prior. Do the changed priors have an effect only when the visual environment is identical or sufficiently similar to

the environment that drove the change or can the new priors generalize to different circumstances? Consider the aforementioned direction expectations formed by exposure to the random-dot stimulus ensemble of [Chalk et al. \(2010\)](#) - would these carry over to random-dot stimuli of different colour or even to different stimuli altogether? [Gekas et al. \(2013\)](#) provided clues to the answer by performing a very similar experiment to that of [Chalk et al. \(2010\)](#) but using two differently coloured random-dot ensembles, red and green, each with its own distribution of motion directions. The authors found that subjects applied the learned prior from the green ensemble to the estimates for the red ensemble (showing similar biases as in [Chalk et al., 2010](#)) but only if there was nothing to learn for the latter, i.e. only when the distribution of the red ensemble was uniform. On the contrary, if both distributions were non-uniform, subjects appeared to be able to learn two priors and apply each to the respective ensemble. Therefore [Gekas et al. \(2013\)](#) showed that learned priors exert their influence on similar but not identical stimulus ensembles to the ones that shaped the priors. Furthermore, when two stimulus ensembles have different and non-uniform statistics, humans are flexible enough to employ separate priors. A case for the ability of humans to maintain different priors for similar stimuli was also made by [Seydell et al. \(2010\)](#) who revisited the prior for isotropy of shapes ([Knill, 2007](#), see above). The authors performed a series of experiments. In the first, they used two groups of stimulus shapes (ellipses and diamonds) giving each group a different distribution of aspect ratios. In the second experiment, the stimuli in the two groups had the same shape (elliptical) but they had different colours (pink and purple). Subjects were able to learn separate distributions of aspect ratios (evidenced by the different changes in the weights they gave to stereoscopic cues to slant, as in [Knill, 2007](#)) for each shape. Interestingly, however, their isotropy prior was shared between the two colour groups - learning of the prior by exposure to the pink stimuli generalized to the purple stimuli. Taken together, the results of [Gekas et al. \(2013\)](#) and [Seydell et al. \(2010\)](#) suggest that a single prior can be used for stimuli of different colour, although the circumstances under which this happens are not yet fully known.

Can long-term expectations also generalize across dimensions such as task or stimulus type? The literature addressing this question is limited and most investigations have focused on the light-from-above prior. [Adams et al. \(2004\)](#), who showed that the light-from-above prior can be changed, also presented evidence that the same

light-from-above prior applies to different tasks and stimuli. The stimuli that subjects were exposed to were groups of shaded hemispheres. However, in order to exclude the possibility that the subjects learned a direct relationship between shape and shading or used a cognitive strategy instead of changing their light-from-above prior, the authors tested the subjects with a different stimulus type that consisted of an orange square flanked by two grey areas. By asking subjects to compare the brightness of the two grey areas at various orientations of the stimulus, the authors found a difference between groups that had learned two different illumination direction distributions (both different from their baseline distribution) during training. This difference suggested that there was a single prior for illumination directions that subjects changed and then applied to a different stimulus and task. Further evidence that there is a single light prior for each subject that applies to all tasks and stimuli was provided by Adams (2007). There was substantial variability in the light prior across subjects (several degrees on either side of the vertical) but approximately the same prior for each subject applied for three different tasks - shape from shading, visual search and reflectance (Figure 6).

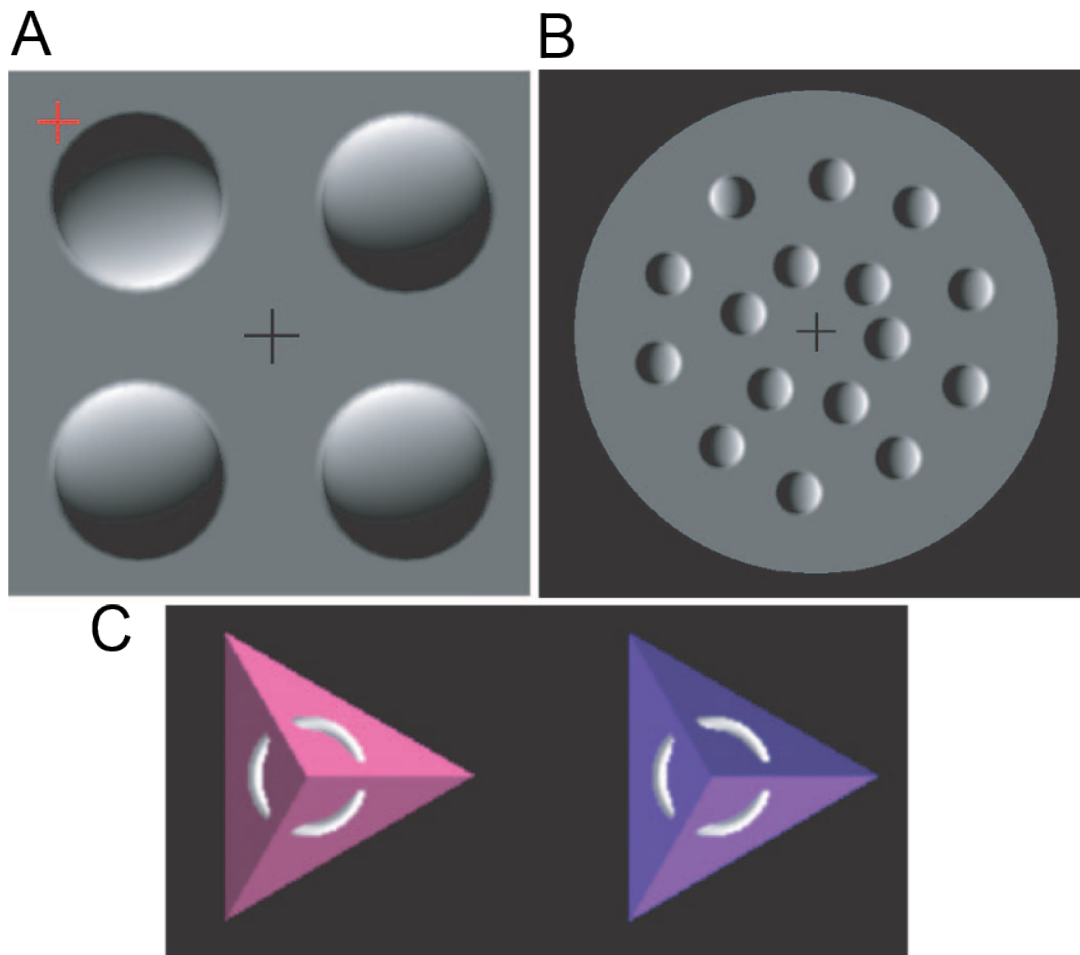


Figure 6: Examples of stimuli used in the three experiments of [Adams \(2007\)](#) that measure the “light-from-above” prior. **A**, shape task: observers usually see the top left object as concave and the others as convex, consistent with the “light-from-above” prior. **B**, visual search: target detection is relatively difficult until the figure is rotated by 90° . **C**, reflectance task: the left object is perceived as a homogeneous object lit from above, whereas the right object appears to have a darker upper surface, consistent with overhead lighting. (Image adapted from [Adams, 2007](#))

Although in most investigations the light-from-above prior appears to be shared across various tasks and stimuli, there is evidence that observers can develop context-specific light priors under certain circumstances. [Kerrigan and Adams \(2013\)](#) performed an experiment similar to the first experiment of [Adams et al. \(2004\)](#) (modification of the light-from-above prior with a shape-from-shading task) but with two distinct stimulus sets, one where the illumination was green and another where it was red. The two sets also differed in the distributions of lighting direction, with the red-light set having a distribution favouring overhead lighting and the green-light set having a distribution that favoured a direction 10° away from overhead. Subjects managed to learn the two different distributions and apply them to the appropriate

context. Interestingly, the amount of learning was greater when the two sets were presented to the subjects in an intermingled fashion rather than one set after the other.

Other priors may also be specific. Indirect clues that the prior for isotropy of shapes is specific to task and/or stimulus properties can be found in the results of Knill (2007), who noted that changes in cue weights (reflecting changes in the isotropy prior) were monotonic across experimental sessions: the prior did not show any re-adaptation effects from session to session, even though the sessions were on separate days and subjects were exposed to the regularities of the world in the meantime. According to the author, this suggests that subjects learned a context-specific prior.

Lack of re-adaptation of a learned prior was also noted by Adams et al. (2010) in the case of the light-from-above prior. Using the same stimulus (shaded hemispheres) as in a previous investigation (Adams et al., 2004), the authors repeatedly tested their subjects over the next days after the end of the training procedure. They found that learned changes in the interpretation of shading, and thus in the light-from-above prior, persisted for at least four days - further evidence that the learned prior was context-specific and thus did not re-adapt during everyday life.

2.2.3 *Research focus: speed prior*

From the above it has become clear that research on the modification of long-term priors via statistical learning is limited to a handful of cases. The light-from-above prior is the prior that has received the majority of attention in literature. However, there is another prior that is assumed to be long-term and appears to underlie a number of diverse, well-studied perceptual phenomena but for which the questions in 2.2.1 and 2.2.2 have not been addressed: the slow speed prior. Before laying out the unexplored aspects of this prior that are addressed in this thesis, I summarize the effects of the slow speed prior on perception.

2.2.3.1 *Effects of the slow speed prior on perception*

A prior favouring slow speeds has been postulated as an explanation for a variety of perceptual phenomena, including certain visual illusions and biases. One of the earliest known effects attributed to the slow speed prior is described in 2.1.2 as the

universal way that humans solve the aperture problem. A related illusion that can also be accounted for by a Bayesian model with a slow speed prior (Weiss and Adelson, 1998) is the “barberpole effect”: a diagonally striped pole rotating around its vertical axis (horizontally) appears as though its stripes are moving in the direction of its vertical axis. Another effect attributed to this prior by a number of authors (Heeger and Simoncelli, 1993; Weiss and Adelson, 1998; Weiss et al., 2002; Hürlimann et al., 2002; Stocker and Simoncelli, 2006) is the “Thompson effect” (Thompson, 1982; Stone and Thompson, 1992; Hawken et al., 1994; Hürlimann et al., 2002) – the decrease in perceived speed of a moving stimulus when its contrast (the difference in luminance between the brightest and darkest parts of the image) decreases – although other authors have suggested alternative, non-Bayesian explanations (Thompson et al., 2006). The Bayesian explanation of the effect of contrast on perceived speed is as follows: given the noisy nature of eye optics and of neural activity, at low contrasts the signal-to-noise ratio is lower than it is at high contrasts, i.e. at low contrasts the stimulus is less salient. This means that the speed measurements that the visual system performs are more noisy at low contrasts, which is reflected by a broader (higher-variance) likelihood function. According to Bayes rule (Equation 1), the prior, which is multiplied by the likelihood to give the posterior, will thus have a greater influence on the posterior than at high contrasts, where the likelihood is sharper. Since the perceived speed is assumed to be determined by the posterior, if the prior favours slow speeds it follows that perceived speed will be lower at low contrasts (Figure 7). A real-world manifestation of the Thompson effect is the well-documented fact that drivers speed up in the fog (Snowden et al., 1998). It should be noted that contrast is not the only factor that can affect uncertainty. Short duration of motion can have a similar effect to low contrast, in terms of increasing uncertainty, by broadening the temporal frequency spectrum (Derrington and Goddard, 1989). Zhang et al. (2013) used stimulus duration as a way of manipulating uncertainty in perceived speed, noting that it has a similar effect to contrast but avoids certain complications that arise from nonlinear effects of contrast changes on motion discrimination. These effects stem from contrast-dependent changes in spatial suppression (Tadin et al., 2003).

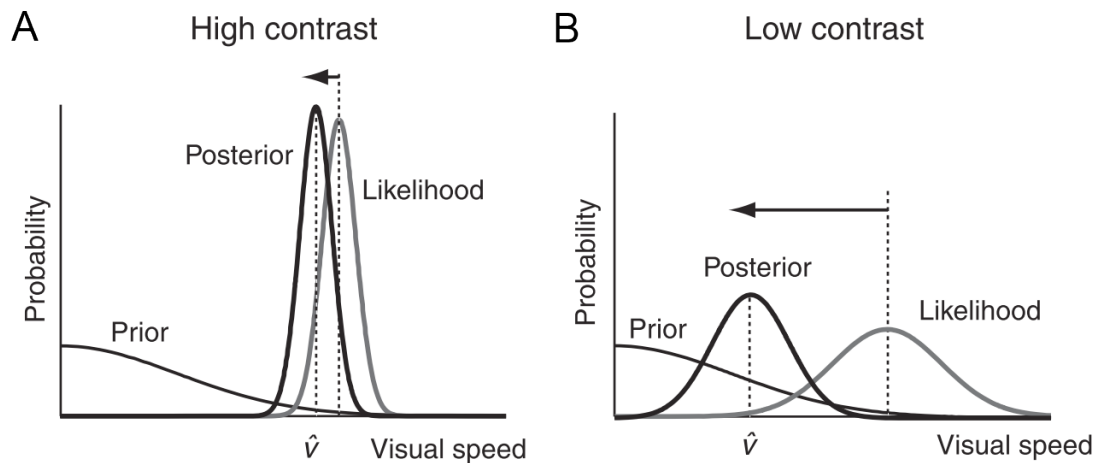


Figure 7: Illustration of a Bayesian estimator accounting for contrast-induced biases in speed perception. **A**, a stimulus with high contrast leads to relatively precise measurements and thus a narrow likelihood. Multiplication by a prior probability for low speeds induces only a small shift of the posterior relative to the likelihood. **B**, a low-contrast stimulus is assumed to produce noisier measurements and thus a broader likelihood. Multiplication by the same prior induces a larger shift and thus the low-contrast stimulus is typically perceived as moving slower. (Image adapted from [Stocker and Simoncelli, 2006](#))

In the case of motion in two dimensions, where the velocity (speed and direction) must be estimated, the slow speed prior may affect not only the perceived speed but also the perceived direction. In certain types of stimuli such as plaids (superposition of moving gratings), contrast has been found to influence perceived direction ([Stone et al., 1990](#); [Yo and Wilson, 1992](#); [Burke and Wenderoth, 1993](#); [Hedges et al., 2011](#)). For these stimuli, models incorporating preferences – either as Bayesian priors ([Weiss et al., 2002](#); [Hedges et al., 2011](#)) or as constraints/regularization factors ([Farid and Simoncelli, 1994](#); [Langley, 1999](#)) – for slow speeds and coherent motion have provided good descriptions of psychophysical data. In the Bayesian model of [Weiss et al. \(2002\)](#), the speed of a rigidly translating stimulus is determined by the integration of local motion signals under the assumptions of measurement noise and a prior that favours slow speeds. The model of [Weiss et al. \(2002\)](#) also accounts for the directional biases that have been observed regarding the motion of lines that are unoccluded (their terminators are visible, cf aperture problem) but their contrast is low. A spectacular instance of the influence of contrast on perceived direction that can be explained by assuming a slow speed prior is the rhombus illusion reported by [Weiss et al. \(2002\)](#) (Figure 8): a thin, high-contrast rhombus that moves horizontally appears to move in its true direction whereas the same rhombus at a low contrast appears to move

diagonally. [Weiss et al. \(2002\)](#) make a convincing case that all these perceptual biases and illusions are actually the result a probabilistically optimal combination of noisy image measurements with a prior for slow speeds, which seems to be ecologically motivated.

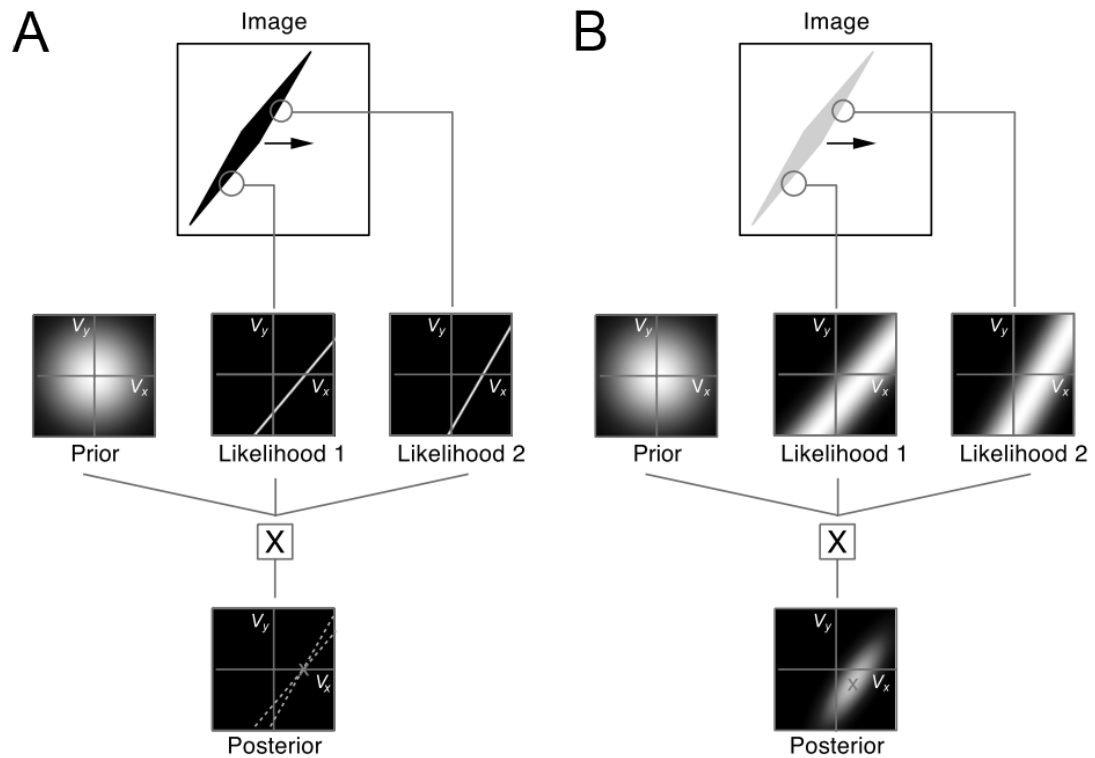


Figure 8: Influence of contrast on the perceived direction of a horizontally moving rhombus. Probabilities are plotted in velocity space: the coordinates of each point represent the magnitude of the horizontal and vertical component of the velocity at a particular image location whereas the brightness of the point is proportional to the probability of the corresponding velocity. **A**, with a high-contrast rhombus the signal-to-noise ratio of the two local measurements (only two are shown for clarity) is high and thus the likelihood in velocity space is sharp, tightly concentrated around a straight line, and dominates the prior, which is broader. The resulting posterior is much more similar to the likelihood and favours the veridical direction. **B**, with a low-contrast rhombus, the constraint line of the likelihood is fuzzy, i.e. the likelihood is broad and the prior exerts greater influence on the posterior, resulting in an posterior that favours an oblique direction. (Image adapted from [Weiss et al., 2002](#))

The slow speed prior has been suggested as the underlying mechanism for certain biases in 3D motion perception as well. Various studies have established that observers overestimate angular trajectories of objects on a collision course with their head, reporting instead that the object will miss them ([Peper et al., 1994](#); [Harris and Dean, 2003](#); [Welchman et al., 2004](#); [Gray et al., 2006](#)). One of the groups that reported this bias showed in a subsequent study ([Welchman et al., 2008](#)) that a Bayesian model

of motion perception including a slow speed prior – similar to the prior reverse-engineered by [Stocker and Simoncelli \(2006\)](#) from human psychophysical data – was able to account for the perceptual biases in 3D motion.

The slow speed prior has even been invoked to explain phenomena in other modalities, such as the cutaneous rabbit illusion, whereby two taps at a particular skin location followed quickly by two taps at another skin location are perceived as a series of taps, two out of which are felt in locations in-between the two veridical ones ([Goldreich and Tong, 2013](#)). The authors propose a Bayesian model of motion perception that includes both prediction and postdiction. Under the assumption that the tactile stimuli originate from a single moving object (a “cutaneous rabbit”), the illusory perception results from predicting and postdicting the distances that the object covers between successive taps, which should be minimal during the perceived inter-stimulus interval due to the slow speed prior.

2.2.3.2 Bayesian models with a speed prior

One of the first and most popular Bayesian models that incorporate a speed prior was proposed by [Weiss and Adelson \(1998\)](#) and was subsequently presented in simplified form by [Weiss et al. \(2002\)](#), mentioned in the previous section. In these models, the visual system estimates velocity (speed and direction) as an ideal Bayesian observer under the assumptions of visual measurement noise and a prior that favours slow motion. The first model also features a preference for smooth motion. In particular, it includes a prior on a velocity field that favours fields with small derivatives – more specifically, the prior is inversely proportional to the sum of squares of a differential operator that measures derivatives of all orders, including the zeroth³, summed over the extent of the image. [Weiss et al. \(2002\)](#) simplified the model for instances of rigid motion, where only the magnitude of the velocity is nonzero (higher-order derivatives are zero). In this simplified model, the prior on speed is a Gaussian centered at zero:

$$p(v) \propto \exp\left(-\frac{\|v\|^2}{2\sigma_p^2}\right) \quad (4)$$

The prior in the model of [Weiss et al. \(2002\)](#) is multiplied by the likelihood of the image sequence to give the posterior over velocities. The likelihood is formulated un-

³ A strong zeroth derivate describes fast velocities whereas strong higher-order derivatives describe nonsmooth velocities.

der the assumptions that the temporal derivative of the image sequence is corrupted by additive Gaussian noise (whereas the spatial derivatives are not). Since both the prior and the likelihood are Gaussian, the posterior is also Gaussian and the velocity estimate (output of the model), which is the MAP solution, has an analytical form (also see Figure 8 and section 3.2.2).

Despite the wide applicability of these models and the elegance by which they explain different phenomena in motion perception, a number of authors have pointed out certain theoretical limitations and have identified cases where the models cannot account for certain subtle phenomena. Hürlimann et al. (2002) analyzed the model of Weiss et al. (2002) for the case of drifting sinusoidal gratings of different contrasts and noted that the estimated (“perceived”) speed \hat{v} of a drifting grating of contrast c moving at (actual) speed v is given by

$$\hat{v} = v \frac{c^2}{k^2 + c^2} \quad (5)$$

where $k = \sigma/\sigma_p$ is the ratio of the standard deviations of the likelihood and the prior. Hürlimann et al. (2002) point out that this derivation rests on the assumption that the contrast response of the visual system is linear – an assumption likely to be wrong as it ignores the saturation effect that has been observed in the responses of motion-processing neurons of cortical area MT (Sclar et al., 1990). They propose a modification to Equation 5 in the following way:

$$\hat{v} = v \frac{c^q}{k^q + c^q} \quad (6)$$

where $q < 2$ would indicate a compressive nonlinear effect of contrast. The authors then tested the model in a speed matching task where the apparent speeds of gratings of different contrasts are compared. They showed that their model provides a better description of the data than the original model of Weiss et al., 2002 does. The best-fitting values of q were less than 2.

A few years later, Stocker and Simoncelli (2006) highlighted further limitations of the model of Weiss et al. (2002). One such limitation is that both the likelihood function and the prior were chosen to be Gaussian merely for mathematical convenience and not due to theoretical or empirical (e.g. compliance with natural scene statistics) considerations. The noise characteristics of the model likelihood are also unlikely to

accurately describe the noise properties of neurons involved in motion perception. We have already seen how the model of [Hürlimann et al. \(2002\)](#) addresses one aspect of the problem, by taking into account contrast saturation effects. [Ascher and Grzywacz \(2000\)](#), who present an alternative Bayesian model with a slow speed prior, take into account other properties of neural responses - namely response normalization and multiplicative noise. [Stocker and Simoncelli \(2006\)](#) have similar considerations: they note that the variability in cortical responses follows a Poisson distribution (that is, the response variance is proportional to the mean firing rate). Combining this observation with a physiologically motivated contrast response function ([Sclar et al., 1990](#)), they propose a likelihood with a width that is a separable function of speed and contrast, $\sigma(v, c) = g(v)h(c)$, and varies slowly with speed. With regards to the prior, [Stocker and Simoncelli \(2006\)](#) adopt a non-parametric approach, assuming no particular functional form but only that the prior is locally linear, i.e. it is expressed as $p(v) = av + b$. In a similar vein, [Zhang et al. \(2013\)](#) express the speed prior in their Bayesian model of peripheral motion perception as a Gaussian process, discretizing the prior function at 25 control points. This formulation allows the approximation of most continuous functions. Both [Stocker and Simoncelli \(2006\)](#) and [Zhang et al. \(2013\)](#) find that the priors fitted to their psychophysical data exhibit considerable departures from normality.

Once the shape of the speed prior is reconstructed using a non-parametric prior, parametric forms of a similar shape can be chosen for subsequent predictions. For example, observing that the extracted priors in the [Stocker and Simoncelli \(2006\)](#) study had significantly heavier tails than a Gaussian and resembled a power-law distribution, [Hedges et al. \(2011\)](#) used a modified power-law prior. Although this is a parametric prior, and more constrained than the priors used by [Stocker and Simoncelli \(2006\)](#) and [Zhang et al. \(2013\)](#), it led to reasonably good quantitative predictions in an experiment involving plaids (superpositions of drifting gratings at an angle).

An important and unique feature of the model of [Stocker and Simoncelli \(2006\)](#) is the modelling of trial-to-trial variability. Previous models ([Ascher and Grzywacz, 2000](#); [Weiss et al., 2002](#); [Hürlimann et al., 2002](#)) only account for average measurements and do not deal with the trial-to-trial variability observed in psychophysical experiments. [Stocker and Simoncelli \(2006\)](#) state that it is exactly this variability that provides the additional information necessary to disambiguate the contributions of

the likelihood and the prior. In the model of [Weiss et al. \(2002\)](#) these contributions cannot be disambiguated even in principle because the likelihood and prior variances only appear in computations as a single parameter – the ratio σ^2/σ_p^2 . The model of [Stocker and Simoncelli \(2006\)](#) is described in more detail in section [4.3.1](#).

2.2.3.3 *Outstanding questions regarding the slow speed prior*

A prior on slow speeds has been proposed as a unifying explanation for a variety of perceptual phenomena⁴. However, despite considerable amount of research in motion perception, it is not known how this prior is shaped and how it compares to other studied long-term priors described earlier (such as the light-from-above one). Are we born with a slow speed prior (like chickens are born with a light-from-above prior) or does it form during development, based on the spatiotemporal statistics of natural scenes? In relation to this, can this prior change through exposure to different statistics or does it become fixed at some point in the person's lifetime? Do people with different visual experiences, such as race car drivers or video game players who are exposed to relatively high speeds, have different speed priors? What are the characteristics of the prior distribution and do these vary among individuals? Assuming the speed prior can be reshaped by exposure to modified environmental statistics, how fast does this happen? Is it fast, such as the learning of the light-from-above prior in [Adams and Mamassian \(2004\)](#) and the direction prior in [Chalk et al. \(2010\)](#)? Or is it as slow as perceptual learning ([Karni and Sagi, 1993](#)) or the learning of the isotropy prior ([Knill, 2007](#)), occurring over days? Would it really be a reshaping of a single prior, so that any modification of the prior in the lab would be partially reverted by re-exposure to the real world? Or would the new speed prior be specific to the task/stimulus that reshaped it (which would imply the ability to maintain multiple context-specific speed priors)? To my knowledge, only the question of interindividual variability has been partially addressed ([Stocker and Simoncelli, 2006](#); [Zhang et al., 2013](#)), although no attempt to date has been made to account for this variability, e.g. in terms of interindividual differences in visual experience.

In the theoretical department, it is clear that all Bayesian models make various assumptions – some of them for mathematical convenience or conceptual simplicity –

⁴ [Weiss and Adelson \(1998\)](#) present an impressive list of perceptual phenomena that their Bayesian model with a prior for slow and smooth motion can account for.

and it has been shown that some of these assumptions must be abandoned in order to account for certain subtle phenomena in motion perception. The model of [Stocker and Simoncelli \(2006\)](#) makes the fewest assumptions regarding the noise characteristics of visual measurements and the form of the speed prior. Is it possible that even this model fails to capture some phenomena in its domain of applicability? Furthermore, are its assumptions about the effect of contrast on cortical responses and perceived speed robust?

These questions regarding the speed prior are the focus of the present work. This list is of course far from exhaustive; for example, my investigation of Bayesian models of speed perception is confined to the computational level ([Marr and Poggio, 1976](#)). I do not address in this thesis the underlying neural mechanisms responsible for the representation of probability distributions and for probabilistic inference and learning. As we saw in the previous section, that is not to say that physiological constraints are ignored; such constraints are taken into account in model design.

SPEED PRIOR CHANGE AND ITS EFFECT ON PERCEIVED MOTION DIRECTION

The work described in this chapter was published in *Current Biology* (Sotiropoulos et al., 2011) and can be accessed online at

<http://www.sciencedirect.com/science/article/pii/S0960982211010153>.

3.1 INTRODUCTION

The “slow speed prior” hypothesis – the hypothesis that our low-level visual system assumes that objects are static or move slowly (typically $0.5^\circ/s$, equivalent to 1 km/h viewed at 30 m (Stocker and Simoncelli, 2006)) rather than quickly – has been postulated because it can elegantly explain a number of perceptual biases observed in situations of uncertainty (see 2.2.3.1). Interestingly, those biases affect not only the perception of speed, but also of motion direction. For example, the motion direction of a line whose endpoints are hidden (as in the previously described “aperture problem” – Figure 4) or poorly visible (e.g. at low contrast, or for short stimulus presentations) is perceived as being perpendicular to the line more often than it is. This perceptual illusion is consistent with expecting that the line moves more slowly than it really does. As is the case with most other structural priors, the slow speed prior is thought to develop over the course of our lifetime, in a world where static or slowly moving objects are more frequent than fast objects. However, in reality, when and how this prior is acquired and whether it plays a causal role in extracting motion directions is still unclear. To progress on these issues, I investigated whether new experiences with visual stimuli that are considerably faster than the majority of those experienced in everyday life could result in a disappearance or reversal of the classically reported direction biases (Lorceau et al., 1993; Wuerger et al., 1996; Weiss et al., 2002).

3.2 METHODS

3.2.1 *Experiment*

3.2.1.1 *Stimulus and apparatus*

Stimuli were presented on a Samsung 2043BW LCD monitor (20 inches diagonal size, 1680×1050 resolution, 75 Hz refresh rate), driven by an Apple Mac Mini 2009 computer (Intel Core2 Duo CPU, OpenGL-capable GPU) running MacOS X. Stimuli were generated in MATLAB with the Psychophysics Toolbox v3 and consisted of a field of parallel lines, translating rigidly and coherently (Figure 9). Line elements had a length of 4° of visual angle, a thickness of $2.4'$ (arcmin) and an orientation of 110° counter-clockwise from horizontal i.e. approximately 11 o'clock. Motion direction was either perpendicular to the line, which means 20° from horizontal (upward motion; "up" condition), or oblique, that is, -20° from horizontal (downward motion; "down" condition). The equal distance of the two directions from the horizontal ensured that any bias towards the horizontal (Loffler and Orbach, 2001) would have the same effect in both conditions. In each trial, the direction was picked randomly without replacement. Stimulus presentation durations, in milliseconds, were uniform randomly chosen from the set {133, 266, 532}. Stimulus speed was either $4^\circ/s$ (test session and training session of low-speed group, see below) or $8^\circ/s$ (training session of high-speed group). With the exception of line length, all above parameters were the same as in the first experiment in Lorenceau et al. (1993). The field of parallel lines was visible through a circular mask of 24° in diameter and was of "infinite" area (as lines moved out of the display, new lines moved in), as in previous work (Lorenceau et al., 1993). The circular mask is much larger than individual lines, so most of the lines (excluding those near the circumference of the aperture) are visible in their entirety (i.e. including their endpoints). The background was black (luminance $0.29cd/m^2$) whereas the lines were grey. Two grey levels were used for the lines: they were either shown with a luminance of $0.342cd/m^2$ corresponding to a Michelson contrast of 8% ("low" contrast) or with a luminance of $0.948cd/m^2$ corresponding to a Michelson contrast of 53% ("high" contrast). Luminance measurements were performed with an ColorCAL colorimeter (by Cambridge Research Systems).

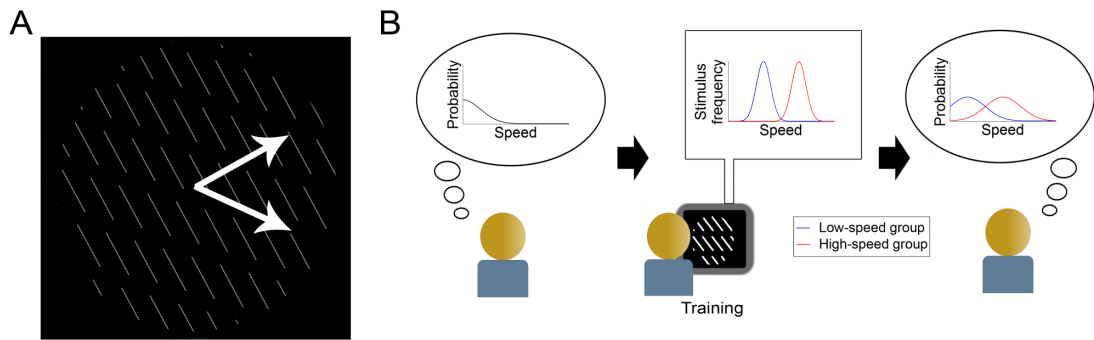


Figure 9: **A**, stimulus used in the experiment. The entire line field translates rigidly along either of the two directions shown by the white arrows (the latter are not part of the stimulus). **B**, cartoon of experimental hypothesis. Left: initially subjects have a prior favouring slow speeds. Middle: the low-speed group is exposed to low speeds (blue), while the high-speed group viewed faster speeds (red). Right: training causes the high-speed group to shift their prior expectations towards higher speeds (red) compared to the low-speed group (blue).

3.2.1.2 Subjects and procedure

Twelve naive subjects (all undergraduate or postgraduate students from the University of Edinburgh) participated in this experiment, evenly divided in two groups, high-speed and low-speed. Subjects performed the experiment in a dimly lit, sound-proof booth and were seated at a distance of 70 cm from the monitor. No chin rest or eye tracker were used; instead subjects were instructed to maintain their head position throughout the experiment. They were also instructed to fixate at the centre of the screen at all times. Each subject completed 5 sessions, each on a separate day, at the same time of day for a given subject. Each session lasted 35-45 minutes and was divided into 3 blocks - a small *test* block, followed by a large *training* block, followed by a small *test* block identical to the first. Stimulus contrasts, durations and directions were randomly interleaved in each block. Motion speed was kept constant and equal to $4^\circ/s$ in the test blocks and in the training block of the low-speed group, while it was fixed at $8^\circ/s$ in the training block of the high-speed group. The number of trials in each test block was chosen so that each condition was presented 18 times. Since there were 2 (equiprobable) contrast levels, 3 (equiprobable) durations and 2 directions, the total number of trials was $2 \times 3 \times 2 \times 18 = 216$. The number of trials in the training block was chosen to be larger than both test blocks. In particular, each condition was presented 60 times, so that each training block consisted of

$2 \times 3 \times 2 \times 60 = 720$ trials. In total, the number of trials in a session was 1152. The structure of an individual trial was as follows:

1. A 200 ms fixation period, during which the static stimulus (line field) was displayed, together with a red central fixation dot (diameter = 24').
2. A stimulus presentation period. The fixation dot disappeared at the beginning of this period and motion of the line field was initiated. All lines moved in the same direction and at the same speed, so the apparent motion is a global translation of the entire field. The duration of this period was either 133, 266 or 532 ms.
3. A response period. The stimulus disappeared and the fixation dot reappeared. Subjects reported the perceived direction of motion by pressing the Up/Down arrow key on the keyboard. No horizontal contours were displayed at any point in the trial and therefore subjects presumably had to rely on internal or external (such as monitor frame) cues to determine the horizontal (and thus deviations from it). When a response was made or when 3 seconds had elapsed, the trial ended.

3.2.1.3 Data analysis

Following [Lorenceanu et al. \(1993\)](#), the proportion of oblique responses (hereafter referred to as p_o) was measured for each duration and contrast level. Performance (proportion of correct responses in the task) before and after training was also calculated separately for the upward- and downward-motion trials. ANOVAs were used to determine whether there was a change in p_o between experimental groups, test blocks (pre and post-training), sessions and durations. and whether there was a difference between high-speed and low-speed groups. Where appropriate, significant results from ANOVA were followed up by t-tests with the Bonferroni correction for multiple comparisons in order to examine these differences.

My hypothesis was that subjects might implicitly update their expectations towards faster speeds, which would lead to a decrease in the direction bias, i.e. an *increase* in p_o towards 0.5 (or perhaps even higher), when tested at the low speed.

3.2.2 Bayesian model

To investigate quantitatively whether the results are consistent with the idea of a changing speed prior, I adapted the Bayesian model of motion perception proposed by Weiss et al. (2002). This model suggests that motion perception can be described as an optimal estimation of object velocities under the assumption of local measurement noise and an a priori preference for slower velocities (in Weiss et al. (2002), the prior is centred at zero). The idea behind my extension of this model is that the speed prior is initially centred close to zero but shifts away from zero from session to session due to exposure. As described below, model parameters were determined to best fit the group-averaged data. When receiving as inputs the moving stimulus image, the model computes a velocity estimate, which is a variant of the solution of Weiss et al. (2002), generalized to include a velocity prior with a non-zero mean. It can be shown that the estimate is:

$$\hat{v} = - \begin{pmatrix} \sum I_x^2 + \frac{\sigma^2}{\sigma_p^2} & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 + \frac{\sigma^2}{\sigma_p^2} \end{pmatrix} \begin{pmatrix} \sum I_x I_t - \frac{\sigma^2}{\sigma_p^2} \mu_x \\ \sum I_y I_t - \frac{\sigma^2}{\sigma_p^2} \mu_y \end{pmatrix} \quad (7)$$

where I_x, I_y, I_t are the spatial (two dimensions) and temporal (partial) derivatives of the image intensity function, σ^2/σ_p^2 is the ratio of the likelihood and prior variances, the only free parameter in the original model of Weiss et al. (2002), and μ_x, μ_y are the means (two dimensions) of the velocity prior.

The sums were computed over the pixels in the stimulus images. The spatial derivatives were computed using MATLAB's gradient function, which performs a simple subtraction of the values of neighbouring pixels along each axis. The temporal derivative is given by:

$$I_t = -(I_x v_x + I_y v_y) \quad (8)$$

where v_x, v_y are the local velocity measurements in the two axes (and equal to the veridical velocity of the stimulus and to each other, since the stimulus translates rigidly). Following Weiss et al. (2002), I simulated the 2-alternative forced choice

experiment by assuming that the decision ('up' or 'down') is corrupted by Gaussian noise. The model response in the presence of this "decision noise" is given by:

$$r = \text{sign}(\hat{\phi} + \eta)$$

where $\hat{\phi}$ is the angle between the estimated velocity vector and the horizontal and η a zero-mean Gaussian random variable of standard deviation σ_D . $\hat{\phi} = 0^\circ$ corresponds to horizontal (rightward) motion perception, $\hat{\phi} > 0^\circ$ corresponds to right-upward motion and $\hat{\phi} < 0^\circ$ to right-downward motion.

3.2.2.1 Updating the prior

The mean of the velocity prior $\boldsymbol{\mu}_p = (\mu_x, \mu_y)$ was allowed to vary within a session and between sessions. I modelled learning using the following three assumptions:

- I. Based on observed performance differences between successive sessions, I assumed that only a proportion $(1 - \gamma)$ of the shift of the prior within a session would be retained in the next session. I call γ the *unlearning rate* parameter.
- II. To describe the trajectory of the prior, two variants of the model were investigated:
 - a) In the first (nonparametric) model, the prior mean is determined independently for each session, by fitting to the data. In total, this model has 10 parameters: 5 for the prior mean locations after training (one for each session), 1 for the 'unlearning rate' γ , 1 for the decision noise standard deviation σ_D and 3 for the ratio σ^2/σ_p^2 for each duration.
 - b) In the second (parametric) model, the post-training prior mean was assumed to vary linearly with time (i.e. session number). This model was chosen after observing a roughly linear relation between prior means and session number using the nonparametric model above. The prior mean was modelled as: $\mu_x(s) = as + b$, where s denotes the session number. This reduces the number of prior-related parameters from 5 to 2.
- III. Only the horizontal component μ_x of $\boldsymbol{\mu}_p$ was allowed to vary; μ_y was fixed at zero. Given that in the model v_y is a Gaussian random variable, a nonzero value of μ_y would correspond to an artificially imposed vertical directional bias that

would mask the “up/down” bias that I wish to investigate and that is a result of a prior favouring high speeds in general (without assuming a preference for either vertical direction). A more accurate alternative would be to use a bimodal distribution for v_y with modes that get progressively closer to the two points $(\mu_x, \mu_y) = (v_H \cos(\theta), v_H \sin(\theta))$ where v_H is the training block speed ($8^\circ/s$) and $\theta = \pm 20^\circ$ are the two possible motion directions. However, the closed-form solution of Equation 7 would no longer apply, as it is based on assumptions of normality.

All aforementioned parameters were determined via Maximum Likelihood estimation: `fminsearch` in MATLAB was used to minimize the negative log-likelihood of the entire low-contrast¹ data set:

$$-\sum_{s,b,d} p_o^{data}(s,b,d) \log p_o^{model}(s,b,d) + (1 - p_o^{data}(s,b,d)) \log(1 - p_o^{model}(s,b,d))$$

where $p_o(s,b,d)$ is the probability of oblique response in the s -th session, the b -th block ($b = 1$ for pre-training and 2 for post-training) and for the duration d . The model probability of oblique response is given by

$$p_o^{model}(s,b,d) = p_o^{model}(s,b,d|\phi = 20^\circ)P(\phi = 20^\circ) + p_o^{model}(s,b,d|\phi = -20^\circ)P(\phi = -20^\circ)$$

Since $P(\phi = 20^\circ) = P(\phi = -20^\circ) = 0.5$, this becomes

$$\begin{aligned} p_o^{model}(s,b,d) &= \frac{p_o^{model}(s,b,d|\phi = 20^\circ) + p_o^{model}(s,b,d|\phi = -20^\circ)}{2} \\ &= \frac{\Phi(-\hat{\phi}_+(s,b,d)/\sigma_D) + \Phi(-\hat{\phi}_-(s,b,d)/\sigma_D)}{2} \end{aligned} \quad (9)$$

where $\Phi()$ is the standard normal cumulative distribution function and $\hat{\phi}_+(s,b,d)$ (respectively $\hat{\phi}_-(s,b,d)$) is the direction estimate of the model when actual motion is up (respectively down).

¹ Fitting including the high-contrast data gave slightly worse results, presumably due to fitting more noise: probabilities at high contrasts are theoretically all close to 0.5 and thus are not informative.

3.3 RESULTS

3.3.1 Experiment

Figure 10 shows the proportion of oblique responses (p_o) as a function of session number. Initial perception of motion direction was accurate for both groups at high contrast, as evidenced by the fact that p_o values are close to the 0.5 line (there is an equal number of trials with perpendicular and oblique motion for every condition and therefore veridical perception corresponds to $p_o = 0.5$). However, consistent with previous findings (Lorenceanu et al., 1993), perception is biased towards perpendicular judgements at low contrast. A 4-way ANOVA of the low-contrast data with factors group, session number, test block number and duration showed a main effect of group on p_o ($F_{1,239} = 10.67$, $p = 0.0013$) as well as an effect of session ($F_{4,239} = 3.8$, $p = 0.0054$) and test block number ($F_{1,239} = 14.84$, $p = 0.0002$) and an interaction of group and session ($F_{4,239} = 4.06$, $p = 0.0035$), indicating that learning was different between groups. Subsequent analysis was thus performed separately for each group. A 3-way ANOVA of the low-speed group with factors session number, test block number and duration showed an effect of test block on p_o ($F_{1,119} = 5.08$, $p = 0.0264$), suggesting a small amount of short-term learning due to the intervening training block; however, perception was unaltered across sessions ($F_{4,119} = 0.28$, $p = 0.893$ for the effect of session). The same 3-way ANOVA of the high-speed group showed that session had a strong effect ($F_{4,119} = 8.06$, $p < 0.0001$). Figure 10 indicates that at low contrast, the initial perpendicular bias gradually diminished until the illusion reversed and the motion direction was most often perceived as being more oblique.

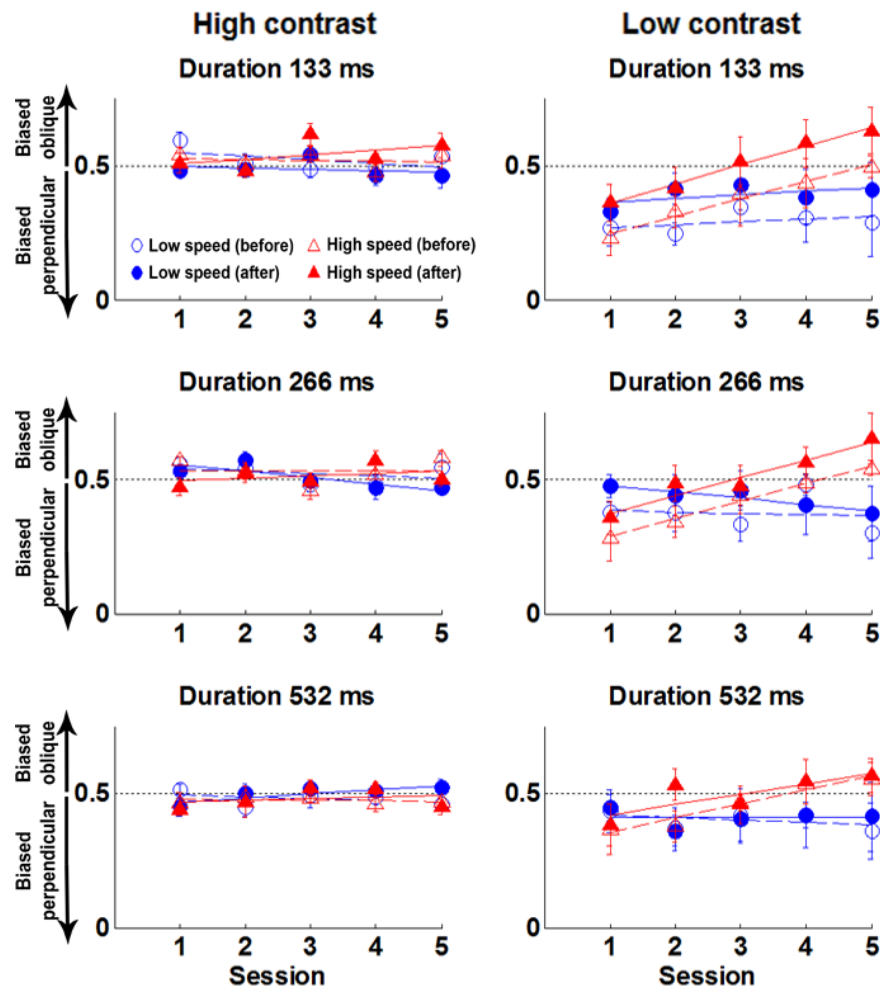


Figure 10: Proportion of oblique responses (p_o) plotted against session number for the low-speed ($n = 6$) and high-speed ($n = 6$) groups and for all durations and contrast levels in the experiment. Lines are best fits to the data. Dashed lines correspond to data from the first test block in each session (before training) whereas full lines correspond to the third block (after training). Error bars represent one standard error of the mean across subjects.

Furthermore, at low contrast, the increase in p_o in the 133 ms (shortest) duration condition across blocks of the same type (pre-training or post-training) appears nearly linear (data points near the best-fit lines). The 266 ms condition is very similar, with both the pre-training and the post-training blocks following an approximately linear trend. To quantify these observations, linear models (both separately for pre- and post-training blocks and with these blocks combined) were fitted to the 133 and 266 ms data (separately and combined) for each group. The fitted models were then evaluated for significance via linear hypothesis tests (`Linhyptest` in MATLAB). Table 1 lists the corresponding p values from the tests showing that linear fits for the high-speed group were all significant at the 5% level.

High-speed group ($n = 6$)	pre-training p ($df = 28$)	post-training p ($df = 28$)	pre & post-tr
133 ms	0.0126	0.0078	0.
266 ms	0.0011	0.0057	0.
133 & 266 ms	0.0023	0.0053	0.

Low-speed group ($n = 6$)	pre-training p ($df = 28$)	post-training p ($df = 28$)	pre & post-tra
133 ms	0.7084	0.6010	0.
266 ms	0.7939	0.3584	0.
133 & 266 ms	0.6748	0.8605	0.

Table 1: p values from linear hypothesis tests on the fitted models for p_o as a function of session number in the low-contrast conditions. Models were fit separately for each of the shortest two durations as well as for both combined. Models were also fit separately for the pre-training and post-training blocks, as well as for both blocks combined.

Interestingly, there is both a fast and a slow learning component reflected in the change in p_o : a 2-way ANOVA on p_o with factors session number and test block number showed a within-session ($F_{1,119} = 11.27$, $p = 0.0011$ for the effect of test block number) and an across-sessions effect ($F_{4,119} = 8.73$, $p < 0.0001$ for the effect of session number). However, that does not mean that there are two types of learning; my hypothesis, which is also reflected in modelling with the introduction of the unlearning rate parameter, is that part of learning occurring in a session reverts by the start of the next session. This may be either due to subjects re-adapting to the statistics of their environment (which presumably is dominated by slow speeds) or due to incomplete consolidation of learning between successive sessions (subjects essentially partially “forget” what they learned in the day). Regarding the effect of session number, t-tests with the Bonferroni correction for multiple comparisons showed that sessions 3, 4 and 5 have a significantly higher p_o than the first session ($p = 0.045$, $p = 0.001$, $p < 0.0001$, respectively) and that session 5 (last) had a significantly higher p_o than session 2 ($p = 0.005$).

Figure 11 indicates that the high-speed group continues to be biased towards perpendicular judgements during the training block, when it is exposed to high speeds, with a tendency for the bias to decrease with time. However, this tendency is not significant: a 2-way ANOVA on training block p_o with factors session number and duration shows neither an effect of session number ($F_{4,89} = 1.18$, $p = 0.3268$) nor an interaction between duration and session number ($F_{8,89} = 0.07$, $p = 0.9998$). These

results provide evidence against the hypothesis that observed changes found in the testing blocks are due to a response bias that develops through the training sessions.

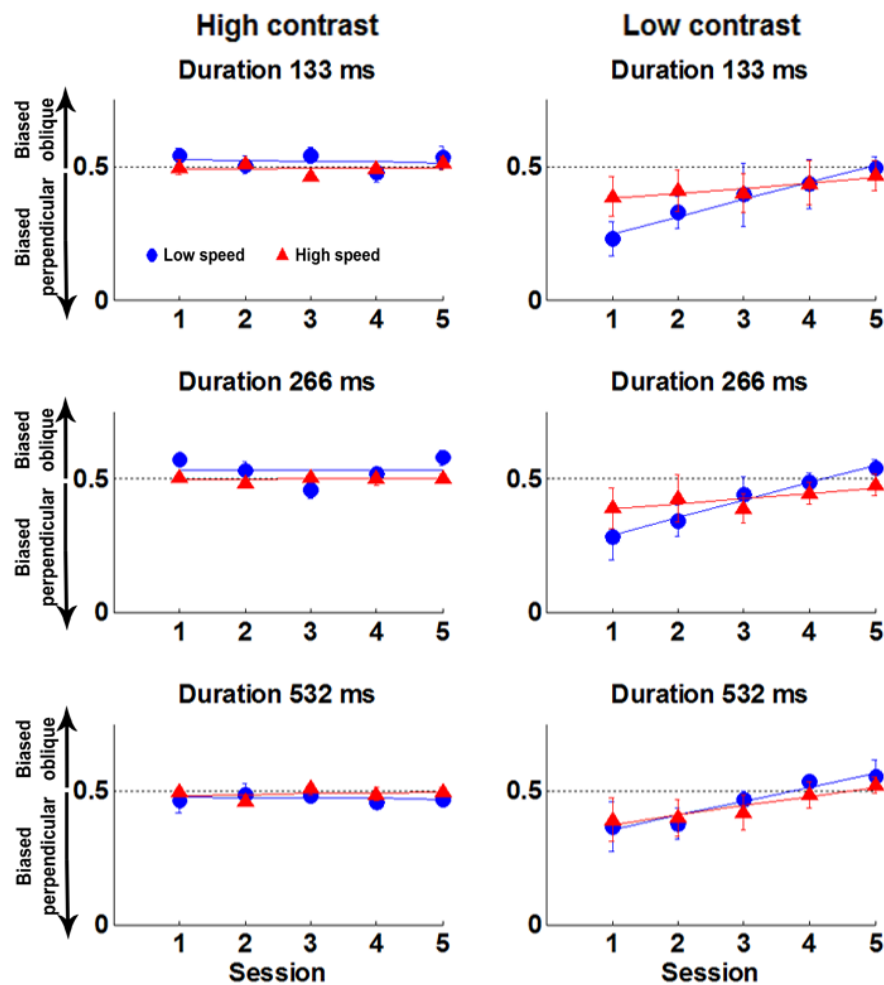


Figure 11: Proportion of oblique responses (p_o) in the training (middle) block for each group (high-speed, $n = 6$ or low-speed, $n = 6$). Error bars represent one standard error of the mean across subjects.

Figure 12 shows that the changes in perception of motion direction occur whether the actual motion is perpendicular or oblique: with exposure to high speeds, subjects show a progressively smaller bias towards perceiving oblique motion as being perpendicular (Figure 12A) but also develop a new bias towards perceiving perpendicular motion as being oblique (Figure 12B). Note that here the dependent variable is performance in the subset of perpendicular and oblique trials.

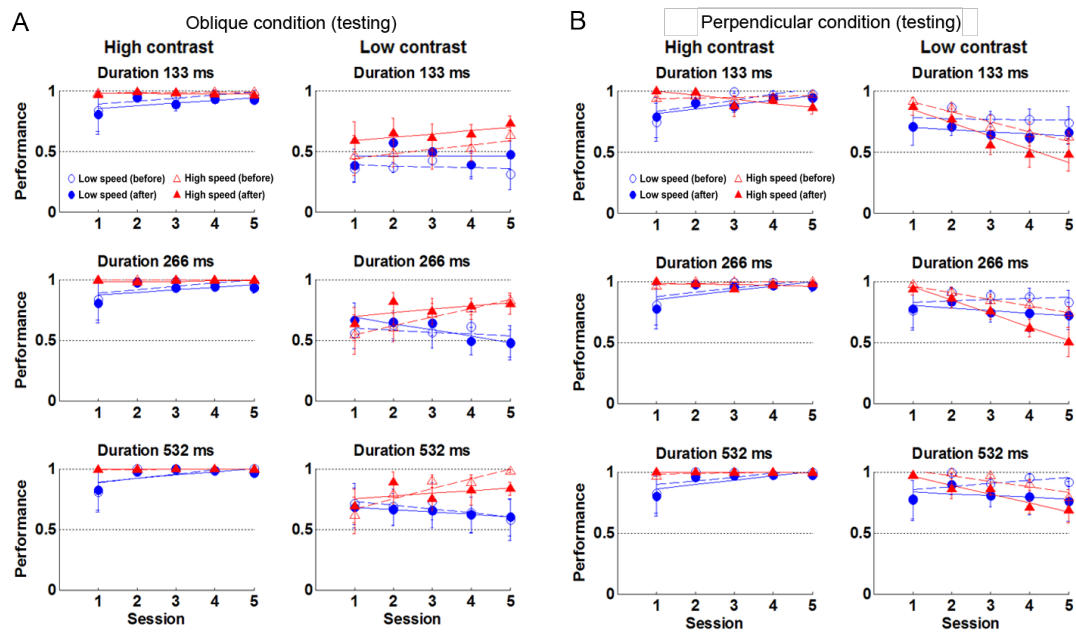


Figure 12: Performance in trials with oblique (A) and perpendicular (B) motion. Data from the low-speed ($n = 6$, blue circles) and the high-speed ($n = 6$, red triangles) groups. Error bars represent one standard error of the mean across subjects.

3.3.2 Modelling

Figure 13 shows the means of the speed prior fitted to the group-averaged data both parametrically and nonparametrically. It is observed that in the high-speed group the prior shifts with training towards higher speeds – reaching a value of $6.2^\circ/s$ at the end of the last session – whereas the prior for the low-speed group remains almost fixed. The fitted σ^2/σ_p^2 values decrease monotonically with duration. This is consistent with the hypothesis that with longer durations, as visual evidence becomes more reliable, the likelihood becomes sharper.

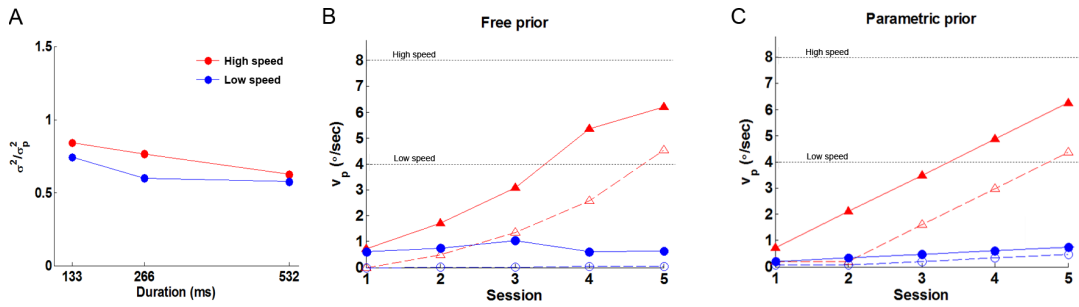


Figure 13: **A**, ratio of the variances of the likelihood (visual evidence) and the prior (expectations) as a function of trial duration. With longer durations, visual evidence becomes more reliable and thus the likelihood is sharper and σ^2/σ_p^2 decreases. **B&C**, evolution of the speed prior through training: mean of the speed prior as a function of session number for the nonparametric and parametric (linear) prior model, respectively. Dashed lines/empty symbols are fits to the pre-training test block; solid lines/filled symbols are fits to the post-training test block. For comparison, the dotted line shows the actual stimulus speed in the test blocks.

The other best-fitting parameters were $\gamma = 0.17$ and 0.99 (high-speed and low-speed group, respectively) and $\sigma_D = 5^\circ$ (both groups). The value of γ is almost 1 in the low-speed group, indicating that the small within-session effect seen in this group vanishes by the next session. In contrast, only 17% of learning is lost between sessions in the high-speed group. The predictions of the fitted model, computed using Equation 9, are represented as triangular and circular points in Figure 14B. Lines are linear fits to the experimental data (fit separately for each duration, block and group), i.e. they are the same as in the top part of the figure (A) which is a replot of the experimental data points and their across-sessions linear fits, shown here for comparison. The model fits the data reasonably well as it can be seen by the fact that the simulated data points fall close to the linear fits. The root mean squared error (RMSE, which has the same units as p_0) of the nonparametric model fit for the high-speed and low-speed group was 0.026 and 0.0403, respectively. The respective values for the parametric variant are 0.0304 and 0.0486. As expected, the nonparametric model fits the data better at the cost of greater complexity (number of parameters). To compare the two models, I used the *corrected Akaike information criterion* (Hurvich and Tsai, 1989), defined as

$$AICc = \frac{2kn}{n - k - 1} - 2L$$

where L is the log-likelihood of the data given the model, k the number of model parameters and n the number of data points. The preferred model according to this

criterion is the one with the lowest score. AICc thus favours models that describe the data well (via the second term) but penalizes complexity (i.e. models with many parameters, via the first term). The model favoured by AICc was the parametric one in both the high-speed ($AICc = 739.2$ parametric vs 745.5 nonparametric) and the low-speed (727.8 parametric vs 736.2 nonparametric) groups.

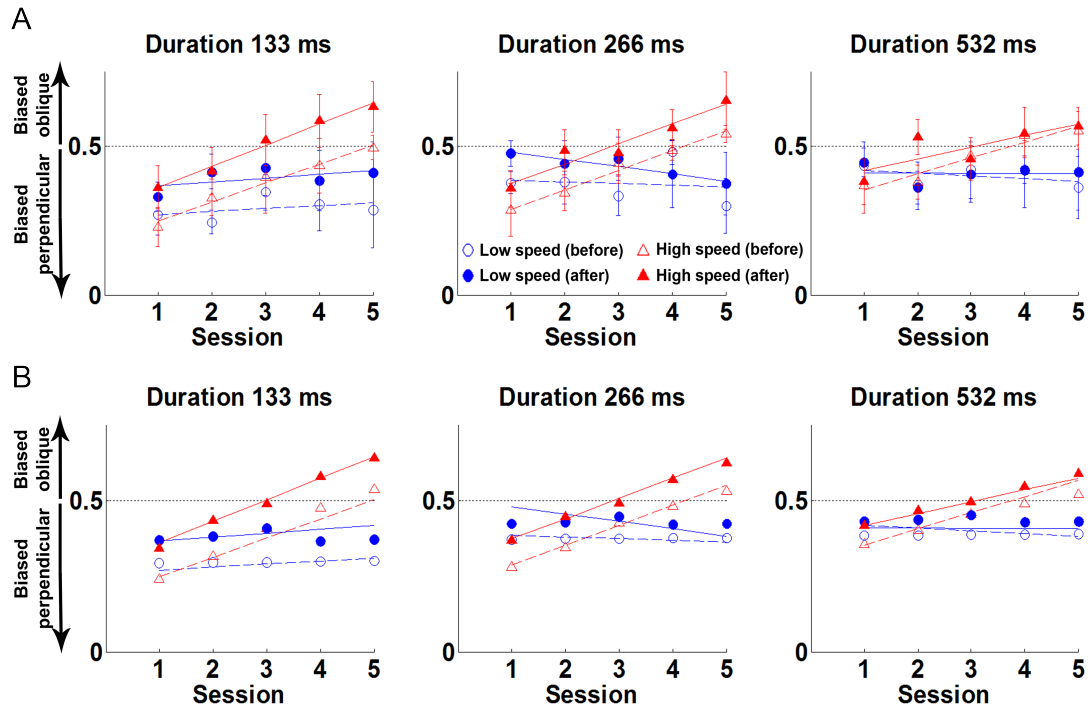


Figure 14: **A**, proportion of oblique perceptions (p_o) in the low-contrast condition, for the 3 trial durations. Each point is the p_o for the first (empty symbols) or last (filled symbols) test block of the session, for the high-speed (red) or the low-speed (blue) group. Lines correspond to linear fits to each block/group combination. Error bars denote between-subjects SEM. **B**, fits from the Bayesian model of motion perception (points) can account for the behaviour of the two groups (lines, corresponding to the linear fits in A).

In theory, not only the mean of the prior could change with exposure, but also its shape. Variants of the model where the prior variance (σ_p^2) was also allowed to vary across sessions were also examined². In the first variant, σ_p^2 is a linear function of session number (requiring 3 additional parameters); in the second variant, σ_p^2 is allowed to vary freely in each session (requiring 6 additional parameters). However, it was found that these models were difficult to fit as the optimization process was prone to local minima, due to the increased number of parameters compared to data points. Essentially, these models were underconstrained by the data; for example, the

² The mean was allowed to vary nonparametrically in these variants to avoid constraining the fitted variance due to constraints in the mean (since these parameters are not independent).

linear-variance model predicts an increase in the mean and a decrease in the variance also in the control (low-speed) group. Why is then no change in p_0 across sessions in the control group? The reason is that, while the mean increases, it remains lower than the test speed; however, the variance decreases, which effectively increases the influence of the prior in the posterior (and thus in perceived speed). Put simply, a given degree of perpendicular bias (perceived speed being lower than the test speed) may have come about either by a prior that favours very low speeds but is broad (compared to the likelihood) or by a prior that favours somewhat higher speeds but is sharp. The design of the experiment is not suitable to differentiate between these two cases and therefore the resulting best-fit values for the (variable) mean and variance of the prior are rendered meaningless.

3.4 DISCUSSION

As seen in 3.3.1, learning in the present experiment appears to have two components – a fast (within-session) and a slow (across-sessions) one. The fast component is a type of perceptual adaptation in which the perceptual system adapts to current conditions (Adams et al., 2004; Chalk et al., 2010) and then is reset. The slow component resembles perceptual learning (Karni and Sagi, 1993; Sagi and Tanne, 1994; Seitz et al., 2005), where changes partially persist after the end of each session, accumulating across days. The lack of a significant effect in the low-speed group is consistent with the need for a learning threshold to be exceeded – a phenomenon that has been observed in perceptual learning (Seitz and Dinse, 2007) and the learning of expectations of object isotropy (Knill, 2007, see 2.2.1).

The results of this experiment demonstrate a clear effect of exposure to fast speeds that is naturally explained in a Bayesian framework incorporating a prior on slow speeds. However, there are certain experimental and theoretical limitations in the present work. Firstly, accuracy in the reconstruction of the prior is limited by the experimental paradigm because the experimentally measured variable that provides information about the prior – proportion of oblique responses – is based on a relatively small number of trials for each condition. Secondly, observers do not report the actual perceived direction but whether direction is above or below a fixed direction (rightward); perceived direction is inferred under assumptions regarding the

stochasticity of responses (decision noise). Thirdly, only two speeds (a lower “test” speed of $4^\circ/s$ and a higher “training” speed of $8^\circ/s$) are used in the experiment; it would be interesting to investigate motion perception at a range of speeds, especially in light of evidence that speed perception is subtly different at low speeds (around $1^\circ/s$), deviating from the Weber-Fechner law (Stocker and Simoncelli, 2006). Fourthly, as discussed in 2.2.3.2, the extended Bayesian model of Weiss et al. (2002) imposes a functional form on the prior that is likely to differ from subjects’ priors. Finally, the variances of the prior and the likelihood appear always as a ratio (σ^2/σ_p^2), i.e. as a single parameter to the model. Therefore, neither the perceptual uncertainty associated with the likelihood nor the width of the prior can be uniquely determined in the current model (also see 2.2.3.2). These issues are addressed in the next two chapters.

In conclusion, the present experimental and modelling results provide, for the first time, evidence in support of the plasticity of the speed prior via an implicit statistical learning procedure (Kim et al., 2009; Di Luca et al., 2010), where no guidance was given to subjects regarding the stimuli’s motion directions. This implicit learning suggests that the brain is constantly revising even its most basic assumptions about the environment even without explicit information regarding the true properties of the stimuli in the world.

SPEED PRIOR RECONSTRUCTION AND INTERACTIVE EFFECTS OF CONTRAST

The work described in this chapter was published in Vision Research (Sotiropoulos et al., 2014) and can be accessed online at <http://www.sciencedirect.com/science/article/pii/S0042698914000248>.

4.1 INTRODUCTION

The experimental paradigm of Chapter 3 provided the first evidence, to my knowledge, in support of the hypothesis that the speed prior is plastic in adults. However, both the experimental paradigm and the Bayesian model used for predictions are limited. These limitations, discussed in 3.4, motivated the work described in this chapter.

An accurate and precise characterization of the speed prior was not the only motivation for the present experiment however. As we saw in Chapter 2, perception has long been known to be susceptible to illusions and biases. Research on visual motion perception in particular has revealed several types of those illusions and biases, such as motion-induced blindness (Ramachandran et al., 1991; Bonnefante et al., 2001), a preference for cardinal directions (Raubert and Treue, 1998; the motion equivalent of the *oblique effect*, Appelle, 1972), illusory “infinite regress” (Tse and Hsieh, 2006) and the dependence of perceived speed on stimulus contrast (Thompson, 1982; Stone and Thompson, 1992; Hawken et al., 1994; Thompson et al., 1996; Blakemore and Snowden, 1999; Hürlimann et al., 2002; Stocker and Simoncelli, 2006, among others). The effect of contrast on perceived speed has been extensively studied in both first (Thompson, 1982; Stone and Thompson, 1992; Hawken et al., 1994; Brooks, 2001; Hürlimann et al., 2002) and second-order (Ledgeway and Smith, 1995) motion; in luminance-based and color-based (Cavanagh et al., 1984; Hawken et al., 1994) motion; using narrowband (Thompson, 1982; Müller and Greenlee, 1994; Thompson et al., 2006) and broadband (Blakemore and Snowden, 1999; Stocker and Simoncelli, 2006)

artificial (such as gratings) as well as natural stimuli (such as a virtual environment simulating the viewpoint of a driver of a road vehicle Snowden et al., 1998). The majority of these studies have found that decreases in contrast cause decreases in perceived speed: a high-contrast stimulus moving at the same speed as a low-contrast one appears faster, and as we saw in Chapter 3 this effect of contrast can also result in changes in perceived direction.

However, a small number of studies have presented evidence for the opposite effect: in certain cases, low contrast results in an *increase* in perceived speed. In these studies, subjects were asked to match the speed of two drifting gratings of different contrasts presented either simultaneously or sequentially. The ratio of the actual speeds of the high- and low-contrast grating at the point of subjective equality (PSE) would be less than 1 if low contrast decreases perceived speed (the high-contrast grating will have to move more slowly in order to have the same apparent speed). Thompson (1982) found that this ratio was indeed less than 1 for temporal frequencies below 8 Hz; above that, the ratio becomes greater than 1 and contrast has the opposite effect. By conducting his experiment at a variety of spatial frequencies, he concluded that this “null point” is invariant with temporal frequency and not speed (which is the ratio of temporal to spatial frequency): regardless of the spatial frequency used, the null point was at 8 Hz. In a later study however, Stone and Thompson (1992) could not replicate this switchover at 8 Hz: in all cases, lower contrast resulted in lower perceived speed. They speculated that their earlier result was a methodological artifact (subjects making judgements other than on speed), supported by the observation that the task became very difficult at high temporal frequencies. In an attempt to settle the issue, Thompson et al. (2006) performed a similar experiment and found evidence for a null point; however, it was invariant in neither temporal frequency nor speed: at a spatial frequency of 2 cycles/°, the null point was 6-8 Hz (i.e. at a speed of 3-4°/s) whereas at a spatial frequency of 8 cycles/°, the null point was 10-14 Hz (or 1.25-1.75°/s). Meanwhile, data from other labs also suggested the existence of a null point at 8 Hz (Hawken et al., 1994; Blakemore and Snowden, 1999). Both studies used 1 cycle/° gratings, corresponding to a speed of 8°/s. At this rate of movement, low contrast slightly increased perceived speed for all four subjects tested by Hawken et al. (1994), whereas Blakemore and Snowden (1999) only found this to be the case

in one of the three subjects tested, although in another subject judgements were more veridical (low contrast did decrease perceived speed but less so).

Despite the wealth of data on the effect of contrast on perceived speed, the issue is not satisfactorily resolved. Is this effect a function of speed? For Bayesian models that assume a monotonically decreasing speed prior, a null point would prove problematic as it would imply a prior that increases beyond that point, at higher speeds. Is such a prior found in human subjects? Furthermore, if the null point were invariant to temporal frequency and not speed, the role of a stimulus-independent speed prior for predictions would be limited.

To address these questions, I performed a speed matching experiment very similar to that by [Thompson et al. \(2006\)](#), using more subjects, each providing a larger number of PSE measurements than in previous studies. Most experimental parameters were the same as those used in existing literature in order to allow for direct comparisons. Additionally, since the experimental paradigm of [Thompson et al. \(2006\)](#) is the same as that of [Stocker and Simoncelli \(2006\)](#), the Bayesian model of the latter can be used in a straightforward manner to reconstruct the subjects' speed priors. Furthermore, by manipulating certain experimental parameters, my aim was to apply the concept of the Chapter 3 experiment – frequent exposure to high speeds driving prior learning – in order to change the speed prior and track its evolution during the course of the experiment.

4.2 PSYCHOPHYSICAL EXPERIMENT

4.2.1 *Methods*

5 subjects (all undergraduate or postgraduate students from the University of Edinburgh) participated in a 5-day experiment similar to [Stocker and Simoncelli \(2006\)](#): a 2-alternative forced-choice (2-AFC) task where subjects were asked to compare the speeds of two rightward-drifting sinusoidal gratings, a reference and a test one, that were presented on either side of a central fixation point on a uniform midgrey background. Stimuli were presented on a Samsung 2043BW LCD monitor (20 inches diagonal size, 1680 × 1050 resolution, 75 Hz refresh rate), driven by an Apple Mac Mini 2009 computer (Intel Core2 Duo CPU, OpenGL-capable GPU) running MacOS

X. Stimuli were generated in MATLAB with the Psychophysics Toolbox v3. Subjects performed the experiment in a dimly lit, soundproof booth and were seated at a distance of 70 cm from the monitor. No chin rest or eye tracker were used; instead subjects were instructed to maintain their head position throughout the experiment. They were also instructed to fixate at the centre of the screen at all times. Each grating was viewed through a sharp-edged circular aperture of 3° (degrees of visual angle) in diameter. The aperture centers were 6° to the left and right of the fixation point. The speed of the reference grating was held constant in each condition tested while the speed of the test grating was adjusted through a QUEST staircase procedure (Watson and Pelli, 1983) until the gratings appeared to move at the same speed. Each staircase terminated after 35 trials, at which point the best estimate (the mode of the posterior pdf of the QUEST algorithm) of the speed of the test grating at the point of subjective equality (PSE) was recorded. The position (left/right) of the reference grating varied randomly but was kept fixed throughout a single staircase (to minimize adaptation effects). The spatial frequency of the gratings was fixed at 2 cycles/° - the lowest of the two values used by Thompson et al. (2006) (2 and 8 cycles/°). Gratings had one of 3 contrast levels (3%, 15%, 95% Michelson contrast). The reference grating had one of 4 speeds (1, 4, 8 and $12^\circ/s$). Each condition corresponded to a unique combination of (reference and test) contrasts and reference grating speeds. Thus there were 12 conditions in total: 3 contrast level pairs (3%/15%, 3%/95% and 15%/95% for reference/test grating, respectively) times 4 reference speeds. 6 of these conditions - the ones corresponding to the lowest two speeds 1 and $4^\circ/s$ - were grouped in a block of 420 trials in total. The remaining conditions - corresponding to the highest two speeds - were grouped in a block of 1680 trials. Thus in each of the 5 sessions, each subject performed 2 staircases for each condition in the low-speed block and 8 staircases for each condition in the high-speed block. Each trial started with a 200 ms fixation period where only the fixation point was visible, followed by a 500 ms stimulus presentation, followed by a response period in which the screen was blank (grey) until the subject pressed the left or right arrow on the keyboard to indicate which of the two gratings appeared to be moving faster. Subjects were offered an optional short break every 10 minutes and a mandatory 5-minute break every 20 minutes. The total duration of a daily session was approximately 1 hour.

The reason that high-speed trials were presented 4 times more often is twofold. First, previous research as well as pilot data indicated that speed judgements are harder at speeds above $8^\circ/s$ and thus there is more variability in subject responses (Stone and Thompson, 1992; Stocker and Simoncelli, 2006), therefore more data is necessary to obtain accurate estimates of the means and variances of the PSE. Second, by presenting high speeds more often, I wished to examine whether the prior favoring slow speeds would gradually change to accommodate the stimulus statistics – i.e. moving some probability mass towards higher speeds, in a similar fashion to previous work (Sotiropoulos et al., 2011).

4.2.2 Results

PSE thresholds from each staircase were examined and average PSEs for each condition are shown in Figure 15. Since in each trial the reference grating had a lower contrast than the test grating, I will refer to the latter as the *high-contrast* grating. If lower contrast results in higher perceived speed, the ratio of speeds of the high and low-contrast grating (hereafter referred to as contrast-dependent bias - CDB) should be greater than one (because the high-contrast grating would have to move faster in order to appear as fast as the low-contrast one).

I first examined whether there was any notable effect of experience on CDB. A 3-way ANOVA on CDB with factors session number, contrast condition and reference speed showed that session had an effect on CDB in all but one subject (all $p < 0.04$). Data from two subjects that exhibited the highest effect of session is shown in Figure 41 (in 8.2 of the Appendix). However, the effect of session was nonspecific and nonmonotonic: there was no consistent change in CDB across sessions and certainly not an increase towards unity. Furthermore, when only the trials with the lowest two reference speeds are considered (where one would expect the greatest effect of exposure to the more frequent high speeds, as in Sotiropoulos et al., 2011), ANOVAs on individual subjects' data failed to show an effect of session (all $p > 0.1$) except in one subject (where $F_{4,30} = 3.59$, $p = 0.017$). Therefore I concluded that there was no consistent perceptual change and thus data from all sessions was pooled, providing a rich data set for subsequent analysis and modelling.

I then examined whether CDB varied as a function of reference speed. Unlike in some previous reports, CDB did not exceed unity in the majority of subjects and conditions; the only exception was one subject (S5) at the lowest contrasts (3% and 15%) and highest speed (12°/s), although CDB was not significantly different from unity ($t_{39} = 1.62$, $p = 0.114$, two-tailed t-test). In all other cases, CDB was less than one, meaning that lowering contrast resulted in a decrease in perceived speed. In other words, apart from the aforementioned single case, there was no “null point” – in contrast to the findings of [Thompson et al. \(2006\)](#).

It is worth noting the considerable variability across subjects, especially with regards to the effect of contrast difference on perceived speed. Furthermore, when data from all subjects is pooled, there is a tendency towards more veridical perception as speed increases in the lowest-contrasts condition, seen as the positive slope of the black line in the leftmost panel of [Figure 15](#). However, the opposite tendency is seen in the highest-contrasts condition (negative slope of the black line in rightmost panel of [Figure 15](#)). The aforementioned 3-way ANOVA showed that there is a marginally significant effect of speed alone ($F_{3,1440} = 2.77$, $p = 0.0402$) but a highly significant interaction of speed and contrast condition ($F_{6,1440} = 9.11$, $p < 0.0001$).

In summary, at contrasts below 15%, results are in qualitative agreement with the finding of [Hawken et al. \(1994\)](#) that as reference speed increased, the effect of contrast on perceived speed diminished and in one case even reversed. However, results at higher contrasts (where both gratings had contrasts at least 15%) show the opposite effect: as reference speed increased, low contrast decreased perceived speed even more strongly.

4.3 MODELLING

4.3.1 *Methods*

To model the relationship between perceived and actual speed under various contrasts and reference speeds and to extract the priors and likelihoods of subjects, the Bayesian model of [Stocker and Simoncelli \(2006\)](#) was used. Briefly, in each trial, an ideal observer computes estimates of the speed of each grating and chooses the grating that has a higher estimated speed. Perceived speed is assumed to be the mode of

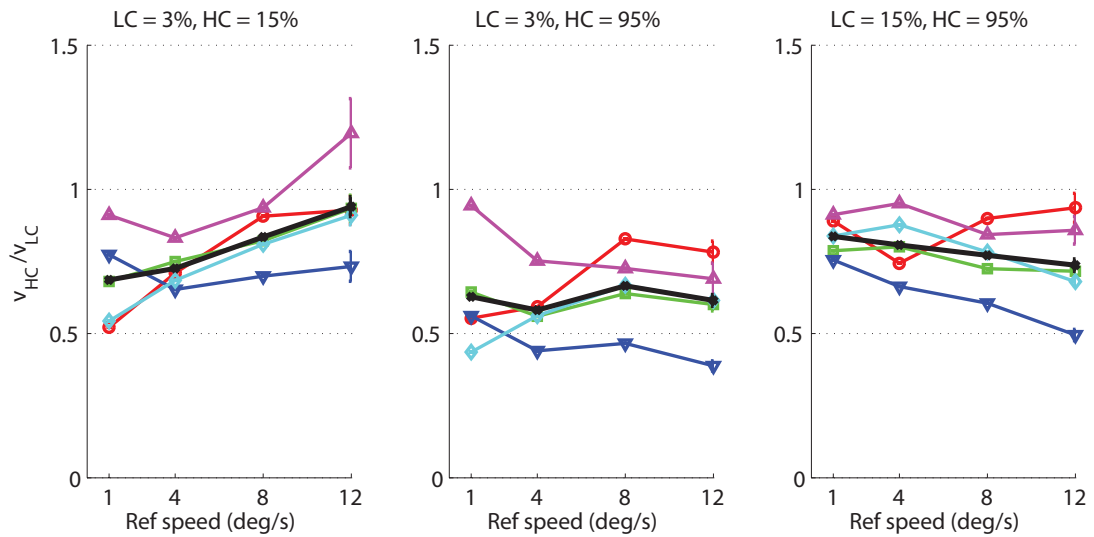


Figure 15: Mean ratio of speeds of the high (v_{HC}) and low-contrast (v_{LC}) gratings at the point of subjective equality (PSE), plotted as a function of speed, separately for each contrast condition. Colored points represent individual subjects ($n = 5$); black points represent the combined data from all subjects. Error bars are standard error of the mean.

the posterior probability density function (pdf) that results from the combination of prior and likelihood. The functional form of the prior is log-linear:

$$p(v) = \exp(av + b) \quad (10)$$

a is the *local* slope of the logarithm of the prior: the log-prior is approximated by a straight line within a narrow speed range but the slope a varies with speed across larger scales (such as across points on a log scale). To reflect the dependence of the slope on speed, it is hereafter denoted by $a(v)$. The likelihood is Gaussian with mean equal to the true stimulus speed and width (standard deviation) separable in speed and contrast:

$$\sigma(v, c) = g(v)h(c) \quad (11)$$

where the dependence on contrast, $h(c)$, obeys a physiologically motivated inverse power law (Stocker and Simoncelli, 2006; Sclar et al., 1990):

$$h(c) = \frac{1}{\sqrt{\frac{c^q}{c^q + c_{50}^q} r_{max} + r_{base}}} \quad (12)$$

The posterior distribution $p(\hat{v}|v)$ of the estimated speed \hat{v} is shown to be Gaussian with mean and variance

$$E(\hat{v}) = v + a(v)\sigma^2(v, c) \quad (13)$$

$$Var(\hat{v}) = \sigma^2(v, c) \quad (14)$$

where v is the true stimulus speed, $a(v)$ is the slope of the logarithm of the prior around v and σ is the standard deviation of the likelihood function, which depends on both speed and contrast. The term $a(v)\sigma^2(v, c)$ is the prior-induced bias of the estimated speed.

The model contains 10 free parameters: 4 for each of $a(v)$ and $g(v)$ (which are not assumed to be any particular function of speed and thus require one parameter for each reference speed used in the experiment) and 2 for $h(c)$. With 10 free parameters, optimization is not trivial and local minima cannot be avoided entirely. [Stocker and Simoncelli \(2006\)](#) exploit the trial-to-trial variability in the data to sufficiently constrain their model by assuming that in each trial the observer samples from the two posterior pdfs and chooses the stimulus whose sample has the highest speed value. They thus derive an expression for the psychometric function

$$p(\hat{v}_2 > \hat{v}_1) = \int_0^\infty p(\hat{v}_2|v_2) \int_0^{\hat{v}_2} p(\hat{v}_1|v_1) d\hat{v}_1 d\hat{v}_2 \quad (15)$$

where \hat{v}_1, \hat{v}_2 are the estimated speeds of the two gratings (reference and test). Equation 15 is fit to the entire dataset via a maximum-likelihood procedure.

I adopt a different, computationally cheaper, approach: given the PSE for a particular condition, the means of the posterior for each grating are equal and thus from Equation 13:

$$v_1 + a(v_1)\sigma^2(v_1, c_1) = v_2 + a(v_2)\sigma^2(v_2, c_2) \quad (16)$$

Since v is known for both gratings, $a(v)$, $g(v)$ and $h(c)$ can be fit to the data but because $a(v)$ and $\sigma(v, c)$ appear in a product there are no unique best-fitting values for them, i.e. the model is not sufficiently constrained as it is. However, the data consists of multiple staircases for each condition and the staircase-to-staircase variability can be exploited to constrain the model. In each session, there are 2 staircases for each of

the low reference speeds and 8 for each of the high reference speeds. Since session number did not have a consistent observable effect on speed perception, data can be pooled, yielding 10 staircases for each of the low-speed conditions and 40 for each of the high-speed conditions. The squared standard error (equivalent to sample variance) σ_{PSE}^2 of the PSE across the 10 (or 40) staircases is informative: it can be shown to be proportional to the variance of the distribution of the test speed v_2 at the PSE and inversely proportional to the number of trials in a single staircase (see 8.1.3). The pdf of the distribution of v_2 (conditioned on the reference speed v_1 and the fact that $\hat{v}_1 = \hat{v}_2$ at the PSE) is Gaussian with variance equal to the sum of variances of the likelihoods of the two gratings. In particular (see 8.1.1) :

$$p(v_2|v_1, \hat{v}_1 = \hat{v}_2) \sim \mathcal{N}(v_1 + a(v_1)(\sigma_1^2 - \sigma_2^2), \sigma_1^2 + \sigma_2^2) \quad (17)$$

where $\sigma(v_1, c_1)$ is written as σ_1 for brevity. Thus the following equation holds:

$$\sigma_{PSE}^2 = \alpha \frac{\sigma_1^2 + \sigma_2^2}{N} \quad (18)$$

where N is the number of trials in a staircase (40 throughout the experiment) and α is a constant of proportionality. By comparing my fitting method against that of [Stocker and Simoncelli \(2006\)](#), using an independent large dataset (21 subjects) obtained with an identical stimulus and task configuration and staircase procedure ([Berbec, 2013](#), see 8.1.2 and Chapter 5), α was found approximately equal to 6.6. Equation 18 thus becomes

$$\sigma_{PSE}^2 = \frac{\sigma_1^2 + \sigma_2^2}{6} \quad (19)$$

Using Equations 16 and 19, the model was fit with a least-squares procedure (`lsqnonlin` function, MATLAB). The (unnormalized) priors were reconstructed as in [Stocker and Simoncelli \(2006\)](#), by numerical integration of the fitted local slope values, according to the following equation (see 8.1.4 for a derivation):

$$p(v) = \exp\left(\int a(v)dv\right) \quad (20)$$

where $a(v)$ (the slope as a function of speed) was linearly interpolated using the slope values at the 4 reference speeds. The maximum and baseline firing rates (r_{max} and r_{base}) in Equation 12 were set to 1 and 0.2, respectively.

4.3.2 Results

The Bayesian model fits the data reasonably well; however, as seen in Figure 16, the model is unable to capture the differential effect of speed on CDB (the interaction between contrast condition and speed described in 5.2.2). This is most evident with subject S5 (magenta triangles), who shows the strongest interaction effect: the model fit is satisfactory in the last two contrast conditions (middle and rightmost panels of Figure 16) but not in the first contrast condition (leftmost panel of Figure 16).

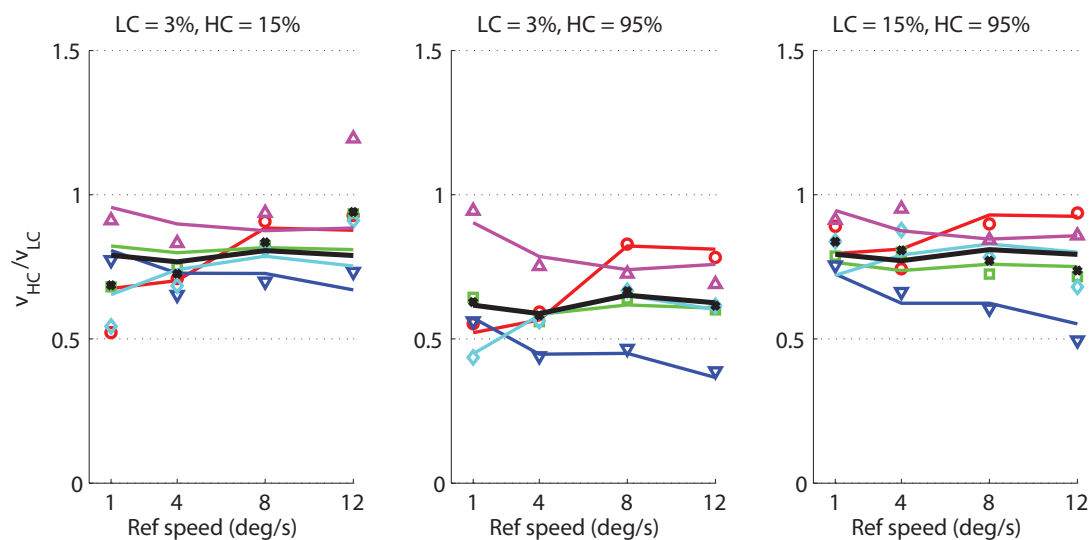


Figure 16: Mean ratio of speeds of the high (v_{HC}) and low-contrast (v_{LC}) gratings at the point of subjective equality (PSE), plotted as a function of speed, separately for each contrast condition. Points represent experimental data (as in Figure 15); lines represent predictions of the fitted Bayesian model. Color represents individual subjects ($n = 5$); black represents the combined data from all subjects.

The extracted priors and likelihood widths (Figure 17) are quantitatively similar to Stocker and Simoncelli (2006). The biggest difference is in the values of $g(v)$, which are somewhat higher for all subjects, compared to the two representative subjects shown in Figure 4 of Stocker and Simoncelli (2006); however, $g(v)$ and $h(c)$ always appear in a product (Equation 11) and therefore there is no unique set of values for either of these functions - in other words, the likelihood model is not fully constrained.

As in [Stocker and Simoncelli \(2006\)](#), $g(v)$ for most subjects is not an increasing function of reference speed, as one might expect from Weber's law; in particular, the group $g(v)$ has a minimum at $4^\circ/s$. This was also reflected in subjects' (and my own) accounts, who found trials with speeds between $4^\circ/s$ and $8^\circ/s$ the easiest. Discrepancies between my extracted components and those of [Stocker and Simoncelli \(2006\)](#) may also be due to the small differences in the stimuli (mainly the different trial duration and spatial bandwidth of the gratings) between my experiment and that of [Stocker and Simoncelli \(2006\)](#). It is also interesting to note the differences in extracted priors among subjects. In particular, S5 exhibits a much shallower prior than S3. Finally, as in [Stocker and Simoncelli \(2006\)](#), the priors for some subjects (S1 and S5) tend to flatten at the lowest and highest speeds.

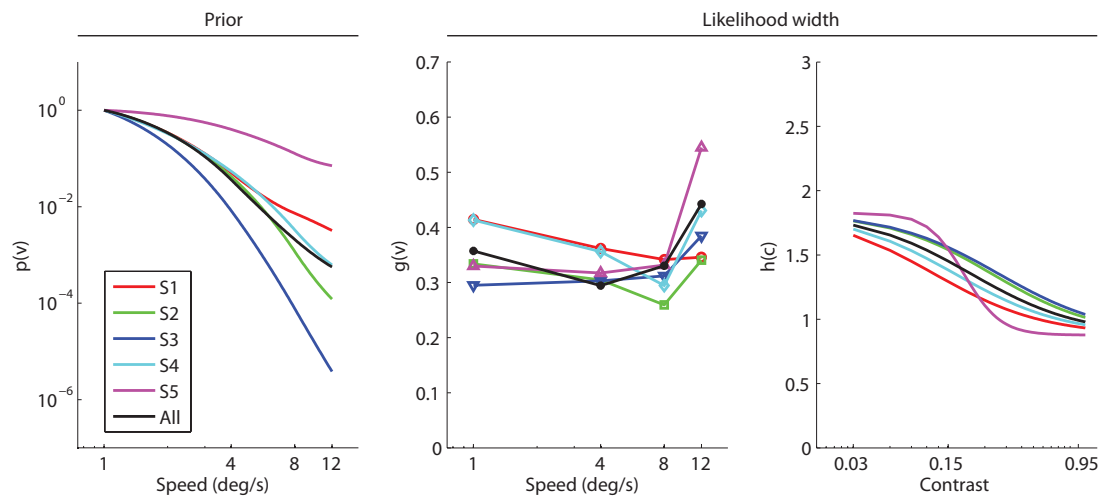


Figure 17: Extracted prior (left panel); speed-dependent $g(v)$ (middle panel) and contrast-dependent $h(c)$ (right panel) components of likelihood width $\sigma(v, c)$ in the Bayesian model. Colored curves are individual subjects ($n = 5$); black curves are all subjects combined.

The Bayesian model of [Stocker and Simoncelli \(2006\)](#) provides a reasonable fit to the data ($R^2 = 0.78$, $SSE = 0.354$) but fails to account for the observed interaction effect: in the model, the effect of speed on the ratio $\hat{v}_{HC}/\hat{v}_{LC}$ is qualitatively the same across all contrast conditions. For example, if the ratio increases with speed in one contrast condition, then it has to also increase in the other contrast conditions. This is because the same prior is used across all contrasts. Clearly, the model needs to be modified to account for the interaction. One approach is the use of a different speed prior depending on contrast level: a prior for high-contrast stimuli that has a smaller slope at low speeds than the prior for low-contrast stimuli (and vice versa

at high speeds). Contrast-modulated priors have been used successfully elsewhere; for example, [Mamassian and Landy \(2001\)](#) used a Bayesian model to describe the way human observers interpret the shape of (2D) drawings of 3D objects based on their shading and on the geometry of parallel dark and bright contours on the objects. Their model included a prior on the tilt and slant of object surfaces and of lighting directions that depended on the contrast of the shading and of the parallel contours. However, their choice rested on the fact that their stimuli had little external noise (their attributes were easily visible) and on the assumption that the internal (observer) noise was equally small. In contrast, internal noise in the form of a likelihood function is an integral part of the present model, which assumes that low contrast broadens the likelihood by increasing the influence of internal noise. Furthermore, to my knowledge, there is no evidence from analyses of natural scene statistics that speed distributions depend on contrast.

Another approach is to model the interaction at the level of the speed measurement, which corresponds to the likelihood mean in the Bayesian model. In particular, an interaction effect would be possible if the average value of the speed measurements depended on the physical stimulus speed and contrast in a nonlinear fashion, such as through a product. One possible choice of such a nonlinearity would be the modification of Equation 13 to

$$E(\hat{v}) = f(v, c) + a(v)\sigma^2(v, c) \quad (21)$$

where $f(v, c)$ is no longer the true stimulus speed but a nonlinear function of true speed and contrast. Such a nonlinearity has been proposed by opponents of Bayesian models of speed perception in an attempt to explain the speed-dependent effect of contrast on perceived speed ([Thompson et al., 2006](#)). In their “ratio model”, itself an extension of the “weighted intersection model” (WIM) of [Perrone and Thiele \(2002\)](#), perceived speed is given by the ratio of a low-pass and a band-pass temporal filter. Since these filters were originally proposed to model speed tuning as a result of the integration of motion-sensitive neurons in area V1, it is natural to apply them at an earlier stage than the Bayesian computations (thought to be carried out in area MT, [Stocker and Simoncelli, 2006](#)). Such an early stage naturally corresponds to modifying

the actual stimulus speed used as input to the Bayesian model - that is, modifying the mean of the likelihood.

The two filters proposed by [Thompson et al. \(2006\)](#), low-pass and band-pass, are inseparable functions of temporal frequency (ω) and contrast (c) and their responses are given, respectively, by:

$$\begin{aligned} p(\omega, c) &= \frac{\bar{p}(\omega)c}{\bar{p}(\omega)c + s_p} \\ m(\omega, c) &= \frac{\bar{m}(\omega)c}{\bar{m}(\omega)c + s_m} \end{aligned}$$

with

$$\begin{aligned} \bar{p}(\omega) &= \sqrt{a^2 + b^2} \\ \bar{m}(\omega) &= \frac{\omega}{k} \bar{p}(\omega) \\ a &= ((2\pi\omega\tau_1)^2 + 1)^{-9/2} \\ b &= ((2\pi\omega\tau_2)^2 + 1)^{-10/2} \end{aligned}$$

τ_1 and τ_2 are time constants, and s_p and s_m are semi-saturation constants of the filters. Perceived speed, as a function of temporal frequency and contrast, is then given by

$$v(\omega, c) = \frac{m(\omega, c)}{p(\omega, c)} \tag{22}$$

Equation 22 thus provides the nonlinearity $f(v, c)$ used in Equation 21.

To avoid adding new free parameters to the model, I used a nested optimization procedure to find the best-fitting values for these parameters and fixed them across all subjects and conditions before fitting the parameters of the Bayesian model. I found that the best-fitting value for both s_p and s_m is 0.5, which is within the range of values used in ([Thompson et al., 2006](#)). The other 3 parameters, τ_1 , τ_2 and k , which in [Thompson et al. \(2006\)](#) (who followed [Perrone and Thiele, 2002](#)) were fixed to 0.0072, 0.0043 (both in units of seconds) and 4 (dimensionless), respectively, had to be changed for our data. In particular, k was set to 0.55 and the time constants were scaled by 4.9, yielding 0.0353 and 0.0211 for τ_1 and τ_2 , respectively. With these parameter values, the output of the ratio model is equal to the true stimulus speed at

all contrasts, except at low speeds (up to $2^\circ/s$), where speed mildly *decreases* with contrast (Figure 18, left panel).

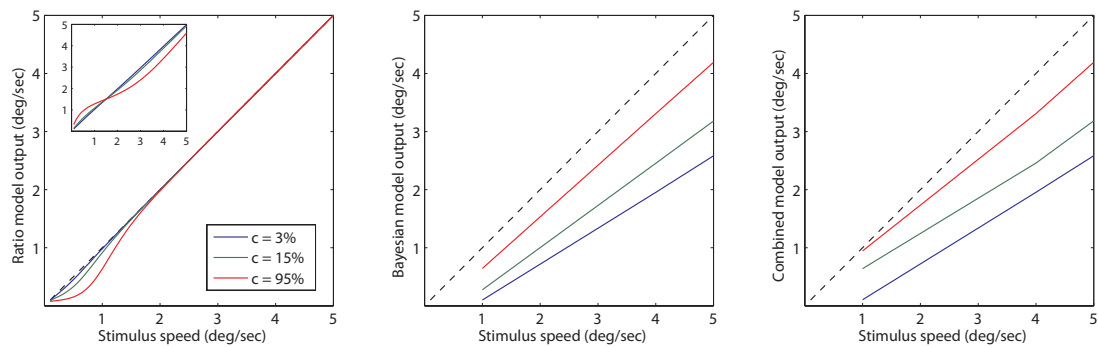


Figure 18: Output of the ratio model as a function of stimulus speed for the 3 contrast values used in the experiment (left panel). Parameter values are from the combined (ratio+Bayesian) model that best fits the entire data. Inset shows the ratio model output with the parameter values used by [Thompson et al. \(2006\)](#); “perceived speed” of the best-fit Bayesian (middle panel) and combined (right panel) models as a function of stimulus speed. Diagonal dashed line corresponds to veridical perception.

By incorporating the ratio model of [Thompson et al. \(2006\)](#) in the Bayesian model of [Stocker and Simoncelli \(2006\)](#), I was able to provide a better description of the data (Figure 19), partially accounting for the interaction effect of speed and contrast and yielding a 31% improvement in the fits ($R^2 = 0.85$, $SSE = 0.243$). This is remarkable given that the number of free parameters in the combined model is the same as in the Bayesian model (namely 10). Note that treating the rest of the ratio model parameters as free resulted in minimal further improvement in fits - too small to justify the increased model complexity.

It should also be noted that the ratio model on its own is not able to account for the data, even if all of its parameters are free. The main reason for this is that there is no set of parameter values that results in an increase in perceived speed with increasing contrast, as is found in my data: up to a certain (low) stimulus speed, which corresponds to the null point reported by [Thompson et al. \(2006\)](#), the output of the ratio model (corresponding to perceived speed) is an increasing function of contrast but beyond that speed the model output is a decreasing or constant function of contrast, across the entire parameter space. In the Bayesian and combined models, however, perceived speed increases with contrast across all stimulus speeds, as seen in my data (Figure 18).

The improved performance of the combined model at low stimulus speeds (around $1^\circ/s$) is due to the opposite effects of the ratio model and of the prior of the Bayesian model on perceived speed. At high contrasts, the prior-induced decrease in perceived speed is attenuated at low speeds, matching the data better (Figure 19, right panel). This attenuation could not have been provided solely by the prior because it would have to apply to all contrast conditions and thus would not fit the data well.

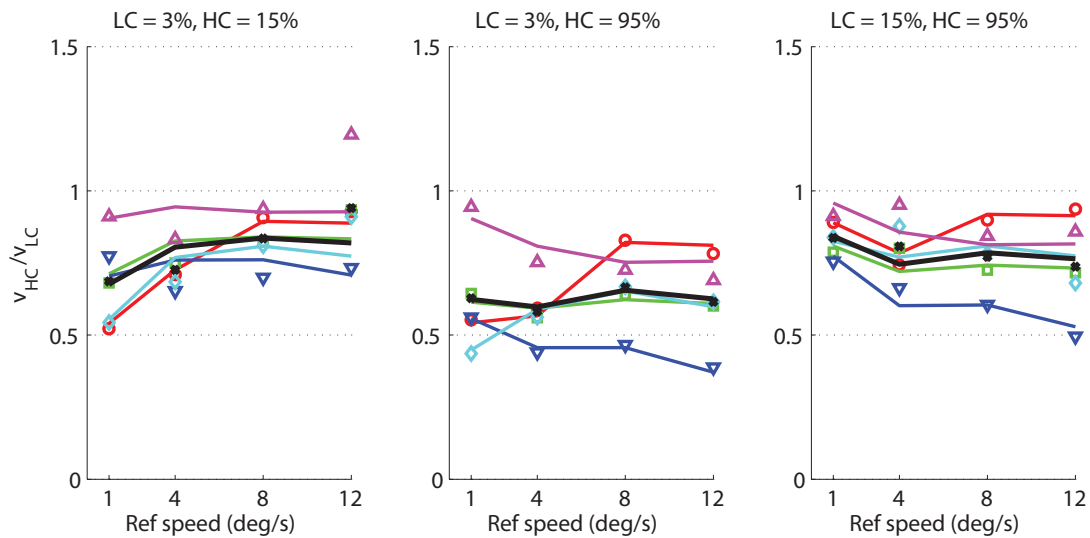


Figure 19: Mean ratio of speeds of the high (v_{HC}) and low-contrast (v_{LC}) gratings at the point of subjective equality (PSE), plotted as a function of speed, separately for each contrast condition. Points represent experimental data (as in Figure 15); lines represent predictions of the fitted combined (Ratio+Bayesian) model. Color represents individual subjects ($n = 5$); black represents the combined data from all subjects.

The extracted priors under the combined model are quantitatively similar to those of the original Bayesian model, with the exception of one subject (S_1), whose prior is significantly steeper under the combined model (Figure 20).

4.4 DISCUSSION

Qualitatively, these results replicate the majority of existing literature in finding that lower contrast decreases perceived speed in all conditions tested. Only in one of the five subjects, at the highest speed ($12^\circ/s$) and only at the lowest contrasts (3% and 15% Michelson contrast for the two gratings) tested was there an inversion of this relationship, although this did not reach significance.

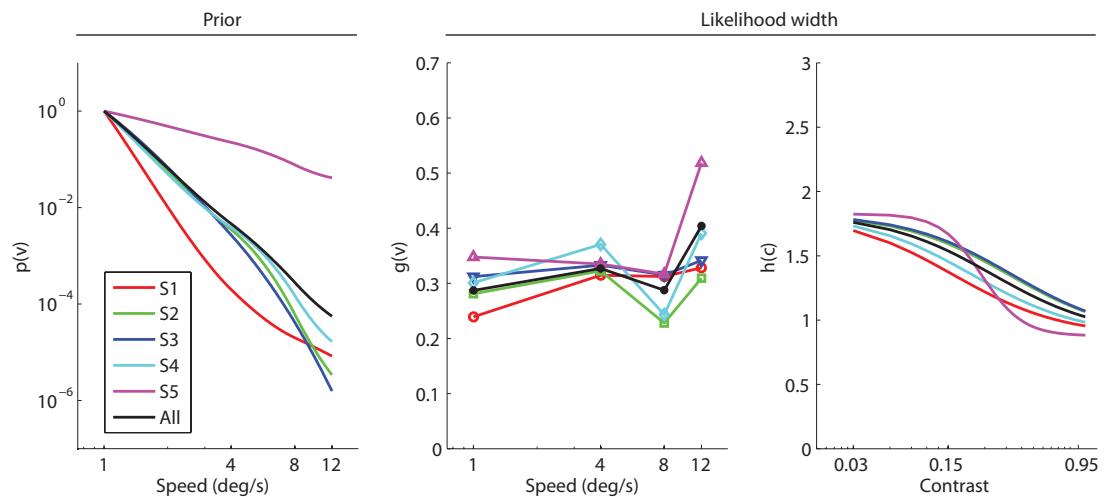


Figure 20: Extracted prior (left panel); speed-dependent $g(v)$ (middle panel) and contrast-dependent $h(c)$ (right panel) components of likelihood width $\sigma(v, c)$ in the combined (Ratio+Bayesian) model. Colored curves are individual subjects ($n = 5$); black curves are all subjects combined.

When the data is quantitatively examined, however, it is apparent that at the lowest tested contrasts, CDB decreases as speed increases (Figure 15, leftmost panel): lowering contrast does not decrease perceived speed as much when speeds are high. Interestingly, the situation is reversed at the highest tested contrasts (15% and 95% Michelson contrast) and CDB becomes more prominent at high speeds.

Does, then, this differential effect of contrast at high speeds depend on absolute contrast levels? Data from existing literature are mixed. Among the studies that have shown evidence that decreasing contrast increases perceived speed (Thompson, 1982; Hawken et al., 1994; Blakemore and Snowden, 1999; Thompson et al., 2006), only that of Hawken et al. (1994) used contrasts as low as the lowest ones used in the present experiment. In the other 3 studies, a reference grating of either 25% (Thompson, 1982), 64% (Blakemore and Snowden, 1999) or 70% (Thompson et al., 2006) was matched against test gratings of lower contrasts. Since these three studies contradict my findings at comparable levels of contrast, a natural question is: are there systematic differences in experimental parameters (other than speed and contrast) between these 3 studies and the rest of the literature (including the present study) that shows evidence of a decrease in perceived speed with decreasing contrast?

I suggest that temporal frequency is not such a parameter; I used temporal frequencies at least as high as those used by all studies that found an increase in perceived speed with decreased contrast (Thompson, 1982; Hawken et al., 1994; Blakemore and

Snowden, 1999; Thompson et al., 2006). If there were a “null point” in the temporal frequency axis, my experiment ought to have hit it. Spatial frequency is likely not a factor either – I used the same value as in one of the conditions in Thompson et al. (2006) (2 cycles/°). The same holds for other stimulus parameters, such as the type, location and drift direction of the gratings – all these parameters were similar in conflicting studies.

Procedural differences are also unlikely to explain why I failed to find the null point. One possible factor could be the different methods of determining the PSE. However, most studies utilized staircase procedures, often very similar - e.g. Blakemore and Snowden (1999) used the same maximum-likelihood-based procedure (QUEST, Watson and Pelli, 1983) that I did. Another possible factor is the task design: the two gratings could be presented simultaneously or successively and there are reports that such manipulations are important (Stone and Thompson, 1992; Blakemore and Snowden, 1999); indeed, two of the studies that conflict with mine used successive presentations (Blakemore and Snowden, 1999; Thompson et al., 2006). However, the other two conflicting studies used simultaneous presentations (Thompson, 1982; Hawken et al., 1994).

It is possible that biases and strategies not directly related to speed perception have an effect, which may also interact with the experimental design: certain biases/strategies may be employed only on certain experimental setups. For example, when subjects are highly uncertain about the relative speed of two gratings, they may be inclined to pick the grating of the higher contrast as being the faster one (because it is also the most salient). Similar biases have been observed in 2-AFC experiments of orientation discrimination (Eero Simoncelli, personal communication). Furthermore, differences in the way subjects are instructed to perform the task may also play a role. For example, subjects sometimes differ in the strategies and response biases they might use when dealing with uncertainty, even if the experimental conditions are identical, and in some cases this is due to subtle differences in instructions. There could also be a “threshold” effect of stimulus uncertainty: at very low contrasts and high speeds (high uncertainty), subjects may switch to a semi-random response strategy, for example by alternating “left” and “right” keypresses.

As for the inability of the experimental procedure to cause a shift in the speed prior via exposure to high speeds more frequently than to low speeds, as evidenced

in the experiment of Chapter 3, I can only speculate about the reasons, based on the differences between the two experiments. One possibility is the difference in stimulus variety between the two experiments: in Chapter 3, there were only two contrast levels and two speeds – one low and one high; in the present experiment each grating had one of three different contrasts, for a total of three contrast conditions. Furthermore, the gratings in the present experiment had a wide range of speeds – one grating had one of four values whereas the speed of the other could have any value from the used speed range. It is thus possible that the large variety of presented speed and contrast combinations made the process of updating the speed prior more difficult. Another conceivable possibility is that subjects are less likely to use the newly learned prior when making explicit speed judgements (as is the case in the present experiment), instead using the default prior that is innate or acquired over their lifetime. In contrast, perceptual judgements in the Chapter 3 experiment were about direction and thus their expectations about speed were only used implicitly (through the relationship between speed and direction during global velocity estimation); it may be the case that the newly learned prior is only applicable during implicit, unconscious estimation. Another related possibility is that people are more accustomed (and thus more sensitive) to estimating direction rather than speed, since precise direction estimation seems like a more common task than speed estimation in everyday life (for example for accurate aiming in sports or hunting).

In regards to modelling, I have presented a significantly faster model fitting procedure than that used by [Stocker and Simoncelli \(2006\)](#); instead of using every trial of every staircase per condition, I used just the final estimate of the staircase and the variability of this estimate across staircases. Effectively, the model is fit by using just $1/N$ th of the data, where N is the number of trials in a single staircase (35 in this case). Using an independent large dataset ([Berbec, 2013](#), see [8.1.2](#) and Chapter 5), I compared my fitting method to that of [Stocker and Simoncelli \(2006\)](#) and, like them, I used the likelihood of the data under the best-fitting model as a performance metric, whereby 100% corresponds to the likelihood of the data when separate Weibull functions for each condition are fit to it and 0% the likelihood under the random (coin-flipping) model. Over the entire dataset, the performance of my fitting method (averaged over subjects) is 87%, compared to 93% of the method of [Stocker and Simoncelli \(2006\)](#). The extracted prior and likelihood components were also very

similar between the two methods. Therefore my method can be useful during model selection/design, allowing rapid iteration between fitting and design, until a suitable model is found, which can then be fit with the method of [Stocker and Simoncelli \(2006\)](#) for slightly more accurate quantitative predictions. My method could also be used in cases where not every trial of the staircase is available, such as when modeling data from existing literature (where usually only the PSE means and standard errors are reported).

By incorporating the ratio model of [Thompson et al. \(2006\)](#) (which in itself is unable to describe the present data, also see [Figure 18](#)) as a pre-processing step in the Bayesian model of [Stocker and Simoncelli \(2006\)](#), I was able to provide an improved account of the interaction effect of contrast and speed. However, this improvement is restricted to low speeds (around $1^\circ/s$), where an increase in contrast causes a mild decrease in speed; there is no improvement at high speeds (around $12^\circ/s$). The observed interaction could be better accounted for by a model in which perceived speed (before the Bayesian computations) decreases with contrast at low speeds but increases with contrast at high speeds. The simple ratio model of [Thompson et al. \(2006\)](#) can only produce the former effect; it cannot produce an increase in perceived speed with contrast at high speeds. It would be interesting to examine whether an extended version of the ratio model, for example one that incorporates band-pass filters tuned to various temporal frequencies, could better account for the interaction; this is left as future work. As a final note, unlike the Bayesian model, the ratio model seemingly constitutes a departure from a normative explanation of speed perception. However, the ratio model was proposed, in the form of the WIM model ([Perrone and Thiele, 2002](#); [Perrone, 2005](#)), as a biologically plausible way of achieving variable speed tuning in MT neurons by using a small number of V_1 neurons tuned not to speeds but to a *limited* range of spatial and temporal frequencies. Thus, while the WIM model does not result in optimal perception, it can be argued that this is due to biological constraints earlier in the visual hierarchy (V_1) rather than an inherent sub-optimality in the model, in much the same way that the Bayesian model is optimal under the assumption of noisy measurements.

5.1 INTRODUCTION

The psychophysical data collected in the experiment of Chapter 4 allowed for the nonparametric reconstruction of subjects' speed priors using the methods of [Stocker and Simoncelli \(2006\)](#). The reconstructed priors had considerably heavier tails than the best-fitting Gaussians would, instead approximating inverse power-law distributions, which are monotonic, as were the priors in [Stocker and Simoncelli \(2006\)](#). This was reflected in their subjects' data by the fact that low contrast decreased perceived speed throughout the entire speed range used. [Stocker and Simoncelli \(2006\)](#) noted that if at higher speeds low contrast *increased* perceived speed, their model would account for this with a prior that increases at these speeds. Could the speed prior be nonmonotonic in some observers, showing an increase at speeds higher than those used by [Stocker and Simoncelli \(2006\)](#)?

On a related note, the results of Chapter 4 confirmed the large interindividual variability in priors that was seen in previous work ([Stocker and Simoncelli, 2006](#); [Hedges et al., 2011](#)). Where could this variability stem from? Innate, genetically determined differences is one possibility; another is differences in visual experience: given that the visual system is very plastic during infancy and even retains significant plasticity in adulthood, it is conceivable that individuals exposed to different environmental motion statistics may differ in their motion estimation biases. For example, professional drivers are routinely exposed to fast optical flow, particularly in the periphery, and video games (especially action games) contain large amounts of fast motion. Avid players who spend several hours per week playing games may be exposed to fast speeds often enough so that, over time, their speed prior has changed to accommodate the different statistical structure of their environment, which includes the virtual environments of games. Video games are known to have diverse effects on vision ([Bavelier et al., 2012](#)), such as the improvement of target identification when

the target is surrounded by distractors in close proximity (Green and Bavelier, 2007) or the enhancement of temporal resolution of the visual system as evidenced by significant reductions in backward masking – whereby a target is harder to see if it is quickly followed by a distractor (Li et al., 2010). Of particular interest here is also the enhancement of contrast sensitivity that has been observed in avid gamers (Li et al., 2009) which, together with the possible increased exposure to fast motion, may lead to measurable differences between gamers and nongamers under the experimental paradigm of Chapter 4. Could avid video gamers have speed priors that favour fast speeds more than those of nongamers?

The work described in this chapter was undertaken to address the above questions. The experimental part of this work includes contributions by Dan Berbec as part of his Honours project (Berbec, 2013) supervised by Peggy Seriès and co-supervised by myself. The present experiments were based on the code for the experiments of the previous two chapters. Dan recruited the experiment participants, performed the experiments and implemented the data cleanup algorithm described in 5.2.2.

5.2 PSYCHOPHYSICAL EXPERIMENTS

5.2.1 *Methods*

21 subjects (all undergraduate or postgraduate students from the University of Edinburgh) participated in two interleaved experiments conducted over two consecutive days – on each day, a session of Experiment 1 (see below) was followed by a session of Experiment 2. Subjects were chosen based on their experience with first-person-shooter (FPS) video games, which typically expose the player to significantly high speeds in the form of fast-moving objects or fast optical flow. Of the 21 subjects, 10 were gamers having played FPS games for at least 5 hours per week in the past year; the remaining 11 subjects rarely (or never) played games, and these were mostly static games. Of the 10 gamers, 2 were female and 8 were male; of the 11 non-gamers, 6 were female and 5 were male. In data analysis, subjects were grouped both according to gender and according to gaming experience; thus there were 4 groups (gamers, non-gamers, males, females) with partial overlap and pairwise unbalanced (more gamers than nongamers and more males than females).

The paradigm of Experiment 1 was identical to the Chapter 3 experiment, i.e. subjects were asked to judge the direction of motion of a field of lines. The paradigm of Experiment 2 was identical to the Chapter 4 experiment, i.e. subjects had to compare the speeds of drifting gratings of different contrasts. Experimental parameters (including the laboratory environment and computer hardware and software) for the two experiments were thus the same as their Chapter 3 and 4 counterparts, with the main difference being that this time the aim was to measure the speed and direction biases (and extract the priors) of the subjects and not attempt to change these biases through extensive exposure to high speeds as in the previous experiments. Thus sessions in the new experiments had to be long enough to provide reliable measurements of subjects' speed and direction judgements but not as long as to significantly change their speed prior through exposure. More specifically, the new experiments differed from their Chapter 3 and 4 counterparts in the following ways:

5.2.1.1 *Experiment 1 (field-of-lines)*

- There was no training block in each of the two sessions; a session consisted of two short test blocks of the exact same structure and stimulus parameters as the pre-training and post-training blocks of the Chapter 3 experiment.

5.2.1.2 *Experiment 2 (gratings)*

- Grating speeds were chosen from the set 2, 4, 8, 12, 16°/s, instead of 1, 4, 8, 12°/s in the Chapter 4 experiment. The highest speed (16°/s) was used in order to explore whether speed biases at such high speeds are comparable to those at lower speed. [Stocker and Simoncelli \(2006\)](#) note that speed judgements above 12 were unreliable for their subjects and thus were not studied. In pilot trials, I found that response variability was significantly higher at 16°/s than at 12°/s however not to the extent that it impairs judgements excessively.
- High speeds (8, 12 and 16°/s) were presented as frequently as low speeds. In total (over the two experimental sessions), each subject performed 6 staircases at each of the 5 speeds and 3 contrast pairs (15 conditions in total), compared to the 10 and 40 staircases performed at each condition at low and high speeds, respectively, in the Chapter 4 experiment.

The experimental setup (computer software and hardware, laboratory environment) was identical to that of Chapters 3 and 4.

5.2.2 Results

5.2.2.1 Experiment 1 (field-of-lines)

Collected data was filtered according to two *pre hoc* criteria, based on pilot data and on the main results from the Chapter 3 experiment. Data that met at least one of these criteria – which was the case for 7 subjects – was excluded from further analysis. Note that only the blocks of trials that met one or both of the *pre hoc* criteria were eliminated from the subject's data; other blocks of the same subject that did not meet any of the criteria were kept. Out of the 20 blocks in total, between 2 and 6 blocks (mean: 3.9, st.dev.: 1.5) were eliminated. The *pre hoc* criteria were the following:

1. Too large proportion of oblique responses ($p_o > 0.7$) in the low-contrast trials (at any duration). Such a high p_o is highly unlikely to correspond to true perception of oblique motion (since the bias in untrained subject is in the opposite direction) and is more likely to be an artifact such as a response bias when the trial is hard.
2. Too small proportion of oblique responses ($p_o < 0.1$) in the low-contrast trials at the longest duration (533 ms). At this duration, oblique motion is correctly perceived by naive observers with a significant probability. Near-absent oblique perception suggests either a response bias (e.g. due to trial difficulty) or unreliable perception at low contrasts.

After data filtering, a 3-way ANOVA on p_o in the low-contrast trials with factors session number, duration and gaming status did not reveal an effect of session (first or second day of the experiment, $F_{1,94} \simeq 0$, $p \simeq 0.99$) or an interaction between any of the factors, thus the data from the two sessions were pooled.

As expected, perception of motion direction was accurate for all 4 (pairwise overlapping; see first paragraph of 5.2.1) groups at high contrast, as the proportion of oblique responses (p_o) was close to 0.5 in all 3 trial durations with small error bars, very similar to the respective high-contrast data of the Chapter 3 experiment (Figure 10, left panels).

At low contrast, in agreement with previous results (Lorenceanu et al., 1993; Sotiropoulos et al., 2011), all 4 groups show a preference for perpendicular motion, as evidenced by the low p_o values seen in Figure 21. The top (A) and bottom (B) panels show results before and after data filtering, respectively. Contrary to expectations, the largest difference is seen not between gamers and nongamers but between males and females. A 2-way ANOVA on p_o with factors duration and gaming status showed no effect of gaming status ($F_{1,54} = 0.03$, $p = 0.867$), duration ($F_{2,54} = 0.87$, $p = 0.426$) or their interaction ($F_{2,54} = 0.34$, $p = 0.717$). The same 2-way ANOVA but with gender instead of gaming status revealed an effect of gender ($F_{1,54} = 4.64$, $p = 0.036$); however, because the gaming and gender groups overlap, they are not independent and therefore a correction for multiple comparisons must be performed. A straightforward way is the Bonferroni correction, which effectively divides α , the significance threshold, by the number of comparisons and thus $\alpha = 0.025$ and the effect of gender becomes borderline insignificant. Figure 21 shows that the greatest difference between males and females is in the 532 ms condition. Indeed, a 1-way ANOVA for the 532 ms data with factors gender shows a borderline significant (after the Bonferroni correction) effect ($F_{1,17} = 6.17$, $p = 0.024$).

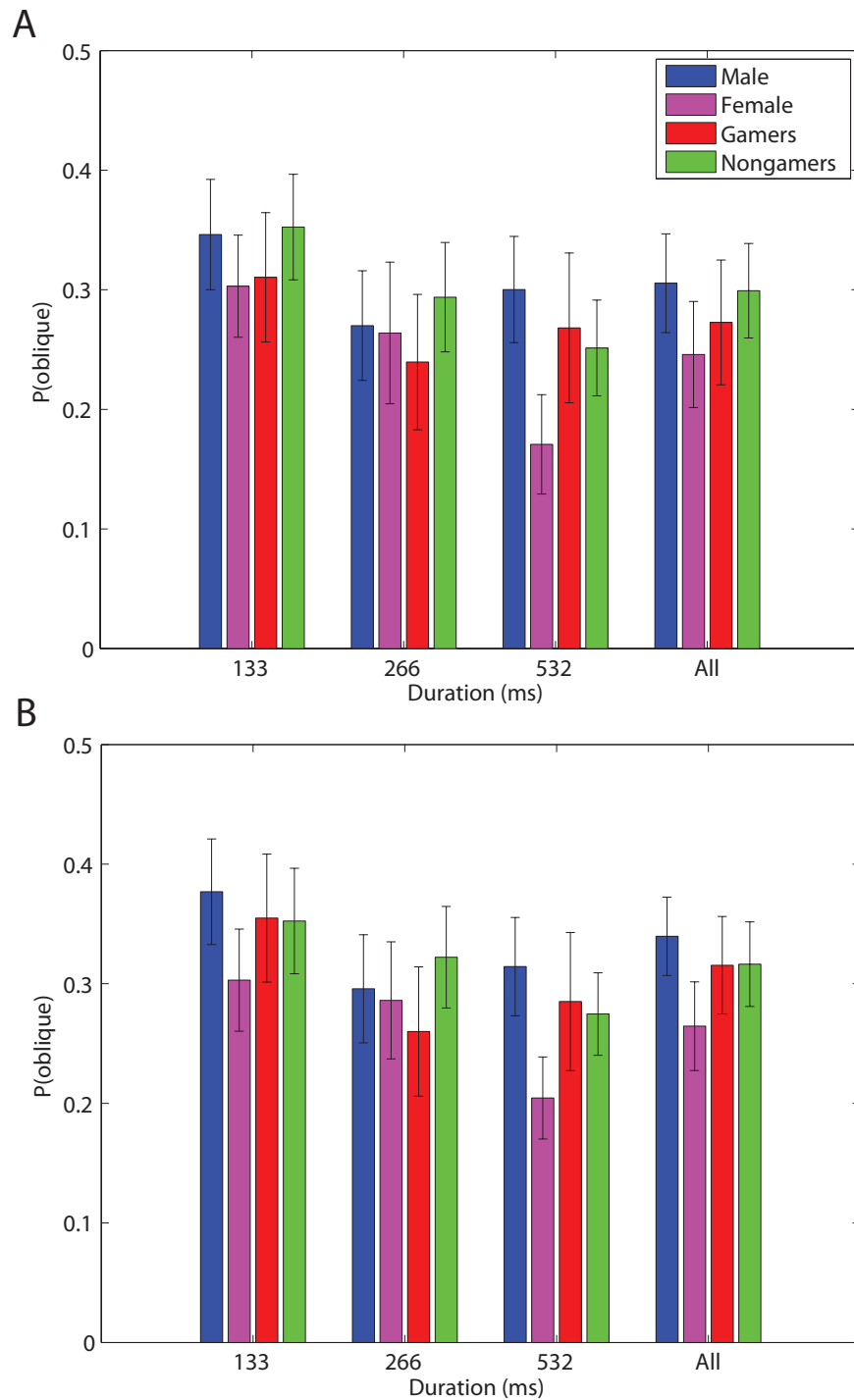


Figure 21: Experiment 1: proportion of oblique responses (p_o) of subjects ($n = 21$) grouped by gaming experience (10 gamers, 11 nongamers) or gender (13 males, 8 females), plotted separately for each duration. **A**, raw data. **B**, filtered data, after a number of trials were excluded from analysis according to the *pre hoc* criteria listed in 5.2.2. Error bars are standard error of the mean.

5.2.2.2 Experiment 2 (gratings)

As in Chapter 5, PSE (point of subjective equality) thresholds were averaged over all staircases for each contrast condition. Since in each trial the reference grating had a lower contrast than the test grating, I refer to the latter as the “high contrast” grating. If lower contrast results in higher perceived speed, the ratio of speeds of the high and low-contrast grating (i.e. the contrast-dependent bias - CDB) should be greater than one. Figure 22A shows the CDBs for the 21 subjects grouped by gaming experience and gender. Contrary to initial expectations but in accordance with the results of the field-of-lines experiment, there is a greater difference in respective CDBs between males and females than between gamers and nongamers. In the trials with the lowest contrasts for the two gratings (leftmost panel), CDB is greater than one for the highest two reference speeds (12 and 16°/s) in the male and gamer group; in all other cases CDB remains below one.

Upon detailed inspection of the psychometric data, it was found that, especially at high reference speeds and low contrasts (which are exactly the conditions where CDB exceeds unity), some staircases did not converge and furthermore, some staircases converged at very different PSE values than the rest (in total 6) for that condition. These abnormal staircases were observed in several subjects and are likely to indicate random or biased responses when trials are hard, as is the case at low contrasts. For example, faced with uncertainty about stimulus speeds, subjects may respond at random or choose the grating with the lowest contrast. To eliminate noise due to these staircases, the data was filtered with an algorithm that discarded all staircases that either did not converge (examining the fluctuations in the last 10 trials of the staircase) or converged to a PSE that was sufficiently different from the average PSE (across all 6 staircases) for that condition. The definition of “sufficiently different” that yielded the best results (verified by manual inspection of the data) is a PSE that differs from the mean PSE by at least 1.15 standard deviations.

Figure 22B shows results from the “clean” data – after all invalid staircases were removed by the algorithm. The main difference from the raw data are that in the lowest contrast condition (leftmost panel) all CDBs at 16°/s are now less than one, with only the CDB of the male group at 12°/s exceeding unity, although nonsignificantly ($t_{53} = 1.45$, $p = 0.152$ for the difference of CDB from 1, t-test). In agreement with the findings from the experiment of Chapter 4, there is no “null point” where CDB

crosses the unity line from below (see 5.2.2). Another difference between the raw and the clean data is that at the highest contrasts (rightmost panel of Figure 22, B) there is a general (in all groups and at all reference speeds) tendency for CDBs in the clean data to be lower than the respective ones in the raw data.

The interaction effect of speed and contrast that was observed in the Chapter 4 experiment is also present here, with the exception of the female group. A 2-way ANOVA on CDB with factors speed and contrast showed that there is a marginally nonsignificant effect of speed ($F_{4,462} = 2.29, p = 0.059$) and no interaction of speed and contrast ($F_{8,462} = 0.58, p = 0.798$) in females. However in males, gamers and nongamers there is a significant effect of speed ($F_{4,782} = 6.53, p < 0.001$; $F_{4,620} = 4.57, p = 0.002$; and $F_{4,624} = 3.5, p = 0.016$, respectively) and an interaction of speed and contrast ($F_{8,782} = 4.95, p < 0.0001$; $F_{8,620} = 2.75, p = 0.011$; and $F_{8,624} = 2.49, p = 0.023$, respectively). Overall, the effect seen in the Chapter 4 experiment is still present for speeds up to $12^\circ/s$, although milder and not in all groups: at contrasts below 15%, as reference speed increases, the effect of contrast on perceived speed diminishes and in one case even reverses. Higher contrasts (where both gratings have contrasts at least 15%) show the opposite effect: as reference speed increases, low contrast decreases perceived speed even more strongly. However, at $16^\circ/s$, a speed not tested in the Chapter 4 experiment, the effect of contrast on perceived speed is again somewhat stronger, terminating the upward trend of the curves in the leftmost panel of Figure 22B. As in Experiment 1, all statistical reporting in the present experiment includes the Bonferroni correction.

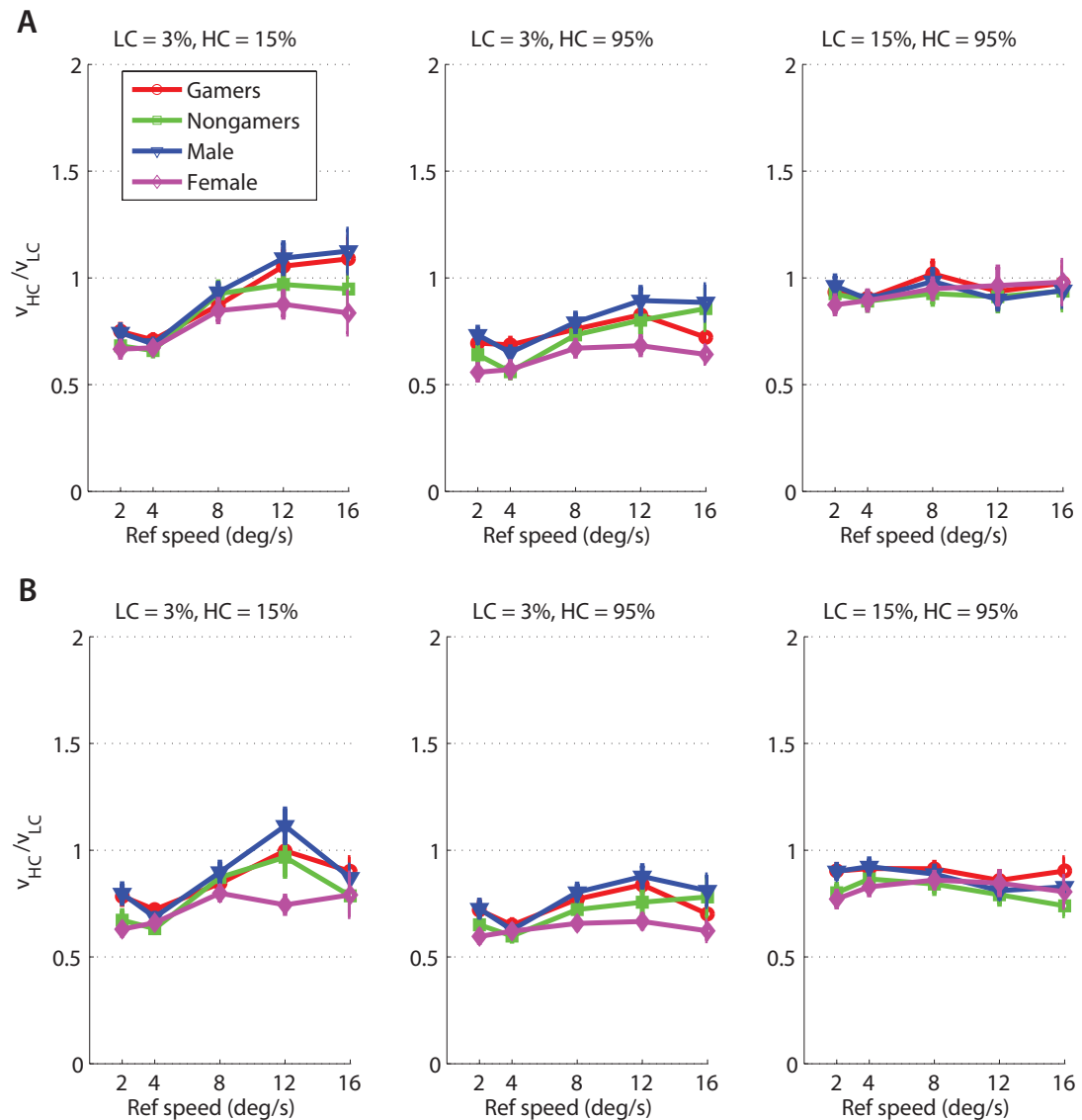


Figure 22: Experiment 2: mean ratio of speeds of the high (v_{HC}) and low-contrast (v_{LC}) gratings at the point of subjective equality (PSE), plotted as a function of speed, separately for each contrast condition. Subjects ($n = 21$) were grouped by gaming experience (10 gamers, 11 nongamers) or gender (13 males, 8 females). **A**, raw data. **B**, filtered data, after invalid staircases (those that did not converge or converged to an abnormal value) were excluded from analysis. Error bars are standard error of the mean.

By comparing the two experiments, it becomes apparent that the differences in biases between males and females are congruent: females exhibit more biased perception at low contrasts, seen by the lower p_o in Experiment 1 (Figure 21, 532 ms condition) and the lower CDB in Experiment 2 (Figure 22, left and middle panels). It is therefore reasonable to ask whether this congruence is seen on a subject-by-subject basis. Assuming that a single speed prior for each subject is used across tasks, in a similar manner to the light-from-above prior, which applies to different tasks

(Adams, 2007; see 2.2.2), it is possible that p_o values from Experiment 1 would correlate with CDB values from Experiment 2 (as a prior that favours slow speeds causes a decrease in both) on a subject-by-subject basis (such correlations were observed by Adams, 2007). Pearson's correlation test, however, did not reveal significant correlations between the p_o values in the 533 ms condition of Experiment 1 and CDB values in any of the conditions of Experiment 2 ($p > 0.13$ for all combinations of reference speeds and contrast conditions).

5.3 MODELLING

5.3.1 *Methods*

The Bayesian model of Stocker and Simoncelli (2006) used in Chapter 4 was also applied to this data set; in fact, Experiment 2 provided a useful validation data set for the purpose of testing my model fitting method, which is different to the method used by Stocker and Simoncelli, 2006. Details of the model and the fitting methods are given in 4.3.1. The model was fit to the raw data – not the filtered (“clean”) data – for two reasons. Firstly, in the clean data all staircases that are considered invalid have been discarded. As there are 40 trials in each staircase, this resulted in the elimination of a significant number of trials, several of which may be valid. Since the fitting method of Stocker and Simoncelli (2006) uses all trials, and not just the last trial of the adaptive staircase (which represents the PSE), this would result in information loss. Secondly, my fitting method relies on the staircase-to-staircase variability. In Experiment 2 the raw data consists of just 6 staircases per condition and this number is smaller in the clean data in some cases. This small number of PSEs per condition reduces the accuracy of the variance of the PSE on which my fitting method relies to estimate the likelihood variance. Indeed, by its very nature, the filtering algorithm is likely to lead to an underestimation of the PSE variance.

5.3.2 *Results*

Priors and likelihood components are plotted in Figure 23, where differences between groups can be seen. $g(v)$ values are comparable between males and females under

both fitting methods; $h(c)$ values are practically identical under my fitting method (solid curves in Figure 23) but there is a difference around the middle of the contrast range under the fitting method of [Stocker and Simoncelli \(2006\)](#) (dashed curves). However, under both fitting methods, there is considerable difference in the priors between males and females as well as between gamers and nongamers: the females prior is very similar to the gamers prior, decreasing more steeply compared to the prior of the males and nongamers (which are also very similar). Furthermore, under the fitting method of [Stocker and Simoncelli \(2006\)](#) the priors of all groups except the gamers increase by a small amount at the highest speed.

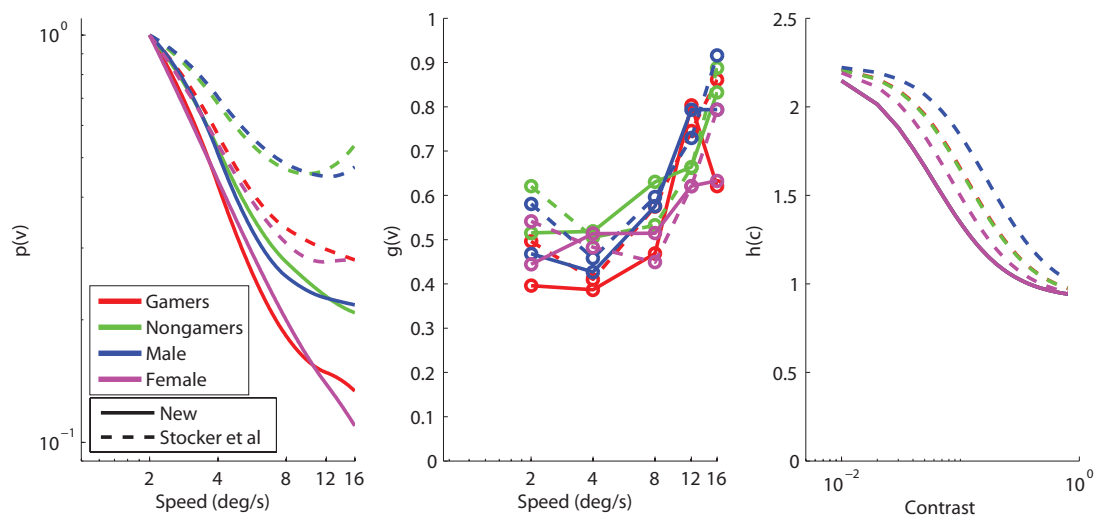


Figure 23: Bayesian model fit to the data of each of the four groups (denoted by different colours). Dashed lines correspond to the fitting method of [Stocker and Simoncelli \(2006\)](#); solid lines correspond to my fitting method. Extracted priors $p(v)$ (left panel); speed-dependent $g(v)$ (middle panel) and contrast-dependent $h(c)$ (right panel) components of likelihood width $\sigma(v, c)$ in the Bayesian model. Priors are unnormalized.

I then compared my method to that of [Stocker and Simoncelli \(2006\)](#) and, like them, I used the likelihood of the data under the fitted model as a performance metric, whereby 100% corresponds to the likelihood of the data when separate Weibull functions for each condition are fit to it and 0% corresponds to the likelihood under the random (coin-flipping) model. Averaged over the four groups, the performance of my fitting method is 96.5% whereas the method of [Stocker and Simoncelli \(2006\)](#) yields a 98.4% performance. The likelihood components are also similar between the two methods, the main difference being that the contrast-dependent component of the likelihood $h(c)$ has lower values with my fitting method. This is offset by the

greater values of the local slopes $a(v)$ of the prior so that the speed biases are very similar between the two methods¹. The greater $a(v)$ values in my fitting method also result in steeper priors. Interestingly, with both methods, the priors of gamers are very similar to those of females, and the priors of nongamers are very similar to those of males, despite the facts that most gamers are male and that the CDB data does not show the same similarities between groups. These results are explained by the group differences in the $g(v)$ component of the likelihood.

As in the Chapter 4 experiment, there was significant interindividual variability in the reconstructed priors but, unlike the Chapter 4 priors, they decrease much less steeply with speed. Figure 24 shows the two subjects with the most extreme priors: one that favours high speeds the most (A), to the point that it is nonmonotonic and starts to increase beyond $6^\circ/s$; and one that favours low speeds the most (B) and is monotonic. The raw psychometric data of these two subjects, as well as model fits, are shown in Figures 42 and 43 in Appendix (8.3). In total, 7 subjects had nonmonotonic priors extracted with the fitting method of [Stocker and Simoncelli \(2006\)](#); the corresponding number with my fitting method was 5. The local minimum of the prior in each of these subjects varied but it was always in the range $2 - 8^\circ/s$. This range contains the “null point” found by [Thompson et al. \(2006\)](#) (the point where low contrast switches from resulting in decreased perceived speed to resulting in increased perceived speed) at around $4^\circ/s$. As for the quantitative difference between the priors in Chapter 4 and the present priors, it is discussed and put in a broader context in Chapter 7.

¹ Equation 13 in Chapter 4 shows that the speed bias at a given speed range is a product of the likelihood width and the local slope of the prior. Thus for a given speed bias, prior slope and likelihood width are inversely related.

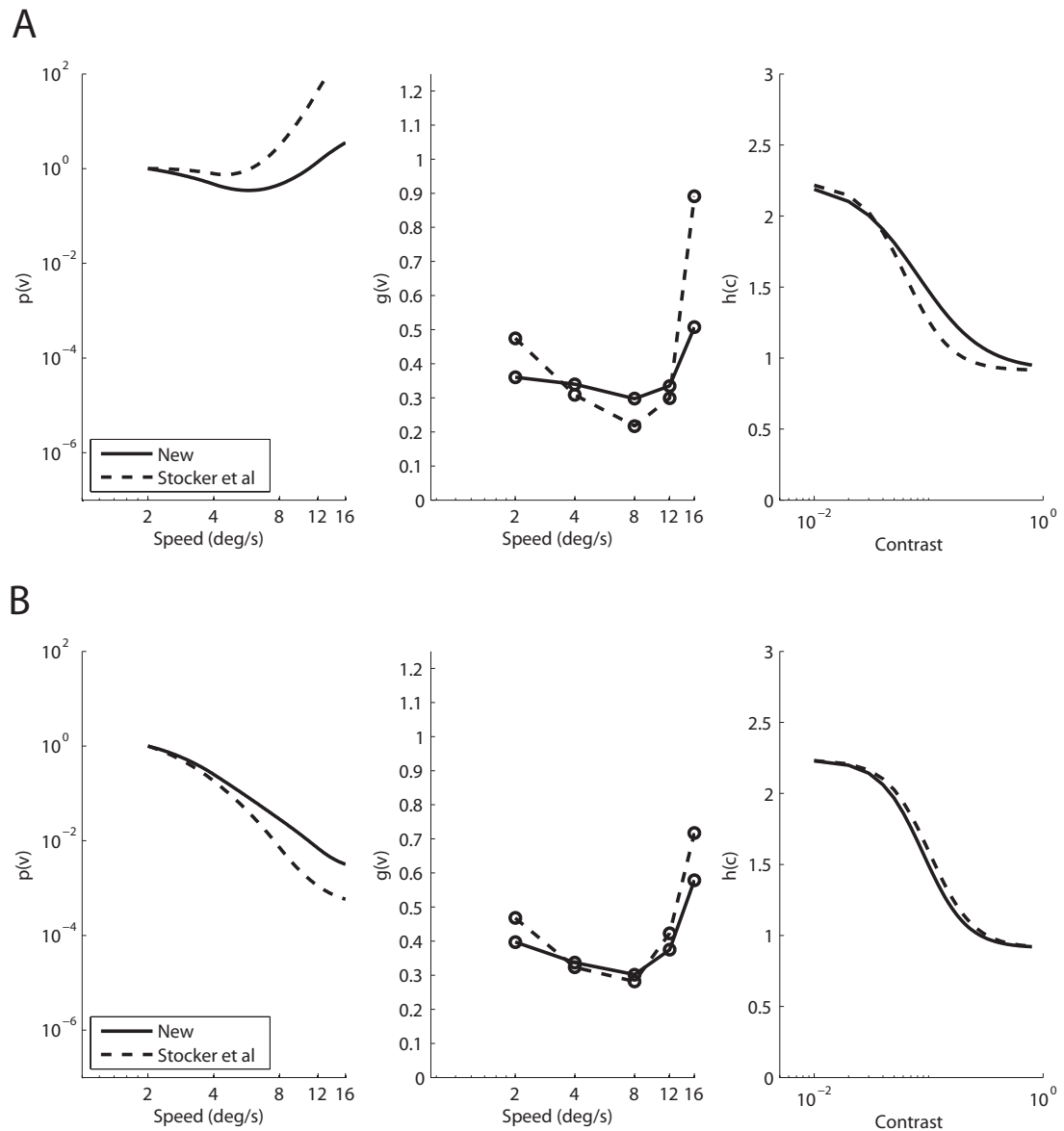


Figure 24: Bayesian model components for the two subjects with the least (A) and most (B) decreasing priors. Dashed lines correspond to the fitting method of [Stocker and Simoncelli \(2006\)](#); solid lines correspond to my fitting method. Extracted priors ($p(v)$) are shown in the left panel; speed-dependent ($g(v)$) and contrast-dependent ($h(c)$) components of likelihood width $\sigma(v, c)$ are shown in the middle and left panel, respectively. Subject A is a male; subject B is a female; both are nongamers. Priors are unnormalized.

5.4 DISCUSSION

In Experiment 1, results are in line with previous investigations ([Lorceau et al., 1993](#); [Sotiropoulos et al., 2011](#)): at low contrast, subjects perceive oblique motion less often. However, the bias towards perpendicular motion direction is not significantly

different between gamers and nongamers. Surprisingly, the only significant difference is in the biases of males and females for the longest duration (533 ms). Assuming that there is indeed a gender difference in contrast-induced biases, it is somewhat surprising that such a difference is not evident at shorter durations, where stimulus uncertainty is higher and biases become more apparent. It is also noteworthy that there is a general (nonsignificant) tendency for p_o values to be highest at the shortest duration, and decrease as duration increases. Both of these counterintuitive observations may be explained by the possibility that at short durations, subjects' responses are partly random. That is, beyond a certain point of perceptual uncertainty, subjects tend to switch to a random response strategy instead of reporting their perception.

As for Experiment 2, the novel interactive effect of reference speed and contrast on CDB seen in Chapter 4 is also apparent in the present data but is not as pronounced in all groups. In particular, there is no such interaction in the female group. One reason for this difference is that in the present experiment, while the number of subjects was greater (6-11 subjects per group versus 6 subjects in the Chapter 4 experiment), much less data was collected from each subject; whereas in the Chapter 4 experiment subjects performed 10 staircases for each condition at low reference speeds and 40 staircases at high speeds, in the present experiment only 6 staircases were performed per condition and some of them had to be excluded as they were found invalid, further decreasing the amount of available data. Especially at high speeds, where speed judgements are less reliable (Stocker and Simoncelli, 2006; also reflected in the Bayesian model of Chapter 4 as high $g(v)$ values above $8^\circ/s$, – see Figure 17), a larger number of staircases than 6 is needed to measure CDB with sufficient accuracy. Thus it is likely that the present data is insufficient to fully reveal the subtleties observed in Chapter 4, which is also suggested by the larger error bars in CDB values in Figure 22 (cf. Figure 15).

A comparison of the results of the two experiments shows that the largest differences are not between gamers and nongamers but between males and females. Although the effect of gender is not particularly strong, it is significant in both experiments and, moreover, it is consistent across tasks and supports the hypothesis of a common speed prior used in both tasks. The reasoning for the latter is as follows: a larger CDB in females, as found in Experiment 2, implies that either their prior is different, favouring slow speeds more than the prior of the males does, and/or their

likelihood is different. The difference in likelihood could be either due to a difference in signal-to-noise ratio, which would be reflected by higher values of the speed-dependent component $g(v)$ of the likelihood width $\sigma(v, c)$, or due to a difference in the contrast response function, reflected in the difference in the contrast-dependent component $h(c)$ of the likelihood width. As seen in 5.3.2, the main difference is in the prior, with the prior of the females favouring slow speed more than the prior of the males. Assuming that a single speed prior is used in all tasks, the differences in the prior found in Experiment 2 are also responsible for the perceptual differences found in Experiment 1, where females perceived oblique motion less often than males (since oblique perception corresponds to a higher velocity than perpendicular perception).

Could the differences in priors between males and females imply a true gender difference in motion perception? Gender differences in vision have been consistently found only in a handful of tasks, and recent literature suggests that the magnitude of differences has been generally overestimated in earlier research (Voyer et al., 1995). The most well-studied cases are 3D mental rotation tasks, where observers are shown two pictures of objects from different viewpoints and are asked to judge whether the two pictures are of different objects or of the same object viewed from different viewpoints. In these tasks, males have been found to have an advantage, such as faster reaction times (Voyer et al., 1995; Kimura, 2000). Mental rotation is a primarily cognitive task, however; in low-level vision, the majority of research has focused on colour perception. The results of these colour studies have been inconsistent and when effects are found, they are small. In general, females have been found to have slightly wider Rayleigh matching ranges (a measure of colour discrimination ability where a wider range indicates lower performance) on the red-green colour axis (Birch et al., 1991). However, as Murray et al. (2012) explain, these differences were mainly due to the presence of female carriers of X chromosome-linked colour deficiencies and the most recent study on the topic found that after these female carriers were excluded from analysis, the sex differences disappeared (Rodriguez-Carmona et al., 2008). A remarkable case of gender differences in colour perception is the existence of a fourth cone type in a minority of women. These women have tetrachromatic vision and a richer colour experience than normal trichromats (Jameson et al., 2001).

Colour differences, however, are unlikely to be related to the differences found in the present experiments, which manipulate luminance differences (contrast) of grey-

scale stimuli. To my knowledge, the only study that is potentially relevant to the present work is by [Abramov et al. \(2012\)](#) who tested subjects with drifting gratings similar to those in Experiment 2 and found gender differences in the spatio-temporal contrast sensitivity function and in spatial acuity. In general, males had higher contrast sensitivity (defined as the reciprocal of contrast difference threshold) across all spatial and temporal frequencies in the measured ranges (0.5-12 cycles/° and 1-24 Hz, respectively). At the spatial frequency used in Experiment 2 (2 cycles/°), [Abramov et al. \(2012\)](#) found that males had a 30% greater contrast sensitivity at stimulus speeds of $0.5^\circ/s$ – a difference that gradually decreased with speed, disappearing at $12^\circ/s$. Males also had approximately 10% greater spatial acuity than females across all tested temporal frequencies. The differences in spatial acuity found by [Abramov et al. \(2012\)](#) are unlikely to be responsible for the differences found in the present work (since the spatial frequency of the gratings was well above acuity thresholds). Differences in contrast sensitivity, however, cannot be ruled out as a factor. Males and females in the present work differed in their priors but not in the likelihood components (where contrast sensitivity differences would be expected to have an effect); however, it is conceivable that differences in the prior were shaped over years from subtle differences in contrast sensitivity: after all, in any Bayesian learning scenario, the current posterior (to which the current likelihood contributes) influences the future prior. Therefore, while the difference in contrast sensitivity may cause a difference in the likelihood that is too small to measure in the present experiment, the cumulative differential effect on the prior throughout life may be measurable. An interesting direction for future experimental work would be to look for correlations between contrast sensitivity and CDB in both sexes.

Lastly, although the differences in biases between males and females were congruent in the two experiments (females demonstrated a more biased perception at low contrasts than males), the biases were not correlated in the two experiments on a subject-by-subject basis. A likely reason for this is the low sensitivity of Experiment 1 which is inherent in its design: a very large number of trials would be needed to obtain p_0 values precise enough to uncover any possible correlations.

In conclusion, these two experiments provided evidence towards a common origin of biases in velocity estimation across different paradigms by showing that differences in biases between groups in one experiment are congruent with differences in the

other experiment. This evidence is not conclusive however; while the differences between males and females are statistically significant in both tasks, there are no subject-by-subject correlations in the observed biases between the two tasks. As a final note, the reconstructed priors in the present work exhibit significant variability across subjects, with a small number of subjects even having nonmonotonic priors.

SPECIFICITY AND ROBUSTNESS OF THE LEARNED SPEED PRIOR

6.1 INTRODUCTION

The experiment described in Chapter 3 presented evidence in support of a plastic speed prior. Observers were able to quickly (in about 20 minutes of exposure) learn a new prior favouring faster speeds than the baseline (start of session) prior did. Furthermore, this new prior partially persisted until the next experimental session, indicating a slower component of learning and a relative robustness of this prior to influences due to exposure to the real world (that presumably has different speed statistics) between sessions. The three experiments described in this chapter are a follow-up to this work with the purpose of investigating whether the changes observed across sessions transfer to other motion directions and stimulus orientations and, in the last experiment, whether the changed prior is long-lasting, as in perceptual learning research and the studies of Knill (2007) and Adams et al. (2010).

6.2 GENERAL METHODS

The experiments described here have the same structure as the experiment in Chapter 3, with the addition of two short test blocks at the beginning of the first (day 1) and last (day 5) sessions. These extra two blocks are similar to the main test blocks (pre- and post-training blocks of the Chapter 3 experiment) in that stimuli move at low speed ($4^\circ/s$) and have the same number of trials but they differ from the main test blocks in that the longest duration of presentation (532 ms) is omitted, so the extra blocks are 33% shorter than the standard test blocks. The rationale for omitting the longest duration was that biases are smaller in long durations and thus adding such trials would not be informative and would make sessions unnecessarily long and possibly undermine any direct comparisons to the original experiment (see 6.6 for a

discussion). The total number of trials in the extra blocks was 144, compared to 216 in the main test blocks.

The extra blocks were added to see how much learning would transfer to a different line orientation (but with the same motion directions) as well as to different motion directions (and stimulus orientation). In one of the extra blocks, where direction was unchanged from the standard blocks, lines were “flipped” vertically so that downward motion now corresponds to perpendicular motion, whereas upward motion corresponds to oblique motion, i.e. the opposite of what holds for the standard blocks. This manipulation also serves as a check for possible response biases. For example, if in cases of uncertainty subjects respond with ‘Up’ all the time, this will show up as an abnormally high p_o (proportion of oblique responses) in the flipped orientation condition whereas in the standard blocks a high proportion of ‘Up’ response would correspond to low p_o . In the other extra block, the whole stimulus, including its direction, was rotated by 90° , so that both orientation and direction were at right angles to their main counterparts (see Figure 25 in Experiment 1). Thus in the second extra block the equivalent of the oblique direction (which in the main block is -20° clockwise from horizontal) is -20° clockwise from vertical. All other stimulus parameters (contrast values, line and aperture dimensions) were as in the Chapter 3 experiment, as were the rest of the experimental parameters, such as laboratory environment and computer hardware and software.

As in the previous experiments, experiment participants were all undergraduate or postgraduate students from the University of Edinburgh.

6.3 EXPERIMENT 1

6.3.1 *Methods*

Figure 25 shows the structure of the first and last (fifth) sessions. Sessions 2, 3 and 4 were identical to the experiment of Chapter 3 and consisted of the last 3 blocks (MAIN condition; red shapes in Figure 25). Sessions 1 and 5 included an additional two blocks at the beginning of the session – a test block (ORTH) with a stimulus whose orientation and two motion directions were at right angles to the respective main orientation and directions (those used in the experiment of Chapter 3), followed by a

test block (FLIP ORI) with a stimulus whose orientation differed by 40° from the main orientation and whose motion directions were the same as the main directions. Thus the oblique direction in the latter block is the same as the perpendicular direction in the main blocks and vice versa. 7 subjects (5 female and 2 male) participated in the experiment, 2 of whom (both female) had to be excluded *post hoc* (see 6.3.2).

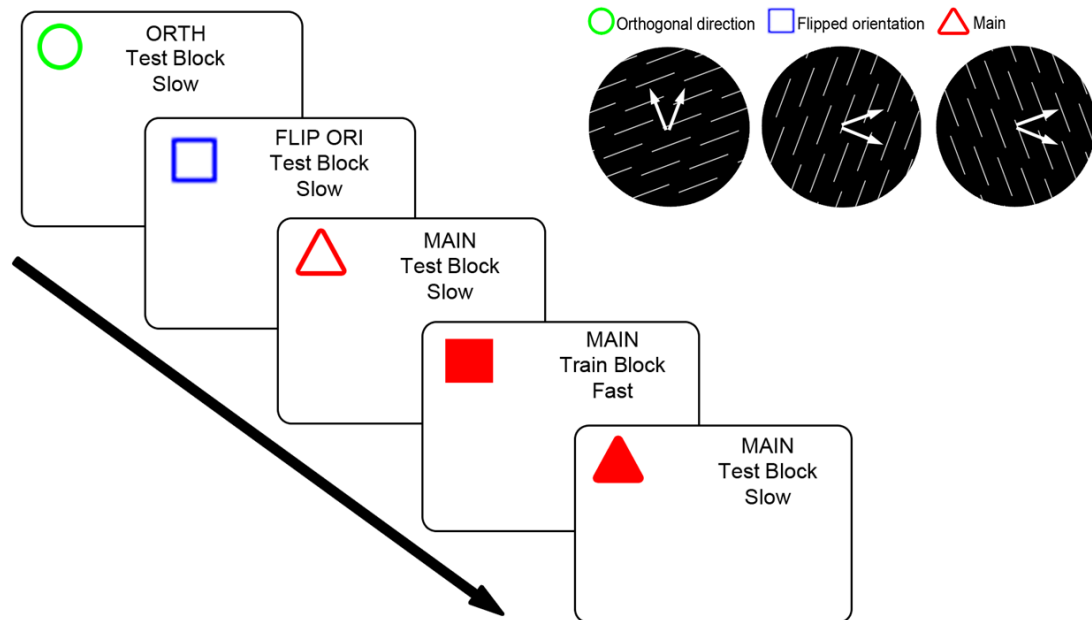


Figure 25: Block sequence in the first and last session in Experiment 1. Sessions 1 and 5 were identical to the respective sessions in the experiment of Chapter 3, with the addition of two extra blocks (ORTH followed by FLIP ORI) at the beginning of the session; sessions 2,3 and 4 were as in the Chapter 3 experiment, i.e. consisted only of the last 3 blocks.

6.3.2 Results and discussion

Figure 26 summarizes the results of Experiment 1. The red triangles (data points from the MAIN condition) correspond to the data points in the Chapter 3 experiment; the green circles and blue squares correspond to the extra blocks (ORTH and FLIP ORI conditions, respectively). At high contrast (top row of Figure 26) perception is veridical, as implied by the fact that p_o values are close to the 0.5 line.

In the low-contrast data of the MAIN condition, a 3-way ANOVA with factors session number, test block number and duration showed an effect of session ($F_{4,139} = 34.31$, $p < 0.0001$) and test block ($F_{1,139} = 13.04$, $p = 0.0004$) but not an effect of duration ($F_{2,139} = 2.05$, $p = 0.155$) or any interactions. Thus data from all durations

was pooled for subsequent analysis. Data from the main test blocks (pre- and post-training; empty and filled red triangles respectively) follows a similar pattern to the Chapter 3 experiment but the change in p_o across sessions is less pronounced. This is also reflected in the statistical analysis: the upward slope of the linear fit to the data is significant in the post-training but not in the pre-training block ($p = 0.006$ and 0.117 respectively, linear hypothesis test). The difference in p_o between first and last session is significant in the post-training block ($t_{12} = -2.489$, $p = 0.028$, t-test) but not in the pre-training block ($t_{12} = -1.09$, $p = 0.113$).

In the extra blocks, neither ORTH (green circles) nor FLIP ORI (blue squares) responses show any significant change across sessions. Note that there are no post-training extra blocks in the experiment (for reasons discussed in the first paragraph of 6.2) and therefore short-term (within-session) effects cannot be observed.

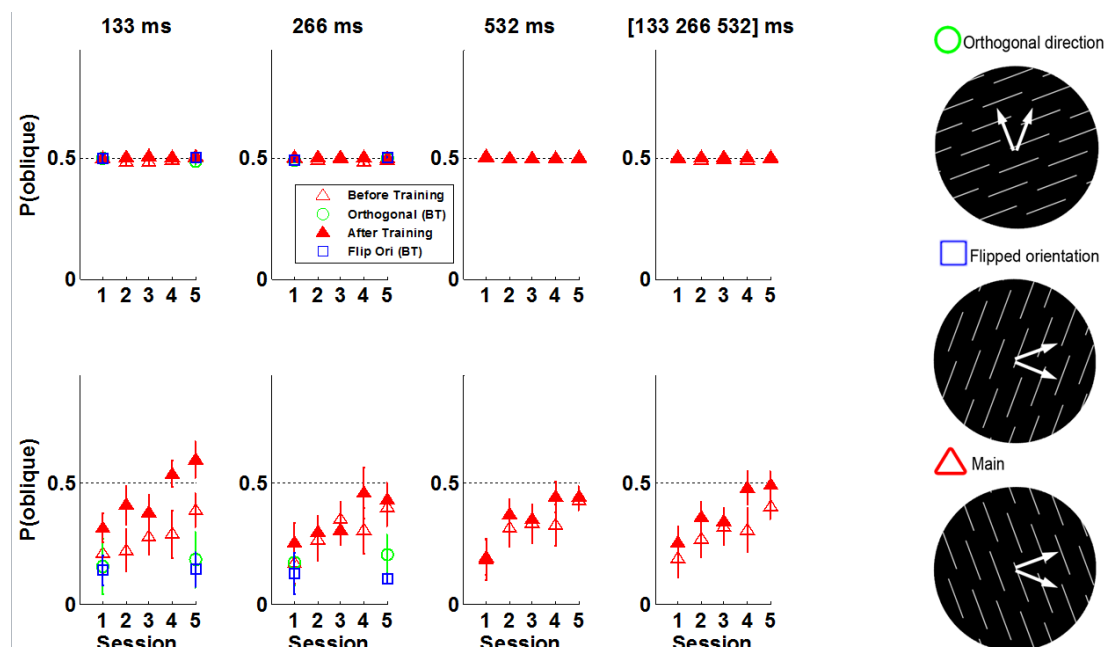


Figure 26: Experiment 1: proportion of oblique responses (p_o) as a function of session number. Data from 5 subjects. Top row, high contrast; bottom row, low contrast. Error bars are SEM (standard error of the mean).

6.3.2.1 Excluded subjects

The above results are after the data from 2 out of 7 subjects were discarded (Figures 27 and 28 show data from these two subjects). The *post hoc* criteria for data exclusion, based on pilot data and on the main results from the Chapter 3 experiment, were

1. Poor performance ($< 85\%$) in at least one block of high-contrast trials, which suggests unreliable perception or inability to perform the task correctly. A short familiarization task and a small performance test prior to the main experiment was performed by all participants and those that failed this *pre hoc* test were excluded from the main experiment. However, trials in the *pre hoc* test had a longer duration, on average, than even the longest (533 ms) duration used in the main experiment. This may explain why these two subjects passed the *pre hoc* test but were excluded *post hoc*.
2. Too large proportion of oblique responses ($p_o > 0.7$) in the low-contrast trials (at any duration). Such a high p_o is highly unlikely to correspond to true perception of oblique motion (since the bias in untrained subject is in the opposite direction) and is more likely to be an artifact such as a response bias when the trial is hard.
3. Too small proportion of oblique responses ($p_o < 0.1$) in the low-contrast trials at the longest duration (533 ms). At this duration, oblique motion is correctly perceived by naive observers with a significant probability. Near-absent oblique perception suggests either a response bias (e.g. due to trial difficulty) or unreliable perception at low contrasts.

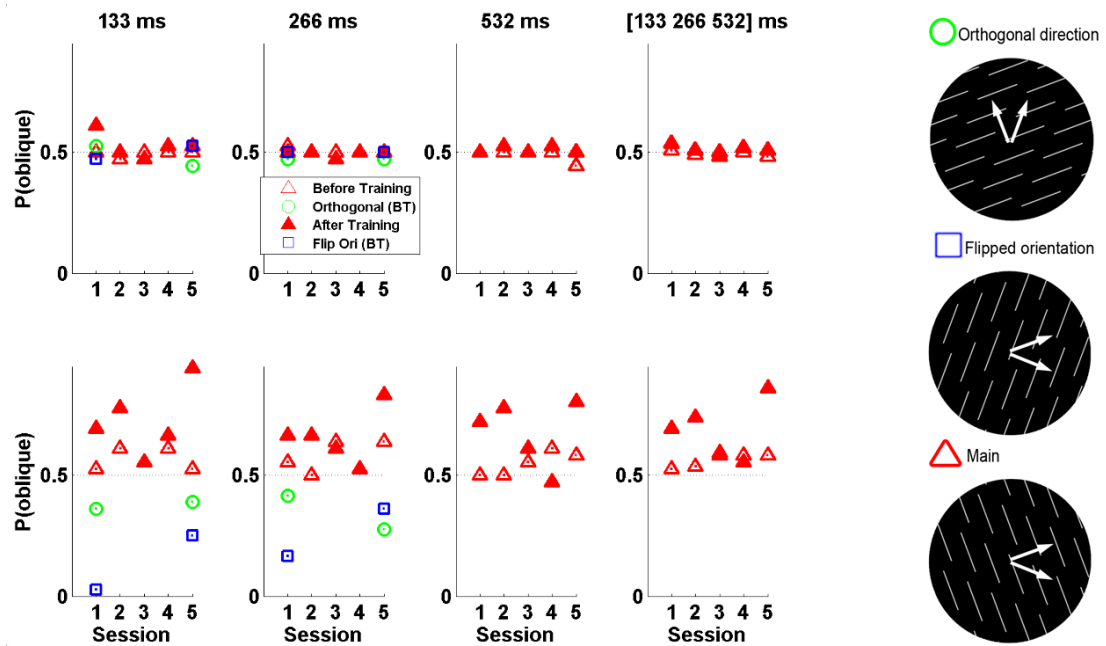


Figure 27: Proportion of oblique responses (p_o) as a function of session number. Data from excluded subject 1 in Experiment 1. Top row, high-contrast trials; bottom row, low-contrast trials.

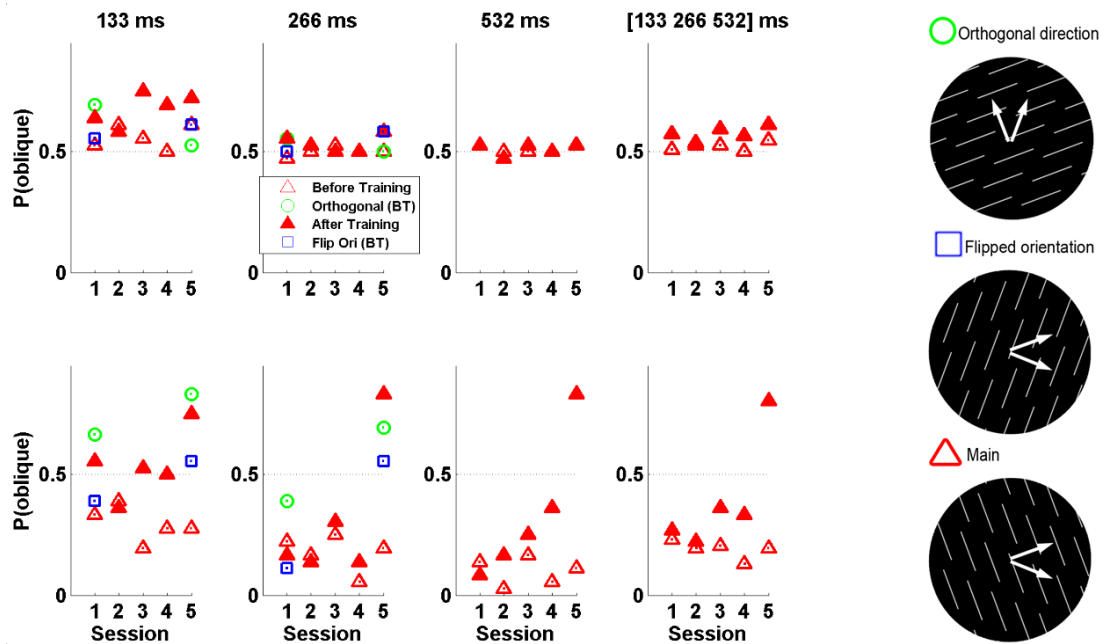


Figure 28: Proportion of oblique responses (p_o) as a function of session number. Data from excluded subject 2 in Experiment 1. Top row, high-contrast trials; bottom row, low-contrast trials.

In conclusion, subjects' responses in the main test blocks (red triangles in Figure 26) are very similar to those of the experiment of Chapter 3: p_o follows an upward trend both within a session (from pre- to post-training) and across sessions. p_o in the last

post-training block exceeds 0.5, i.e. the “perpendicular motion” illusion is reversed and oblique motion is perceived more often, as seen in the Chapter 3 experiment. However, the results in the present experiment are statistically weaker. One likely factor for this is the smaller number of subjects (5) after data from two subjects were excluded. In contrast to the changes seen in the main blocks, there are no changes in the extra blocks and thus no transfer of learning to either FLIP ORI or ORTH (although there is an effect on FLIP ORI in the data from the two excluded subjects).

6.4 EXPERIMENT 2

The motivation for Experiment 2 was twofold. First, in the previous experiment the difference between the first and last session in the extra blocks (FLIP ORI and ORTH) was not significant. I thus decided to repeat the experiment and gather additional data that might shed more light on the situation. Second, I wanted to investigate whether there is a systematic difference in perception between the vertical and horizontal directions. Anisotropies between horizontal and vertical motion direction have been reported previously: [Loffler and Orbach \(2001\)](#) found a strong attractive bias towards horizontal directions (directions with a small deviation from horizontal were perceived as horizontal). Such a bias is expected to blur the difference between perceived directions very close to either side of the horizontal and thus influence results in a different way than it would if motion directions were around the vertical instead. Thus repeating Experiment 1 with the entire moving stimulus (line orientation, direction of motion) rotated by 90° could conceivably yield different results.

6.4.1 *Methods*

Figure 29 shows the structure of the first and last sessions. The line orientation and motion directions in the ORTH block are now identical to the orientation and directions of the main blocks in Experiment 1. Line orientation and directions in the FLIP ORI block and main blocks are 90° counter-clockwise to their respective counterparts in Experiment 1. As before, sessions 2, 3 and 4 consisted of the last 3 blocks in Figure 29 (main blocks). 9 subjects (5 female, 4 male) participated in the experiment; data

from 3 subjects (2 female, 1 male) had to be discarded according to the same criteria as those in Experiment 1.

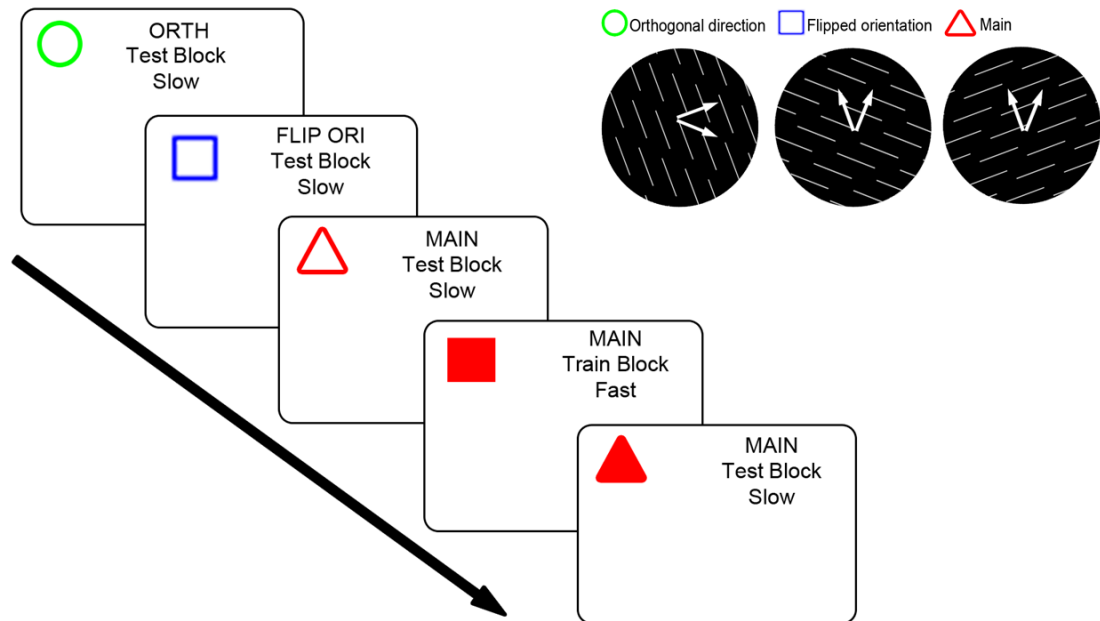


Figure 29: Block sequence in the first and last session in Experiment 2. Sessions 2 through 4 were as in the Chapter 3 experiment, i.e. consisted only of the last 3 blocks.

6.4.2 Results and discussion

Figure 30 summarizes the results of Experiment 2. In high-contrast trials (top panel), results are as expected and very similar to those of Experiment 1. In low-contrast trials of the MAIN condition, a 3-way ANOVA with factors session number, test block number and duration showed an effect of session ($F_{4,209} = 7.15, p < 0.0001$) and test block ($F_{1,209} = 15.28, p = 0.0001$) but not an effect of duration ($F_{2,139} = 1.12, p = 0.328$) or any interactions. Thus data from all durations was pooled for subsequent analysis. Data from the main test blocks (pre- and post-training) shows both short-term (within session) and long-term (across sessions) learning. Results follow a similar pattern to Experiment 1, where the change in p_o across sessions is less pronounced in the pre-training block. This is also reflected in the statistical analysis: the upward slope of the linear fit to the data is significant in the post-training but not in the pre-training block ($p = 0.005$ and 0.095 respectively, linear hypothesis test). The change in p_o between first and last session is significant in the post-training ($t_{12} = -3.35, p = 0.006$) but not in the pre-training ($t_{12} = -1.53, p = 0.152$) block.

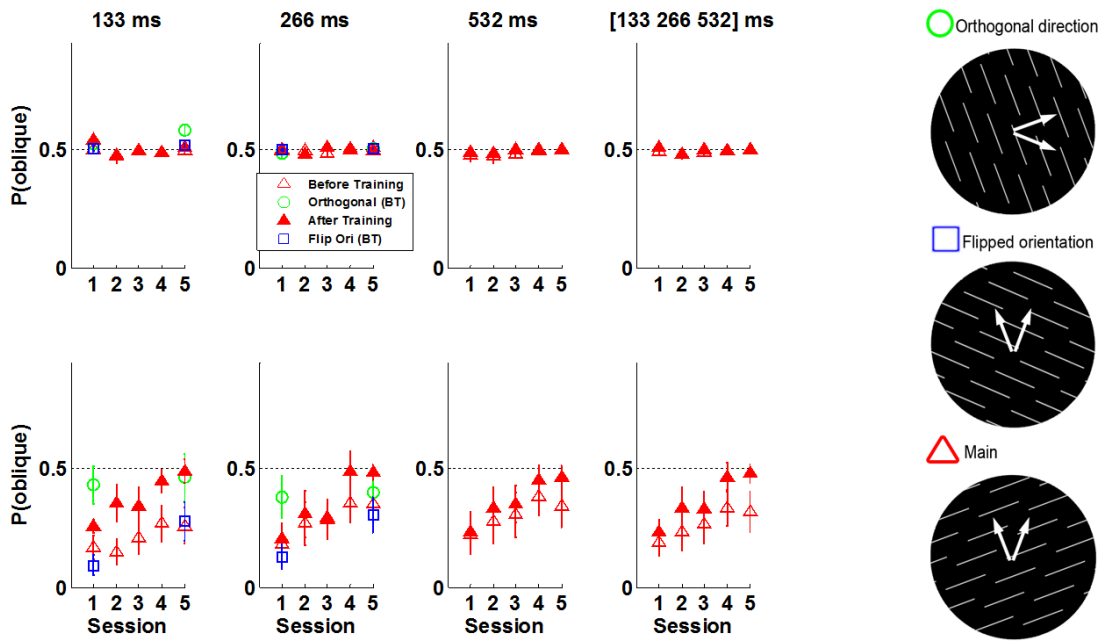


Figure 30: Experiment 2: proportion of oblique responses (p_o) as a function of session number. Data from 6 subjects. Top row, high contrast; bottom row, low contrast. Error bars are SEM.

In the extra blocks, the results are different from Experiment 1. In contrast to the absence of change in p_o in FLIP ORI of Experiment 1, there is now a visible increase in p_o from session 1 to session 5 in FLIP ORI. A 2-way ANOVA with factors session number and duration shows an effect of session ($F_{1,27} = 11.78$, $p = 0.002$) but no effect of duration ($F_{1,27} = 0.36$, $p = 0.553$), therefore data from both durations was pooled. A t-test showed that the difference in p_o between session 1 and 5 is significant ($t_{12} = -2.68$, $p = 0.019$). The p_o values in FLIP ORI (blue squares in Figure 30) are comparable to the respective (first and last session) values in the main pre-training block (empty red triangles in Figure 30). Thus, in contrast to Experiment 1, there is evidence that the learned prior was applied to the different stimulus orientation.

Data in the ORTH block is quite different from main and FLIP ORI blocks, as p_o is unusually high in both sessions (1 and 5). Looking at individual subject data, these high p_o values are due to two subjects, shown in Figures 31 and 32. It is not clear why these subjects reported mostly oblique perception in the ORTH block; block ordering effects (e.g. due to motion adaptation) are unlikely given that the ORTH block was presented first. As before, the results in the ORTH block may reflect a response bias employed in trials that the subjects found too hard, although this does not explain why they only used this strategy in that block. Figure 33 shows the data

remaining after excluding the ORTH block data from these two subjects. p_o values in the ORTH block are now comparable to the respective values of the main and FLIP ORI blocks, although there is a (nonsignificant) tendency towards higher values. A 2-way ANOVA with factors session number and duration showed no effect of session number ($F_{1,19} = 0.15$, $p = 0.705$), duration ($F_{1,19} = 0.28$, $p = 0.606$) or an interaction ($F_{1,19} = 0.03$, $p = 0.863$).

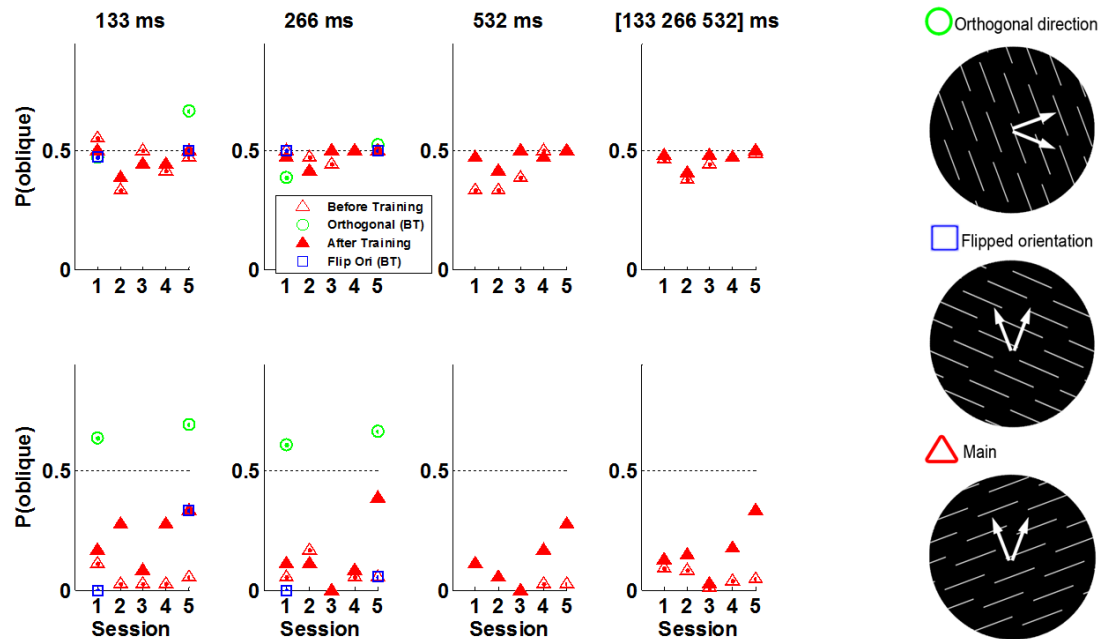


Figure 31: Experiment 2: proportion of oblique responses (p_o) as a function of session number. Data from the first subject with an unusually high p_o in the ORTH block. Top row, high contrast; bottom row, low contrast.

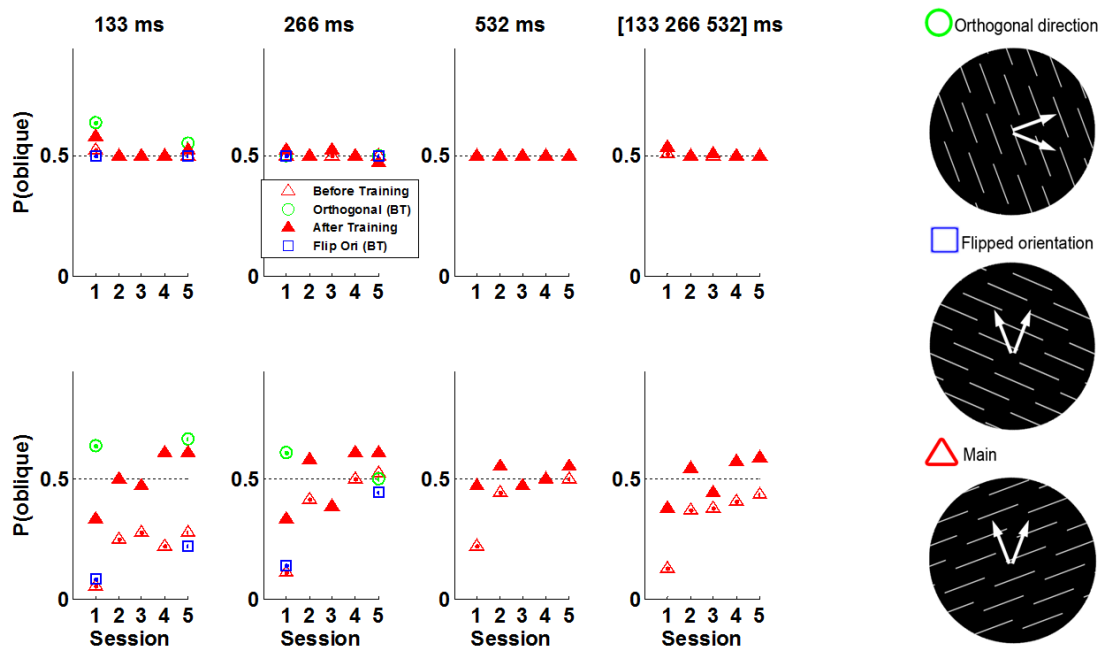


Figure 32: Experiment 2: proportion of oblique responses (p_o) as a function of session number. Data from the second subject with an unusually high p_o in the ORTH block. Top row, high contrast; bottom row, low contrast.

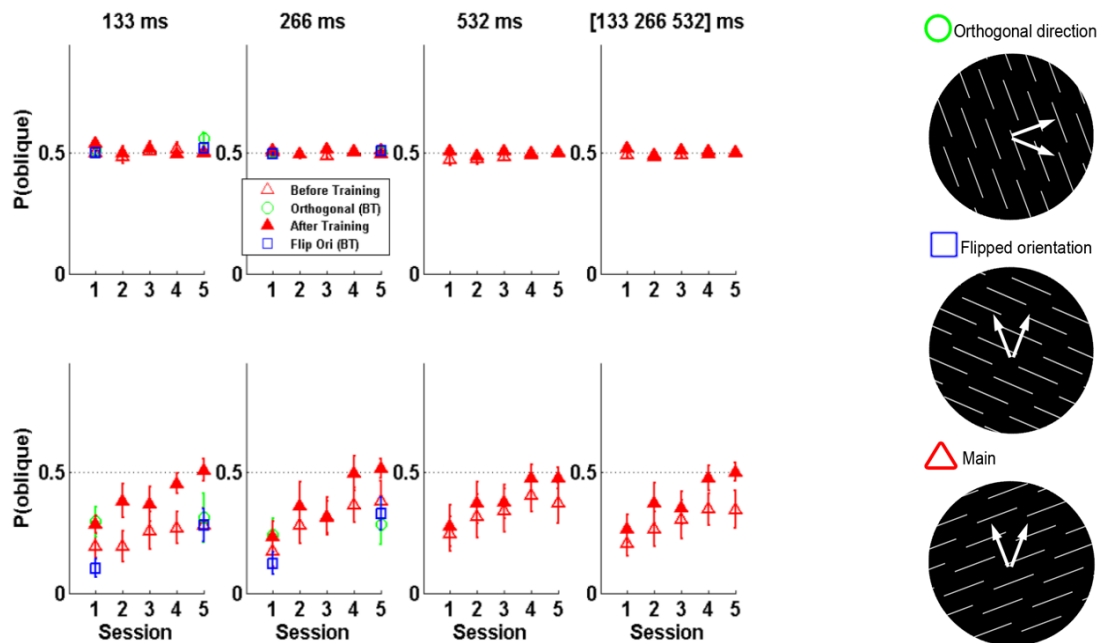


Figure 33: Experiment 2: proportion of oblique responses (p_o) as a function of session number. Data from 6 subjects, after excluding the ORTH block data from 2 subjects with unusually high p_o in that block (leaving 4 subjects in that block). Top row, high contrast; bottom row, low contrast. Error bars are SEM.

6.5 EXPERIMENT 3

In light of the somewhat inconsistent results of Experiments 1 and 2, I decided to collect more data with the same stimulus setup as Experiment 2 but with the order of ORTH and FLIP ORI blocks swapped (so that FLIP ORI is performed first, then ORTH and then the main blocks). Would the counterintuitive results of the ORTH block in two subjects (Figures 31 and 32) still be observed? Furthermore, if there is indeed an ordering effect, would we see a difference in the FLIP ORI block between the two experiments? Conceivable reasons for an ordering effect are motion adaptation or greater familiarization with the task between the first and second extra blocks.

Additionally, the specificity of learning suggested by these two experiments is reminiscent of perceptual learning. Apart from specificity, another hallmark of perceptual learning is its persistence in time. Is the learning observed in Experiments 1 and 2 similarly persistent? If subjects were tested again a week after the fifth session, how would their responses compare to those of the first week?

6.5.1 *Methods*

The experiment consisted of 6 sessions: 5 consecutive ones (Monday-Friday) as in Experiments 1 and 2; and a 6th session a week later (next Friday), identical to the first session. In sessions 1, 5 and 6 (Figure 34) the extra blocks (FLIP ORI and ORTH) were identical to the respective blocks in Experiment 2 except that their order was swapped: FLIP ORI was presented first, followed by ORTH, followed by the main blocks. Sessions 2, 3 and 4 were identical to the respective sessions of Experiment 2. 7 subjects (3 female, 4 male) participated in the experiment; data from 2 subjects (1 male, 1 female) had to be discarded according to the criteria described in 6.3.2.

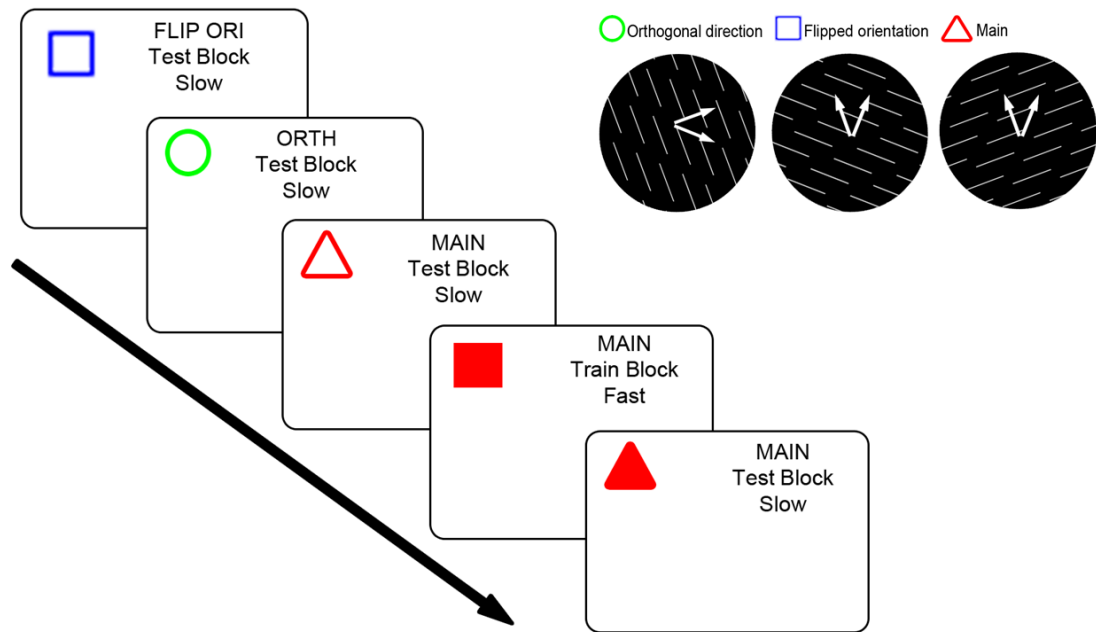


Figure 34: Block sequence in the first, fifth and last (week-after) session in Experiment 3. Sessions 1 and 5 were identical to the respective sessions in Experiment 2 except that the order of the two extra blocks was swapped (thus FLIP ORI was followed by ORTH). Sessions 2,3 and 4 were identical to the respective sessions of Experiment 2.

6.5.2 Results and discussion

Figure 35 summarizes the results of Experiment 3. Results for the main test blocks (red triangles) are similar to Experiment 2, with the change in p_o being less pronounced in the short term (within session, between pre- and post-training) in the longest two durations (266 and 532 ms). In the long-term (across sessions) there is an upward tendency in p_o across sessions for both test blocks. A 3-way ANOVA with factors session number, block number and duration showed an effect of session ($F_{4,149} = 9.72$, $p < 0.0001$) and block ($F_{1,149} = 5.61$, $p = 0.019$), as in Experiments 1 and 2; however, this time there was also an effect of duration ($F_{2,149} = 6.45$, $p = 0.002$) as well as a borderline nonsignificant interaction of block number and duration ($F_{2,149} = 2.08$, $p = 0.129$). Further analysis (repeating the ANOVA on subsets of the {133,266,532} ms data set) showed that it was the shortest (133 ms) duration that differed from the rest and that test block number had an effect only on that condition. Thus linear hypothesis tests were performed separately for the 133 ms and for the pooled 266+532 ms data. In the 133 ms condition, there is a significant upward

tendency in the pre-training block ($p = 0.004$, linear hypothesis test) and a borderline significant tendency in the post-training block ($p = 0.049$). In the 266+532 ms conditions, the upward tendency is highly significant in both the pre-training ($p = 0.0001$) and post-training ($p = 0.0004$) blocks. Therefore, long-term learning occurred in both the 133 ms and the pooled 266+532 ms conditions.

Notably, there is also a significant difference in p_o between session 1 and 6 (the week-later session) in the pre-training blocks of the 133 ms ($t_8 = -2.42$, $p = 0.042$) and the 266+532 ms ($t_8 = -3.61$, $p = 0.007$) conditions. As for the post-training blocks, the difference between session 1 and 6 is significant at 266+532 ms ($t_8 = -2.64$, $p = 0.029$) and borderline nonsignificant at 133 ms ($t_8 = -1.78$, $p = 0.113$). Overall, these results indicate that learning persisted for at least a week after the fifth session.

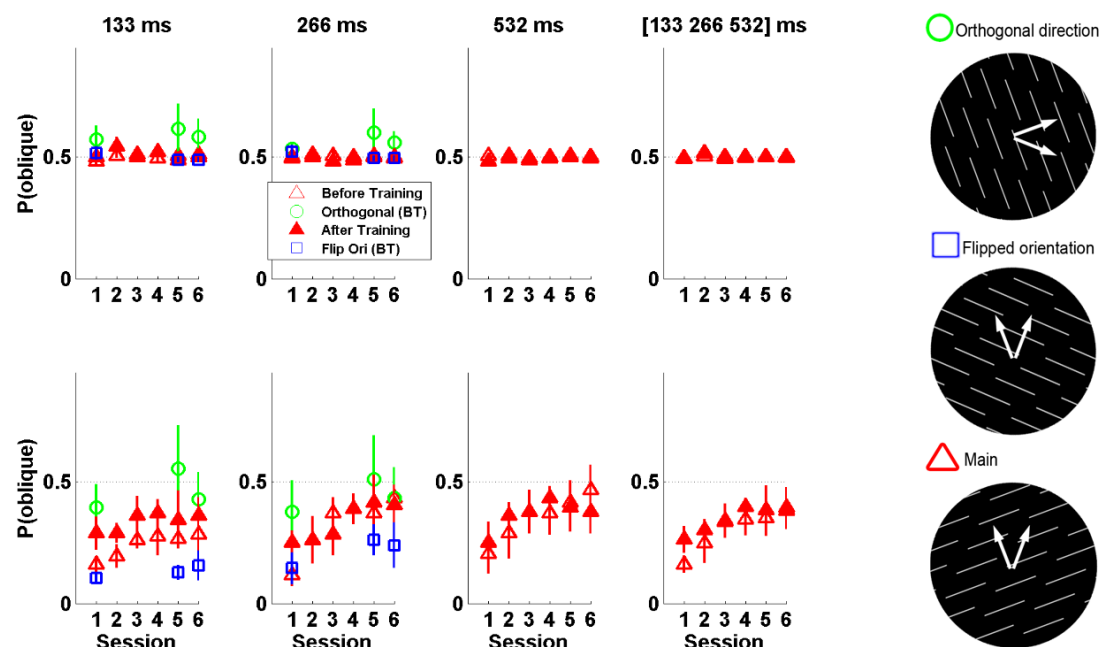


Figure 35: Experiment 3; proportion of oblique responses (p_o) as a function of session number. Data from 5 subjects. Top row, high contrast; bottom row, low contrast. Error bars are SEM.

In the extra blocks, there is a small tendency for increase in p_o in the FLIP ORI blocks but the values remain generally low. However, p_o in the ORTH blocks are unusually high (and with high variance), as was the case in Experiment 2. Looking again at individual data, these results were due to two subjects (both male), shown in Figures 36 and 37.

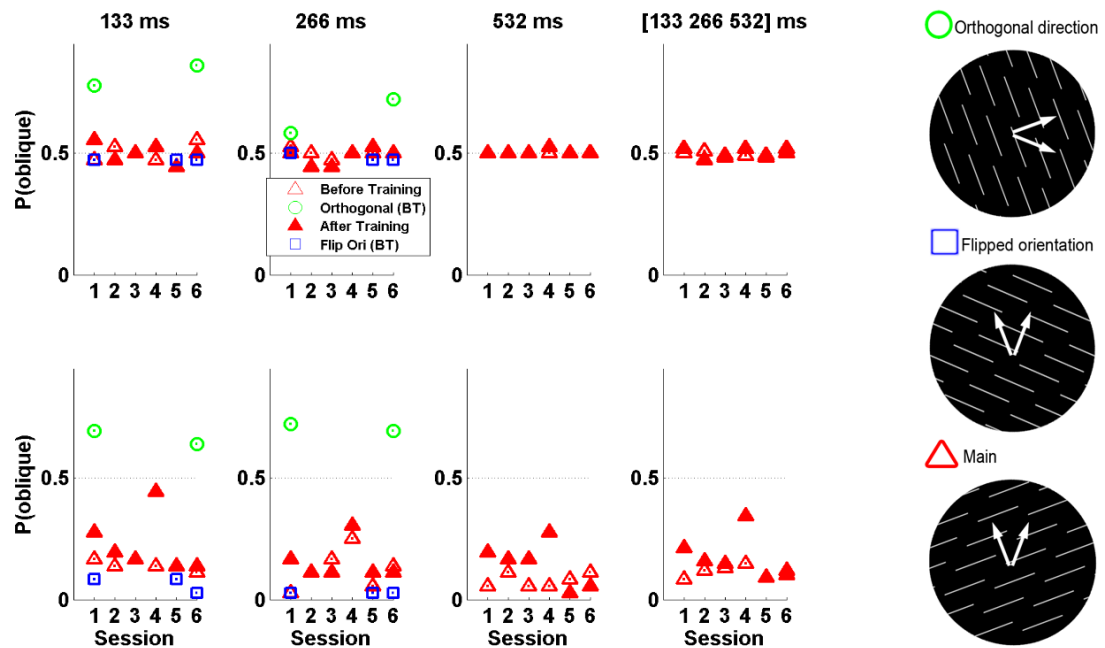


Figure 36: Experiment 3: proportion of oblique responses (p_o) as a function of session number. Data from the first subject with an unusually high p_o in the ORTH block. Top row, high contrast; bottom row, low contrast.

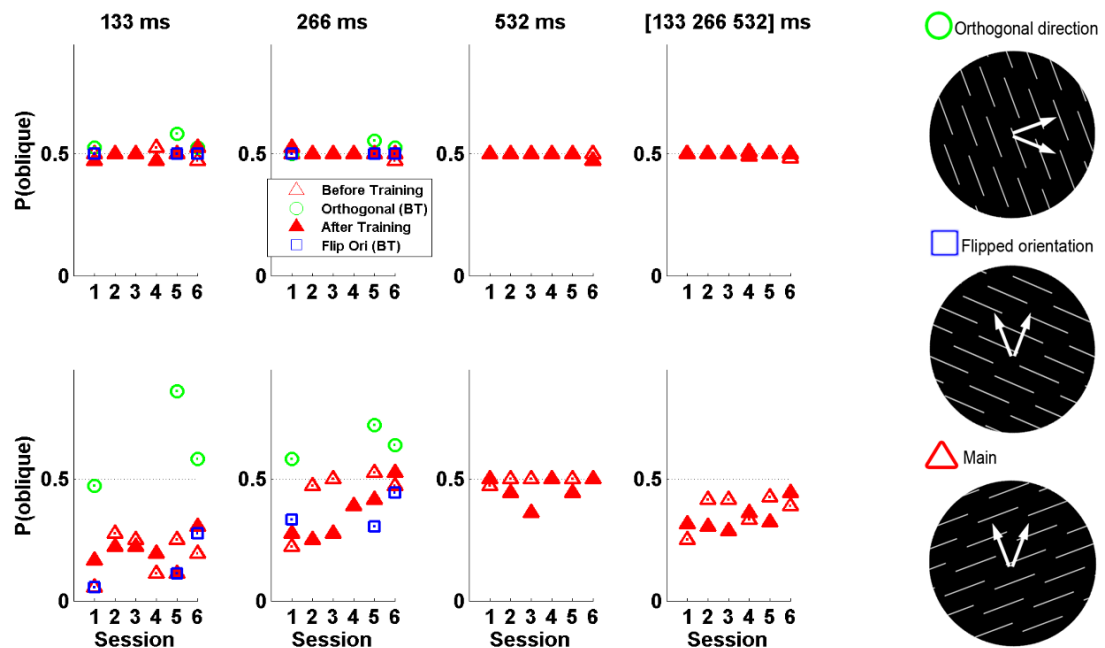


Figure 37: Experiment 3: proportion of oblique responses (p_o) as a function of session number. Data from the second subject with an unusually high p_o in the ORTH block. Top row, high contrast; bottom row, low contrast.

The ORTH block data is again counterintuitive, especially considering that no such effect is seen in the FLIP ORI block and, furthermore, that the ORTH block in this experiment (and in Experiment 2) corresponds exactly to the main test blocks in the

Chapter 3 experiment. Order of presentation is a highly unlikely explanation for the unusually high p_o , considering that in this experiment the ORTH block is presented *after* the FLIP ORI block.

Figure 38 shows the remaining data, after excluding the ORTH block data from the two subjects shown in Figures 36 and 37. For the latter subject (Figure 37), the FLIP ORI data was also excluded as it failed the third *post hoc* criterion described in 6.3.2 (too small proportion of oblique responses). Thus Figure 38 shows data from 3 subjects in the ORTH condition and 4 subjects in the FLIP ORI condition.

2-way ANOVAs on the remaining data (separately for the FLIP ORI and ORTH conditions) with factors session number and duration showed no main effects or interactions (all $p > 0.122$). However, Figure 38 shows a moderate change in the FLIP ORI condition between session 1 and 5/6 in the 266 ms condition, suggesting possible transfer of learning from main to FLIP ORI. Therefore the ANOVA was followed up by t-tests for the difference between session 1 and 5 and session 1 and 6 in the 266 ms condition (using the Bonferroni correction for multiple comparisons). No significant differences were found between any of the comparisons ($t_8 = 0.68$, $p = 0.513$ for the difference between session 1 and 5 and $t_8 = 0.52$, $p = 0.619$ for the difference between session 1 and 6).

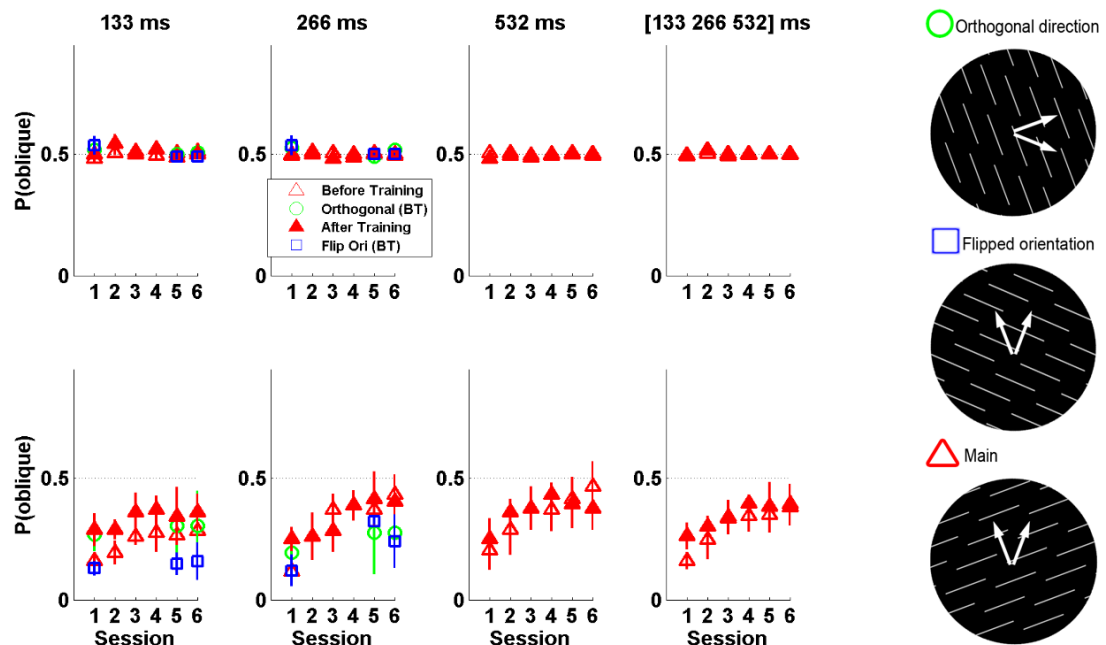


Figure 38: Experiment 3: proportion of oblique responses (p_o) as a function of session number. Data from 5 subjects, after excluding the ORTH block data from 2 subjects with unusually high p_o in that block (leaving 4 subjects in that block). Top row, high contrast; bottom row, low contrast. Error bars are SEM.

6.6 GENERAL DISCUSSION

In general, the three experiments presented in this chapter yielded quantitatively similar results and reproduced the findings of Chapter 3: subjects were able to internalize the speed statistics of the stimulus ensemble, which led to a shift in perceived direction. Furthermore, the present experiments extended previous findings through the inclusion of additional testing conditions.

The results of Experiments 2 and 3 suggest that the order of presentation of the two extra blocks is not important. These results were also similar to Experiment 1; there was a qualitative similarity in the direction of the observed biases between near-horizontal (Experiment 1) and near-vertical (Experiments 2 and 3) directions. This implies that, unlike the horizontal/vertical anisotropies observed by [Loffler and Orbach \(2001\)](#), the perceptual biases observed with this task do not depend on the absolute orientation or directions of the stimulus, at least among the cardinal axes.

In contrast to the main blocks, there is no change between the first and the last session in the ORTH blocks in any of the three experiments. A possible explanation for this is that different priors are used for different stimulus configurations, i.e. the learned prior is highly specific to stimulus details. Evidence for the specificity of learned priors comes from the work of [Kerrigan and Adams \(2013\)](#), where subjects learned a different light prior for different contexts (where context was the illumination colour). Another possible reason for the absence of transfer of the learned prior to the extra blocks in some subjects may be the relatively small size (i.e. number of trials) of the (main and extra) test blocks. This might also explain the lack of significant change in the main test blocks in some cases: the number of trials in these blocks may be too small to allow for the discrimination of real effects from noise. This is an unavoidable trade-off in the experiment: the size of the test blocks must be large enough to provide sufficient data but not so large as to have a significant training effect (by exposing subjects to low speeds too frequently and reverting the high-speed bias induced by the training block).

The absence of transfer to FLIP ORI in some subjects suggests that while most subjects consider the two orientations (main and FLIP ORI) similar enough, some do not. The criteria for similarity of different observers are unclear at this point; it is likely though that a number of different, interacting criteria are used. For example, [Gekas](#)

et al. (2013) found that the learned prior on motion direction generalized across stimulus colours but only when there were no statistical regularities in the non-trained colour. Additionally, Seydell et al. (2010) found transfer of the shape-isotropy prior across colours but not across shapes (subjects were able to learn separate priors for each shape). Future research studying a large number of subjects exposed to an array of different stimulus and task configurations should be able to shed more light in the complexities of specificity in the learning of priors.

One potential complication in interpreting differences between conditions is the fact that the stimulus used in these experiments (and that of Chapter 3) exhibits a large-scale spatial structure that it was not apparent to me at the time of the experiments. In particular, the stimulus has the potential to evoke the perception of illusory contours. These are perceived visual borders which are not associated with luminance or colour changes across the border (Petry and Meyer, 1987). A famous example of an illusory contour is Kanizsa's triangle (Figure 39), named after the Italian psychologist who first described it (Kanizsa, 1955). The illusory contours in the present stimulus can be recognized by drawing imaginary lines that connect the endpoints of adjacent lines along the oblique direction, for the stimuli in the main and ORTH conditions. The black space between these imaginary lines can be seen as such an illusory contour. The orientation of the contour is the same as the oblique motion direction. For example, in the stimulus depiction in the main condition of Figure 38, one such illusory contour is just to the left of the northeast-pointing arrow, parallel to it. In contrast, in the FLIP ORI condition, the orientation of the illusory contour is perpendicular to the oblique direction. Therefore, in the main and ORTH conditions, the illusory contour is stationary in trials with oblique motion and could be used as a frame of reference (whereas in trials with perpendicular motion the contour moves obliquely to its orientation). In the FLIP ORI condition, the contour moves in all trials and is less likely to be perceived. Whether such an effect explains the differences between FLIP ORI and ORTH in the 133 ms condition in Experiments 2 and 3 (Figures 33 and 38) is unknown at this point, although somewhat unlikely, given that the FLIP ORI data, in the cases that it differs from the ORTH data, is similar to the (pre-test) main data. It would be straightforward and potentially interesting to compare the results obtained with the current stimulus and with a stimulus where illusory contours are avoided (e.g. by adding positional jitter to the line elements).

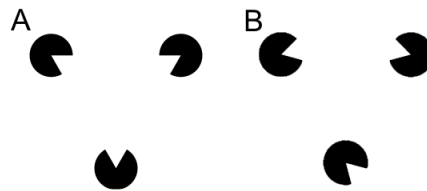


Figure 39: Kanizsa's triangle showing illusory contours. When the relative position and orientation of 3 "pacmen" is as shown in panel **A**, the resulting percept is a white triangle superimposed on 3 black circles. The triangle's sides are "seen" even in the gaps between the pacmen. In panel **B**, no such percept is produced due to the different orientation of the pacmen.

Although the degree of specificity of the learned prior with respect to stimulus configurations is debatable and likely varies among observers, the learned prior appears to be context-specific and may not be used outside the laboratory. This is suggested by the results of Experiment 3, where exposure to the motion statistics of everyday life over a week did not revert the experimentally induced prior. The observed persistence of learning is in accordance with the majority of the perceptual learning literature as well as with the finding of [Adams et al. \(2010\)](#) that the changed light-from-above prior persisted for at least 4 days after the end of the experiment.

The results in this chapter are thus broadly consistent with the hypothesis that subjects use the learned prior in a highly specific manner: not only the context must be the same (as suggested by the absence of transfer to ORTH and the persistence of learning after a week); stimulus parameters (such as orientation) must also be sufficiently similar.

DISCUSSION

In this chapter, I review the experimental and theoretical work that has been presented in this thesis and put it in perspective. I also discuss the limitations of this work and outline possible future directions that may address these limitations.

7.1 LIMITATIONS

7.1.1 *Reconstruction of the changing prior*

The limitations of the experimental paradigm of Chapters 3 and 6 were discussed in 3.4, where it was argued that the experimental paradigm and the associated Bayesian model limit accuracy and precision in reconstructing the speed prior and studying its dynamics – how the prior changes within and across sessions. For example, increased precision in prior reconstruction would improve our understanding of how the prior partially “reverts” between the end of a training block and the beginning of the next (day-after), i.e. as subjects are exposed to the statistics of their everyday life; or, alternatively, how the rapidly (within-session) learned prior is partially consolidated between training blocks (the consolidation hypothesis seems more likely in light of the discussion in 7.2.2). Precisely tracking the changing prior was one of the aims of the experiment described in Chapter 4; however, no learning was observed in this experiment. Possible reasons for this were discussed in 4.4. Briefly, these reasons are based on the differences between the experimental paradigms of Chapters 3 and 4: the greater stimulus variety in terms of speed and contrast; and the explicit (and therefore more prone to cognitive interference) nature of speed judgements in Chapter 4.

Further investigation into the reasons for this failure could provide useful insights and help design experiments that combine the best of both worlds: an effective training procedure to change the prior (Chapter 3) within an experimental and theoretical

framework that enables the accurate reconstruction of the prior (Chapter 4). An experiment that might satisfy these criteria would be one using plaid (superpositions of gratings) stimuli. The slow speed prior has been successfully used to model how perceived direction of moving plaids depends on the speed, contrast and direction of the component gratings (Weiss and Adelson, 1998; Weiss et al., 2002; Hedges et al., 2011). By exposing subjects to fast-moving plaids, it might be possible to induce a shift in the prior similar to that of Chapter 3, which would cause a shift in the perceived direction of slow-moving plaids. This experiment would avoid explicit speed judgements and, at the same time, would allow the application of the model of Stocker and Simoncelli (2006) and thus the precise reconstruction of the changing speed prior. If this experiment proved successful in changing the speed prior, it would also be interesting to test (but not train) these subjects with the field-of-lines stimuli of Chapter 3. This could shed more light in the matter of specificity of the speed priors.

7.1.2 Bayesian models

The Bayesian models of motion perception examined in the present work provide a satisfactory account of the experimental findings quantitatively but also conceptually, as these models view perception as a principled process of unconscious inference and recast certain motion illusions as a result of the workings of a visual system that performs optimally on average. However, as any model, they are not without limitations. The limitations of the Bayesian model of Chapter 3, which is an extension of the Weiss et al. (2002) model, are discussed in 2.2.3.2, where I also describe how the model of Stocker and Simoncelli (2006) was designed to overcome several of these limitations and, importantly, to allow a less constrained reconstruction of the speed prior.

One of the main motivations for the development of the model of Stocker and Simoncelli (2006) was to naturally (without additional assumptions about decision noise) account for trial-to-trial variability. In fact, the model uses that variability to separate the effects of the likelihood and the prior on the posterior (which corresponds to perception). The experimental data collected from the speed matching experiments in Stocker and Simoncelli (2006) and in Chapters 4 and 5 can be summarized in the form of a psychometric function, which is modelled by Equation 15.

For a given stimulus speed, the contrast-dependent bias (CDB) in perceived speed determines the location of the psychometric function whereas the trial-to-trial variability determines the slope. The perceived speed biases and variances are functions of stimulus speed (v), slope of the prior (a) and likelihood width ($\sigma(v, c) = g(v)h(c)$) given by Equations 13 and 14. Relying on the speed bias alone, which is a product of the likelihood width and the prior slope, would make impossible the disambiguation of the contribution of the prior and likelihood. Such disambiguation is made possible by also taking into account the slope of the psychometric function, which can be shown to only depend on the likelihood width (see 8.1.3 in Appendix).

In theory, the aforementioned considerations should allow the reliable reconstruction of the prior. However, the results of Chapters 4 and 5 show that the reconstructed priors can be very different from one experiment to another: whereas the means and variances of CDBs are less than 50% different on average between the two experiments (at the reference speeds that are common to the two experiments, i.e. 4, 8 and $12^\circ/s$), the group prior in Chapter 4 is almost three orders of magnitude steeper than the priors of the four groups in Chapter 5 (see Figures 17 and 23). The priors in both experiments are also less steep than those in [Stocker and Simoncelli \(2006\)](#) and [Hedges et al. \(2011\)](#). A possible reason for the differences between the present work and the aforementioned two studies is the differences in stimuli: in my experiments I followed [Thompson et al. \(2006\)](#) and used only one spatial frequency whereas [Stocker and Simoncelli \(2006\)](#) used broadband gratings and [Hedges et al. \(2011\)](#) used plaids. However, stimulus differences cannot explain the differences between the priors of Chapters 4 and 5; stimuli were identical, with the exception of the reference speed range, which in Chapter 5 extended up to $16^\circ/s$, compared to $12^\circ/s$ in Chapter 4. It is thus remarkable that the extracted priors differ to such degree. The difference between the priors of Chapter 5 on one hand and the priors of Chapter 4 and those in [Stocker and Simoncelli \(2006\)](#) and [Hedges et al. \(2011\)](#) on the other hand suggest that the model is sensitive to noise in the data (and thus to the amount of collected data, which was significantly less in Chapter 5 compared to Chapter 4) but also to experimental parameters, such as the number and range of reference speeds. At this point, it is worth noting that the experiments of Chapters 4 and 5 differed from the experiment of [Stocker and Simoncelli \(2006\)](#) in another way: the direction (left or right) of drift of the gratings was the same for the two gratings in each trial but varied

randomly across trials. In contrast, in my experiments, gratings drifted always to the right. It is possible that subjects' perception of speed is influenced by the knowledge of direction of motion; even in one-dimensional motion, such as in the present experiment, knowing that the gratings always drift to the right could lead an observer to ignore all evidence for the opposite direction. [Stocker and Simoncelli \(2008\)](#) argue that a similar phenomenon underlies certain repulsive biases in motion perception, such as those observed in [Jazayeri and Movshon \(2007\)](#). However, this issue is unlikely to be of relevance in the present experiment, in which the lowest reference speed used is $1^\circ/s$: such speed is well above the speed discrimination threshold in humans (typically $0.3\text{-}0.4^\circ/s$, [Stocker and Simoncelli, 2006](#)) and thus noise in visual measurements would be extremely unlikely to lead to a negative measurement as that would imply an error greater than $1^\circ/s$.

Thus, although the model of [Stocker and Simoncelli \(2006\)](#) offers a substantial improvement over that of [Weiss et al. \(2002\)](#), and although the combination of the ratio model of [Thompson et al. \(2006\)](#) with that of [Stocker and Simoncelli \(2006\)](#) presented in Chapter 4 can better account for certain interactive effects of speed and contrast, care should be taken in the choice of experimental parameters for the models to provide reliable quantitative estimates of subjects' speed priors, especially in cases when psychophysical data is not abundant. Additionally, in my opinion, there is need for further work (which I would have done given more time) towards a more complete model that would combine the strengths of the models of [Stocker and Simoncelli \(2006\)](#) (a non-parametric prior; realistic noise and contrast response models; the use of signal detection theory), [Weiss and Adelson \(1998\)](#) (global integration of complex, non-rigid motion) and [Thompson et al. \(2006\)](#) (physiologically grounded model of speed tuning).

7.2 RELATION TO OTHER WORK

7.2.1 *Motion direction perception and locus of the speed prior*

The series of experiments presented in this thesis fall into two paradigms: one where observers are asked to judge the direction of a field of moving lines and one where observers are asked to compare the speeds of drifting gratings. In Chapters 3 and 6

we saw that directional judgements of low-contrast (and therefore high-uncertainty) low-speed stimuli can change after exposure to a large number of high-speed stimuli. This change occurs in a way consistent with the hypothesis that observers adjust their internal statistics of motion speed (their speed prior) to better match the stimulus statistics and that this adjustment causes a shift in perceived direction (through the relationship between speed and direction in two-dimensional visual motion). The exposure-related change attenuates (and in some cases even reverses) the well-known tendency of observers to perceive the direction of a moving line as being perpendicular to the line when perceptual uncertainty is high, which is the case when the duration of motion or the contrast of the moving stimulus are low. Previous investigations have attributed this perpendicular bias to the fact that the receptive fields (RFs) of motion-sensing neurons (such as those found in areas V₁ and MT of the visual cortex) are small and thus cannot fit the entire moving line, instead “seeing” part of the line (Pack and Born, 2001). This is the neuronal manifestation of the aperture problem, depicted in Figure 4, whereby a neuron’s RF is akin to an aperture: the direction of a moving line longer than the receptive field diameter is ambiguous because its endpoints fall outside the RF. To recover the true direction, ambiguous motion signals from such neurons must be integrated with signals from neurons whose small RF contains one of the line terminators or from neurons with large RFs that are *end-stopped*, i.e. their RF consists of an excitatory (facilitatory) region flanked by inhibitory (suppressive) regions (Figure 40). Lorenceau et al. (1993) suggested that such integration occurs at different time scales for different contrasts and in particular is slower at low contrasts. This hypothesis accounts for the authors’ findings, as well as the findings at the beginning of the experiments of Chapters 3 and 6 in the present thesis: perceived direction tends to be perpendicular to line orientation at the beginning of the trial (or throughout the trial if it is short) but gradually changes to the veridical direction. This change is reflected in observers’ responses (as well as their eye movements, Pack and Born, 2001) at longer trial durations.

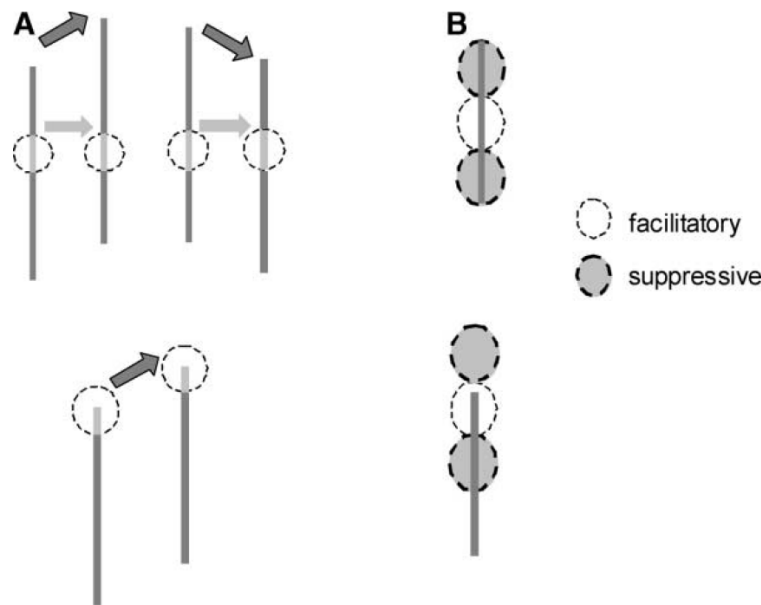


Figure 40: **A**, a small aperture (dotted circle) positioned at the centre of a moving bar allows detection of only the rightward component of motion (light grey arrow), regardless of the actual motion direction (dark grey arrows). However, an aperture positioned at the endpoint of the bar permits measurement of the correct motion direction. **B**, an end-stopped cell responds poorly to a long bar centered on the receptive field, because the bar stimulates both the excitatory (dotted oval) and inhibitory (shaded oval) regions of the receptive field. However, the same end-stopped neuron responds better to the endpoint, because only part of the inhibitory region is stimulated. Adapted from [Pack et al. \(2003\)](#)

However, we saw in Chapters 3 and 6 that with exposure to high speeds, this initial bias towards perpendicular direction progressively decreases. Can such a decrease be explained within the framework of [Lorenceanu et al. \(1993\)](#)? Would the effect of contrast on the time scale of motion integration depend on prior exposure to fast motion? This question, to my knowledge, has not been satisfactorily answered. The Bayesian explanation for the results in Chapters 3 and 6 is that the brain combines sensory input from motion-sensitive neurons with prior expectations to infer the stimulus velocity. This explanation accounts for the fact that it is *fast* motion in particular that can cause the change in bias (because exposure to fast motion changes the prior on slow speeds). However, the Bayesian models described in this work are not formulated on the neural implementation level ([Marr and Poggio, 1976](#)) and thus make no mention of where exactly the priors are encoded; as [Serriès and Seitz \(2013\)](#) note, it is not clear “whether (all) prior expectations correspond to top-down signals modulating early visual pathways, whether they reside entirely in higher-level areas, or on the contrary whether they are formed in sensory cortex itself”. Thus it is not pos-

sible at present to exclude the possibility that a mechanism akin to that suggested by [Lorenceanu et al. \(1993\)](#) is in operation; for example, the prior may be encoded in the way that the non-end-stopped neurons from various channels (i.e. tuned to various speeds or temporal frequencies) are pooled. Pooling of neuronal responses is in fact thought to occur in the middle temporal area (area MT aka V5) and account for the high perceptual sensitivity that is seen in laboratory animals: for example, monkeys have typically lower psychometric thresholds in speed discrimination tasks than do individual neurons or even entire populations of neurons in MT ([Liu and Newsome, 2005](#)). Pooling and integration of non-end-stopped neurons may happen earlier than the integration with the unambiguous signals from the end-stopped neurons that process line terminators ([Lorenceanu et al., 1993](#)). The changing prior may correspond to changes in early integration of non-end-stopped neurons. This may at least partly explain the observation that, at low contrasts, the direction bias is attenuated due to exposure to fast speeds, even at the shortest durations used in [Lorenceanu et al. \(1993\)](#) and in Chapters 3 and 6. Alternatively, or complementarily, the speed prior may be encoded in the tuning properties of individual neurons; for example, [Liu and Newsome \(2005\)](#) found that the proportion of high-pass neurons (i.e. neurons that prefer high speeds) in MT increased in a monkey after extensive exposure to high-speed stimuli. This shift in preferred speeds was hypothesized to help primarily with speed discrimination but a role of such a shift in the encoding of the speed prior cannot be ruled out.

While the encoding of the speed prior in the responses of early sensory areas (such as V1) is not unlikely, the two-stage model presented in Chapter 4, which is able to account for certain subtleties of speed perception (see [4.3.2](#)), fits more naturally to the assumption that the prior is encoded at a level higher than V1. The first stage, which is an instantiation of the ratio model of [Thompson et al. \(2006\)](#), has been proposed to model speed tuning of MT neurons as a result of the integration of motion-sensitive V1 neurons; the Bayesian model of [Stocker and Simoncelli \(2006\)](#) is then used to combine the output of the ratio model (which is taken to represent the likelihood) with the prior in order to derive a speed estimate based on the posterior. This latter stage presumably takes place in higher areas, such as the MT or later. It must be noted, however, that the conceptual order of the ratio and Bayesian models proposed in Chapter 4 need not correspond to the actual order in which computations

are performed in the visual hierarchy. In short, the prior and the likelihood may or may not be encoded in the same area; the available evidence to date is inconclusive. An overview of the possible loci of learned priors in the visual cortex, based on results from electrophysiological studies as well as studies employing electroencephalography (EEG) and functional magnetic resonance imaging (fMRI), is given by [Seriès and Seitz \(2013\)](#).

As a final note, the Bayesian models presented in this thesis make no assumptions regarding the way inferences are implemented in the brain. This is an area of active research, mostly divided between two main theories. One of them, called *probabilistic population codes*, maintains that the activity of a population of neurons represents the posterior probability over a particular stimulus property ([Ma et al., 2006, 2008](#); [Beck et al., 2008](#)). The other theory maintains that posterior probabilities are represented in the brain with a *sampling code*, where neuronal responses correspond to samples of the posterior ([Hoyer and Hyvarinen, 2003](#); [Fiser et al., 2010](#)). While both theories have received experimental support, the evidence is far from conclusive for either theory. It would be interesting to study whether priors are encoded in one of these ways. In the case of the speed prior, the most promising area to look for such populations would be MT, as speed-tuned neurons are primarily found in this area (whereas motion-sensitive neurons in V1 are mostly tuned to temporal frequency and direction).

7.2.2 Specificity of priors

Whether learned priors are highly specific to stimulus configurations or whether (and to what extent) they can be applied to different configurations was examined in Chapter 6. It was found that subjects, on average, can use the learned prior in configurations similar to the trained one – in particular, with stimuli with a small difference in orientation but moving in the same directions as the trained ones (FLIP ORI condition – see [6.3.1](#)). Transfer of learning to the FLIP ORI condition was not seen in all subjects however; the possible reasons for this were discussed in [6.6](#).

What does partial transfer of learning imply for the priors represented in the brain? On one hand, the generalization of the learned prior to the FLIP ORI condition suggests a common prior that is applied to different (but not too different) configurations. This is in agreement with results from the literature on the light-from-above prior,

where it was found that the same prior is used in a variety of stimuli and tasks, such as in a shape-from-shading task with two different stimuli (Adams et al., 2004) as well as a shape-from-shading, a visual search and reflectance task (Adams, 2007, also see Figure 6). The experiment of Chapter 5 also provides evidence for the hypothesis of a single prior: the biases of the female group were different from those of the male group in the same way in both experiments. In particular, contrast had a greater effect on perceived motion speed and direction in females than in males and the modelling of these biases revealed differences in the prior (and not the likelihood).

On the other hand, the prior learned in the Chapter 6 experiments did not transfer to stimuli with orientation and directions that were orthogonal to the trained one (ORTH condition), and for some subjects it did not even transfer to FLIP ORI. If the new speed prior is stimulus-specific, does this mean that there is a “default” speed prior (which favours slow speeds) that is applied to novel stimuli and which can be overridden in a context-specific manner? The evidence from Chapter 6 points towards an affirmative answer in two ways. First, the specificity of the learned prior, which generalizes at most to the FLIP ORI condition, implies that the initial prior continues to apply to the other conditions (ORTH and, for some subjects, FLIP ORI) while observers are learning the specific prior. Second, the finding that the learned prior persists for at least a week after the last training session (Experiment 3 of Chapter 6) implies that exposure to the speed statistics of everyday life did not revert the prior, even partially – in agreement with the persistence of learned priors observed in previous work (the isotropy prior in Knill, 2007 and the light-from-above prior in Adams et al., 2010). If there was a single speed prior we would expect an interference of everyday experience on the experimentally induced prior during the course of a week, given the results in Chapter 3 and 6, which suggest that experience *can* modify the speed prior. For these two reasons, it is likely that the brain can represent multiple, context-specific priors – a hypothesis that is consistent with previous work on other types of priors, such as the study of Gekas et al. (2013), who found that observers can represent two different stimulus-specific priors for motion direction, and the study of Seydell et al. (2010), who presented similar findings in the case of the isotropy prior (see 2.2.2).

In summary, a plausible hypothesis that reconciles the specificity and lack of interference found in Chapter 6 and in previous work (Knill, 2007; Adams et al., 2010;

Seydell et al., 2010; Gekas et al., 2013; Kerrigan and Adams, 2013) with the congruence between the priors in the two experiments of Chapter 5 as well as in previous work (Adams et al., 2004; Adams, 2007) is that there is a “default” prior, possibly innate or shaped during development, that is nevertheless able to spawn context-specific priors that the brain can represent alongside the default one. The present work provides preliminary evidence towards this hypothesis; further work will be necessary to confirm or refute it.

7.2.3 *Relation to other types of learning*

Both short-term (within-session) and long-term (across sessions) learning effects were observed in Chapters 3 and 6. The short-term effects are consistent with the dynamics of statistical learning, which can be very fast, occurring over minutes (Fiser and Aslin, 2001). Such effects are also seen in the learning of priors for motion direction (Chalk et al., 2010; Gekas et al., 2013) and lighting direction (Adams et al., 2004). The long-term learning seen in Chapters 3 and 6 is reminiscent of perceptual learning, which typically progresses over days (Karni and Sagi, 1993), and of the study of Knill (2007) on the plasticity of the object isotropy prior. These different timescales of prior learning and its similarities to other learning types raise interesting questions about the relationship between prior learning, statistical/associative learning and perceptual learning. For example, specificity, which has been considered the hallmark of perceptual learning, is also observed in the learning of priors. At the same time, however, the learned priors can generalize to other stimuli/contexts under certain conditions. Could this be the case for perceptual learning too? Interestingly, recent research suggests that it could. For example, retinal location specificity of perceptual learning has been challenged by Xiao et al. (2008), who used a double-training paradigm that employed training with a particular primary feature at one retinal location coupled with additional training with an irrelevant feature/task at another location. They found that this additional location training enabled a complete transfer of learning of the primary feature to the second location. Perceptual learning has also been used to improve vision in amblyopic patients, with a review by Levi and Li (2009) suggesting that practising on a combination of different tasks and stimuli improves general visual acuity. Generalization of skills acquired through perceptual learning methods

was also recently demonstrated by [Deveau et al. \(2014\)](#), who managed to improve the vision of baseball players with a combination of multiple perceptual learning approaches.

As research continues, links between various types of learning become increasingly clear. The unification of all learning under a common framework that is physiologically grounded and includes Bayesian inference as a component is an exciting prospect and would be a major step towards understanding visual perception and applying its principles in practice, for vision rehabilitation or enhancement. I intend to pursue this line of research in the future.

7.3 CONCLUSIONS

In this thesis, I have presented experimental and theoretical work suggesting that the brain represents a prior probability on speed and that this prior, far from being innate and fixed, can be changed by visual experience. Visual motion perception is viewed here as a form of Bayesian inference, which is the process of combining sensory evidence (likelihood) with expectations (prior). Moreover, this process is dynamic: the prior is continually updated through exposure to environmental statistics. The data from the experiment of Chapter 3 demonstrates a clear shift in direction perception towards oblique directions in the course of the experiment. Furthermore, these perceptual findings are naturally explained by a Bayesian model with a speed prior that favours slow speeds progressively less. In Chapter 4, I provided a more accurate characterization of the speed prior and showed that it varies considerably among individuals. Along the way, I uncovered a novel interactive effect of contrast on perceived speed, observing that contrast biases speed perception less at high speeds than at low speeds when absolute contrast levels are low, whereas the opposite effect was seen at high contrast levels (contrast-induced biases are smaller at high speeds). This interactive effect falls beyond the explanatory power of the presented Bayesian models; however, I showed that a combination of a Bayesian model with a biologically plausible model of speed tuning in visual cortex can account for these effects. Interindividual variability in the speed prior was further explored in Chapter 5, where I presented preliminary evidence for a gender difference in the speed prior, with females having priors that decrease more steeply with speed. Finally, Chapter 6

presented evidence for the partial specificity of the learned speed prior and a related phenomenon – its persistence in time. Both these phenomena are also seen in perceptual learning, hinting towards common mechanisms of apparently different types of learning.

Overall, the present work supports the hypothesis that motion perception is a dynamic, adaptive process that combines past and present experience in a principled way and is responsible both for the remarkable efficiency of human vision and for certain motion illusions and biases. Not all experimental results in this thesis are conclusive and whereas steps were taken towards answering important questions regarding the mutability of the brain's assumptions about motion perception, other questions arose. It is hoped that the present work will stimulate further research to address these matters through a combination of psychophysical and electrophysiological/imaging experiments, thus bringing us one step closer towards understanding how our visual world shapes and is shaped by our expectations and how these expectations are represented in the brain.

BIBLIOGRAPHY

- Abramov, I., Gordon, J., Feldman, O., and Chavarga, A. (2012). Sex & vision i: Spatio-temporal resolution. *Biology of Sex Differences*, 3:1–20. (Cited on page 80.)
- Adams, W. J. (2007). A common light-prior for visual search, shape, and reflectance judgments. *Journal of Vision*, 7(11):11. (Cited on pages 9, 19, 20, 74, 109, and 110.)
- Adams, W. J., Graf, E. W., and Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience*, 7(10):1057–1058. (Cited on pages 7, 16, 18, 20, 21, 44, 109, and 110.)
- Adams, W. J., Kerrigan, I. S., and Graf, E. W. (2010). Efficient visual recalibration from either visual or haptic feedback: the importance of being wrong. *The Journal of Neuroscience*, 30(44):14745–14749. (Cited on pages 16, 21, 82, 100, and 109.)
- Adams, W. J. and Mamassian, P. (2004). Bayesian combination of ambiguous shape cues. *Journal of Vision*, 4(10):7. (Cited on pages 7 and 28.)
- Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The " oblique effect" in man and animals. *Psychological Bulletin*, 78(4):266. (Cited on page 46.)
- Ascher, D. and Grzywacz, N. M. (2000). A bayesian model for the measurement of visual velocity. *Vision Research*, 40(24):3427–3434. (Cited on page 27.)
- Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5(8):617–629. (Cited on page 9.)
- Battaglia, P. W., Kersten, D., and Schrater, P. R. (2011). How haptic size sensations improve distance perception. *PLoS Computational Biology*, 7(6):e1002080. (Cited on page 4.)
- Bavelier, D., Green, C. S., Pouget, A., and Schrater, P. (2012). Brain plasticity through the life span: learning to learn and action video games. *Annual Review of Neuroscience*, 35:391–416. (Cited on page 65.)

- Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., and Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, 60(6):1142–1152. (Cited on page 108.)
- Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., and Pouget, A. (2012). Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron*, 74(1):30–39. (Cited on page 2.)
- Berbec, D. F. (2013). A priori beliefs about speed in video game players. Honours thesis,, University of Edinburgh. (Cited on pages 54, 63, 66, and 126.)
- Birch, J., Young, A., and David, S. (1991). Variations in normal trichromatism. In *Colour vision deficiencies X*, pages 267–272. Springer. (Cited on page 79.)
- Blakemore, M. R. and Snowden, R. J. (1999). The effect of contrast upon perceived speed: a general phenomenon? *Perception*, 28:33–48. (Cited on pages 46, 47, 61, and 62.)
- Bogadhi, A. R., Montagnini, A., Mamassian, P., Perrinet, L. U., and Masson, G. S. (2011). Pursuing motion illusions: A realistic oculomotor framework for bayesian inference. *Vision Research*, 51(8):867–880. (Cited on page 8.)
- Bonneh, Y. S., Cooperman, A., and Sagi, D. (2001). Motion-induced blindness in normal observers. *Nature*, 411(6839):798–801. (Cited on page 46.)
- Brooks, K. (2001). Stereomotion speed perception is contrast dependent. *Perception*, 30(6):725–732. (Cited on page 46.)
- Burge, J., Fowlkes, C. C., and Banks, M. S. (2010). Natural-scene statistics predict how the figure–ground cue of convexity affects human depth perception. *The Journal of neuroscience*, 30(21):7269–7280. (Cited on pages 8 and 12.)
- Burke, D. and Wenderoth, P. (1993). The effect of interactions between one-dimensional component gratings on two-dimensional motion perception. *Vision Research*, 33(3):343–350. (Cited on page 23.)
- Calow, D. and Lappe, M. (2007). Local statistics of retinal optic flow for self-motion through natural sceneries. *Network: Computation in Neural Systems*, 18(4):343–374. (Cited on page 14.)

- Cavanagh, P., Tyler, C. W., and Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *Journal of the Optical Society of America A*, 1(8):893–899. (Cited on page 46.)
- Chalk, M., Seitz, A. R., and Seriès, P. (2010). Rapidly learned stimulus expectations alter perception of motion. *Journal of Vision*, 10(8):1–18. (Cited on pages 10, 15, 18, 28, 44, and 110.)
- Coppola, D. M., Purves, H. R., McCoy, A. N., and Purves, D. (1998). The distribution of oriented contours in the real world. *Proceedings of the National Academy of Sciences*, 95(7):4002–4006. (Cited on page 12.)
- Derrington, A. M. and Goddard, P. A. (1989). Failure of motion discrimination at high contrasts: Evidence for saturation. *Vision Research*, 29(12):1767–1776. (Cited on page 22.)
- Deveau, J., Ozer, D. J., and Seitz, A. R. (2014). Improved vision and on-field performance in baseball through perceptual learning. *Current Biology*, 24(4):R146–R147. (Cited on page 111.)
- Di Luca, M., Ernst, M. O., and Backus, B. T. (2010). Learning to use an invisible visual signal for perception. *Current Biology*, 20(20):1860–1863. (Cited on page 45.)
- Dong, D. W. and Atick, J. J. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems*, 6(3):345–358. (Cited on page 14.)
- Downing, C. J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. *Journal of Experimental Psychology: Human Perception and Performance*, 14(2):188. (Cited on page 10.)
- Farid, H. and Simoncelli, E. (1994). The perception of transparency in moving square-wave plaids. In *Investigative Ophthalmology & Visual Science*, volume 35, pages 1271–1271. (Cited on page 23.)
- Finney, D. J. (1971). *Probit Analysis: 3d Ed.* Cambridge University Press. (Cited on page 127.)
- Fiser, J. and Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, 12(6):499–504. (Cited on page 110.)

- Fiser, J., Berkes, P., Orbán, G., and Lengyel, M. (2010). Statistically optimal perception and learning: from behavior to neural representations. *Trends in Cognitive Sciences*, 14(3):119–130. (Cited on page 108.)
- Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, 59:167–192. (Cited on pages 11 and 15.)
- Geisler, W. S. and Perry, J. S. (2009). Contour statistics in natural images: Grouping across occlusions. *Visual Neuroscience*, 26(1):109–121. (Cited on pages 12 and 13.)
- Gekas, N., Chalk, M., Seitz, A. R., and Seriès, P. (2013). Complexity and specificity of experimentally induced expectations in motion perception. *BMC Neuroscience*, 14(Suppl 1):P355. (Cited on pages 18, 98, 109, and 110.)
- Gershman, S. J., Vul, E., and Tenenbaum, J. B. (2012). Multistability and perceptual inference. *Neural Computation*, 24(1):1–24. (Cited on page 4.)
- Girshick, A. R., Landy, M. S., and Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, 14(7):926–932. (Cited on pages 8 and 12.)
- Gogel, W. C. (1965). Equidistance tendency and its consequences. *Psychological Bulletin*, 64(3):153. (Cited on pages 8 and 12.)
- Goldreich, D. and Tong, J. (2013). Prediction, postdiction, and perceptual length contraction: a bayesian low-speed prior captures the cutaneous rabbit and related illusions. *Frontiers in Psychology*, 4. (Cited on page 25.)
- Gray, R., Regan, D., Castaneda, B., and Sieffert, R. (2006). Role of feedback in the accuracy of perceived direction of motion-in-depth and control of interceptive action. *Vision Research*, 46(10):1676–1694. (Cited on page 24.)
- Green, C. S. and Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science*, 18(1):88–94. (Cited on page 66.)
- Haijiang, Q., Saunders, J. A., Stone, R. W., and Backus, B. T. (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proceedings of the National Academy of Sciences*, 103(2):483. (Cited on page 6.)

- Harris, J. M. and Dean, P. J. (2003). Accuracy and precision of binocular 3-d motion perception. *Journal of Experimental Psychology: Human Perception and Performance*, 29(5):869. (Cited on page 24.)
- Hawken, M. J., Gegenfurtner, K. R., and Tang, C. (1994). Contrast dependence of colour and luminance motion mechanisms in human vision. *Nature*, 367(6460):268–270. (Cited on pages 22, 46, 47, 51, 61, and 62.)
- Hedges, J. H., Stocker, A. A., and Simoncelli, E. P. (2011). Optimal inference explains the perceptual coherence of visual motion stimuli. *Journal of Vision*, 11(6). (Cited on pages 8, 23, 27, 65, 102, and 103.)
- Heeger, D. J. and Simoncelli, E. P. (1993). Model of visual motion sensing. *Spatial Vision in Humans and Robots*, 19:367–392. (Cited on page 22.)
- Hershberger, W. (1970). Attached-shadow orientation perceived as depth by chickens reared in an environment illuminated from below. *Journal of Comparative and Physiological Psychology*, 73(3):407. (Cited on page 16.)
- Horswill, M. S. and Plooy, A. M. (2008). Reducing contrast makes speeds in a video-based driving simulator harder to discriminate as well as making them appear slower. *Perception*, 37(8):1269. (Cited on page 10.)
- Hoyer, P. O. and Hyvarinen, A. (2003). Interpreting neural response variability as monte carlo sampling of the posterior. *Advances in Neural Information Processing Systems*, pages 293–300. (Cited on page 108.)
- Hürlimann, F., Kiper, D. C., and Carandini, M. (2002). Testing the bayesian model of perceived speed. *Vision Research*, 42(19):2253–2257. (Cited on pages 8, 22, 26, 27, and 46.)
- Hurvich, C. M. and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307. (Cited on page 42.)
- Jameson, K. A., Highnote, S. M., and Wasserman, L. M. (2001). Richer color experience in observers with multiple photopigment opsin genes. *Psychonomic Bulletin & Review*, 8(2):244–261. (Cited on page 79.)

- Jazayeri, M. and Movshon, J. A. (2007). A new perceptual illusion reveals mechanisms of sensory decoding. *Nature*, 446(7138):912–915. (Cited on page 104.)
- Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di psicologia*, 49(1):7–30. (Cited on page 99.)
- Kanizsa, G. and Gerbino, W. (1976). Convexity and symmetry in figure-ground organization. *Vision and Artifact*, pages 25–32. (Cited on pages 7 and 12.)
- Karni, A. and Sagi, D. (1993). The time course of learning a visual skill. *Nature*, 365(6443):250–252. (Cited on pages 28, 44, and 110.)
- Kerrigan, I. S. and Adams, W. J. (2013). Learning different light prior distributions for different contexts. *Cognition*, 127(1):99–104. (Cited on pages 20, 98, and 110.)
- Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as bayesian inference. *Annual Review of Psychology*, 55:271–304. (Cited on page 2.)
- Kim, R., Seitz, A., Feenstra, H., and Shams, L. (2009). Testing assumptions of statistical learning: Is it long-term and implicit? *Neuroscience Letters*, 461(2):145–149. (Cited on page 45.)
- Kimura, D. (2000). *Sex and cognition*. MIT press. (Cited on page 79.)
- Knill, D. C. (1998). Discrimination of planar surface slant from texture: human and ideal observers compared. *Vision Research*, 38(11):1683–1711. (Cited on page 7.)
- Knill, D. C. (2007). Learning Bayesian priors for depth perception. *Journal of Vision*, 7(8):13. (Cited on pages 1, 7, 16, 17, 18, 21, 28, 44, 82, 109, and 110.)
- Knill, D. C. and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719. (Cited on page 2.)
- Langer, M. S. and Bulthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception*, 30(4):403–410. (Cited on page 7.)
- Langley, K. (1999). Computational models of coherent and transparent plaid motion. *Vision Research*, 39(1):87–108. (Cited on page 23.)

- Ledgeway, T. and Smith, A. T. (1995). The perceived speed of second-order motion and its dependence on stimulus contrast. *Vision Research*, 35(10):1421–1434. (Cited on page 46.)
- Levi, D. M. and Li, R. W. (2009). Perceptual learning as a potential treatment for amblyopia: a mini-review. *Vision Research*, 49(21):2535–2549. (Cited on page 110.)
- Li, R., Polat, U., Makous, W., and Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience*, 12(5):549. (Cited on page 66.)
- Li, R., Polat, U., Scalzo, F., and Bavelier, D. (2010). Reducing backward masking through action game training. *Journal of Vision*, 10(14):33. (Cited on page 66.)
- Liu, J. and Newsome, W. T. (2005). Correlation between speed perception and neural activity in the middle temporal visual area. *Journal of Neuroscience*, 25(3):711–722. (Cited on page 107.)
- Loffler, G. and Orbach, H. S. (2001). Anisotropy in judging the absolute direction of motion. *Vision Research*, 41(27):3677–3692. (Cited on pages 31, 88, and 98.)
- Lorceau, J., Shiffrar, M., Wells, N., and Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research*, 33(9):1207–1217. (Cited on pages 30, 31, 33, 37, 69, 77, 105, 106, and 107.)
- Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 200(11):1432–1438. (Cited on pages 2 and 108.)
- Ma, W. J., Beck, J. M., and Pouget, A. (2008). Spiking networks for bayesian inference and choice. *Current Opinion in Neurobiology*, 18(2):217–222. (Cited on page 108.)
- Mamassian, P. and Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition*, 81(1):B1–B9. (Cited on page 7.)
- Mamassian, P. and Landy, M. S. (1998). Observer biases in the 3d interpretation of line drawings. *Vision Research*, 38(18):2817–2832. (Cited on pages 8 and 9.)
- Mamassian, P. and Landy, M. S. (2001). Interaction of visual prior constraints. *Vision Research*, 41(20):2653–2668. (Cited on pages 7, 9, and 57.)

- Mamassian, P., Landy, M. S., and Maloney, L. T. (2002). Bayesian modelling of visual perception. In *Probabilistic models of the brain: Perception and neural function*. MIT Press. (Cited on page 4.)
- Mareschal, I., Calder, A. J., and Clifford, C. W. (2013). Humans have an expectation that gaze is directed toward them. *Current Biology*, 23(8):717–721. (Cited on page 8.)
- Marr, D. and Poggio, T. (1976). From understanding computation to understanding neural circuitry. *MIT AI Memos (1959-2004)*, 357. (Cited on pages 29 and 106.)
- McKee, S. P., Klein, S. A., and Teller, D. Y. (1985). Statistical properties of forced-choice psychometric functions: Implications of probit analysis. *Perception & Psychophysics*, 37(4):286–298. (Cited on pages 127 and 128.)
- Montagnini, A., Mamassian, P., Perrinet, L., Castet, E., and Masson, G. S. (2007). Bayesian modeling of dynamic motion integration. *Journal of Physiology-Paris*, 101(1-3):64–77. (Cited on page 8.)
- Moreno-Bote, R., Knill, D. C., and Pouget, A. (2011). Bayesian sampling in visual perception. *Proceedings of the National Academy of Sciences*, 108(30):12491–12496. (Cited on page 5.)
- Mueller, S. T. and Weidemann, C. T. (2008). Decision noise: An explanation for observed violations of signal detection theory. *Psychonomic Bulletin & Review*, 15(3):465–494. (Cited on page 4.)
- Müller, R. and Greenlee, M. W. (1994). Effect of contrast and adaptation on the perception of the direction and speed of drifting gratings. *Vision Research*, 34(16):2071–2092. (Cited on page 46.)
- Murray, I. J., Parry, N. R., McKeefry, D. J., and Panorgias, A. (2012). Sex-related differences in peripheral human color vision: a color matching study. *Journal of Vision*, 12(1):18. (Cited on page 79.)
- Owens, D. and Leibowitz, H. (1976). Oculomotor adjustments in darkness and the specific distance tendency. *Perception & Psychophysics*, 20(1):2–9. (Cited on pages 8 and 12.)

- 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(6823):1040–1042. (Cited on page 105.)
- Pack, C. C., Livingstone, M. S., Duffy, K. R., and Born, R. T. (2003). End-Stopping and the Aperture Problem: Two-Dimensional Motion Signals in Macaque V1. *Neuron*, 39(4):671–680. (Cited on page 106.)
- Pei, Y. C., Hsiao, S. S., and Bensmaia, S. J. (2008). The tactile integration of local motion cues is analogous to its visual counterpart. *Proceedings of the National Academy of Sciences*, 105(23):8130. (Cited on page 11.)
- Peper, L., Bootsma, R. J., Mestre, D. R., and Bakker, F. C. (1994). Catching balls: How to get the hand to the right place at the right time. *Journal of Experimental Psychology: Human Perception and Performance*, 20(3):591. (Cited on page 24.)
- Perrone, J. A. (2005). Economy of scale: A motion sensor with variable speed tuning. *Journal of Vision*, 5(1). (Cited on page 64.)
- Perrone, J. A. and Thiele, A. (2002). A model of speed tuning in mt neurons. *Vision Research*, 42(8):1035–1052. (Cited on pages 57, 58, and 64.)
- Peterson, M. A. and Salvagio, E. (2008). Inhibitory competition in figure-ground perception: Context and convexity. *Journal of Vision*, 8(16):4. (Cited on pages 7, 9, and 12.)
- Petry, S. E. and Meyer, G. E. (1987). *The perception of illusory contours*. Springer-Verlag Publishing. (Cited on page 99.)
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, 331(6152):163–165. (Cited on page 7.)
- Ramachandran, V. S., Gregory, R. L., et al. (1991). Perceptual filling in of artificially induced scotomas in human vision. *Nature*, 350(6320):699–702. (Cited on page 46.)
- Rauber, H. J. and Treue, S. (1998). Reference repulsion when judging the direction of visual motion. *Perception*, 27:393–402. (Cited on page 46.)

- Rodriguez-Carmona, M., Sharpe, L., Harlow, J., and Barbur, J. (2008). Sex-related differences in chromatic sensitivity. *Visual Neuroscience*, 25(03):433–440. (Cited on page 79.)
- Roth, S. and Black, M. J. (2005). On the spatial statistics of optical flow. In *Tenth IEEE International Conference on Computer Vision*, volume 1, pages 42–49. IEEE. (Cited on page 14.)
- Sagi, D. and Tanne, D. (1994). Perceptual learning: learning to see. *Current Opinion in Neurobiology*, 4(2):195–199. (Cited on page 44.)
- Sclar, G., Maunsell, J. H., and Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research*, 30(1):1–10. (Cited on pages 26, 27, and 52.)
- Seitz, A. R. and Dinse, H. R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology*, 17(2):148–153. (Cited on page 44.)
- Seitz, A. R., Nanez, J. E., Holloway, S. R., Koyama, S., and Watanabe, T. (2005). Seeing what is not there shows the costs of perceptual learning. *Proceedings of the National Academy of Sciences*, 102(25):9080. (Cited on page 44.)
- Sekuler, R. and Ball, K. (1977). Mental set alters visibility of moving targets. *Science*, 198(4312):60–62. (Cited on pages 6 and 10.)
- Seriès, P. and Seitz, A. R. (2013). Learning what to expect (in visual perception). *Frontiers in Human Neuroscience*, 7. (Cited on pages 6, 7, 106, and 108.)
- Seydell, A., Knill, D. C., and Trommershäuser, J. (2010). Adapting internal statistical models for interpreting visual cues to depth. *Journal of Vision*, 10(4):1. (Cited on pages 18, 99, 109, and 110.)
- Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology*, 13(2):144–149. (Cited on pages 11 and 15.)
- Snowden, R. J., Stimpson, N., and Ruddle, R. A. (1998). Speed perception fogs up as visibility drops. *Nature*, 392:450. (Cited on pages 10, 22, and 47.)

- Sotiropoulos, G., Seitz, A. R., and Seriès, P. (2011). Changing expectations about speed alters perceived motion direction. *Current Biology*, 21(21):R883–R884. (Cited on pages 30, 50, 69, and 77.)
- Sotiropoulos, G., Seitz, A. R., and Seriès, P. (2014). Contrast dependency and prior expectations in human speed perception. *Vision Research*, 97:16–23. (Cited on page 46.)
- Stocker, A. and Simoncelli, E. P. (2008). A bayesian model of conditioned perception. In *Advances in neural information processing systems*, pages 1409–1416. (Cited on page 104.)
- Stocker, A. A. and Simoncelli, E. P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nature Neuroscience*, 9(4):578–585. (Cited on pages 8, 10, 12, 14, 22, 23, 25, 26, 27, 28, 29, 30, 45, 46, 48, 50, 51, 52, 53, 54, 55, 56, 57, 59, 63, 64, 65, 67, 74, 75, 76, 77, 78, 102, 103, 104, 107, 127, 130, and 131.)
- Stone, L., Watson, A., and Mulligan, J. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research*, 30(7):1049–1067. (Cited on page 23.)
- Stone, L. S. and Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research*, 32(8):1535–1549. (Cited on pages 22, 46, 47, 50, and 62.)
- Sun, J. and Perona, P. (1998). Where is the sun? *Nature Neuroscience*, 1(3):183–184. (Cited on page 7.)
- Tadin, D., Lappin, J. S., Gilroy, L. A., and Blake, R. (2003). Perceptual consequences of centre-surround antagonism in visual motion processing. *Nature*, 424(6946):312–315. (Cited on page 22.)
- Thompson, P. (1982). Perceived rate of movement depends on contrast. *Vision Research*, 22(3):377–380. (Cited on pages 22, 46, 47, 61, and 62.)
- Thompson, P., Brooks, K., and Hammett, S. T. (2006). Speed can go up as well as down at low contrast: Implications for models of motion perception. *Vision Research*, 46(6–7):782–786. (Cited on pages 22, 46, 47, 48, 49, 51, 57, 58, 59, 61, 62, 64, 76, 103, 104, and 107.)
- Thompson, P., Stone, L. S., and Swash, S. (1996). Speed estimates from grating patches are not contrast-normalized. *Vision Research*, 36(5):667–674. (Cited on page 46.)

- Tomassini, A., Morgan, M. J., and Solomon, J. A. (2010). Orientation uncertainty reduces perceived obliquity. *Vision Research*, 50(5):541–547. (Cited on pages 8 and 11.)
- Tse, P. and Hsieh, P. J. (2006). The infinite regress illusion reveals faulty integration of local and global motion signals. *Vision Research*, 46(22):3881–3885. (Cited on page 46.)
- Vilares, I. and Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences*, 1224(1):22–39. (Cited on page 2.)
- von Helmholtz, H. (1962). *Helmholtz's treatise on physiological optics*. Dover Publications. (Cited on page 2.)
- Voyer, D., Voyer, S., and Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. *Psychological Bulletin*, 117(2):250. (Cited on page 79.)
- Watson, A. B. and Pelli, D. G. (1983). Quest: A bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics*, 33(2):113–120. (Cited on pages 49 and 62.)
- Weiss, Y. and Adelson, E. H. (1998). Slow and smooth: a bayesian theory for the combination of local motion signals in human vision. Technical report. (Cited on pages 8, 22, 25, 28, 102, and 104.)
- Weiss, Y., Simoncelli, E. P., and Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, 5(6):598–604. (Cited on pages 4, 8, 10, 14, 22, 23, 24, 25, 26, 27, 28, 30, 34, 45, 102, and 104.)
- Welchman, A. E., Lam, J. M., and Bühlhoff, H. H. (2008). Bayesian motion estimation accounts for a surprising bias in 3d vision. *Proceedings of the National Academy of Sciences*, 105(33):12087–12092. (Cited on page 24.)
- Welchman, A. E., Tuck, V. L., and Harris, J. M. (2004). Human observers are biased in judging the angular approach of a projectile. *Vision Research*, 44(17):2027–2042. (Cited on page 24.)

- Wozny, D. R., Beierholm, U. R., and Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS Computational Biology*, 6(8):e1000871. (Cited on page 4.)
- Wuerger, S., Shapley, R., and Rubin, N. (1996). On the visually perceived direction of motion by Hans Wallach: 60 years later. *Perception*, 25:1317–1368. (Cited on pages 10 and 30.)
- Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., and Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Current Biology*, 18(24):1922–1926. (Cited on page 110.)
- Yang, Z. and Purves, D. (2003). A statistical explanation of visual space. *Nature Neuroscience*, 6(6):632–640. (Cited on page 12.)
- Yo, C. and Wilson, H. R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast and eccentricity. *Vision Research*, 32(1):135–147. (Cited on page 23.)
- Zanker, J. and Zeil, J. (2005). Movement-induced motion signal distributions in outdoor scenes. *Network: Computation in Neural Systems*, 16(4):357–376. (Cited on page 13.)
- Zhang, R., Kwon, O.-S., and Tadin, D. (2013). Illusory movement of stationary stimuli in the visual periphery: evidence for a strong centrifugal prior in motion processing. *The Journal of Neuroscience*, 33(10):4415–4423. (Cited on pages 14, 22, 27, and 28.)

APPENDIX

8.1 SUPPLEMENTARY PROCEDURES OF CHAPTER 4

8.1.1 *Derivation of Equation 17*

The perceived speeds of the two gratings, \hat{v}_1 and \hat{v}_2 are normally distributed random variables (Equations 13 and 14) and thus can be written as

$$\begin{aligned}\hat{v}_1 &= v_1 + a(v_1)\sigma_1^2 + \sigma_1^2 s_1 \\ \hat{v}_2 &= v_2 + a(v_2)\sigma_2^2 + \sigma_2^2 s_2\end{aligned}$$

where s_1 and s_2 are independent standard Gaussian random variables ($\sim \mathcal{N}(0, 1)$). At the PSE, $\hat{v}_1 = \hat{v}_2$ i.e. $v_2 = v_1 + a(v_1)\sigma_1^2 - a(v_2)\sigma_2^2 + \sigma_1^2 s_1 - \sigma_2^2 s_2$. Thus v_2 is a (weighted) sum of independent normal random variables plus a constant and therefore it is itself a normal random variable with mean $v_1 + a(v_1)\sigma_1^2 - a(v_2)\sigma_2^2$ and variance $\sigma_1^2 + \sigma_2^2$. Since the logarithm of the prior is assumed locally linear in the model, $a(v_1) = a(v_2)$ and Equation 17 follows directly.

8.1.2 *Model validation dataset*

Following up to the present study, additional data was collected by Dan Berbec as part of his Honours project under the supervision of Peggy Seriès (Berbec, 2013), with the view to looking at different aspects of the topic of speed priors. This dataset consists of a larger number of subjects (21) than the present dataset; however it includes significantly fewer sessions (at most 6 staircases for each reference speed and contrast condition) per subject, particularly in the conditions with the highest two reference speeds (which show the greatest variability and thus require more data for an accurate estimate of CDB). Although some experimental parameters, such as reference

speeds, differ from this experiment and thus this dataset cannot be pooled with that of the present study, it was useful for validating my method of fitting the Bayesian model of [Stocker and Simoncelli \(2006\)](#).

8.1.3 PSE Variability

Here I derive a lower bound on the variance of the test speed v_2 at the PSE. I first note that the psychometric function (Equation 15) presented in [Stocker and Simoncelli \(2006\)](#) can be simplified if I slightly rearrange the lhs as $p(\hat{v}_2 - \hat{v}_1 > 0)$. Since \hat{v}_1 and \hat{v}_2 are Gaussian variables, their difference is also Gaussian with mean $\mu_2 - \mu_1 = v_2 + a(v_2)\sigma_2^2 - v_1 - a(v_1)\sigma_1^2$ and variance $\sigma_1^2 + \sigma_2^2$. Therefore $p(\hat{v}_2 > \hat{v}_1) = p(\hat{v}_2 - \hat{v}_1 > 0) = p(\Delta\hat{v} > 0)$ is the complement of a cumulative normal distribution

$$p(\Delta\hat{v} > 0) = \int_0^\infty p(\Delta\hat{v})d\Delta\hat{v} = \Phi\left(\frac{\mu_2 - \mu_1 - \Delta\hat{v}}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \quad (23)$$

Therefore the Bayesian observer model is a type of probit analysis and thus I can use standard theoretical ([Finney, 1971](#)) and simulation ([McKee et al., 1985](#)) results from probit analysis. The simplest analytical formula for the standard error (SE) of the staircase threshold estimate (in this case T_{50}) is

$$\sigma_{PSE} = \frac{\sigma}{\sqrt{\sum wn}} \quad (24)$$

where $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$ is the slope of the psychometric function, n is the number of trials at each stimulus level and w are the probit weights. As [McKee et al. \(1985\)](#) point out, this is a variation of the common statistical formula s/\sqrt{N} used to estimate the standard error of the mean. The similarity between the two formulas makes intuitive sense if the staircase procedure is thought of as an estimator of the mean of the underlying distribution of the test speed v_2 at the PSE (Equation 17). In their simulations, [McKee et al. \(1985\)](#) found that the SE of a staircase estimate for T_{75} (the threshold at 75%) in a 2-AFC paradigm is approximately $3\sigma/\sqrt{N}$ (in the variable-slope case) and also that the SE of T_{50} in a yes-no paradigm is approximately half of that. Furthermore, they assert that under certain conditions of trial placement, these are lower bounds for any possible staircase procedure. Since the staircase procedure

was designed to estimate T_{50} and $p(\hat{v}_2 > \hat{v}_1)$ corresponds to a yes-no paradigm, it follows that

$$\sigma_{PSE} \geq \frac{3\sigma}{2\sqrt{N}} = \frac{3\sqrt{\sigma_1^2 + \sigma_2^2}}{2\sqrt{40}}$$

and thus

$$\sigma_{PSE}^2 \geq \frac{\sigma_1^2 + \sigma_2^2}{17.7} \quad (25)$$

This lower bound is tighter when the number of trials per staircase is relatively small, typically less than 60 (McKee et al., 1985), such as in this experiment. My empirically determined SE, whose square is given by Equation 19, is in broad agreement with these results.

8.1.4 Derivation of Equation 20

The logarithm of the prior is assumed a locally linear function of speed (see Equation 10), that is

$$f(v) = \ln p(v) = a(v)v + b$$

Thus, within a small range of speeds Δv , the following holds:

$$f(v + \Delta v) = f(v) + a(v)\Delta v$$

or

$$a(v) = \frac{f(v + \Delta v) - f(v)}{\Delta v} \quad (26)$$

Letting $\Delta v \rightarrow 0$, the right-hand side of Equation 26 is simply the derivative of $f(v)$, i.e. $a(v) = \frac{df(v)}{dv}$. Integrating both sides, this becomes

$$f(v) = \int a(v)dv + C$$

and thus

$$p(v) = \exp(f(v)) = \exp\left(\int a(v)dv + C\right) = C' \exp\left(\int a(v)dv\right)$$

where $C' = \exp(C)$ plays the role of a normalizing constant, i.e. its value is such that $\int_0^\infty p(v)dv = 1$.

8.2 SUPPLEMENTARY FIGURES OF CHAPTER 4

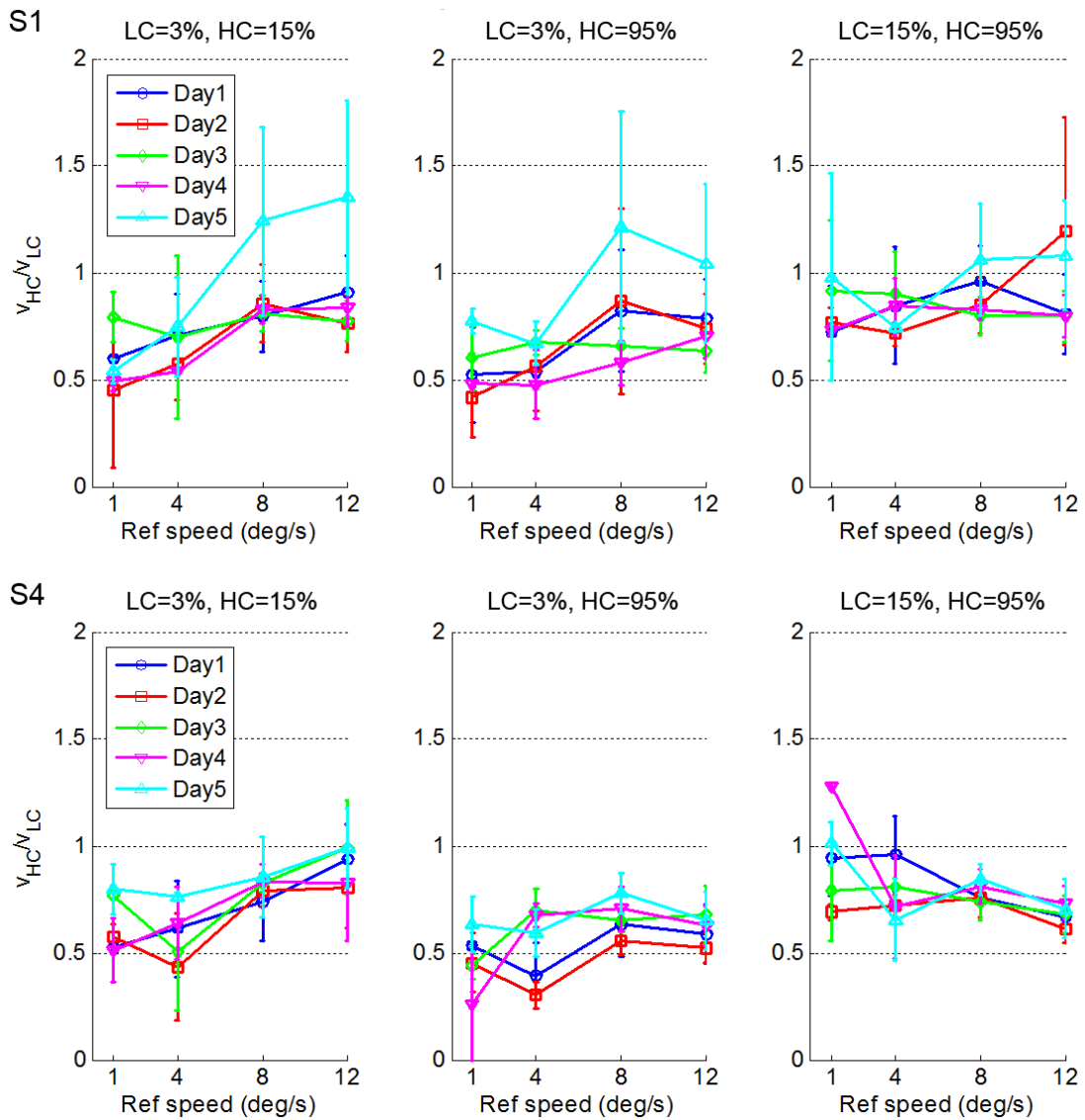


Figure 41: Mean ratio of speeds of the high (v_{HC}) and low-contrast (v_{LC}) gratings at the point of subjective equality (PSE), plotted as a function of speed, separately for each experimental session and contrast condition. Data from two subjects (S1 and S4) who showed the greatest across-sessions variability. Error bars are std.dev.

8.3 SUPPLEMENTARY FIGURES OF CHAPTER 5

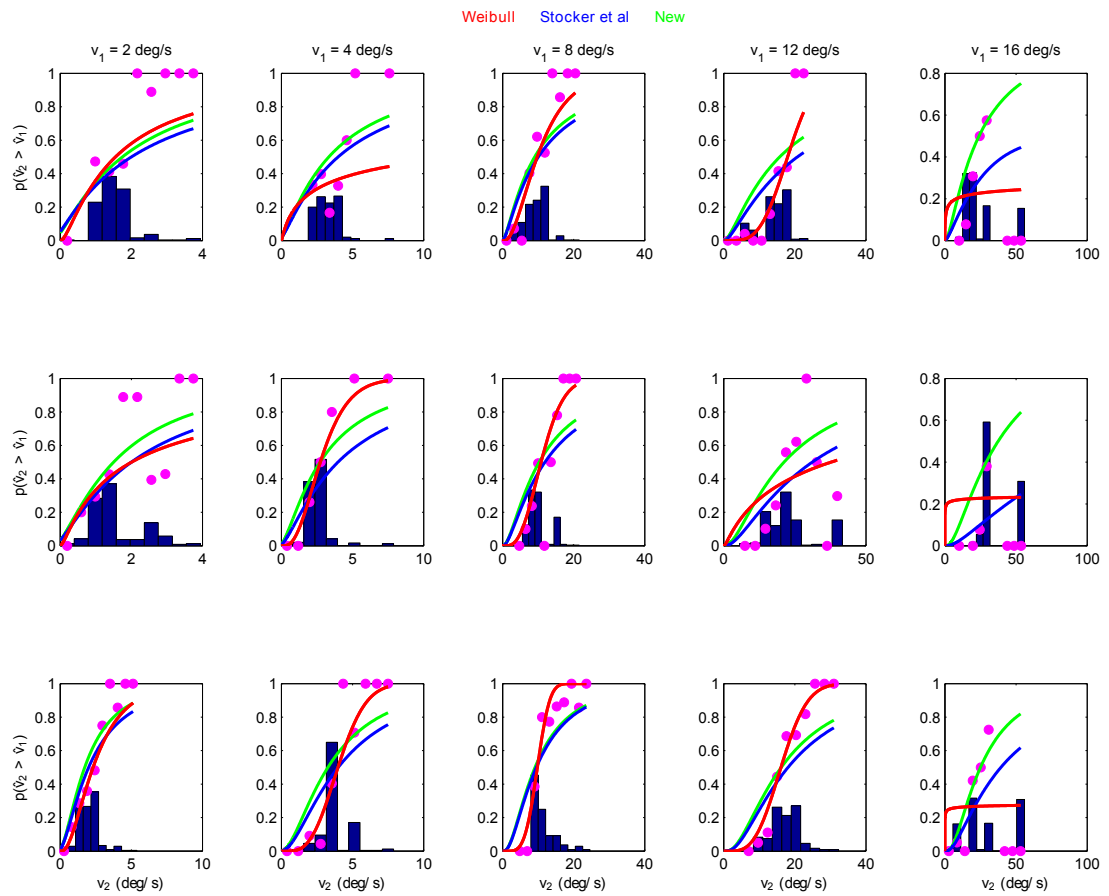


Figure 42: Raw psychometric data collected from the subject with the least decreasing speed prior in Experiment 2 of Chapter 5 (Figure 24A). Rows represent contrast conditions: in the top row, the contrasts of test and reference grating were 3% and 15%, respectively. Horizontal axis is speed of test grating (v_2). Histograms show the distribution of test speeds presented to the subject. Circles are proportion of trials in which the test grating is perceived faster than the reference grating ($p(\hat{v}_2 > \hat{v}_1)$), i.e. they represent the subject's psychometric function. Curves are fits to the psychometric function: green and blue are Bayesian model fits using my method and the method of [Stocker and Simoncelli \(2006\)](#), respectively (see 4.3.1); red are Weibull functions fit individually to each condition. Cf Supplementary Figure 1a-d in [Stocker and Simoncelli \(2006\)](#).

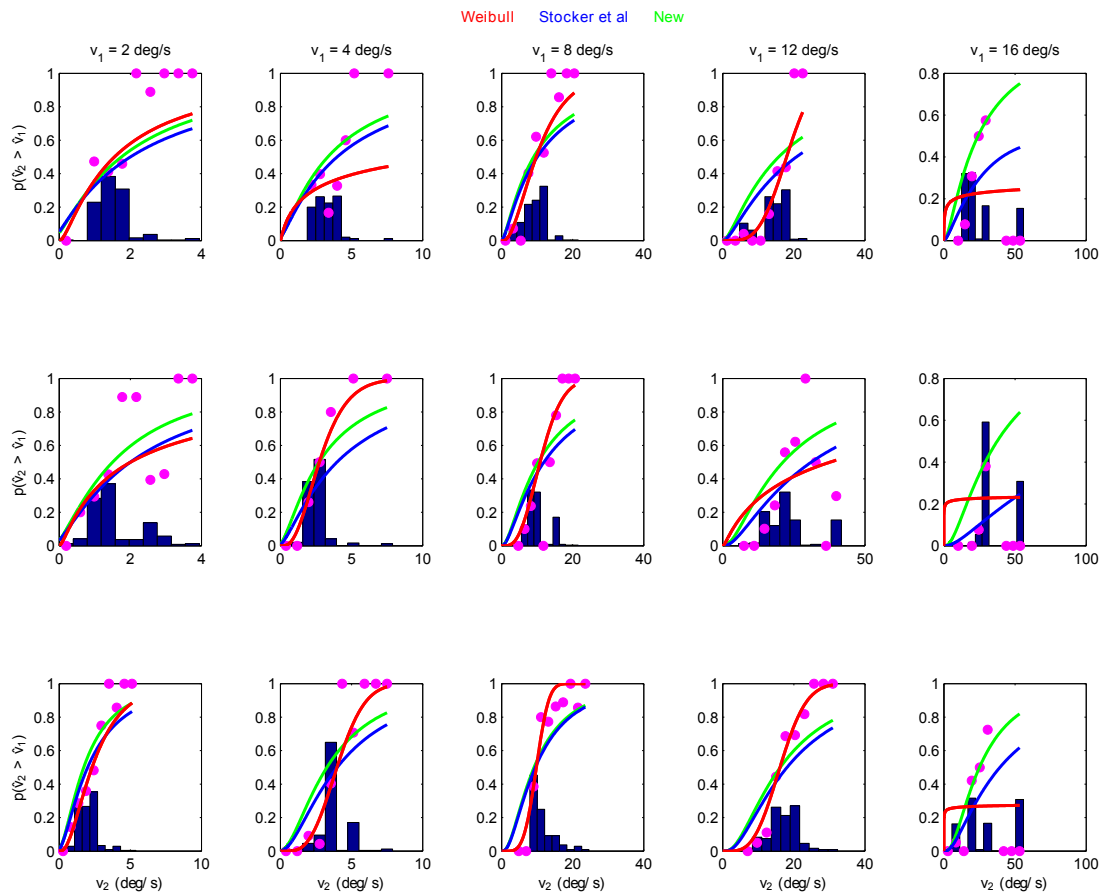


Figure 43: Raw psychometric data collected from the subject with the most decreasing speed prior in Experiment 2 of Chapter 5 (Figure 24B). Rows represent contrast conditions: in the top row, the contrasts of test and reference grating were 3% and 15%, respectively. Horizontal axis is speed of test grating (v_2). Histograms show the distribution of test speeds presented to the subject. Circles are proportion of trials in which the the test grating is perceived faster than the reference grating ($p(\hat{v}_2 > \hat{v}_1)$), i.e. they represent the subject's psychometric function. Curves are fits to the psychometric function: green and blue are Bayesian model fits using my method and the method of [Stocker and Simoncelli \(2006\)](#), respectively (see 4.3.1); red are Weibull functions fit individually to each condition. Cf Supplementary Figure 1a-d in [Stocker and Simoncelli \(2006\)](#).