The Perceptual Mechanisms of Probability Effects

by
Syaheed Jabar

A thesis
presented to the University of Waterloo
in the fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Psychology

Waterloo, Ontario, Canada, 2018

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Matthijs van der Meer Assistant Professor
Supervisor	Dr. Britt Anderson Associate Professor
Internal Member	Dr. Roxane Itier Associate Professor
	Dr. Evan Risko Associate Professor
Internal-external Member	Dr. Ben Thompson Associate Professor

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

All data analyses were conducted by me. Chapters 2,3,4 and 5 are based on published articles in which I was the lead author, and which were co-authored with my supervisor, Dr. Britt Anderson. Alex Filipowicz contributed to the experimental design and data collection for Chapters 4 and 5.

Abstract

Environmental statistics impact human behaviour. The more likely something is to occur, the faster and more accurate we are at detecting it. This *probability effect* has been studied in numerous forms. However, there is no clear account of the mechanisms driving the effect. While attention and decision-making has been implicated, these interpretations largely hinge on the task employed. Instead, probability might have an earlier effect, one that is perceptual in nature. This thesis explores the idea that feature (e.g. orientation, color, etc.) probability shapes perception through selective tuning of the relevant neurons. Particularly, where orientation probability is involved, V1 neurons preferring the likely orientations are selectively sharpened. To test this hypothesis, a mixture of established tasks (Chapter 2) and novel behavioural paradigms (Chapters 3/4) were utilized. An electrophysiological examination specifically aimed at V1 was also carried out (Chapter 5). Neural modelling was then done to link the behaviour to a concrete neural mechanism, which generated predictions that could be evaluated by additional behavioural data (Chapters 6/7). These diverse methods provide converging evidence for the tuning hypothesis of feature probability, and argue for an interdisciplinary approach in cognitive research.

Acknowledgements

I’d first like to express my gratitude to the thesis committee, to the people who have reviewed my works for publication, and to anyone who has provided feedback for my research.

This work would not have been possible without the masterful supervision of Dr. Britt Anderson. My main priority in choosing a lab in which to pursue graduate studies was one in which I could learn the most from, and being in Britt’s lab has definitely delivered on that. Because of the ‘do-it-yourself’ spirit that Britt inculcates in his students, there was ample room to learn and make use of new techniques and approaches, which I likely would not have picked up anywhere else.

Appreciation also goes to past and current lab members, particularly those who have helped with the data collection. Special thanks goes to (now Dr.) Alex Filipowicz. I can think of no finer example of a graduate student from which to learn from, and with whom it has been a joy collaborating/ chatting about research with.

Because I tend to write papers / code with my headphones on, it would be remiss of me not to also to thank all the composers whose tunes have made them the unsung heroes behind this research (pun intended). And to everyone else who has distracted me from thinking too much about research or has just been there for moral support, my sanity and I thank you.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. Probability Effects are Widespread

Objects that occur frequently, or that are likely to occur, are generally reacted to faster and/or more accurately, as compared to objects that are rarer. Such *probability effects* has been noted in a wide variety of tasks ranging from simple detection of glyphs, i.e. letter and digits (Hon, Yap, & Jabar, 2013[1]; Laberge & Tweedy, 1964; Miller & Pachella, 1973) to more complex tasks like visual search (Rich et al., 2008; Wolfe et al., 2007). Probability effects are also easily and rapidly evoked (Hasher & Zacks, 1984). Because we are constantly inundated with environmental information, *how* an object's implicit probability (or frequency of occurrence) affects behavioral outcomes and the mechanisms by which it does so is a fundamental cognitive process that we should try to get a grasp on.

Unfortunately, the mechanisms underlying probability effects are ill-understood. While some have attributed it to decisional criterion shifts (Menneer, Donnelly, Godwin, & Cave, 2010; Wolfe & Van Wert, 2010) or to attentional mechanisms (Hon & Tan, 2013), others posit a perceptual locus (Dykes & Pascal, 1981). Lack of agreement on what probability does is particularly disconcerting when one considers how probability manipulations have been used for the express purpose of studying the effects of attention (e.g. Schwartz & Kirsner, 1982; MacLeod et al., 2002; Mounts & Tomaselli, 2005; Iarocci et al., 2006). Supporting the idea that probability effects are due to low-level perceptual changes rather than due to attention are studies that utilise perceptual estimation tasks. By manipulating the orientation probability of spatial gratings, the precision with which participants make their orientation estimation can be modulated (Anderson, 2014; Jabar & Anderson, 2015[2]).

---

[1] Letter-based probability effects was something I researched when I was an undergraduate.

[2] Jabar & Anderson (2015) was comprised of the studies that were done for my Master's thesis. While the findings are referenced in this PhD thesis, they are not part of the experimental chapters.

1.2. Probability and V1

In Jabar & Anderson (2015), it was further observed that probability effects were particularly pronounced for the cardinal orientations, specifically the near-vertical tilts. Orientation anisotropy, referring to the phenomena where cardinal orientations are better perceived than oblique ones, is well-studied (Appelle, 1972), and is likely due to differences in orientation tuning properties across the primary visual cortex (V1).

V1 gets majority of its input from the lateral geniculate nucleus (LGN) in the thalamus, and represents the initial site of visual cortical processing. V1 is responsible for simple feature processing such as color and orientation (Johnson, Hawken & Shapley, 2008). Within any localized patch of V1 cortex, orientation-selective neurons that are arranged in a columnar fashion based on their stimulus preference. Neurons that preferentially fire to 45 degree stimuli are found in proximity to neurons that prefer a 44 degree stimulus. V1 firing also tapers off gradually: A neuron that maximally prefers 45 degrees will fire to a 44 degree stimulus, but to a smaller extent. It will fire even less to a 42 degree stimulus, and so on. A tuning curve is a useful illustration of how much a neuron will fire given a range of possible stimuli. The more broadly-tuned a neuron is, the greater the range of possible stimuli it is responsive to, and the less selective it is. *Figure 1.1* shows example tuning curves for a V1 population that does not have anisotropy.

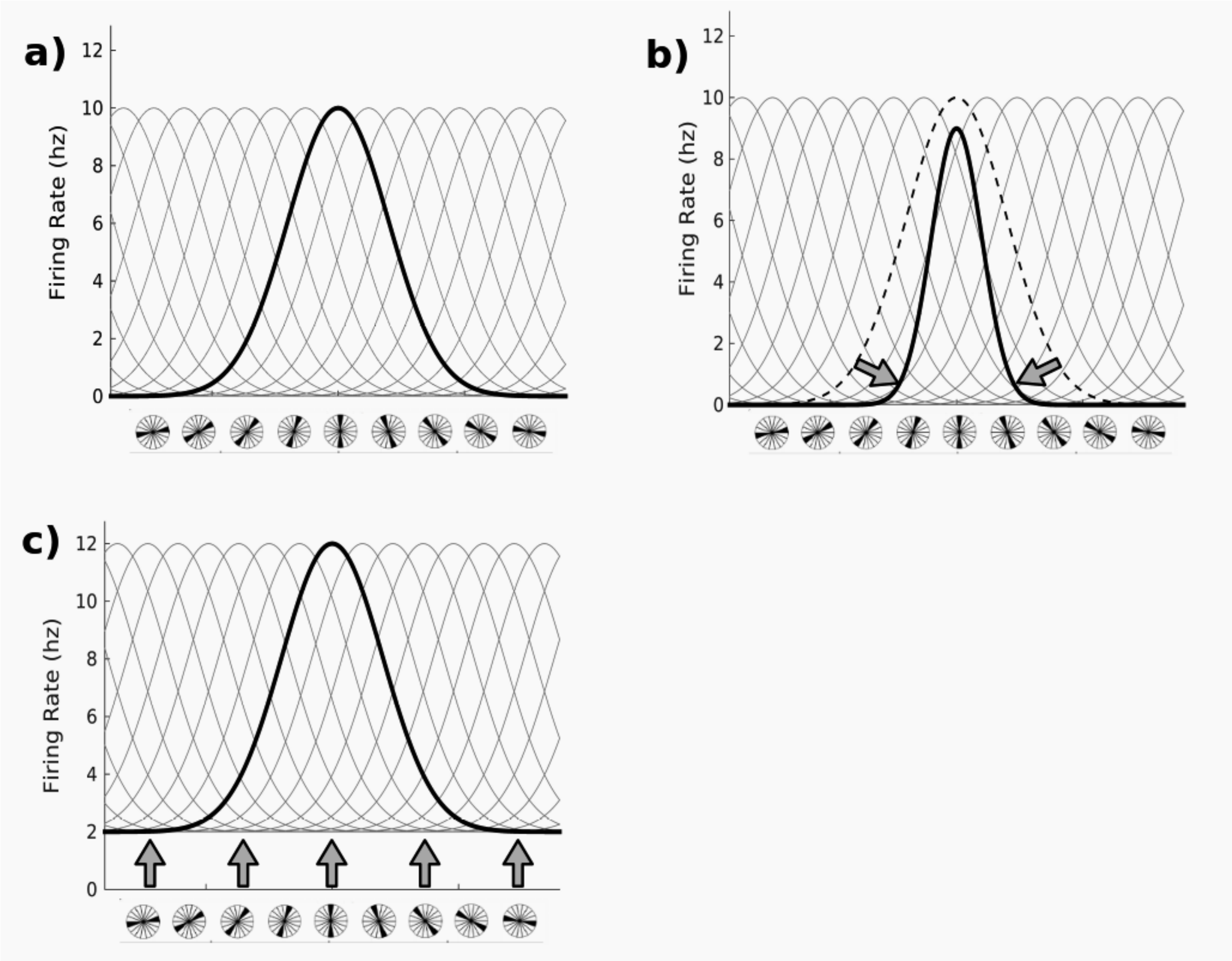

Figure 1.1. Depiction of gain vs. tuning. (a) The default tuning curves of a population of orientation-selective neurons. Y-axis: firing rate, X-axis: orientation shown within the population's receptive field. The bold line represents the tuning curve of a vertical-preferring neuron. (b) Tuning curves if the vertical-preferring neuron is *selectively* tuned. (c) Tuning curves after a *baseline* increase due to gain. Note that no orientation is particularly privileged in this instance.

In actuality, there are more V1 neurons coding for cardinal (vertical/horizontal) than oblique orientations, with vertical-preferring neurons being more broadly tuned than the oblique and horizontal preferring neurons (Li, Peterson, & Freeman, 2003). If probability interacts with anisotropy, it is an indication that orientation probability effects might have a similar locus: Perhaps the frequency of occurrence of stimuli affects the tuning properties of the V1 orientation-selective neurons that prefer it.

If the neural mechanism behind how probability impacts V1 activity can be mapped, it would provide some insight as to why probability effects are difficult to eliminate or counteract with instructions (Lau & Huang, 2010) or reward (Wolfe et al., 2007). Changes in decision-making could simply be a downstream effect of earlier changes in perceptual tuning (Eckstein, Peterson, Pham, & Droll, 2009), rather than a cause of the probability effects.

1.3. Effects of Attention on V1

Can ‘attention’ provide an account for how orientation probability affects perception? While the core notion of attention is that it is a selection process of competing representations (Allport, 1993), this is a nebulous definition (Anderson, 2011), particularly since there are different types of ‘attentional’ manipulations which may elicit different effects (Carrasco, 2011). Of importance is that while attentional manipulations might typically be thought of as involving the frontal and parietal cortices, these regions can modulate perceptual processing areas in a ‘top-down’ manner (e.g. Bressler et al., 2008).

Attentional cuing demonstrably modulates V1 in both human functional imaging (e.g. Brefczynski & DeYoe, 1999; Gandhi, Heeger & Boynton, 1999; Tootell et al., 1998), and in monkey single-cell recording (e.g. Sharma et al., 2015; Wang, Chen, Yan, Zhaoping, & Li, 2015) studies. The basis for these perceptual effects of attention are broadly of two types. While space-based attention is thought to involve a gain-based process (*Figure 1.1c*), feature-based attention is thought to additionally involve a tuning-process (*Figure 1.1b*) (Martinez-Trujillo & Treue, 2004; Ling, Liu, & Carrasco, 2009; Carrasco, 2011).

For current purposes, ‘features’ refer to non-spatial dimensions such as orientation or color. Supporting this feature/spatial distinction is the direction of V1 modulation in fMRI studies: Cuing a space results in an *increase* of activation in V1 (Gandhi, Heeger, & Boynton, 1999), which would be line with a gain-based account, since neurons would fire more. Cuing orientations *decreases* V1 activation (Kok et al., 2012), in line with a neural tuning/sharpening account, since neurons fire less.

1.4. Effects of Perceptual Learning on V1

While the attention literature offers some insights as to how perceptual tuning in V1 might occur, neural tuning can also be observed without the use of explicit ‘attentional’ cues. By showing monkeys the same oriented stimulus repeatedly, V1 neurons preferring that orientation are selectively ‘sharpened’ in their tuning width (Ringach, Hawken, & Shapley, 1997; Schoups, Vogels, Qian, & Orban, 2001). While V1 primarily receives its input from the LGN, there exist inhibitory and excitatory connections that allow for interactions local to the visual cortex (Tsumoto, Eckart, & Creutzfeldt, 1979; Sillito, Kemp, Milson, & Berardi, 1980; Ahmed, Anderson, Douglas, Martin, & Nelson, 1994; Sato et al., 2016). In contrast to the idea that V1 is modulated ‘top-down’ by the fronto-parietal cortices (e.g. Bressler et al., 2008), mechanistic accounts of orientation training posit that these local cortical circuits are responsible for differences in tuning (Carandini & Ringach, 1997; Teich & Qian, 2003).

Given that probability-effects have also been suggested to be experience-dependent rather than due to top-down influence (Lau & Huang, 2010), and that a ‘trained’ stimulus is axiomatically an extremely probable stimulus, orientation probability effects could operate in a manner similar to orientation training in monkeys.

1.5. Event-related potentials and feedback

Is it possible to distinguish between V1 tuning due to top-down attention, from V1 tuning due to a more bottom-up, experience-dependent process? One method is to look at the time-course of *when* the V1 modulation occurs. Neural firing creates an electrical potential which is conducted through the brain matter, skull and scalp. While electroencephalography (EEG) offers great temporal resolution, spatial resolution is lacking. Because measurements are made on the scalp, inferences have to be made about how shape/volume of the head distorts these electrical potentials. Because event-related potentials (ERPs) are a particularly valuable way to look at perception and attention (Woodman, 2010), there are studies on the cortical sources of these potentials that we can leverage on.

C1 is an early ERP component that has been source localized to the banks of the calcarine sulcus, where V1 is located (Di Russo, Martínez, Sereno, Pitzalis, & Hillyard, 2002). This component is modulated by low-level stimulus characteristics (brightness, eccentricity, etc.), and occurs too early for significant top-down trial-specific ‘attentional’ modulation, unlike the P1, a later visual ERP (Clark & Hillyard, 1996; Di Russo, Martinez, & Hillyard, 2003). These studies also source-localized P1 to the extrastriate cortex rather than to the striate cortex. While Chapter 5 goes into more detail, the hypothesis we can derive from these ERP studies is that while experience-dependent tuning should cause sharpening in V1 and hence a *decrease* in C1 amplitudes (e.g. Pourtois, Rauss, Vuilleumier, & Schwartz, 2008), top-down attention to a feature would likely feedback via the extrastriate cortex, causing an *increase* in P1 amplitudes (e.g. Zhang & Luck, 2009).

1.6. Thesis Outline

Through converging behavioural, electrophysiological and computational methodologies, this thesis explores the hypothesis that *feature* probability (e.g. orientation or color probability) can result in an experience-dependent tuning of neurons involved in perceptual processing. Since there is information to be gained from ruling out competing hypotheses, the behavioural experiments in Chapters 2-4 pitted orientation probability against manipulations with better understood underlying mechanisms. Chapter 2 used an orientation probability manipulation concurrently with a traditional 'attentional' manipulation: spatial exogenous cuing/ peripheral cuing. While the two manipulations created similar effects, these effects were non-interactive, suggesting that separate mechanisms are involved. Chapters 3 contrasted orientation probability against spatial probability in perceptual estimation tasks, with the main finding that only orientation probability affected perceptual precision. This idea was expanded in 2-alternative forced choice (2AFC) tasks in Chapter 4, which demonstrated clear differences between feature (orientation/color) probability and spatial probability both in *when* the effects occur and *what* the effects are. These chapters suggest that the mechanisms involved with feature probabilities are necessarily distinct from those involved in spatial manipulations, such as neural gain (Carrasco, 2011; also see *Figure 1.1c* for an illustration).

With alternative mechanisms ruled out, Chapters 5 and 6 sought to provide evidence for the tuning hypothesis. Chapter 5 introduced an electrophysiological investigation to verify that early V1 activity can be modulated by feature probability. The direction of the modulation was also consistent with the tuning hypothesis, but inconsistent with results looking at the same components with 'attentional' manipulations. Chapter 6 used computational modelling to bridge between the neurophysiological and behavioural data. Not only is it demonstrated that probability effects can be produced with simple recurrent cortical circuits within V1 without additional assumptions on the role of ‘attention’, but it is also demonstrated that this perceptual model accurately *predicts* new behavioral effects. Further demonstrating that probability effects can develop passively, Chapter 7 showed that probability effects develop without the need to attend to the relevant feature.

While varying in their approach, all the studies presented in this thesis support the tuning hypothesis. One important implication is that probability mechanisms *can* be independent of mechanisms involved in ‘attention’, and likely have distinct neural signatures as well, *depending on how ‘attention’ is understood.* For example, although both effects look similar on the surface, distinctions can be drawn between ‘top-down’ attention and ‘bottom-up’ expectations (Summerfield & Egner, 2009), and we should not equate one with the other. Chapter 2 replicates and contrasts the effects from two separate past studies. It will also directly highlight why it is more fruitful to think about the neural mechanisms involved than to have a semantic argument about whether a manipulation is ‘attentional’ or not.

Chapter 2: Contrasting Orientation Probability and Spatial Exogenous Cuing[3]

2.1. Introduction

Probable stimuli are detected faster and more accurately than non-probable stimuli. These behavioural outcomes of probability manipulations resemble those of typical ‘attentional’ manipulation. With exogenous spatial cuing (e.g. Posner & Cohen, 1984), cued targets are reacted to more quickly and are reported more accurately than non-cued targets. As such, probability and spatial cuing might be suspected to reflect a common ‘attentional‘ locus (e.g. Hon & Tan, 2013). However, that might not say much about the actual mechanisms involved as ‘attention’ refers to a class of effects, not a single unitary construct (Anderson, 2011).

While attentional cues and probability biases both result in response facilitation, there are differences. For example, Wyart, Nobre, & Summerfield (2012) and Cheadle, Egner, Wyart, Wu, & Summerfield (2015) showed that they could distinguish probability-cuing of target identity from trial-by-trial attentional cuing when using a reverse correlation procedure. One possibly important methodological detail is that in these studies the probability information was explicitly provided to participants, as opposed to previous orientation probability experiments (e.g. Anderson, 2014; Jabar & Anderson, 2015) where participants were naive to the manipulation and remained so throughout the experiment, as verified by post-experiment questionnaires. It could be that on-line, implicit probability learning leads to faster and more precise estimates of the orientation of probable stimuli by the same mechanisms as spatial cuing (Anderson & Druker, 2013), but that additional and distinct mechanisms contribute to behavioral benefits when information is explicitly provided.

---

[3] A version of this chapter has been published as Jabar, S. B. & Anderson, B. (2017). Orientation probability and spatial exogenous cuing improve perceptual precision and response speed by different mechanisms. *Frontiers in psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00183

On the other hand, it could also be the case that while the effects of probability and cuing are similar, they arise due to different mechanisms. What might these mechanisms be? In Chapter 1, it was hypothesized that exposure to orientation probability results in neural tuning, i.e., changes in the range of orientations to which a neuron responds (refer to *Figure 1.1*). Tuning changes occur when features are cued (Çukur, Nishimoto, Huth, & Gallant, 2013; David, Hayden, Mazer, & Gallant, 2008; Ling, Jehee, & Pestilli, 2015; Paltoglou & Neri, 2012). Tuning changes also occur when there is no explicit cue. Ringach, Hawken & Shapley (1997) and Schoups et al. (2001) showed that when monkeys receive orientation training, the tuning width of V1 neurons selective for the trained orientations are ‘sharpened’. This is consistent with the suggestion that learnt likelihoods are reflected in the early phase of sensory processing (Summerfield & Egner, 2009).

Although V1 neural responses to the subsequent target are also modulated by exogenous spatial cues (Wang, Chen, Yan, Zhaoping, & Li, 2015), spatial manipulations are thought to recruit a gain rather than a tuning mechanism (Carrasco, 2011; Ling, Liu, & Carrasco, 2009; *Figure 1.1*). Spatial manipulations are necessarily non-selective for features. Unlike with orientation probability (Ringach, Hawken & Shapley, 1997) or feature-based cues (Martinez-Trujillo & Treue, 2004) where the feature-relevant neurons are facilitated, spatial cues provide no information about the features of the upcoming target stimuli.

Concurrently examining the effects of orientation probability *and* spatial exogenous cuing within a single perceptual estimation task would provide us with information as to the possible mechanisms involved. The current study therefore combined the designs of Anderson (2014) and Anderson & Druker (2013). If the two manipulations interact, as is the case with attentional and memory manipulations (Haskell & Anderson, 2015; Liu & Becker, 2013), they likely share the same mechanism. If they create additive, independent effects (Sternberg, 1969), that provides support for the idea that underlying mechanisms of orientation probability and spatial exogenous cuing are separable.

2.2. Experiment 2.1: Methods

Participants

For all the studies reported in this thesis, participants were recruited from the University of Waterloo, approved by the University's Office of Research Ethics. Each participant only took part in one experiment. These participants took part in exchange for course credits, unless otherwise specified (e.g. Chapter 5).

In this study twenty participants (*median* age = 19 years) were recruited (16 females, 4 males), in exchange for course credits. Nineteen reported themselves right-handed. All participants had normal or corrected-to-normal vision.

Stimuli

Gabors, oriented grayscale sine-wave gratings with a circular Gaussian mask (*Figure 2.1a*), were used as the target stimuli in this experiment[4]. These Gabors were shown at 50% maximum contrast with an average measured luminance of 39cd/m^2. The Gabors had a spatial frequency of 4 cycles per degree of visual angle, and were presented on a gray background with a similar luminance of 40cd/m^2. When viewed from a distance of 60cm, the Gabors subtended approximately 4 degrees of visual angle both vertically and horizontally. Gabors appeared either left or right of the black fixation cross, with a distance of 4 degrees from the center of the cross to the center of the Gabor. Lines, used as feedback and for participants to report their estimations, had a length of 4 visual degrees and always

---

[4] Experiments in this thesis were programmed in *Python* (version 2), using the *PsychoPy* library (Peirce, 2008), on a Linux system. Luminance was measured using a ColorCAL MKII colorimeter. All the programming and data analyses were done by me.

occurred in the same location as the Gabor for that trial. Response lines always started with a horizontal orientation.

Gabors were equally likely to appear at each location. Only one Gabor was presented on each trial. *Collapsed across these two locations*, all orientations were equally likely. The critical manipulation was the occurrence-rate of the *probability-location conjunctions*. Half the participants saw the conjunction depicted in *Figure 2.1b*: When a Gabor appeared on their left, its orientation was more likely to be left-tilting, and this was reversed on the right. Tilts were uniformly distributed across each quadrant. High-probability orientations accounted for 80% of the trials. The lines in *Figure 2.1b* depict the distribution observed by the first participant. The location-orientation conjunctions were counterbalanced across participants. The use of color in the figures is illustrative, blue is used for for high-probability trials and plots and red for low-probability. This color scheme is consistent across *all* experiments in this thesis.

Probability distributions were maintained throughout the experiment. For every set of 20 trials, there were 8 left-tilting Gabors on the left, 8 right-tilting Gabor on the right, etc. Participants were *not* informed about these probability distributions. Practice trials had uniformly distributed orientations. Spatial exogenous cuing was done by having a white (72cd/m^2) box surround the Gabor location. These cues were uninformative (50% congruent with Gabor location, 50% incongruent), and there was always one cue on each trial. A gamma-corrected CRT monitor that refreshed at 89Hz was used, and stimulus timings were programmed as numbers of frames (*mean* refresh = 11.27ms, *SD* = 0.07ms).

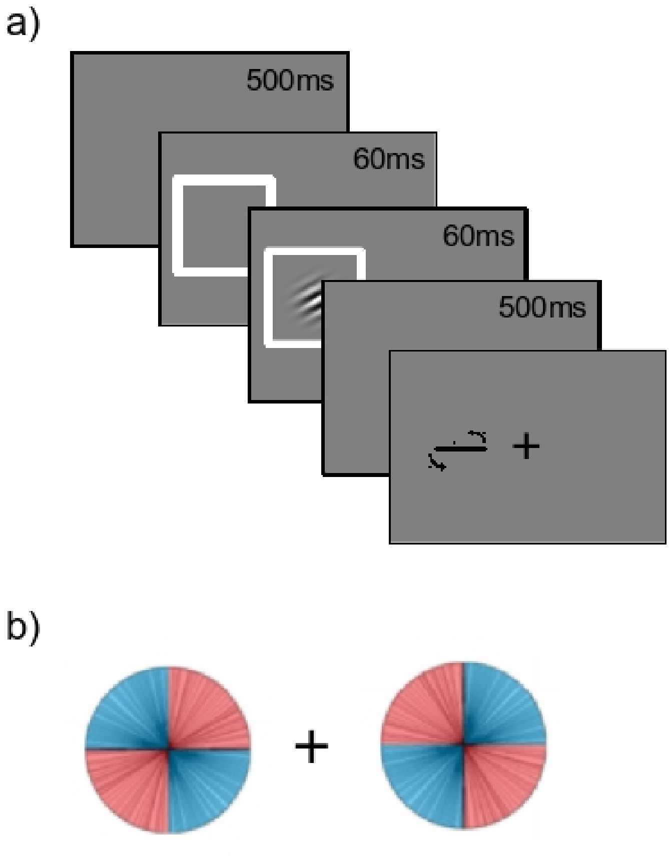

Figure 2.1. Experiment paradigm. a) Participants fixated onscreen for 500ms. The left or right location was spatially cued for 60ms, then a Gabor appeared in one of the two locations. The (uninformative) cue and Gabor offset following an additional 60ms. After a delay of 500ms, a response line appeared for participants to make their estimations. b) Gabor orientations were manipulated based on location. For example, when the Gabor appeared on the left, it was more likely (80%) to be left-tilting. This was reversed on the right. The location-orientation conjunction was counterbalanced across participants.

Auditory feedback was given after each response to maintain motivation. A high pitched sound indicated an error of less than 12 degrees. A lower pitch indicated an error greater than 12 degrees. Participants were not informed of the exact error threshold.

Procedure

Prior to the task, participants were instructed to make their estimations of the Gabor orientations as accurately as possible. They were *not* told that they needed to be fast. Responses were made with a QWERTY keyboard using their dominant hand. Participants were also instructed to fixate on the central fixation symbol at the start of each trial. This fixation phase lasted 500ms. A spatial cue then appeared at either the left or right location for approximately 60ms (5 screen refreshes), after which a Gabor appeared. The cue and the Gabor disappeared together approximately 60ms later (again 5 frames). After a delay of 500ms a response line appeared. Participants made their orientation estimations by rotating this line counter-clockwise or clockwise by pressing “Z” or “C” on the keyboard. This rotation was at a maximum of 1 angular degree per frame refresh of the monitor. Participants pressed the “X” key to confirm their estimations. The auditory feedback was then given. There were 40 practice trials where a white feedback line with the actual orientation was displayed after the participant’s response. For the main 400 trials, visual feedback was not given. The task was separated into two blocks, with a break in-between. A post-experiment open-ended questionnaire was given to each participant that probed with increasing specificity whether they had been aware of or could report the probability by location manipulation.

Post-experiment Questionnaire

The questionnaire consisting of the following four open-ended questions:

1. Did anything about the experimental task stand out to you?

2. Please describe any strategies you may have used.
3. Did you feel that you perceived some stimuli better or differently than others, or in certain cases? Did you notice any change over time in your experience?
4. Do you think that some orientations are more likely at certain times? If yes, please elaborate.

Analysis

All analyses, unless otherwise stated (e.g. *Chapter 5*), were done using the *R* statistical software package. Bayesian inference testing was conducted using the BayesFactor R package (Morey, Rouder, & Jamil, 2015). Initiation time (IT) was taken as the time from the appearance of the response line to when the participants' first adjusted the orientation of response line. This measure was used because it has proven in the past to be robust (Jabar & Anderson, 2015), and since it is uncontaminated by the amount of movement required to report the orientation that is a function of the angle of the stimulus. Vacillations, the number of times that participants changed directions when generating their reports, were also analyzed. Vacillations might be linked to the confidence of the decision, which could be post-decisional (e.g. Hilgenstock, Weiss, & Witte, 2014), and is not tightly coupled to perceptual precision (Jabar & Anderson, 2015; Samaha, Barrett, Sheldon, LaRocque, & Postle, 2016).

Angular error was measured as the difference (in degrees) between what was presented and what was estimated. The possible range of angular errors occupied an axial (half-circular) space ranging from -90 (anti-clockwise error) to +89 degrees (clockwise error), after which the report wraps back as an an error of -90 degrees. The median of the absolute magnitude of these errors is the primary dependent variable (see Gorard, 2005; Prinzmetal, Amiri, Allen, & Edwards, 1998).

The kurtosis of the (signed) angular error distribution was also looked at. The kurtosis of a distribution is a function of the fourth moment of the data (DeCarlo, 1997). By definition, normal distributions have zero *excess*[5] kurtosis. In previous studies, orientation probability affected kurtosis, leading to non-normal error distributions (see Anderson, 2014; Jabar & Anderson, 2015). Higher probability orientations shift the weight from the 'shoulders' of the distribution to the 'tails' and/or the 'peak' of the distribution. This results in a higher computed kurtosis (*Figure 2.2*). Changes in kurtosis therefore reflect changes in the shape of the error distribution, with different kurtoses suggestive of different types of perceptual errors. For example, if an effect only works to reduce the number of small errors, we should see the shoulders move into the peaks, while the tails (large errors) might stay relatively intact. This should reduce both median absolute error and increase kurtosis. However, if an effect acts uniformly across the whole range of errors, we expect a reduction in the median absolute angular error without a change in the 'shape' (kurtosis) of the distribution.

---

[5] Mathematically, $Kurtosis = \frac{\sum_i (X_i - \bar{X})^4 / n}{\left(\sum_i (X_i - \bar{X})^2 / n\right)^2}$,
where n is the number of samples in the distribution, X_i are the individual observations, and X is the sample mean (DeCarlo, 1997). Because normal (Gaussian) distributions always have a kurtosis of three, *excess* kurtosis is taken as *kurtosis – 3*. This thesis uses the *excess* measure.

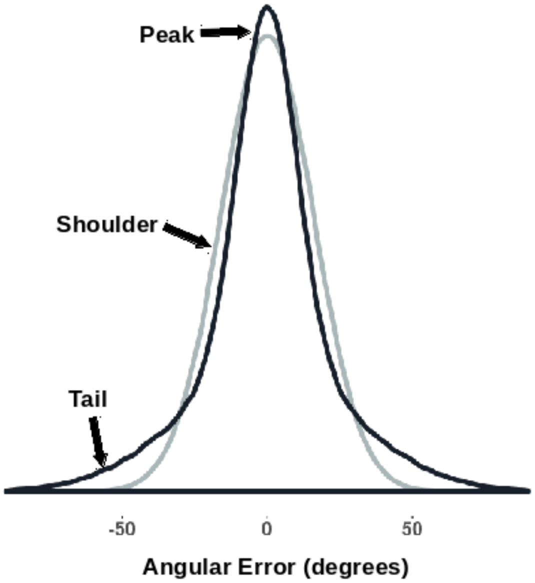

Figure 2.2. Kurtosis as a description of the shape of a distribution. The two generated error distributions have the same median absolute value (10.8 degrees) while having significantly different *excess* kurtoses (darker curve = 1.85; lighter (normal) curve = 0.0). Note how the distribution with the increased kurtosis has a higher peak, narrower shoulders, and broader tails.

2.3. Experiment 2.1: Results

Data from one participant was removed because the participant failed to respond to most trials. For the remaining nineteen participants, two-way repeated measure ANOVAs were computed for median angular error, angular error kurtosis, mean vacillations, and median ITs. Across the thesis, alpha cut-offs were taken at p=.05. A summary of the raw results are shown in *Table 2.1*.

Table 2.1. Summary of Means

	Spatial Cuing			
	Congruent		Incongruent	
Orientation Probability	High: 80%	Low: 20%	High: 80%	Low: 20%
Vacillations	0.093 (0.071)	0.105 (0.078)	0.088 (0.073)	0.178 (0.091)
Initiation Time (ms)	189 (116)	243 (120)	216 (127)	278 (125)
Median Angular Error (deg)	8.02 (2.49)	8.74 (2.69)	8.63 (2.09)	9.53 (3.00)
Kurtosis	0.84 (1.25)	-0.08 (0.72)	1.61 (1.96)	-0.18 (1.01)

Note. Values given are means across participants (± 1 SD)

Vacillations and initiation times

There was a significant main effect of orientation probability on vacillations ($F(1,18) = 9.73$, $MSE < 0.01$, $p = .006$). Lower probability led to more vacillations. There was also a significant main effect of spatial cuing ($F(1,18) = 13.19$, $MSE < 0.01$, $p = .002$). Invalid spatial cues led to more vacillations. A significant interaction effect (*see Figure 2.3*) was also noted ($F(1,18) = 21.48$, $MSE < 0.01$, $p = .001$). Post-hoc tests suggested no effect of spatial cuing on vacillations for high probability orientations ($t(18) = 0.64$, $p = .532$), but there was one with low-probability orientations ($t(18) = 4.70$, $p < .001$). Chiefly, congruently cued trials ($M = 0.11$) had less vacillations than incongruent ones ($M = 0.18$).

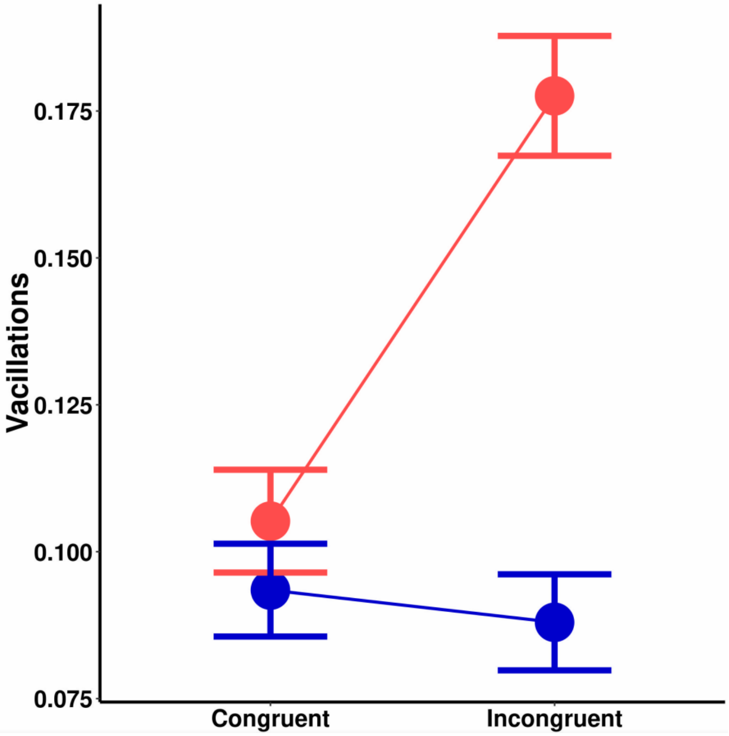

Figure 2.3. Effects on vacillations (change in the direction of motion) as a function of spatial cuing (x-axis) and orientation probability (blue = high-probability, red = low-probability).

A significant main effect of spatial cuing on IT, (*F*(1,18) = 19.62, *MSE* = 938, *p* < .001), a significant main effect of orientation probability (*F*(1,18) = 73.48, *MSE* = 855, *p* < .001), and no interaction (*F*(1,18) = 0.6, *MSE* = 578, *p* = .433) was observed (*Figure 2.4*). There was no effect of cuing on the size of the orientation probability effect (*t*(18) = 0.80, *p* = .433). Bayesian hypothesis testing returned a Bayes Factor (*BF*) of 0.32 (±0.01%), which is moderate evidence in favour of the *null* hypothesis that the effects of spatial cuing and stimulus probability are independent. The same trends remained even if trials where participants had vacillated in making their responses were removed.

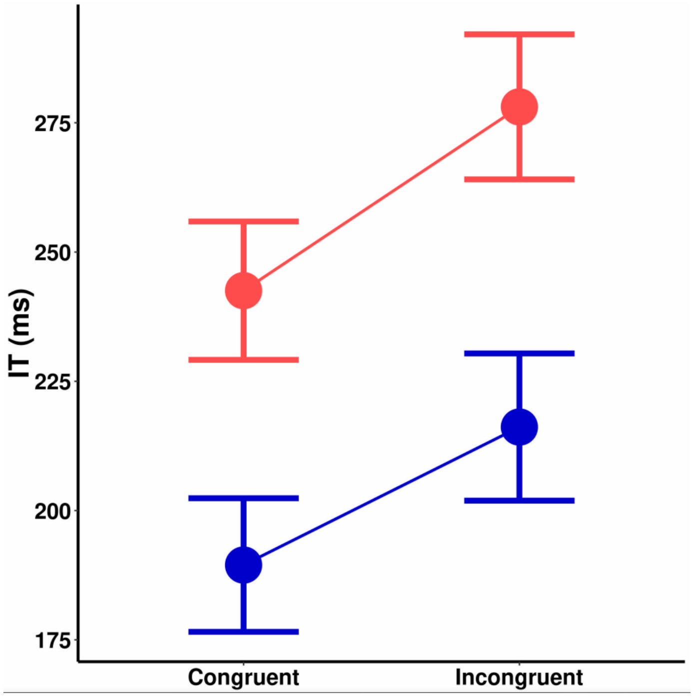

Figure 2.4. Initiation time (ms) as a function of spatial cuing and orientation probability (blue = high-probability, red = low-probability).

Median angular errors and kurtosis

The trends are identical for angular errors (*Figure 2.5b*). There was a significant main effect of spatial cuing on angular precision, ($F(1,18) = 5.2$, $MSE = 1.8$, $p = .035$), a significant main effect of orientation probability ($F(1,18) = 4.9$, $MSE = 2.5$, $p = .039$), and no significant interaction ($F(1,18) = 0.1$, $MSE = 1.5$, $p = .744$). As before, there was no effect of cuing on the size of the orientation probability effect ($t(18) = 0.33$, $p = .744$), with a corresponding *BF* of 0.25 (±0.02%). This is moderate evidence in favour of the null hypothesis. In short, while both spatial cuing *and* orientation probability significantly impacted ITs *and* perceptual precision, these effects were strictly additive.

For kurtosis, there was a significant main effect of orientation probability ($F(1,18) = 19.9$, $MSE = 1.7$, $p < .001$) without an effect for spatial cuing ($F(1,18) = 0.3$, $MSE = 1.6$, $p = .260$), nor a significant interaction effect ($F(1,18) = 2.2$, $MSE = 1.6$, $p = .155$). Trials that were spatially cued congruent and incongruent to the Gabor location were compared separately for each probability condition, with Bayesian testing again revealing evidence for the null hypothesis (high-probability *BF* = 0.60 (±1%); low-probability *BF* = 0.25 (±2%)). In contrast, high and low-probability trials were compared separately for both congruent and incongruently cued trials, (congruent *BF* = 7.0 (±1%); incongruent *BF* = 15.4 (±1%)): Strong evidence in favour of the alternative hypothesis. These results suggest that while orientation probability causes a change in the shape (kurtosis) of the error distributions, spatial cuing does not.

One potential issue is that there were four times as many high probability trials as there were low-probability ones. As with previous studies (e.g. Jabar & Anderson, 2015), sub-sampling and bootstrapping with matched numbers of trials was conducted, yet the kurtosis effects remained (ps < . 05). Kurtosis effects by orientation probability were not an artifact of an uneven number of trials.

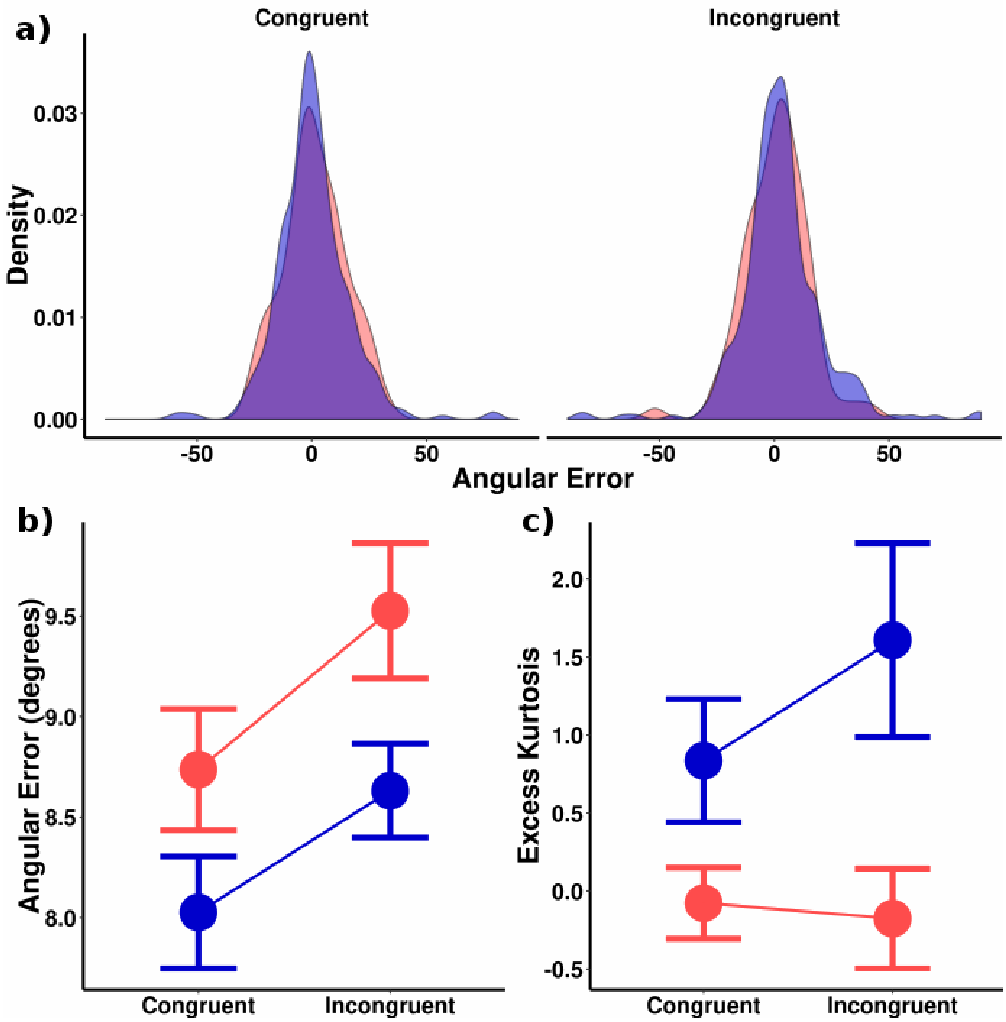

Figure 2.5. Measurements of perceptual precision. a) Sample distribution of signed angular errors for one participant. Left plot: Congruently- cued locations. Right plot: Incongruently-cued locations. b) Median angular error (degrees) as a function of spatial cuing (x-axis) and orientation probability. c) Kurtosis as a function of spatial cuing (x-axis) and orientation probability. (Blue = high-probability, red = low-probability).

Post-experiment questionnaire

For the post-experiment questionnaire, no participant reported that probability was being manipulated, even when directly asked “Do you think that some orientations are more likely at certain times? If yes, please elaborate.” This was the same as our previous experiments using these probabilities for orientation estimation tasks. Some participants volunteered that they found the task more difficult when ‘incorrectly’ cued.

2.4. Experiment 2.1: Discussion

Probability and spatial cuing manipulations generally result in similar behavioural effects, even within orientation estimation tasks: Speed and accuracy are facilitated (Anderson, 2014; Jabar & Anderson, 2015; Anderson & Druker, 2013). As both these effects resemble what might traditionally be labelled as ‘attentional’, one could surmise a common mechanism.

Both orientation probability and spatial cuing were conjointly manipulated within a single perceptual estimation task. Effects from previous reports where either one or the other manipulation was done were replicated. Both manipulations improved initiation times *and* perceptual precision. However, the manipulations act independently: There were *additive* effects in *both* initiation times and perceptual precision. If these two manipulations share the same ‘attentional’ mechanism, one would expect only high-probability orientations *in the cued space* to benefit. Instead, the size of the orientation probability effect was unaffected by the spatial cue. Further, spatial cuing did not affect the shape (kurtosis) of the angular error distribution while orientation probability did (and as has been shown previously Anderson, 2014; Jabar & Anderson, 2015). In addition, a report that predated our discovery of kurtosis effects (Anderson & Druker, 2013) had used spatial cues. Although that study did not look at kurtosis, a simple reanalysis of that data (not reported here) suggests that spatial cuing alone does not affect kurtosis. This is consistent with the findings from *Experiment 2.1*.

Had the the same perceptual mechanisms been invoked by both spatial cuing and orientation probability, there should have been common effects on error distribution shape. Instead the results suggest that spatial cuing and orientation probability affect perceptual precision and speed of response via different mechanisms. This is concordant with the argument made in the introduction, that while orientation probability affects precision through selectivity tuning neurons preferring the probable orientations, exogenous spatial cuing recruits a non-selective gain process. This gain process might increase apparent contrast either by increasing the input baseline of neural responses (Cutrone, Heeger, & Carrasco, 2014) or by speeding up information accrual (Carrasco, Giordano, & McElree, 2004; Carrasco & McElree, 2001).

If the two cue-types modulate perceptual precision through separate perceptual mechanisms, how can we understand the interaction on vacillations? It might be important to note that there could be multiple processing stages (Sternberg, 1969) involved in perceptual discrimination (e.g. Sun & Landy, 2016), and what we might be observing are additive effects of exogenous spatial cuing and orientation probability in the *stimulus processing stage*, but interactions in the downstream response or *decision-making stage*. If vacillations are indicative of the confidence associated with the decision-making process (e.g. Petrusic & Baranski, 2009), then they are not necessarily tied to the quality of perceptual processing. Precision-based probability effects were seen even across trials equated for self-reported confidence (Jabar & Anderson, 2015). Perhaps in the congruently-cued case, participants were explicitly expecting targets in the ‘cued’ location, which reduced the impact that probability had on their confidence? It could be that *repeats* in orientation are driving confidence more than probability *per se*. It is important that future studies seeking to understand the differences between the two mechanisms not simply equate perception with detection, since the act of decision-making involves additional factors.

If orientation probability effects are distinct from other attentional manipulations, should they still be called ‘attentional’? If attention is understood as a class of *effects* (Anderson, 2011), then that label still applies. Where confusion arises is when attention is understood as a general *causal* mechanism. Orientation probability mechanisms are distinct from spatial exogenous cuing mechanisms. They lead to independent effects, and we suggest that this is due to orientation probability causing neural tuning differences. Whether that also applies for other attentional manipulations is an open question. It is worth keeping in mind that effects that look behaviorally similar are not necessarily indicative of common neural and psychological mechanisms.

Still, orientation probability differs from spatial cuing across multiple dimensions. The latter is explicit and occurs trial-by-trial, whereas the former is implicit and necessarily has to be learnt via an accumulation of experience. The information they provide is also different: One pertains to spatial information, while the other pertains to feature-based information. If instead, feature probability was contrasted with spatial probability, would there be dissociable behavioural effects? Since distinctions have been drawn between spatial and feature-based attention (Carrasco, 2011), a hypothesis is that this distinction might hold true for probability as well, and this was explored in Chapter 3.

Chapter 3: Feature and Spatial Probabilities are not Equivalent[6]

3.1. Introduction

Chapter 2 demonstrated that probability effects can be independent of attentional mechanisms. If probability effects are simply due to participants ‘attending’ to the probable object, we might expect that when they learn about the probability of a target, all aspects of their perceptions and judgments regarding that target are affected similarly. For example, one should get better at both perceiving the orientation of the probable object, and at spatially locating that object. However, one common dissociation in attentional paradigms is between spatial and feature-based attention (Carrasco, 2011; Fink, Dolan, Halligan, Marshall, & Frith, 1997; Ling et al., 2015; Yantis & Serences, 2003), such as having different time courses (Liu, Stevens, & Carrasco, 2007). Given the dissociation between feature and spatial-based attention, and that feature probability and spatial attention are seemingly independent, one could extend the argument and hypothesize that feature and spatial-based probability effects should also be distinct.

On the surface, performance benefits from manipulations of spatial probability mirror that of orientation probability. When the probability that a target will appear at a particular location is adjusted, reaction times change (Gekas, Seitz, & Seriès, 2015; Geng & Behrmann, 2005; Vincent, 2011; Walthew & Gilchrist, 2006). Jiang, Sha, & Remington (2015) used a T-among-L search display with the target appearing 3 times more likely in one particular quadrant. The targets in the probable location were detected significantly faster. Spatial probability effects are also obtainable when the probability distribution is made more complex, and varies across space continuously (Druker & Anderson, 2010).

---

[6] A version of this chapter is published as Jabar, S. B. & Anderson, B. (2017). Not all probabilities are equivalent: Evidence from orientation versus spatial probability learning. *Journal of Experimental Psychology. Human Perception and Performance, 43*, 853-867. https://doi.org/10.1037/xhp0000354

However, while the hypothesis for orientation probability is that it affects neural tuning, spatial manipulations have been suggested to be gain-based (Carrasco, 2011). If this is true, orientation and spatial probability might lead to dissociable outcomes, since it has been suggested that optimal tuning functions of sensory neurons are context and environmental-specific rather than being universal (e.g. Dean, Harper, & McAlpine, 2005; Hosoya, Baccus, & Meister, 2005), and therefore that adaptive-tuning (e.g. *Figure 1.1b*) is *required* for optimal performance of the sensory system (Yaeli & Meir, 2010). If feature probability affects V1 tuning, then it potentially explains why orientation precision could be improved by orientation probability (Anderson, 2014; Jabar & Anderson, 2015): The probable (trained) orientations are simply better perceptually encoded, due to the relevant neurons now being more optimally tuned (Summerfield & Egner, 2009). By contrast, an additive gain without associated tuning (e.g. *Figure 1.1c*) might not lead to optimal perceptual encoding since the signal-to-noise ratio remains constant. If spatial probability is gain-based, orientation precision should remain unaffected across high and low-probability *locales*. However, the ability to *detect* target occurrences might improve, since average firing of the neural population would be increased. While spatial exogenous cuing apparently also increases orientation precision, this might be because the presence of an explicit cue increased V1 firing of neurons selective to the space *prior* to the target onset (Wang et al., 2015).

To reiterate, the crux of this chapter is to understand how manipulations of spatial and orientation probabilities affect their *joint* reporting, measured by both speed and accuracy. *Experiment 3.1* begins by testing if previous findings that orientation probability increases orientation precision are robust to the change in procedure that required each participant to report location and orientation on every trial, and where participants used a mouse instead of a computer keyboard (Anderson, 2014; Jabar & Anderson, 2015; *Figure 2.1*). *Experiment 3.2* was similar, but manipulated spatial probability.

Experiment 3.3 was a defence against ceiling effects for spatial localization, and used a preliminary staircase procedure to find a stimulus contrast that made locating the stimulus challenging. The upshot of these experiments is primarily that the two probability manipulations decompose, as they do for attention, and that spatial probability affects primarily target *detection* while orientation probability primarily mediates changes in discrimination *precision*.

3.2. Experiment 3.1: Methods

Experiment 1 served to test if the probability effects on orientation judgments reported in Jabar & Anderson (2015) and in Chapter 2 could be replicated when the response method was changed. Rather than the keyboard based response used before, a novel mouse-based paradigm was developed. Participants reported both a location and an orientation on every trial using a mouse. After a brief presentation of a single oriented grating, participants first localized where they perceived the center of the grating to have been by clicking on it with a computer mouse. They then moved the mouse away from this anchor point to ‘draw’ an orientation (finalized with another mouse click).

Participants

Twenty-one participants (13 females, 8 males) took part in the study. 18 participants were right-handed. The 3 left-handed participants were told that they could move the mouse to their left hand if they wanted, but they opted to use their right hand. All participants had normal or corrected-to-normal vision, and did not declare any auditory deficits.

Stimuli

The centre of the display was marked by a black fixation symbol subtending approximately 1.2 degrees in both axes. Spatial gratings were presented to participants on each trial. Unlike Chapter 2, these were oriented grayscale sine-wave gratings with a circular *hard-edge*[7] mask (*Figure 3.1a*), with an average measured luminance of 39 cd/m^2. They had a spatial frequency of 3.16 cycles per degree of visual angle, and were presented on a grey background with a similar luminance of 40 cd/m^2. When viewed from a distance of 60 cm, the gratings subtended approximately 2 and a half degrees of visual angle. The center of the grating was displaced 4 degrees either to the left or right of the centre of the display and four degrees up or down, thus they fell at the corners of a virtual square. The center of the grating was marked with a small black dot to aid participants in their location reports. Lines, used as feedback, always occurred in the same location as the grating for that trial and matched its diameter.

Because the effects of repetition are a possible confound when examining probability effects (e.g. Hale, 1969), the occurrence-rate of the various *probability-location conjunctions* was manipulated. Collapsed across the four *equiprobable locations*, all orientations were equally likely. Repeats of orientation are therefore as likely to occur for 'low-probability tilts' as they are for 'high-probability' ones. Further, the possible orientations were continuous, reducing the odds of an exact repeat. Half the participants saw the conjunction depicted in *Figure 3.1d*: When a grating appeared on the top-left or bottom-left, its orientation was more likely to be left-tilting, but this high-probability tilt was reversed if the grating appeared on the top-right or bottom-right.

Probability distributions were maintained throughout the experiment. In every set of 20 trials, all 4 locations appeared 5 times each, and there were 4 occurrences of one set of orientations (high-probability: 80%), and 1 occurrence of the other set (low-probability: 20%). The lines in *Figure 3.1d* depict an example distribution. The location-orientation conjunctions were counterbalanced across participants. Participants were *not* informed about these probability distributions.

---

[7] Gabors have a Gaussian mask (edges are blurred out). When stimulus contrast is decreased, the edges are harder to perceive, making Gabors appear smaller. Since *Experiment 3.3* adjusted contrast, hard edges were for the gratings to minimize this apparent size change.

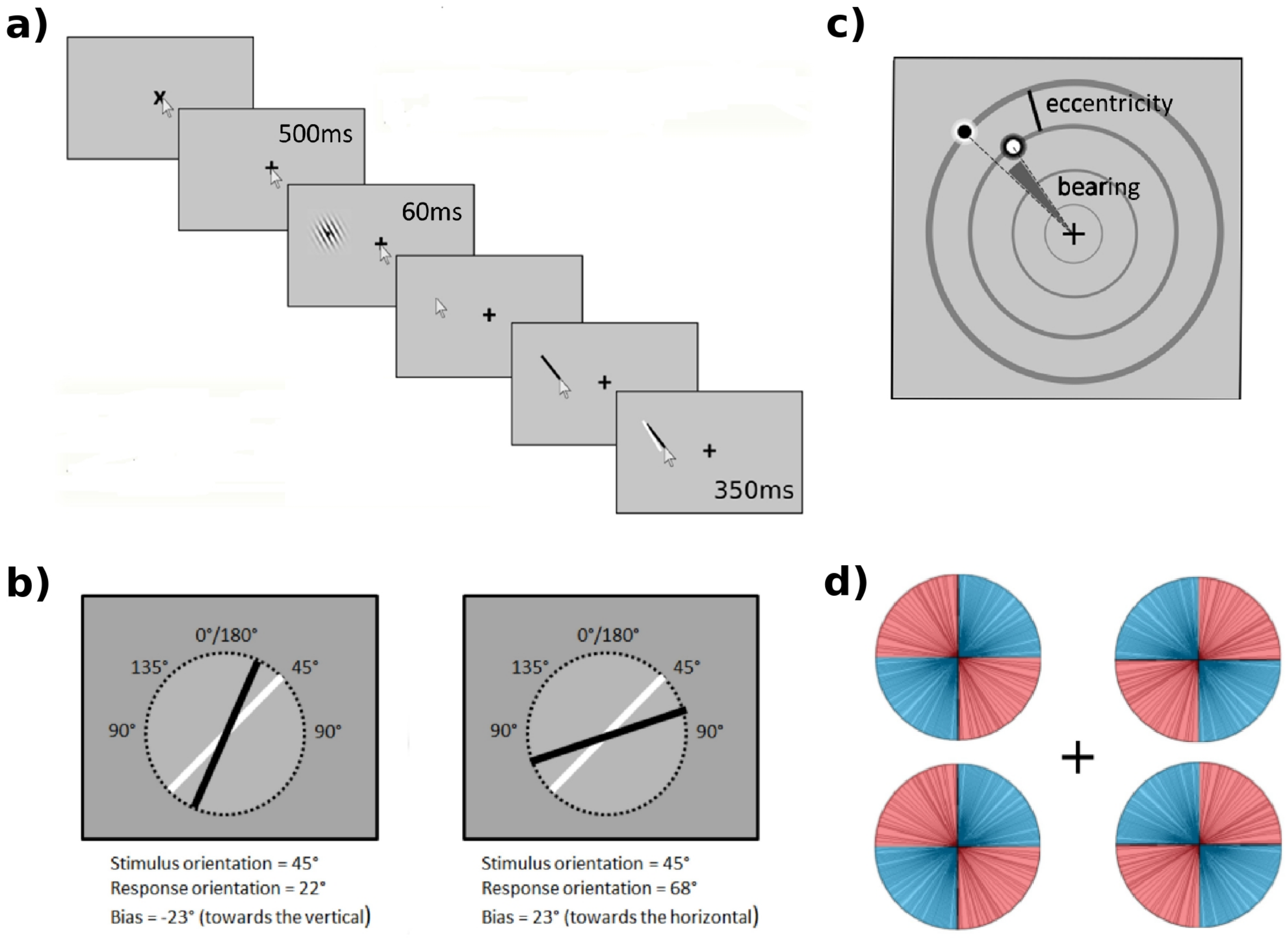

Figure 3.1. Paradigm and precision measurements. (a) General paradigm used for all experiments reported. Participants initiated trials by clicking on the central fixation symbol. After a delay of 500ms, a spatial grating (with a central dark spot) appeared for a brief 60ms. Participants then moved the mouse to click on where they thought the center of the grating was. After this click they moved the mouse a way to draw a line that best represented the orientation of the grating, then gave a second click to confirm the orientation estimate. Visual and auditory feedback was then provided. (b) Example of how orientation angular errors are calculated. White lines: Actual orientation. Black lines: Estimated orientation. Note that between the example trials, the bias will be zero, but the median absolute error would be 23 degrees. (c) Example of how spatial eccentricity and bearing errors are calculated. White marker: Actual location. Black marker: Estimated location. (d) Orientation probability distribution used in *Experiment 1*. Note all blue lines/symbols will correspond to high probability trials, red for low-probability trials.

Auditory feedback was given after each trial to maintain motivation. A high pitched sound indicated an orientation error of less than 12 degrees and a spatial (Euclidean) error within 0.5 degrees. The spatial error threshold was chosen after pilots tests, such that participants crossed the spatial threshold about as often as they did the orientation threshold (about 85% of the time). Any error beyond these cut-offs caused a lower pitch sound to be played. Participants were not explicitly informed of the error threshold.

Procedure

Participants sat approximately 60 cm from a 32 cm x 24 cm gamma-corrected CRT monitor that refreshed at 89 Hz. Responses were made with an optical mouse. The position of the on-screen mouse cursor was recorded every frame of the experiment. Participants were instructed to fixate at the center of the screen.

Prior to the task, participants were instructed to make their estimations of the location and then the orientation of the gratings as precisely as they could. They were *not* told that they needed to be fast. Participants were given 40 practice trials in which the orientations occurred uniformly. These data were not included in the analysis. The main task consisted of 400 trials, which were sectioned into two blocks. Participants were given the option to take a break in-between the blocks. At the end of the computerized task, participants were given a short, open-ended, questionnaire to examine whether they could explicitly report the spatial and orientation distribution of the gratings that they had seen. The experiment took approximately 20-25 minutes to complete.

Participants initiated each trial by clicking on the central fixation symbol, which changed from an “x” to “+” to indicate that the trial had started. Participants were shown the “+” fixation symbol for

500ms after this initial click. The grating then appeared in one of the four locations for 60ms (5 frame refreshes). Participants were then free to move the mouse. They first had to click at the center of the location where they perceived the grating to have appeared. After this, participants could move the mouse away from the location of the grating in any direction and a line would appear that followed the mouse and grew to the diameter of the grating. After the participant had drawn a line that reproduced the orientation of the grating, they made another click. The auditory feedback was then given, along with a white feedback line representing the actual grating displayed.

Post-experiment Questionnaire

After the 400 trials and before participants were debriefed, participants were given a short questionnaire. This was the same as the one given in *Experiment 2.1*, with the addition of a fifth question: *Do you think that some locations are more likely at certain times? If yes, please elaborate.*

Analysis Procedure

As with Chapter 2, angular errors were calculated for each trial in an axial space. These data were also used to get a measure of bias on the cardinal axes. Vertical-biased estimations, when participants estimated the orientation more vertically then it should have been, were coded as negative. Horizontal biases were coded as positive (*Figure 3.1b*).

Angular error analyses were done on both the *bias* and the *mean error* measures. The bias measurement gives the average of these signed errors across trials, such that a non-biased participant should approach a mean of ‘0’ bias. To look at the mean error made, the median of the *absolute values* of the estimation errors were taken. As an example of these calculations, there would be a ‘0 degree’ bias and ’23 degree’ magnitude of angular error across the trials depicted in *Figure 3.1b*.

Spatial error was broken down into the two polar axes: Eccentricity and bearing (*Figure 3.1c*). Since the targets were always equidistant from the fixation, people were highly precise for this measure, and no manipulation affected it. Therefore, spatial bearing error was the main measure of spatial precision, and was calculated in an analogous way to how orientation errors were measured. Spatial bearing errors occupied a circular instead of an axial space.

Reaction time (RT) for each trial was decomposed into three components. Initiation time (IT) was the time taken from the grating onset to when a participant first moved the mouse. The time between initiation and the click indicating location was labelled the *sRT*. Time from this click to the click indicating orientation was labelled the *oRT*. Total distance moved in both the spatial estimation and orientation estimation phases were also recorded.

Responses were quick and fluid, and appeared “automatic”. Therefore, trials in which participants reacted very slowly (IT > 1000ms: 0.2% of trials) were excluded. Also excluded were trials where the Euclidean distance between the response and the central fixation was less than 2 degrees (2% of trials). This was so that precision metrics did not include trials where participants either accidentally clicked again on the fixation symbol, or where participants purposely clicked the central location again because they did not know the actual location.

3.3. Experiment 3.1: Results

Effects of orientation probability on RT, on orientation precision, and on spatial precision were examined. Only slight effects on RT were observed. More importantly, orientation probability affected orientation precision, and these effects manifested quickly. Orientation probability did not affect spatial precision.

RT Analyses

The 400 trials were binned into 50-trial blocks to look at effects of time on the experiment. A two-way fully-within ANOVA was carried on these data. For the IT measure (*Figure 3.2b*), there was a main effect of probability, (*F*(1, 19) = 5.74, *MSE* = 381, *p* = .027), a main effect of trial bin, (*F*(7,139) = 18.86, *MSE* = 1330, *p* < .001), but no significant two-way interaction, (*F*(7, 139) = 0.50, *MSE* = 546, *p* = .830). In general, the difference in IT between high (*M* = 246ms) and low-probability (*M* = 249ms) was small, although quite consistent across the trial blocks. These initiations are not only fast, but also tended to be accurate, with 77% of them falling within a 45 degree bearing of the actual stimulus bearing (see *Figure 3.2a* for example mouse movements).

The same analysis on sRT (*Figure 3.2c*) suggested no main effect of probability, (*F*(1, 19) = 0.18, *MSE* = 14152, *p* = .680), no main effect of trial bin, (*F*(7,139) = 1.09, *MSE* = 48924, *p* = .376), and no significant two-way interaction, (*F*(7, 139) = 0.95, *MSE* = 27287, *p* = .473). The analysis on oRT (*Figure 3d*) suggested a marginal main effect of probability, (*F*(1, 19) = 4.04, *MSE* = 5348, *p* = .059), a main effect of trial bin, (*F*(7,139) = 16.73, *MSE* = 32568, *p* < .001), but no significant two-way interaction, (*F*(7, 139) = 0.39, *MSE* = 8903, *p* = .905).

In sum, the IT measures proved to be the most robust. Initiating movements to gratings with a probable orientation were made slightly faster. Drawing the orientation was also marginally faster. Orientation probability had no apparent effect on the time taken to complete the spatial localization post-initiation.

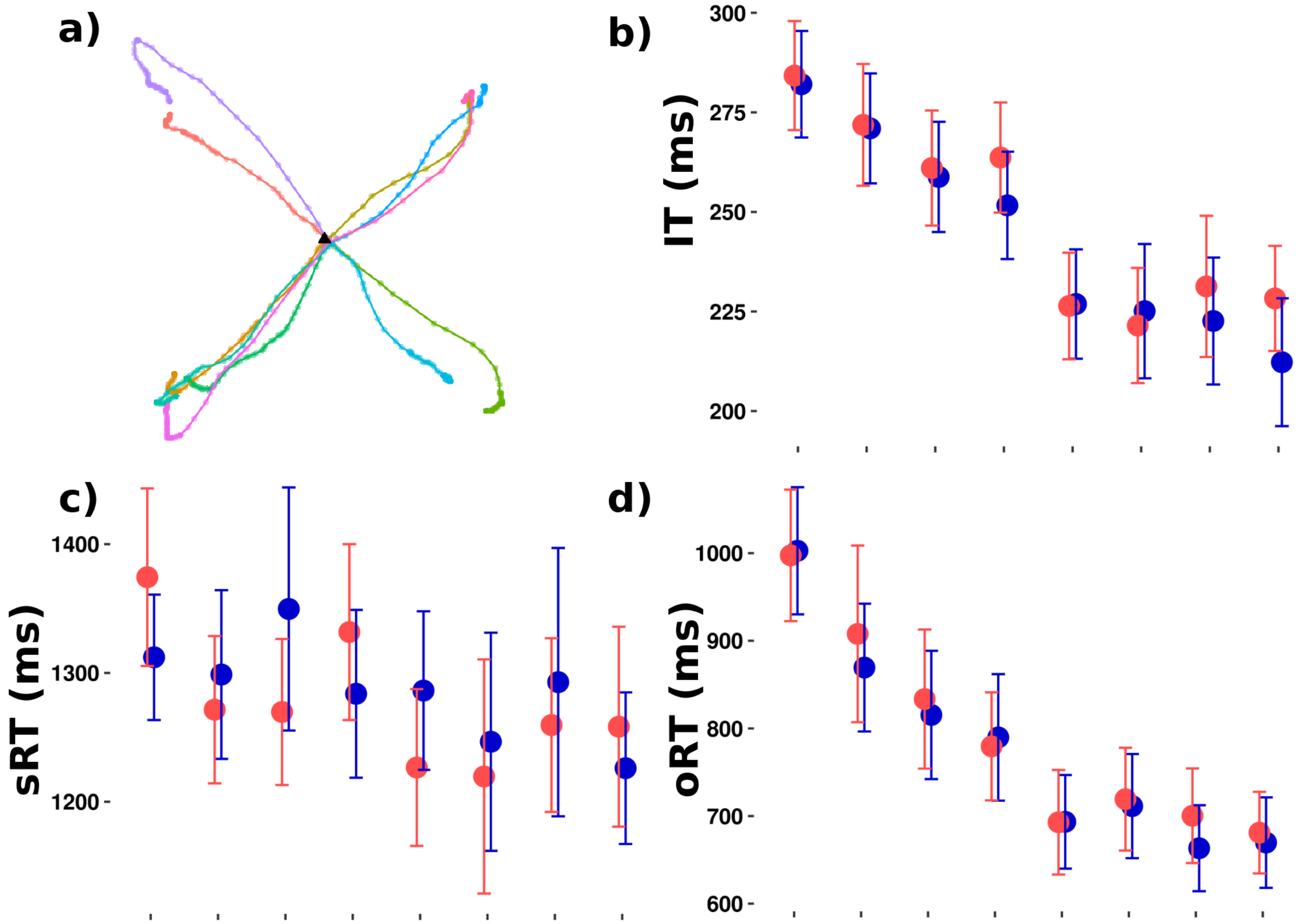

Figure 3.2. Reaction time data from *Experiment 3.1*. (a) Sample mouse movement during the initiation and spatial estimation phase. Note how initial movements from centre are ballistic, and then they slow down (spacing between markers indicates 1 frame: 11ms) as the mouse is adjusted towards the stimulus locations. (b) Initiation time (IT) across experimental trials (each bin represents 50 trials). (c) Time from initiation to spatial localization (sRT) across experimental trials. (d) Time from spatial localization to completion of orientation estimate (oRT). Note: Blue markers represent high-probability trials, red for low-probability.

Orientation (Angular) Error Analysis

Paired two-tailed t-tests were run on the bias measure. This revealed that the high-probability tilts were significantly vertically-biased ($t(20) = 2.47$, $p = .023$), whereas low-probability tilts were not, ($t(20) = 0.56$, $p = .58$). Compared against each other, there was a significant effect of probability on bias, ($t(20) = 3.44$, $p = .003$), with high-probability tilts being more vertically-biased ($M = -1.13$, $SD = 2.09$) than low-probability tilts ($M = 0.32$, $SD = 2.59$). Bias was not affected by stimulus location: Across possible comparisons, all ps> .05.

The error distributions from high versus low-probability trials also differ (*Figure 3.3a*): The median absolute error measure suggests a significant effect of probability, ($t(20) = 5.44$, $p < .001$), with high-probability tilts associated with an error of smaller magnitude ($M = 7.07$ deg, $SD = 1.99$ deg) than low-probability tilts ($M = 8.09$ deg, $SD = 1.73$ deg). Of the 21 participants, 19 showed this trend, with the 2 showing the opposite trend also being the worst performers (*Figure 3.3b*). This strong trend also remains if the error distributions are bias-corrected, ($t(20) = 6.03$, $p < .001$). Partitioning this data into 50-trial bins also suggest demonstrates that precision effects are observable even within the 1st bin, ($t(20) = 2.61$, $p = .017$). Although the significance fluctuates in later bins, high-probability orientations were always more precisely estimated (*Figure 3.3c*). Taken individually, this orientation probability effect on precision is also significant in all four locations (all ps < .05).

Orientation anisotropy analysis

Possible orientations were binned into 20-degree segments and angular error was looked at for each bin (*Figure 3.3d*). A two-way ANOVA revealed that there was a significant main effect of probability, ($F(1,20) = 24.39$, $MSE = 6.10$, $p < .001$). There was a main effect of orientation bin, ($F(8,160) = 3.59$, $MSE = 17.06$, $p < .001$). There was also a significant interaction, ($F(8,160) = 2.45$, *MSE* = 7.18, *p* = .016). Size of the orientation probability effect (low-high probability) in the bin containing the near verticals was greater than in the ones containing the 45-degree and 135-degree obliques (*p*s < .05), but not the one containing the near-horizontals (*p* >.05). The mean absolute angular error was smallest in the vertical bin as compared to the oblique and horizontal bins (*p*s < .05). In sum, probability effects were largest at the cardinals, but vertical orientations were most precisely represented overall.

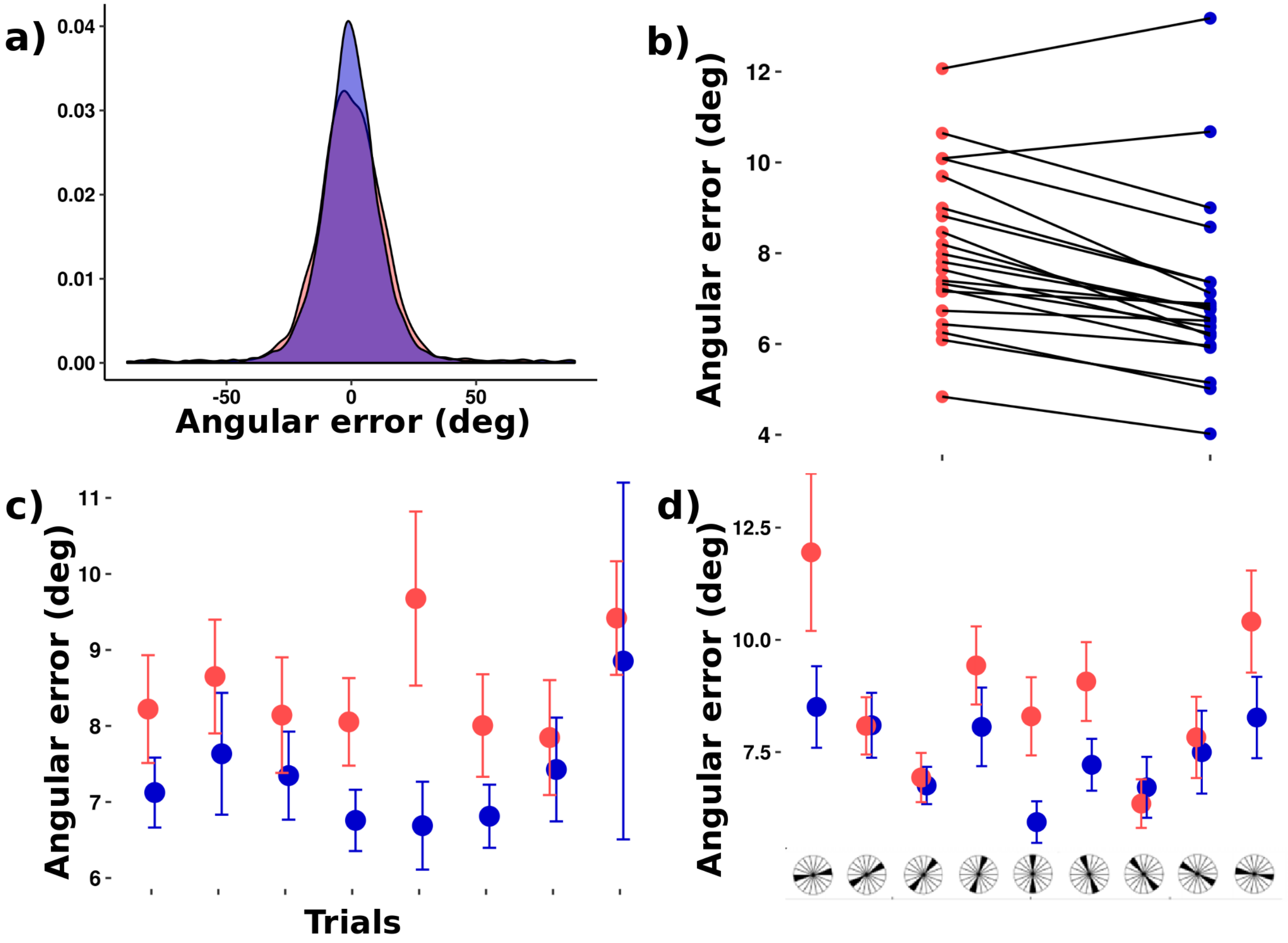

Figure 3.3. *Experiment 3.1* orientation precision. (a) Orientation error distribution. Negative indicates an anti-clockwise error, positive is a clockwise error. (b) By-subject median absolute errors across probability. Note that 19/21 cases show a decrease in error (increase in precision) for the high-probability trials. (c) Median absolute errors as a function of experiment trials. Each bin represents 50-trials. (d) Median absolute errors by orientation segments, indicated by figures on the x-axis. Note: Blue represent high-probability trials, red for low-probability.

Standard Deviation and Kurtosis

As mentioned in Chapter 2, a mixture of perceptual and non-perceptual effects of probability would lead to changes in error distribution shapes (e.g. *Figure 3.3a*) that could be quantified by the kurtosis.

While there is a trend for kurtosis to be higher for high-probability tilts (M = 3.73, SD = 4.37) than for low-probability tilts (M = 2.44, SD = 4.26), the difference is not significant ($t(20) = 1.33$, $p = .197$). However, the kurtoses are significantly above zero for high, ($t(20) = 3.91$, $p < .001$), and low probability orientations ($t(20) = 2.62$, $p = .016$). Since normal distributions have an excess kurtosis of zero, these data confirm that the error distributions are not normally distributed.

Repetition Effects analysis

Repetition effects are possible sources of confounds in probability-related studies since high-probability targets are more likely to be repeated (Hale, 1969), while rare targets are not. If repetitions or sequences are driving performance benefits for the high-probability orientation estimations (e.g. Fischer & Whitney, 2014), then we might observe a significant correlation between absolute inter-trial differences in orientation and estimation performance (small changes lead to small errors; large changes lead to large errors). In fact the correlation is negligible ($r < .01$, $p = .507$). Comparing the extremes, e.g. trials where the inter-trial difference was 10 degrees or less (mean angular error = 9.6 deg) versus those that were 80 degrees or larger (mean = 9.3 deg), also reveals no significant effect ($t < 1$, $p = .541$). Lastly, ignoring trials with small (10 degrees) inter-trial differences, the effect of orientation probability on orientation precision still holds ($t(20) = 3.25$, $p = .004$). The effect of repeating locations (approximately 25% of the trials) versus non repeats was not found to significantly affect either orientation precision or the size of the orientation probability effect (ps > .05).

Spatial error analysis

Participants demonstrated precise bearing reports: 77% of the estimates were made within 3 degrees of the actual bearing (*Figure 3.4a*). This spatial bearing error was not significantly different between high (M = 2.79, SD = 0.59) and low (M = 2.84, SD = 0.63) probability orientations, ($t(20)$ = 0.39, p = .702). These was also no significant linear correlation between bearing error and orientation error for either high probability trials (r = .41, p = .065) or low-probability trials (r = .08, p = .722).

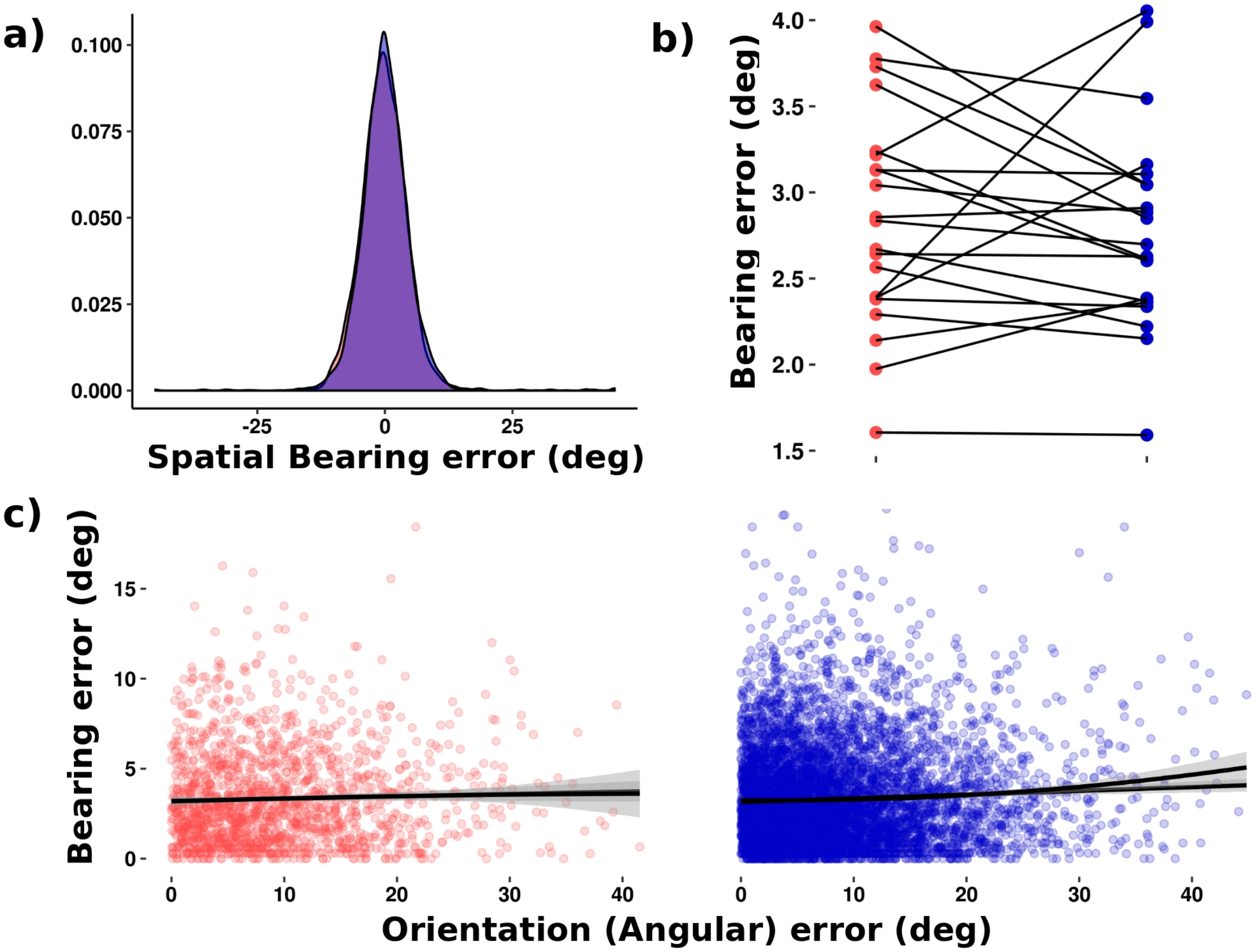

Figure 3.4. *Experiment 3.1* spatial precision. (a) Distribution of spatial bearing error. Negative indicates an anti-clockwise error, positive is a clockwise error. (b) Median absolute spatial bearing errors by-subject. (c) Scatter plot of spatial bearing error versus orientation angular errors. Straight lines indicate best fit line, the curved lines indicate the best fit quadratic. Note: Blue represent high-probability trials, red for low-probability.

As mentioned in the *Methods* section, because targets always appeared equidistant from the central fixation, the spatial eccentricity error was at floor (High-probability: 0.55 deg; Low-Probability: 0.54 deg). Neither this measure, nor the mean Euclidean distance between estimated and actual location, were significant across orientation probability (both *p*s > .05). As with orientation probability, repeating versus non-repeats of locations was not found to significantly affect spatial precision *(t*(20) = 0.51, *p*=.615).

Post experiment questionnaire

While a majority of participants demonstrated behavioural differences contingent on orientation probability, consistent with Chapter 2, *none* correctly indicated the orientation probability distribution.

3.4. Experiment 3.1: Discussion

Experiment 3.1 largely replicated previous findings. Probable orientations were estimated more precisely, despite participants showing no signs of explicit awareness of the distribution. The orientation anisotropy seen previously (Jabar & Anderson, 2015) was also replicated: Near-vertical tilts were most precisely represented, especially when they were probable (*Figure 3.3c*). It is important to note that in prior versions of this task, the response line was always presented at a static fixed orientation, but in this experiment there was no response line until participants initiated their response. Thus, these effects are not due to anchoring. As mentioned in Chapter 1, anisotropy represents an innate perceptual bias (Appelle, 1972), possibly due to V1 neurons' tuning (Li, Peterson & Freeman, 2003).

Compared with the previous paradigm using the keyboard responses (e.g. Jabar & Anderson, 2015; *Experiment 2.1*), participants in this experiment were more precise in their orientation estimation while the size of the orientation probability effect was similar. This is despite having to process the spatial dimension and report it *before* making the orientation estimation. Having to process the spatial information does not seem to impede the orientation reports, implying some separability of the two estimation processes.

Kurtosis did not seem to capture the differences in orientation error distribution between high and low-probability conditions here (*Figure 3.3a*) as well as it did in Jabar & Anderson (2015) or in *Experiment 2.1*, although the trend was still here. This was most likely due to the change in response method: Whereas the previous experiments had participants use a keyboard to rotate a response line, participants in this set of experiments had to draw the orientations. While the distributions are still non-normal – excess kurtoses are above zero – reduced non-perceptual error might have led to less deviations from normality. For these reasons, the median absolute error was still used as the metric for precision here.

In addition to being reported more precisely, responses to high probability tilts were also initiated more quickly. Spatial RT did not vary as a function of probability, perhaps because it reflects time required to make precise motor movements rather than reflect time to detect. While orientation probability affected orientation precision it did not influence spatial precision. This is consistent with the hypothesis that spatial and orientation effects are separable similar to spatial and feature attentional effects.

Results from *Experiment 3.1* are in line with the suggestion that feature probability, similar to feature-based ‘attention’, affects neural tuning. The second part of the hypothesis is that spatial probability, like spatial attention, is linked to neural gain. If true, there should be no spatial probability effects on orientation precision, because a gain mechanism (e.g. *Figure 1.1c*) would predict that orientation-selective neurons would be stable over the range of orientations.

3.5. Experiment 3.2: Methods

Using the same paradigm as *Experiment 3.1*, *Experiment 3.2* examined the effects of spatial probability instead of orientation probability. If the gain hypothesis is true, orientation precision should not be affected by spatial probability. If the dimension-specific hypothesis is true, unlike in *Experiment 3.1,* we should also expect to observe spatial probability effects on spatial precision.

Participants

Two additional sets of participants were recruited. Twenty undergraduate students (18 females, 2 males) took part in *Experiment 3.2a*. 19 participants were right-handed. Twenty undergraduate (11 females, 9 males) took part in *Experiment 3.2b*. 17 participants were right-handed. All participants used their right hand, and had normal or corrected-to-normal vision, and did not declare any auditory deficits.

Stimuli and Procedure

The procedure for *Experiment 3.2a* was the same as before. The only difference was that orientation was kept equiprobable at each location, and spatial probability was manipulated. Two locations on opposite corners were made more probable than the other two. These locations were counterbalanced across participants. Probabilities were maintained in 20-trial blocks. In each set of 20 trials, 8 appeared in one location and another 8 appeared in the opposing corner (high-probability: 80%). The remaining 4 were split across the other 2 locations (low-probability: 20%). As before, participants were not informed about these distributions. They were only instructed to be as precise as they could for both the orientation and spatial estimations.

In *Experiment 3.2b*, instead of having 4 discrete locations, the gratings appeared anywhere in a ring around the central fixation, with the same eccentricity as the other experiments. Instead of biasing two locations, two quadrants were biased. *Figure 3.5* depicts a sample distribution of where participants end up locating the gratings, for both *Experiment 3.2a* and *3.2b*.

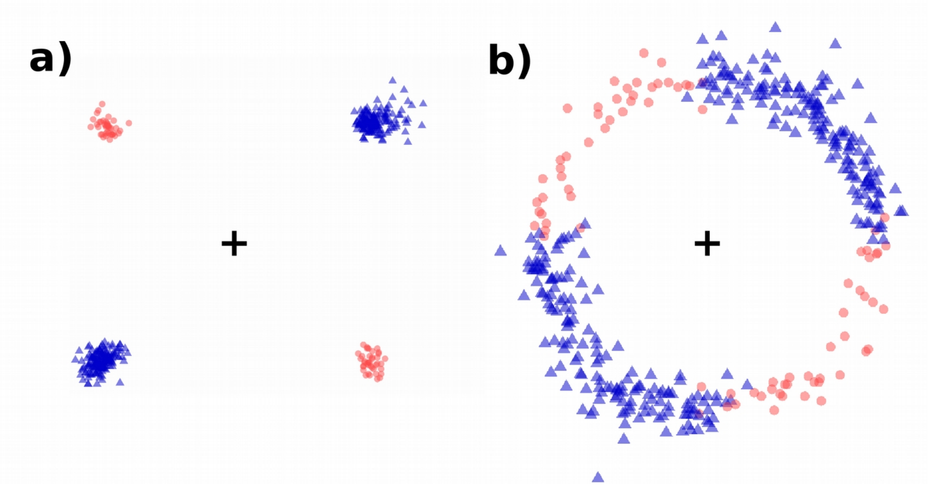

Figure 3.5. *Experiment 3.2* spatial probability distribution. (a) For *Experiment 3.2a*, targets appear in only 1 of 4 discrete locations on any given trial. The markers indicate where the first participant ended up clicking in response to these four locations. Blue markers indicates target in high-probability locales. Two *discrete* locations are biased. (b) For *Experiment 3.2b*, targets appear in *any bearing*, with the same eccentricity from the fixation. Two quadrants are biased. Note that the probable locations were counterbalanced across subjects for both experiments.

3.6. Experiment 3.2: Results

RT analysis

As with *Experiment 3.1*, the RT data from both sets of *Experiment 3.2* was split into IT, sRT and oRT. None of these showed significant effects of spatial probability, in either experiment, (all *p*s > .05).

Orientation error analysis

For *Experiment 3.2a*, in both high (*M* = -1.25, *SD* = 2.24) and low-probability locales (*M* = -1.11, *SD* = 2.23), there was a significant vertical bias (both *p*s <.05) in orientation estimation. However, probability did not modulate this bias (*t*(19) = 0,33, *p* = .742). For *Experiment 3.2b,* although both high (*M* = -0.48, *SD* = 1.90) and low-probability (*M* = -0.35, *SD* = 1.91) had the same trend towards vertical bias, this was not significant (*p*s >.05), and there was no significant difference across the probability conditions (*t*(19) = 0.47, *p* = .641).

The median absolute angular error measure was not significantly different across high (*M* = 7.75, *SD* = 2.12) and low-probability (*M* = 7.96, *SD* = 2.46) locales in *Experiment 3.2a (Figure 3.6a)*, (*t*(19) = 0.81, *p* = .426). This was the same in *Experiment 3.2b (Figure 3.6b)*: There was no significant different across high (*M* = 7.63, *SD* = 1.38) and low-probability (*M* = 7.35, *SD* = 1.33) locales in orientation precision, (*t*(19) = 1.14, *p* = .270). There was also no significant difference in orientation precision across the experiments, for both probability conditions (*p*s> .05).

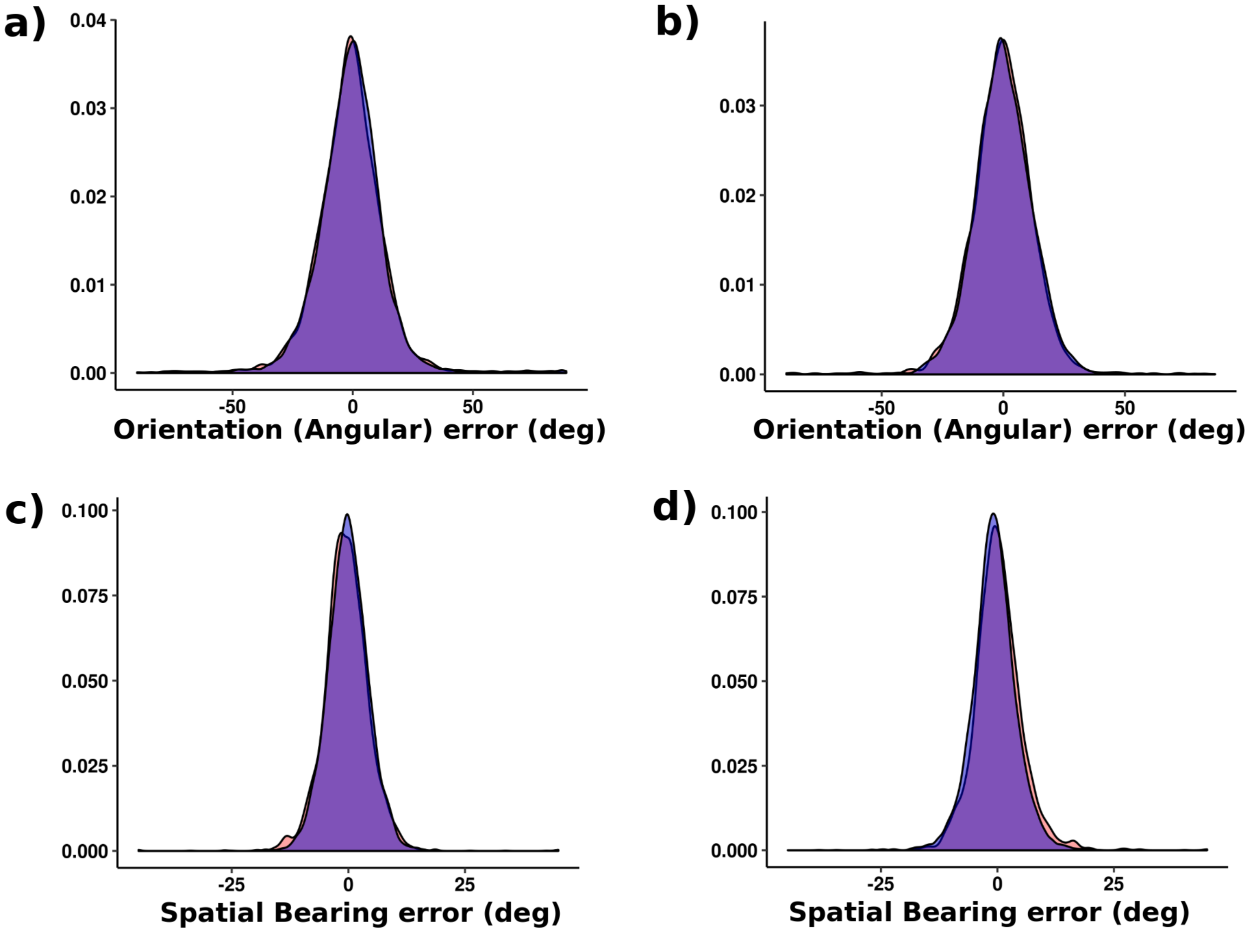

Figure 3.6. *Experiment 3.2* error distributions. (a) Orientation errors for *Experiment 3.2a*. (b) Orientation errors for *Experiment 3.2b*. (c) Spatial errors for *Experiment 3.2a*. (d) Spatial errors for *Experiment 3.2b*. Note that in each panel, there were two distributions: high (blue) and low-probability (red). There is a very high level of overlap between the two in all cases.

Spatial error analysis

The median absolute spatial bearing error was also not significantly different across high (*M* = 2.91, *SD* = 0.93) and low-probability (*M* = 3.05, *SD* = 0.92) locales in *Experiment 3.2a (Figure 3.6c)*, (*t*(19) = 0.70, *p* = .491). This was the same in *Experiment 3.2b (Figure 3.6d)*: There was no significant difference across high (*M* = 2.93, *SD* = 0.69) and low-probability (*M* = 3.04, *SD* = 0.81) locales in orientation precision, (*t*(19) = 0.79, *p* = .439). There was also no significant difference in orientation precision across the experiments, for both probability conditions (*p*s > .05), even if comparing to *Experiment 3.1* (all *p*s > .05). As with *Experiment 3.1*, repeating versus non-repeats of locations was not found to significantly affect spatial precision (*t*(20) = 0.86, *p*=.400) in *Experiment* 3.*2a*. Since locations were continuously distributed in *Experiment* 3.*2b*, quadrants of space were compared. There are two high probability location quadrants for each participant. Trials where there was a switch from one high-probability quadrant to another (40% of trials) were compared to when there was a repeat in the same quadrant (40%). There was no effect of repeated quadrants in either orientation or spatial precision (both *p*s > .05).

Post experiment questionnaire

Three of the participants in each experiment correctly indicated the spatial probability distribution.

3.7. Experiment 3.2: Discussion

Experiment 3.2 suggested that the effects of *spatial* probability are different from the effects of *orientation* probability (e.g. *Experiment 3.1*). When some *orientations* are more probable (*Experiment 3.1*; Anderson, 2014; Jabar & Anderson 2015) orientation judgments were more precise, but when some *locations* were more probable the precision of orientation estimations at the probable locations was unchanged. In addition, spatial probability did not affect spatial precision. The spatial precision in both *Experiment 3.2a* & 3*.2b* matched that from *Experiment 3.1*: The findings are incompatible with the idea of probability causing a domain-general "attentional" effect. The findings also pose trouble for domain-specificity hypothesis as that would have predicted that spatial probability would have improved spatial precision.

Perhaps the spatial estimation task was too easy? This is not the case for *Experiment 3.2b*, where spatial uncertainty was substantial. However, to further probe this account *Experiment 3.3* was run, where contrast was reduced to make the estimation task harder. Since spatial probability has been shown to affect detection (e.g. Walthew & Gilchrist, 2006; Vincent, 2011; Jiang, Sha & Remington, 2015), *Experiment 3.3* was also designed to obtain a measurement of detection rate in addition to precision. This was specifically to show that the effects of spatial probability are on detection, but *not* on precision.

3.8. Experiment 3.3: Methods

Experiment 3.3 introduced two main modifications. First, a '3 down, 1 up' staircase procedure was employed to bring detection down to approximately 80% (Leek, 2001) and remove the possibility of a ceiling effect on performance. The second modification was the inclusion of a confidence report at the end of each trial. If participants fail to detect the target they will guess, and they will report lower confidence. Guessing results in uniformly distributed errors. Using the combination of confidence self-reports and behavioural precision, we can separate guess trials from the detected trials, and better decompose measured imprecision: How much of spatial probability effect is in the change in perceptual estimation precision per se, and how much is in the change in the proportion of trials the target is undetected and participants guess?

Participants

Twenty-one additional participants (13 females, 8 males) took part in *Experiment 3.3*. 19 participants were right-handed. All participants used their right hand, and had normal or corrected-to-normal vision, and did not declare any auditory deficits.

Stimuli and Procedure

Prior to the main task, a “3 down, 1 up” staircase procedure was done to determine what contrast led to an 80% detection rate (absolute bearing error < 22.5 degrees). These trials were qualitatively similar to those used later, but spatial and orientation probabilities were uniform. The first 10 trials were designated as practice and contrast was not altered regardless of performance. Average luminance of the grating was not affected by the contrast reductions (always maintained at 39 cd/m^2).

There were 400 test trials. Spatial probability was manipulated as in *Experiment 3.2b*. After the spatial and orientation estimations, but before the feedback, participants were asked to indicate a response to the question “How confident are you that you *saw* the stimulus?”. The continuous scale ranged from 0-100, but had descriptive labels at 3 points: 0 (Definitely did not see), 50 (Maybe I saw), 100 (Definitely saw). Specifically, this measure was aimed at figuring out when participants perceived or did not perceive the stimulus, and not whether the participant was confident about their responses.

3.9. Experiment 3.3: Results

Staircase and Confidence

Sample staircase results are shown in *Figure 3.7a*. There was a spread of values participants ended up with, but it was mostly in the 4% to 8% range (*median* = 6%, *Figure 3.7b*).

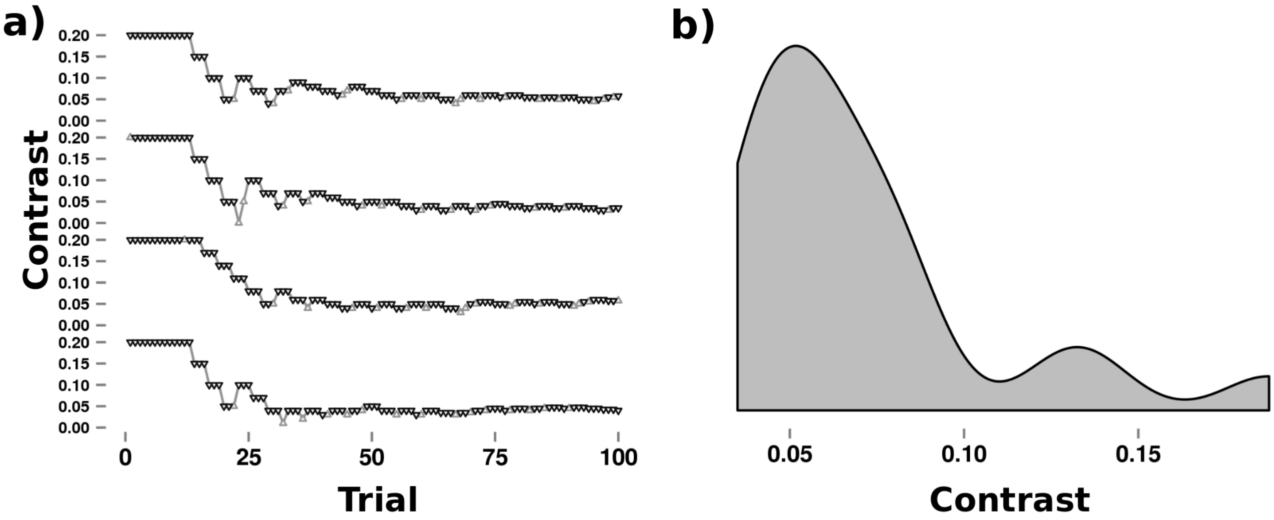

Figure 3.7. *Experiment 3.3* staircase procedure. (a) Sample stair-casing of stimulus contrast using the spatial bearing error via a '3 down, 1 up procedure'. Each panel represents data from one participant. Light markers represents the points where participants crossed the error threshold. The first 10 trials were for practice and were not stair-cased. (b) Distribution of threshold contrasts found by the staircase procedure. Contrast was calculated as the luminance difference between the brightest and darkest 'band' of the grating divided by the sum.

There was a significant effect of spatial probability on reported confidence, (*t*(20) = 2.97, *p* = .008). Participants reported being more confident of detecting the grating in high-probability locales (*M* = 60.2, *SD* = 17.1) than in low-probability locales (*M* = 57.6, *SD* = 17.9). Separating the data into 50-trial bins, high-probability locales were associated with higher confidence reports at all bins.

Samples of the distributions of confidence reports given by the participants are shown in *Figure 3.8a*. There is bimodality: Most of the reports tend to be high, but some are near zero. Since participants were asked to rate how confident they are that they *saw* the stimulus, the peaks could correspond to *detected* (high confidence) versus *non-detected* (low-confidence) trials. What is clear is that different participants use the confidence scale differently: Using arbitrary cut-offs on the raw values to classify the detected trials would be problematic.

Expectation-Maximization Clustering

Instead of using raw confidence values as cut-offs, 2-dimensional distribution of spatial errors and confidence were examined. Because the confidence scale was explicitly related to participants' ability to *detect* the stimulus, when reported confidence is high, participants should have detected/perceived the stimulus and therefore should have high spatial precision (smaller bearing errors). When reported confidence was low, participants likely failed to detect the stimulus and therefore made random guesses, leading to a fairly uniform distribution of spatial errors. Essentially, the 2-dimensional distribution of spatial errors and confidence should end up as an inverted ‘T’ shape (*Figure 3.8b*). We can cluster these points into guesses versus detection. These expected clusters were used as *starting* parameters to cluster participants’ data, although the proportions, centres and size of the clusters were free to vary. This *expectation-maximization clustering* was done using the *EMCluster R* package (Chen & Maitra, 2015).

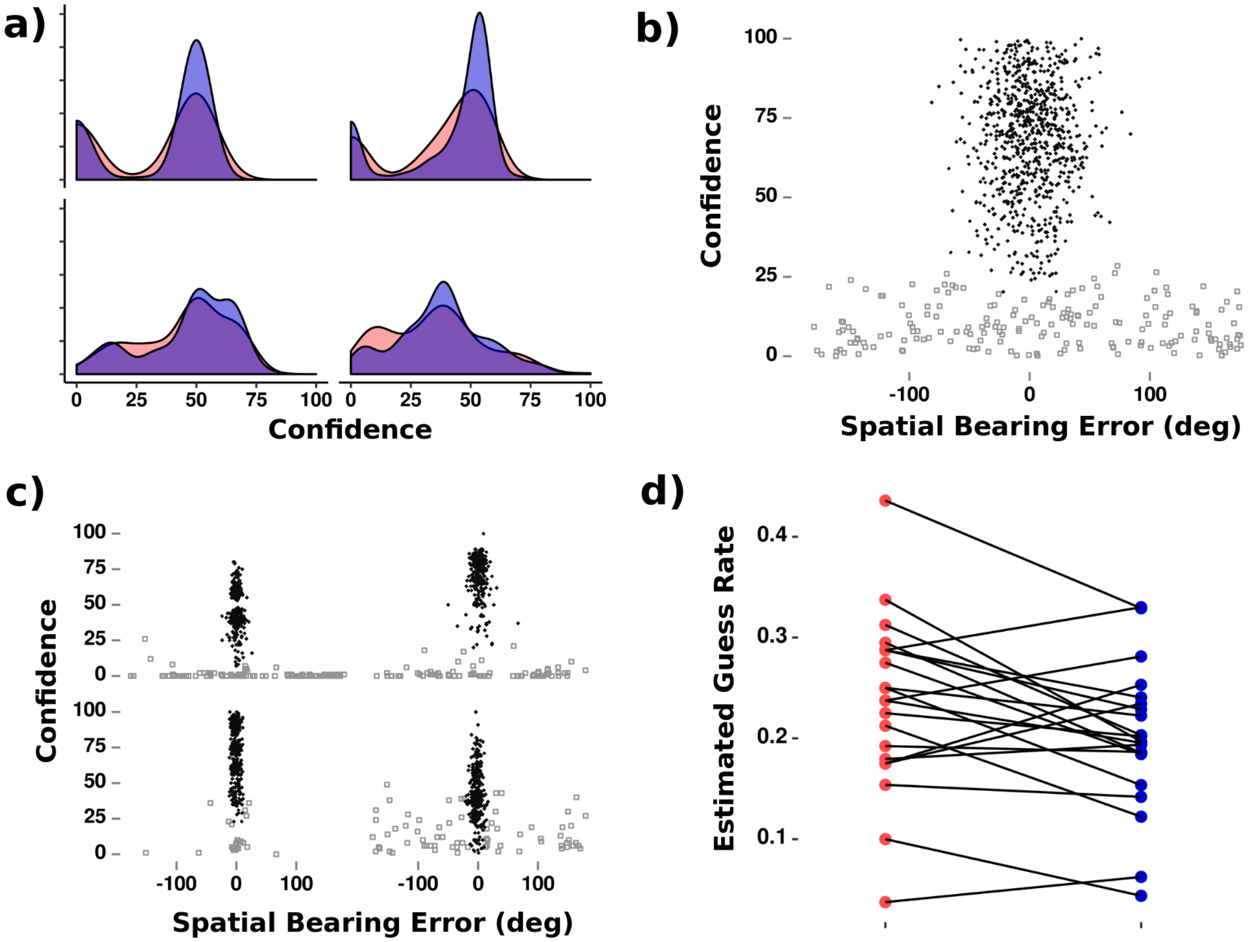

Figure 3.8. *Experiment 3.3* guesses vs. detections. (a) Sample distribution of confidence reports by probability condition. Each panel represents data from one participant. (b) Expected error by confidence distribution. At low-confidence, participants should make uniform errors (indicating guessing) while at high-confidence, the same participant should make precise spatial estimations. (c) Sample of clustering using the expectation-maximization (EM) technique. The darker markers represent trials that were most likely detected/perceived. (d) Guess rates by subject, based on the results of the EM clustering. Note: Blue represent high-probability trials, red for low-probability.

Sample participants' 2-dimensional distributions are depicted in *Figure 3.8c*. The expected inverted 'T' shape is apparent, and the clustering fits as expected. After clustering each participants' *full* dataset, they were separated into high and low-probability trials, the proportion of trials in each condition that belongs in the 'guess' (low reported confidence with random spatial error) cluster (*Figure 3.8d*) were examined. Out of the 21 participants, 15 showed decreased guessing rates, i.e. better detection, for high (mean = 20.5%) versus low probability (mean = 23.7%) locales. This difference was significant across participants, ($t(20) = 2.19$, $p = .041$). Note that the exact initial/expected values used for the clustering (*Figure 3.8b*) had slight effects on the exact values of the guessing rates. To confirm that the results were not dependent on starting values for the optimization, each clustering procedure was run 1000 times, slightly jittering the expected clusters. All tests consistently indicated that there was a significant 3% to 4% difference in guessing rate between probability conditions. The guessing rates also hovered around the 20% to 25% range, which expected given the use of the '3 down, 1 up' procedure (Leek, 2001). The EM clustering also gave similar results to visually (manually) clustered data.

As a check of the clustering method, a mixture model was also used to to estimate the proportion of perceived vs. guessed trials. Similar to the method used by Zhang & Luck (2008), the spatial bearing error data from each participant was fitted with a mix of uniform and Gaussian, where the three parameters – proportion (*lambda*), mean and variance mean of the Gaussian component – were free to vary. Proportion of the uniform was just 1 minus *lambda* and thus was not a free parameter. The idea here is that the uniform distribution represents the guessed trials, and therefore the lambda parameter should match the results of the clustering analysis. The mixture model suggested that participants perceived the grating (lambda) 84.2% (*SD* = 4.7%) of the time in the high probability trials, but only 80.5% of the time (*SD* = 6.1%) in the low-probability condition, with there being a significant (*t*(20) = 2.20, *p* = .034) difference between the two. However, there was no difference in either the mean or variance of the mixture models across high and low spatial-probability trials (*p*s >.05). Both the clustering and mixture models therefore are in agreement that spatial probability modulate guess/detection rates, but not the precision with which estimations are made. The guess rates (high= 15.8%, low= 19.5%) estimated from the mixture model also significantly correlated (*p*s <.05) with the guess rates from the clustering model, for both high (r = .771) and low (r=.723) probability data, further suggesting convergence between the two methodologies at the level of individual subjects.

RT analysis

The IT was faster for perceived (*M* = 387ms) than for guessed (*M* = 685ms) trials, (*t*(20) = 7.06, *p* < .001). The sRT was not different between perceived (*M* = 1564ms) and guessed (*M* = 1686ms) trials, (*t*(20) = 0.94, *p* = .353). However, the oRT was slower for perceived (*M* = 866ms) than for guessed (*M* = 669ms) trials, (*t*(20)= 2.29, *p* = .028). It is possible that participants were caught off guard during the spatial localization phase of the experiment if they did not detect the stimuli, leading to longer ITs, but clicked randomly for their orientation judgments leading to faster oRTs.

Looking only at perceived trials, there was still an effect of spatial probability on IT: Targets appearing in high probability locales (*M* = 386ms, *SD* = 57ms) were reacted to faster than for those appearing in low-probability locales (*M* = 393ms, *SD* = 63ms), (*t*(20) = 2.60, *p* = .017).

Precision analyses

Orientation errors were much larger for the guess trials (*M* = 34.4) than the perceived ones (*M* = 10.1), (*t*(20) = 13.4, *p* <.001). This amount of error highlights that we should only take into account trials that were clustered as ‘detected’ or ‘perceived’, when we are interested in perceptual precision.

Orientation error for the perceived trials was not any better for the high (*M* = 7.75, *SD* = 1.52) versus the low-probability (*M* = 7.56, *SD* = 2.12) conditions, (*t*(20) = 0.72, *p* = .482). The orientation bias was not significantly different from zero for either probability condition (*p*s > .05).

Spatial error for the perceived trials was also not any better for the high (*M* = 3.91, *SD* = 0.90) versus the low-probability (*M* = 3.95, *SD* = 1.20) conditions, (*t*(20) = 0.72, *p* = .482). However, compared to *Experiment 3.2b*, where stimuli were displayed with higher contrast, spatial error in *Experiment 3.3* was increased (*t*(20) = 4.70, *p* < .001), which is an indication that spatial precision can be affected by contrast manipulations. In comparison, orientation precision was not affected across these experiments, (*t*(20) = 0.466, *p* = .642).

To address a lack of statistical power as a potential concern over the lack of precision effects in the spatial probability tasks, the data from *Experiment 3.2a*, *3.2b*, and the perceived trials of *Experiment 3.3* were pulled and placed into a Bayesian comparison (high vs. low spatial probability) treating the 61 participants as a random factor. The Bayes Factor (*BF*) obtained from comparing the spatial bearing error was 0.063 (±1.31%). The *BF* obtained from comparing the orientation error was 0.029 (±2.28%). These numbers support the hypothesis that spatial probability has no effect on either spatial or orientation precision.

As comparison, data from the 21 participants from *Experiment 3.1* was also looked at. The *BF* from comparing the spatial error was 0.116 (±22.86%), while the *BF* obtained from comparing the orientation error was 378.755 (±0.81%). These numbers support the hypothesis that while orientation probability does not affect spatial precision, it very clearly affects orientation precision.

Repetition analysis

Since there was an effect of spatial probability on detection rates, the effects of repeated locations were looked at for the guessed or non-perceived trials. Particularly, for the high-probability locales the proportion of trials from repeated location (mean = 40.2%) compared to trials followed by the other high-probability quadrant (mean = 40.1%), was not significantly different, (*t*(20) = 0.065, *p* = .949), suggesting that repetitions are not driving the probability effect on guess rates.

Post experiment questionnaire

Of the 21 participants, 3 correctly indicated the spatial probability distributions.

3.10. Experiment 3.3: Discussion

Experiment 3.3 replicated *Experiment 3.2*. For trials where the grating was perceived, spatial probability affected neither spatial nor orientation precision, but initiation times were faster for the high-probability locales.

Reducing contrast did reduce spatial precision, and precludes the possibility that the absence of an effect in *Experiments 3.2a* and 3*.2b* were due to a ceiling effect. While spatial probability did not affect spatial precision, it did affect detection: While detection rates on average were about 80% – a rate predicted by the type of staircase procedure employed – detection rates were lower when the grating appeared in the low-probability locale, as compared to the high-probability locales. This lower guess rate for high probability locales was not due to repeats in location (Geng & Behrmann, 2005), suggesting that spatial probability effects in this case is not due to sequential dependencies (c.f. Walthew & Gilchrist, 2006), perhaps because locations were continuously distributed. In addition, even for trials where the stimulus was perceived, the IT was still significantly faster for high-probability locales, consistent with the idea that spatial probability affects detection. Given the size of these detection effects, is difficult to get an estimate of how quickly the changes in detection due to spatial probability occurs. However, reported confidence was higher (trend-wise) for high-probability locales *throughout* the experiment, even within the first fifty trials. It is likely that spatial probability learning, like orientation probability learning, occurs very quickly.

In sum, the three experiments outlined in this chapter consistently indicate that while orientation probability affects orientation precision, spatial probability only affects target detection. The following section discusses the implications of these findings.

3.11. Discussion

The goal of these experiments was to evaluate how generic the effects of probability manipulations are. Does manipulating the probability of where a target will appear (or whether it bears a probable feature) affect the speed and accuracy with which the target can be located, and does it also affect the fineness with which features can be discriminated? A second goal was to infer whether probability effects and attentional effects are likely to index common underlying psychological processes or neural bases.

The principal result was that probability cues, whether spatial or featural, shorten reaction time, just as do conventional cues in spatial and feature attention experiments. Because of the experimental procedure, these speed effects could be decomposed into initiation and movement components. The speed-up was most apparent for movement initiation. This was true for both orientation (*Experiment 3.1)* and spatial probability *(Experiment 3.3).* This effect was larger for spatial probability, but perhaps this was due to the increased difficulty associated with a reduction in contrast. The most obvious reason for this improvement was better target detection (Rich et al., 2008; Wolfe et al., 2007; Walthew & Gilchrist, 2006; Vincent, 2011; Jiang et al., 2015; Fecteau, Korjoukov & Roelfsema, 2009; Geng & Behrmann, 2005).

In Jabar & Anderson (2015) and in Chapter 2, participants were shown a dial after each grating, and asked to rotate it using keyboard buttons, to best represent the orientation they saw. By manipulating feature (orientation) probability, it was found that the initiation time (time taken to initiate an anticlockwise or clockwise dial rotation) required was decreased for probable tilts. Orientation precision was also improved. These results were replicated in *Experiment 3.1* of the current study, despite the changes to the response method. As with feature probability, initiation time (time taken to move the mouse away from central fixation) was also affected by spatial probability. However, in neither *Experiment 3.2a, 3.2b or 3.3* did spatial probability affect orientation precision, even when spatial probability improved detection (*Experiment 3.3*). Consistent with the studies outlined in this chapter's introduction, both spatial and feature probability manipulations create effects that are 'attention-like'. However, this chapter also clearly indicates a difference between the two.

Instead of a general 'probability effect', we should be concerned with probability effect*s (plural)*. Given that the task is not one of forced choice, there is less opportunity for probability effects to be mediated at the level of a decision process (Wolfe & Van Wert, 2010) and one might think that decision effects would more likely affect judgment time than initiation time. Motor (or response) biases are also unlikely to explain the results as the effects of feature probability were location-contingent (this will be explicitly shown in Chapter 4). All movements were equiprobable and all angles were shown equally often. Thus, the most parsimonious explanation is that the basis for the orientation probability effect is perceptual (Dykes & Pascal, 1981), perhaps at the level of orientation tuning in early visual areas. On the other hand, spatial probability affected neither the precision of spatial errors nor orientation estimation errors. This decomposition of spatial and orientation probability effects maps on nicely to the decomposition of attention into spatial and feature specific varieties (e.g. Carrasco, 2011).

A plausible locus for the shared effects of spatial and feature probability and spatial and feature attention is on the primary visual cortex, where attention has been noted to have consistent effects (Brefczynski & DeYoe, 1999; Gandhi, Heeger & Boynton, 1999; Sharma et al., 2015; Tootell et al., 1998). The basis for these effects are broadly of two types (Martinez-Trujillo & Treue, 2004; Ling, Liu, & Carrasco, 2009): first, sharpening of neuronal tuning function in early sensory neurons, and second increasing the gain of neuronal firing curves for early sensory neurons.

Feature attention has been suggested to be primarily due to the first mechanism, changes in neuronal tuning curve width (Paltoglou & Neri, 2012). V1 neurons show changes in tuning functions with an increased experience of particular orientations (Ringach, Hawken & Shapley, 1997; Schoups et al., 2001). By contrast, spatial attention's effects are primarily accompanied by increases in neuron firing generally, a phenomenon referred to as response gain (Kastner & Ungerleider, 2000; Ling et al., 2009; *Figure 1.1c*). Our suggestion that feature probability effects are also mediated by neuronal tuning effects is consistent with recent work on statistical decision theory (Eckstein et al., 2009). Probabilistic information for features might tune relevant perceptual channels, which *then* affects decision-making.

The analogy between categories of attentional cues and probability manipulations is provisional. There are differences as well. Chief among them is the temporal dynamics. Probability learning is acquired across trials while spatial cuing is typically a brief trial-specific signal. In addition, cues provide a non-specific preparatory signal that is absent in pure probability learning, resulting in different effects. For example, explicit spatial cues cause increased V1 firing of neurons selective to the location *prior* to the target onset (Wang et al., 2015). This might explain why spatial *cuing* affects orientation estimation (Anderson & Druker, 2013; Chapter 2), while spatial *probability* (Chapter 3) might not.

In summary, the main empirical result in this chapter is that probability effects, like attention, can dissociate between feature and spatial biases. Feature probability affects precision, while spatial probability only affects detection. Based on these findings, and previous work on the mechanisms underlying attentional cuing, there likely are conceptual and mechanistic links between probability and attention, chiefly that they may both affect perceptual processing by similar neural mechanisms. Cues and probability provide participants with information about *what* stimuli might appear, and *where* they might appear (Anderson, 2011). This information leads to biases in early perceptual systems that show a common reliance on adjusting the gain and tuning of sensory neuronal populations. This hypothesis does predict further dissociations between feature and spatial probability, such as the domains under which they operate. This issue is explored in Chapter 4.

Chapter 4: Feature and spatial probability effects occupy different domains[8]

4.1. Introduction

The experiments in Chapter 3 found that while orientation probability increased the precision of orientation estimations, spatial probability did not. Instead, spatial probability improved the ability to perceive the gratings detected as a decrease in guessing events but it did not change the precision of orientation reports for the perceived proportion of trials. The objective of the present set of experiments is to further explore the similarities and differences between spatial and feature probability. Of particular interest are two questions: 1) Do these effects manifest under the same situations? 2) When they do manifest, do they affect responses the same way?

Feature-specific versus domain-general effects

If orientation probability has its effect by selective tuning of orientation-sensitive V1 neurons, akin to what occurs with orientation training in monkeys (Ringach, Shapley Hawken, 1997; Schoups et al, 2001), one would expect that the probability benefits should not extend to other features that are not processed by that pool of neurons. For example, colour-selective neurons also exist in V1 (Hanazawa, Komatsu & Murakami, 2000). While there are sub-populations of color selective V1 neurons that are also selective for orientation, color and orientation selectivity (Johnson, Hawken & Shapley, 2008) neurons are, for the most part, different pools. If blue objects are probable, under the current hypothesis, V1 neurons selective for blue might be sharpened and perceptual sensitivity to blue features improved. But this might not necessarily improve sensitivity to the *orientations* of objects presented in blue, depending on which neuronal sub-population is being tuned by the probability manipulation.

---

[8] A version of this chapter has been published as Jabar, S. B., Filipowicz, A. & Anderson, B. (2017). Knowing where is different from knowing what. Distinct response time profiles and accuracy effects for target location, orientation, and color probability. *Attention, Perception & Performance, 79*(8), 2338-2353. https://doi.org/10.3758/s13414-017-1412-8

An alternative argument could be made regarding feature-probability as a form of feature-based attention, which has also been suggested to result in neural tuning (Carrasco, 2011; David et al., 2008; Paltoglou & Neri, 2012; Çukur et al., 2013; Ling, Jehee, & Pestilli, 2015). However, this makes the same prediction about the independence of feature probability effects. Both accounts, which rely on domain specific tuning mechanisms, lead to the conclusion that feature probability effects should be specific to the feature whose probability is being manipulated. This is opposed to an object-based attentional account (e.g. Egly, Driver, & Rafal, 1994). If attention is deployed to the object containing the probable feature, then that object-based attention should affect the processing of multiple features.

In contrast to feature-based manipulations which are linked to neural tuning, space-based manipulations are more related to gain-mechanisms (Carrasco, 2011), such as increasing the input baseline of neural responses (Cutrone, Heeger, & Carrasco, 2014). If these baseline increases occur for all neurons coding for that space, it accounts for how spatial probability has general effects. Spatial probability helps in locating objects in search tasks (Jiang et al., 2015), in improving color detection (Druker & Anderson, 2010), and in speeding up orientation judgments (Chapter 3).

If feature probability results in feature-specific tuning changes while spatial probability results in general baseline changes, differences across these manipulations should generalize beyond estimation tasks. For example, in a 2-alternative forced choice (2AFC) task, participants can be asked to report whether a target stimulus has a left (\) or right-tilting (/) orientation. If spatial probability results in domain-global effects, it should facilitate discrimination of orientations. However, since orientation probability also improves perceptual precision where spatial probability does not (Chapter 3), orientation probability should show bigger effects on discrimination accuracy than spatial probability, especially when the discrimination task is perceptually difficult. In addition, if feature probability *only* results in feature-specific effects, one might expect that while manipulations of orientation probability might aid in orientation discrimination, manipulations of color probability should not. If this specificity versus generality hypothesis is true, the reverse situation must also hold: While both spatial and color probability should aid in color discrimination, orientation probability should not.

Initiation versus Movement times

While 2AFC tasks can provide much information about the effect of spatial/featural manipulations on perception across a series of trials (Cutrone et al., 2014; Ling et al., 2009), on any single trial it typically leads to binary data – either one or the other button is pushed. Such binary classifications can miss important details. Prior data from estimation tasks suggest that initiation times (ITs) are more closely linked to perceptual precision while movement times (MTs) are more closely linked to confidence (Jabar & Anderson, 2015). However, standard keyboards only register a button press when the key travel reaches some threshold, and not when the key has started to move. Therefore it would be impossible to distinguish how much time was required to initiate the button press from the time used in the actual motion of the button press. Decomposing of RTs into ITs and MTs can instead be achieved through the use of levers or triggers (e.g. Smeets, Wijdenes, & Brenner, 2016), which can report the state of the response in a continuous fashion.

Experiments in this chapter implemented a novel method for 2AFC tasks that uses the triggers on an Xbox controller to enrich the measuring of the dynamics that goes into making a ‘choice’. Each of the two options is tied to a separate trigger. IT can be taken as the period from the time of stimulus onset to the moment a trigger pull is initiated. MT is taken from the moment the initiation occurs to the time the trigger state reaches a pre-determined threshold value. If both feature and spatial probability affect the time taken to perceive targets, both should affect ITs. MTs on the other hand might be more indicative of confidence than accuracy (Jabar & Anderson, 2015). MT is dependent on other aspects of the response profile, such as the baseline level of preparation, the force of the response, and whether the participant vacillates between the two options. Pre-stimulus baseline levels of preparation should be constant if the probability of either response is kept the same. However, response force has been suggested to be affected by *stimulus* probability, even when *response* probability is controlled for (Mattes, Ulrich, & Miller, 2002), suggesting that probability also has non-perceptual effects. By looking at which component(s) of RT is affected by our manipulations, we can have a better idea of the potential mechanisms involved.

In summary, the aim was to examine differences between feature and spatial probability across two dimensions. Do the two manipulations differ in how domain general they are? Do they have similar response profiles? The current chapter outlines a series of modified 2AFC experiments that are procedurally similar to allow for comparisons. *Experiments 4.1-ori* and *4.1-col* looked at the effects of orientation and color probability when that feature whose probability was manipulated was relevant for the detection task. *Experiments 4.2-ori* and *4.2-col* performed identical probability manipulations but the task depended on detecting the feature whose probability was *not* manipulated. In each experiment there was also a block where spatial probability was manipulated. Despite always being ‘irrelevant’, spatial probability had a consistent effect on ITs and accuracy. In contrast, feature probability effects manifested only when the probable feature was relevant for the discrimination (*Experiment 4.1*), but not when irrelevant (*Experiment 4.2*). In addition, only feature probability affected MTs. *Experiment 4.3* looked at orientation and spatial probability with a more difficult orientation discrimination task, with the result that orientation probability showed an increased effect on task accuracy.

4.2. Experiment 4.1: Methods

Participants

For *Experiment 4.1-ori*, sixteen participants (*median* age = 21 years) were recruited (8 females, 8 males). All reported themselves right-handed. For *Experiment 4.1-col*, eighteen additional participants (*median* age = 20 years) were recruited (9 females, 9 males). Seventeen reported themselves right-handed. All participants had normal or corrected-to-normal vision, and were not colorblind.

Stimuli

Oriented square-wave gratings with a circular mask were the target stimuli for the current series of experiments. These were shown at 50% maximum contrast and colored either blue or green (*Figure 4.1*). These green and blue patches were isoluminant at 40cd/m^2. The gratings had a spatial frequency of 4 cycles per degree of visual angle, and were presented on a gray background with a similar luminance. When viewed from a distance of 60cm, the gratings subtended approximately 4 degrees of visual angle both vertically and horizontally. Targets were presented either above or below a black fixation cross, at an eccentricity of 4 degrees.

Only one grating was presented on each trial. These gratings always had a spatial attribute (top or bottom), a color attribute (green or blue) and an orientation attribute (right-titling or left-tilting). Both *Experiment 4.1-ori* and *4.1-col* had a feature probability block and a spatial probability block. In the spatial block, gratings occurred in one of the two locations 80% of the time (counterbalanced across participants), and the features (color and orientation) was equiprobable for each location (see *Figure 4.1b* for sample sequences). The practice block had uniformly distributed elements (location, color, orientations).

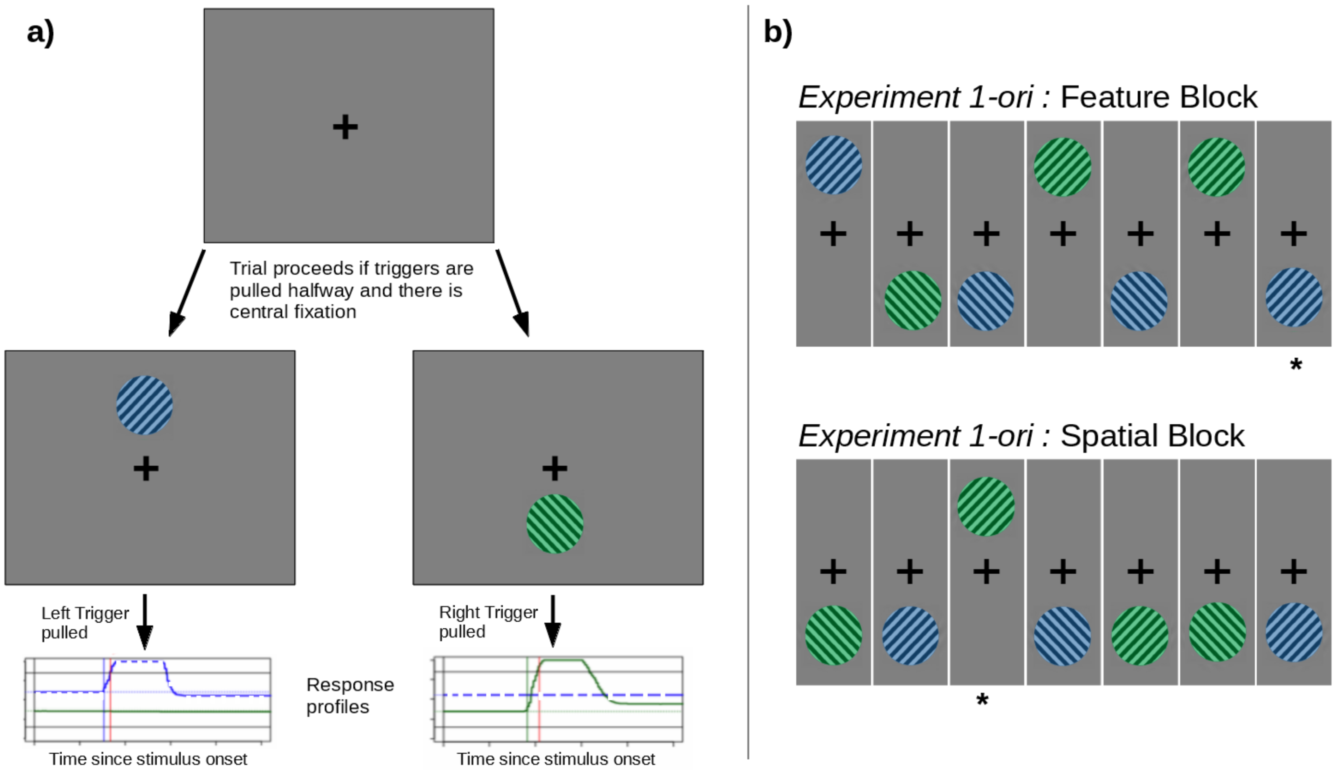

Figure 4.1. General Paradigm. a) All experiments followed the same basic outline, and there were only two possible stimulus locations and colors. In *Experiment 4.1* and *4.2*, there were only two possible orientations, but in *Experiment 4.3*, orientations followed a continuous distribution. The grating on the left example is what is referenced as having a 45-degree tilt. In this case, the participant was responding to orientation with one trigger being assigned to each of 45 and 135 degrees. Note how the two triggers can be operated independently and simultaneously. b) Trial distribution for *Experiment 4.1-ori* for the feature (orientation) and spatial block. The asterisks indicate a low-probability trial for that block. The probable orientation/space was counterbalanced across participants.

For *Experiment 4.1-ori*, the feature block had equiprobable locations and color, but orientation probability was manipulated. Feature probability was manipulated in *a location-contingent manner.* When the grating occurred in one location, it was 80% likely to be right-tilting (/ : 45 degrees). In the other location, the left-tilt (\ : 135 degrees) was 80% likely. Note how the distribution for the feature block differs from that of the spatial block (see *Figure 4.1b*). The location-orientation contingencies were counterbalanced across participants. The participants' task was to report the tilt. Trigger assignments were also counterbalanced across participants. Because of the location-contingent probability mapping, the probability of response was always equal across the two triggers. This location-contingent probability mapping also made the feature probability uniform *if space is ignored*.

For *Experiment 4.1-col*, the feature block had equiprobable locations and orientations, but color probability was manipulated, again in *a location-contingent* manner. When the grating occurred in one location, it was 80% likely to be blue. In the other location, green was 80% likely. This was counterbalanced across participants. Participants were told to detect color and were assigned one trigger for green and the other for blue. Trigger assignment was counterbalanced across participants.

Within a block, probability distributions were maintained. For example, in the spatial block, for every set of 20 trials, there were 16 gratings on the top, 4 gratings the bottom, etc. Participants were *not* informed about these probability distributions. While the probability type was changed across blocks, this was unannounced, and the task from a participant's view remains the same: In *Experiment 4.1-col*, the task was always to detect color; in *Experiment 4.1-ori*, the task was always to detect orientation. Auditory feedback was given after each response to maintain motivation. A high pitched sound indicated a correct response. A lower pitch indicated an error.

Equipment

Stimuli were displayed on a gamma-corrected CRT monitor that refreshed at 89Hz (*mean* refresh = 11.27ms, *SD* = 0.07ms). An Xbox controller was connected via USB and the *xboxdrv* package (https://aur.archlinux.org/packages/xboxdrv/) allowed access to the controller for recording trigger force and excursion. Responses were made using the two back triggers of the Xbox 360 wired controller. Participants were instructed to hold the triggers half-depressed with their index fingers. This allowed us to detect small changes in force at baseline or preceding responses. A continuous measure of trigger depression (recorded at 2000Hz) was obtained for each trigger. The output scaled from -1 (no pressure) to +1 (fully depressed). A value of +0.7 was fixed as the detection threshold and a value of -0.7 or less caused the motors in the handles of the controller to rumble giving the participant haptic feedback indicating to increase their pressure on the trigger. Stimuli were only displayed when both triggers values were within the -0.7 to +0.7 range, i.e. the next trial began only after the participant relaxed their response from the previous trial. See *Figure 4.2* for example response profiles.

Eye-tracking used an Eyelink1000 recording the dominant eye at 2000Hz and using both pupil and corneal signals. Participant head position was stabilized using a chin and forehead rest. Drift corrections were done every 100 trials (approximately every 2-4 minutes). A re-calibration was done if the fixation was lost. Re-calibrations were also done in between blocks (at approximately the 12 minute mark). The next trial began only when the participant fixated on the fixation cross, within a radius threshold of 0.5 degrees visual angle. This was set-up so that stimuli in the same on-screen locations were presented at roughly the same peripheral location retinotopically. This also prevented anticipatory saccades prior to a trial.

Procedure

Prior to the task, participants were instructed as to the trigger assignment, and to maintain pressure at about a halfway point unless making a response, in which case they were to pull the appropriate trigger as fast as possible. Regardless of whether participants were in the practice, spatial, or feature block the task was always the same. Either report orientation in *Experiment 4.1-ori*, or color in *Experiment 4.1-col*. Participants were not informed of the probability manipulations.

The experiment began with 20 practice trials. The next 800 trials were split into two blocks. The first block was either the spatial or feature block, and after a short break the other block type was administered. Trials began with a fixation phase, where only the black fixation cross was shown. If there was central fixation (0.5 degree radius) and the appropriate trigger levels (between -0.7 and +0.7) were detected, the grating appeared between 250 and 500ms (uniform distribution) later. The stimulus remained on-screen until a response was made, at which point auditory feedback was given, and the stimulus disappeared (after 500ms). Participants were asked to minimize blinking when the gratings were on the screen, preferably blinking during the feedback period.

During the drift correction (every 100 trials), participants were told to let go of the controller and to relax their hands and blink a few times. This was to minimize fatigue. During the break (after 400 trials), participants were free to remove their head from the chin-rest. The experiment lasted approximately 20-25 minutes, after which a questionnaire was administered.

Post-experiment Questionnaire

Before debriefing, participants were given a short questionnaire consisting of the following six open-ended questions. This questionnaire was modified from the one used in Chapter 3 to include a question on color, e.g. “Do you think that some colors are more likely at certain times? If yes, please elaborate.”

Analysis

Analysis was done using the *R* statistical software package. With the exception of comparisons across experiments, all statistical tests were done within-subjects. Processing was done by running a smoothing spline (*Figure 4.2a*). Response baselines were taken as the median trigger value between 50ms prior and after stimulus onset. Velocity profiles were taken as a first derivative, and acceleration as the second. Trigger acceleration was taken as an indication of force, since *Force = Mass x Acceleration*, and mass can be assumed constant (the units of the force/ acceleration is therefore in trigger distance per millisecond squared). Vacillations were identified from the points where acceleration goes from positive to zero to negative (the point where the trigger is released). A trial could have multiple initiations, either within or across triggers. Initiation times (IT) were back-calculated from these vacillations/turning points (e.g. where there is a +0.05 increase over the baseline within 100ms prior to the turning point). Reaction times (RT) were taken as when a trigger crossed the +0.7 threshold. Note that participants could have tried to correct their responses either unsuccessfully (*Figure 4.2b*) or successfully (*Figure 4.2c*). Number of initiations takes into account these ‘unsuccessful corrections’. Movement times (MT) were taken as the difference between the RT and the first IT.

Accuracy measures were taken at two points. Initial accuracy was taken at first initiation (whether or not the participant started by pulling the correct trigger). Final accuracy was taken at the point one of the triggers crossed the threshold, where participants were given feedback on their response. Change in accuracy due to vacillations was taken as a difference between these two points.

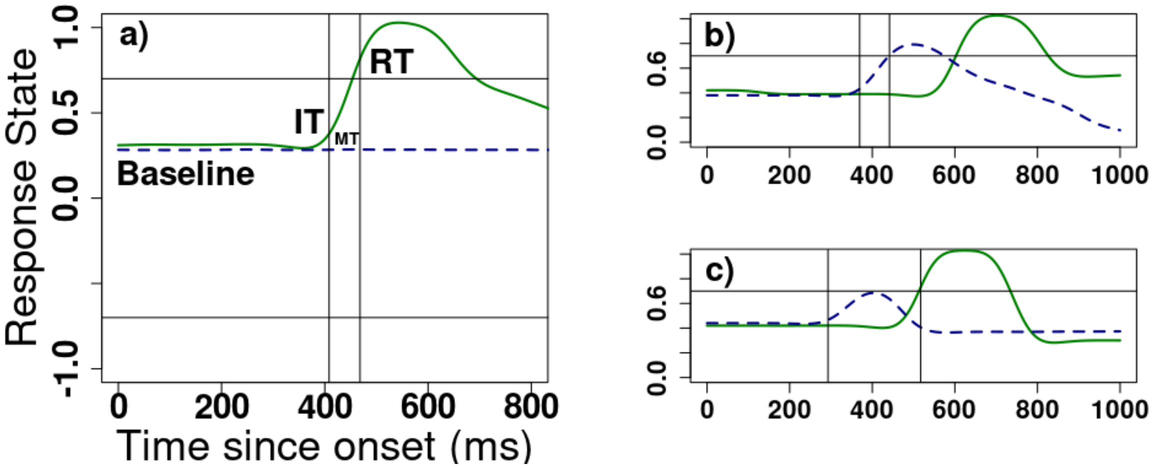

Figure 4.2. Trigger profiles of actual sample trials from one participant. a) Typical trial with only one response. b) Trial with a vacillation in responses that was too late. c) Trial with a vacillation in responses that was in time. The two lines indicate the 2 possible response options (left trigger = blue (dotted) line; right trigger = green (solid) line). The upper black line indicate the detection threshold, the bottom black line indicates the pressure threshold (the controller rumbles below this value). Note how the time required to cross the detection threshold (RT) might be affected by the vacillations. The two vertical lines represent the initiation time and RT respectively.

4.3. Experiment 4.1: Results

Response baselines (*M* = 0.0, *SD* = 0.3), as per instructions, were halfway between a complete pull (+1) and a complete release (-1), with some variability across participants. Importantly, these baselines were equivalent across all blocks of all experiments (all *p*s >.05). This was likely because in all cases, the probability of the left and right response was 50/50, suggesting that there was no *pre-stimulus* response preparation. It is conceivable that participants could have been inclined to respond with the probable trigger *once the stimulus has been shown*. To rule out a *post-stimulus* response-bias account, for each trial, the trigger states just (10ms) prior to the point of response initiation were also noted and compared against the response baselines for that trial. For both the initiated and non-initiated triggers, and in all blocks of all the experiments, there was no change in response states between the time of stimulus presentation and the time of initiation (all *p*s >.05).

There was also no clear bias towards the left or right trigger (*p*s >.05), despite most participants being right-handed. Handedness is unlikely to affect the results because all stimuli-to-response associations were counterbalanced across participants. Potential order effects (e.g. whether the spatial or feature block was first) were examined, and the data were found to be independent of order (*p*s > .05).

Experiment 4.1-ori (spatial + orientation probability, orientation detection)

Reaction Time

For *Experiment 4.1-ori*, in the block where spatial probability was unequal there was an effect of spatial probability on RT (*Figure 4.3*). Orientations at the high-probability location were reported faster (*M* = 564ms, *SD* = 96ms), than at the low-probability location (*M* = 631ms, *SD* = 112ms), (*t*(15) = 9.45 , *p* < .001). Breaking down RT into the IT and MT components revealed that ITs were significantly affected by spatial probability, with orientations in the high-probability location being initiated faster (*M* = 451ms, *SD* = 86ms), than at the low-probability location (*M* = 510ms, *SD* = 88ms), (*t*(15) = 10.07 , *p* < .001). However, there was no effect of spatial probability on MTs. MTs at high-probability locations (*M* = 114ms, *SD* = 55ms) were not significantly different from the low-probability location (*M* = 121ms, *SD* = 63ms), (*t*(15) = 1.48 , *p* = .160). Force of the trigger pull was not affected across high (*M* = 0.002) and low (*M* = 0.002) probability locations (*t*(15) = 1.62 , *p* = .126).

For the feature (orientation) probability block there was also an effect of probability on RT. High-probability orientations were detected faster (*M* = 546ms, *SD* = 64ms), than the low-probability orientations (*M* = 614ms, *SD* = 67ms), (*t*(15) = 10.93 , *p* < .001). ITs were significantly affected by feature probability, with responses to high-probability orientations being initiated faster (*M* = 438ms, *SD* = 74ms), than low-probability orientations (*M* = 485ms, *SD* = 84ms), (*t*(15) = 7.55 , *p* < .001). Unlike with spatial probability, there was a significant effect of feature probability on MTs. MTs associated with high-probability orientations (*M* = 108ms, *SD* = 47ms) were significantly faster than for low-probability orientations (*M* = 129ms, *SD* = 56ms), (*t*(15) = 3.58 , *p* = .003). Force of the trigger pull was not affected across high (*M* = 0.003) and low (*M* = 0.002) probability orientations (*t*(15) = 1.38 , *p* = .188).

For comparison, the effect of spatial probability on RT (*M* = 67ms, *SD* = 28ms) and the effect of feature probability on RT (*M* = 68ms, *SD* = 25ms) were not significantly different from one another, (*t*(15) = 0.17 , *p* = .870). Bayesian testing (*BF* = .259 [± <0.01%]) supports the *null* hypothesis: The effect of spatial and feature probability on total RT was equivalent, even though the effect on the ITs and MTs components differed.

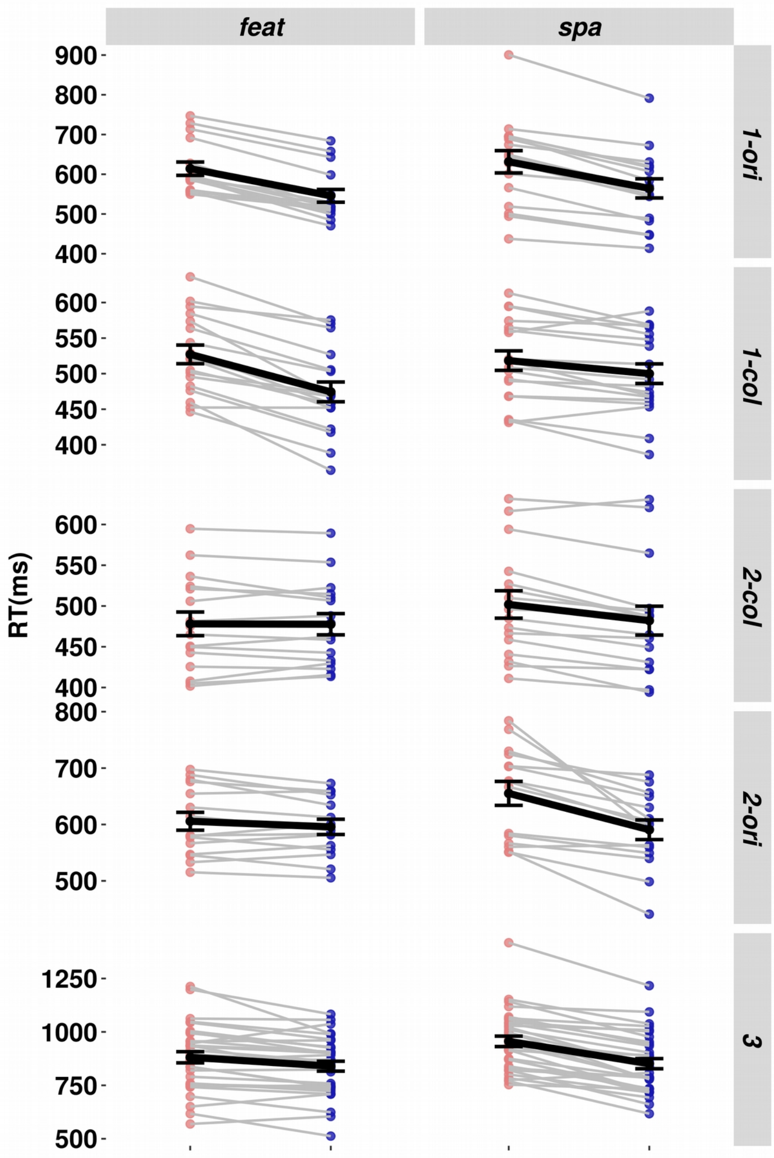

Figure 4.3. Mean reaction time, and trends for each participant. Blue (dark) = High-probability; Red (light) = Low-probability. Panel rows refer to a different experiment, panel columns refer to either the spatial or feature block within the experiment.

Accuracy and Vacillations

For *Experiment 4.1-ori*, in the spatial block there was an effect of spatial probability on final accuracy (*Figure 4.4*), with orientations in the high-probability locations being responded to more accurately (*M* = 96.7%, *SD* = 2.3%), than in the low-probability locations (*M* = 95.0%, *SD* = 2.9%), (*t*(15) = 3.37 , *p* = .004). There was no effect of spatial probability on the amount of vacillations, with orientations in the high-probability locations being as directly responded to (*M* = 0.174, *SD* = 0.250), as in the low-probability locations (*M* = 0.174, *SD* = 0.241), (*t*(15) = 0.02 , *p* = .987). Bayesian testing (*BF* = .255 [± <0.01%]) supports the *null* hypothesis. Vacillations improved initial to final accuracy (*p*s < .05), this improvement was marginally different across high (*M* = 4.4%) and low (*M* = 5.1%) probability locations, (*t*(15) = 1.96 , *p* = .069). Bayesian testing (*BF* = 1.17 [± 0.01%]) suggests that this trend is of minimal impact.

In the feature block there was an effect of orientation probability on final accuracy, with high-probability orientations being responded to more accurately (*M* = 97.7%, *SD* = 2.0%), than in the low-probability orientations (*M* = 89.3%, *SD* = 6.6%), (*t*(15) = 6.08 , *p* < .001). There was no significant effect of orientation probability on the amount of vacillations, with high-probability orientations being responded to more directly (*M* = 0.169, *SD* = 0.263), than in the low-probability orientations (*M* = 0.199, *SD* = 0.260), (*t*(15) = 1.74 , *p* = .102). Bayesian testing (*BF* = 0.875 [± 0.01%]) supports the *null* hypothesis. Again, vacillations improved initial to final accuracy (*p*s < .05), this improvement was marginally different across high (*M* = 4.0%) and low (*M* = 7.0%) probability orientations, (*t*(15) = 2.06 , *p* = .057). Bayesian testing (*BF* = 1.36 [± 0.01%]) suggests that this trend is of minimal impact.

For comparison, the effect of spatial probability on final accuracy (*M* = 1.7%, *SD* = 2.0%) was significantly smaller than the effect of orientation probability on final accuracy (*M* = 8.4%, *SD* =

5.5%), ($t(15) = 4.94$, $p < .001$). The increase in accuracy due to vacillations was lower in the spatial probability block ($M = 0.7\%$) than in the orientation probability block ($M = 3.1\%$), although this was statistically insignificant ($t(15) = 1.70$, $p = .110$).

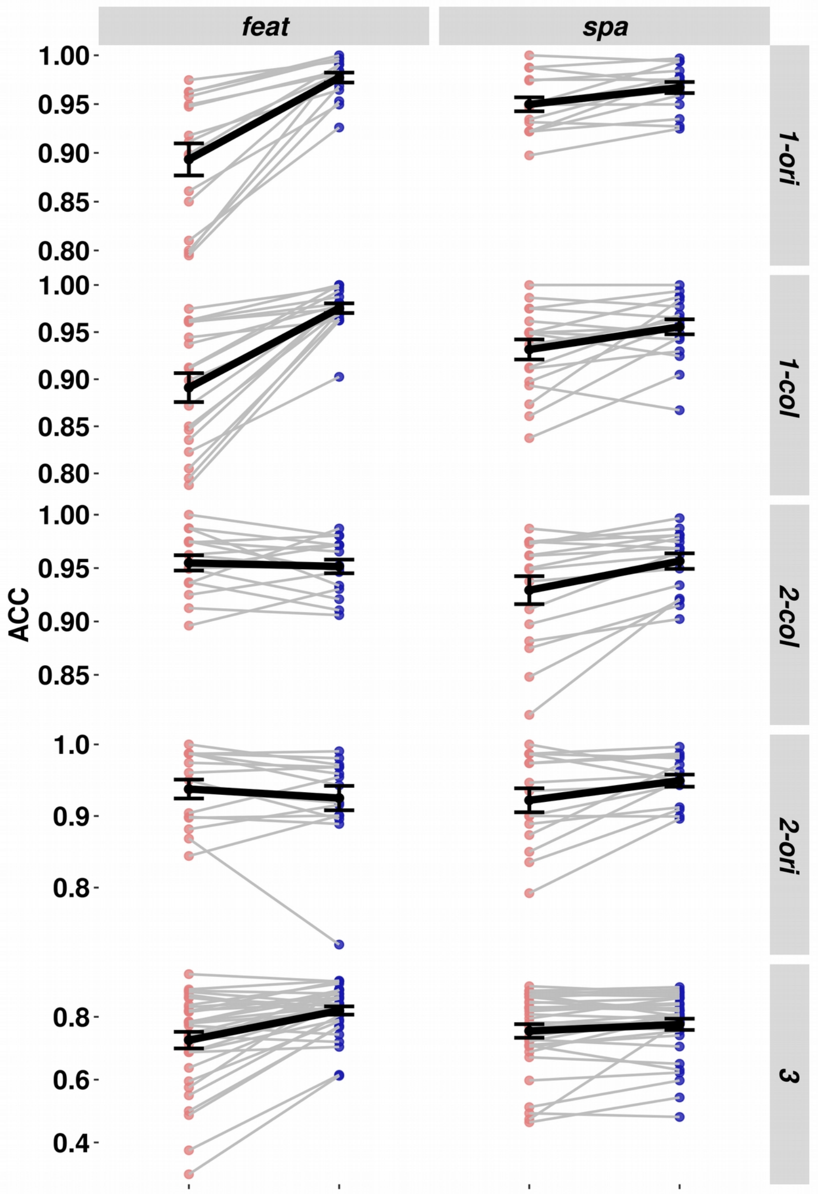

Figure 4.4. Mean accuracy, and trends for each participant. Blue (dark) = High-probability; Red (light) = Low-probability. Panel rows refer to a different experiment, panel columns refer to either the spatial or feature block within the experiment.

Experiment 4.1-col (spatial + color probability, color detection)

Reaction Time

For *Experiment 4.1-col*, in the spatial block there was an effect of spatial probability on RT, with colors at the high-probability location being detected faster (*M* = 500ms, *SD* = 58ms), than at the low-probability location (*M* = 518ms, *SD* = 58ms), (*t*(17) = 3.63 , *p* = .002). ITs were significantly affected by spatial probability, with responses to colors in the high-probability locations being initiated faster (*M* = 400ms, *SD* = 68ms), than at the low-probability locations (*M* = 414ms, *SD* = 77ms), (*t*(17) = 2.87 , *p* = .011). However, there was no effect of spatial probability on MTs, with MTs associated with high-probability locations (*M* = 99ms, *SD* = 26ms) not significantly different from low-probability locations (*M* = 104ms, *SD* = 33ms), (*t*(15) = 1.06 , *p* = .304). Force of the trigger pull was not affected across high (*M* = 0.002) and low (*M* = 0.001) probability locations, (*t*(17) = 1.02 , *p* = .321).

For the feature (color) probability block there was an effect of probability on RT, with high-probability colors being detected faster (*M* = 474ms, *SD* = 59ms), than low-probability colors (*M* = 527ms, *SD* = 56ms), (*t*(17) = 8.89 , *p* < .001). ITs were significantly affected by color probability, with response to high-probability colors being initiated faster (*M* = 380ms, *SD* = 66ms), than low-probability colors (*M* = 416ms, *SD* = 74ms), (*t*(17) = 8.10 , *p* < .001). Unlike with spatial probability, there was a significant effect of color probability on MTs, with movement times associated with high-probability colors (*M* = 95ms, *SD* = 25ms) being significantly faster than for low-probability colors (*M* = 111ms, *SD* = 33ms), (*t*(17) = 3.20 , *p* = .005). Force of the trigger pull was not affected across high (*M* = 0.002) and low (*M* = 0.002) probability colors (*t*(17) = 0.54 , *p* = .596).

For comparison, the effect of spatial probability on RT (*M* = 18ms, *SD* = 21ms) and the effect of color probability on RT (*M* = 52ms, *SD* = 25ms) were significantly different from one another, (*t*(17) =

4.48 , $p < .001$). Across the two experiments, it was also evident that discriminating of orientations (*Experiment 4.1-ori*) was associated with a larger overall RT than the discrimination of colors (*Experiment 4.1-col*), for the spatial and feature blocks (both *ps* < .001).

Accuracy and Vacillations

For *Experiment 4.1-col*, in the spatial block there was an effect of spatial probability on final accuracy, with colors in the high-probability locations being responded to more accurately ($M = 95.6\%$, $SD = 3.3\%$), than in the low-probability locations ($M = 93.2\%$, $SD = 4.5\%$), ($t(17) = 2.60$, $p = .019$). There was no effect of spatial probability on the amount of vacillations, with colors in the high-probability locations being as directly responded to ($M = 0.143$, $SD = 0.147$), as in the low-probability locations ($M = 0.155$, $SD = 0.188$), ($t(17) = 0.83$, $p = .420$). Bayesian testing ($BF = .328$ [± 0.01%]) supports the *null* hypothesis. Vacillations improved initial to final accuracy (*p*s < .05), this improvement was marginally different across high ($M = 4.8\%$) and low ($M = 5.4\%$) probability locations, ($t(17) = 0.80$, $p = .434$). Bayesian testing ($BF = 0.323$ [± 0.01%]) supports the null hypothesis.

In the feature block there was an effect of color probability on final accuracy, with high-probability colors being responded to more accurately ($M = 97.5\%$, $SD = 2.2\%$), than in the low-probability colors ($M = 89.1\%$, $SD = 6.5\%$), ($t(17) = 6.23$, $p < .001$). There was a marginally significant effect of color probability on the amount of vacillations, with high-probability color being responded to more directly ($M = 0.143$, $SD = 0.174$), than the low-probability colors ($M = 0.177$, $SD = 0.160$), ($t(17) = 2.04$, $p = .057$). Bayesian testing ($BF = 1.31$ [± <0.01%]) suggests that this trend is of minimal impact. Again, vacillations improved initial to final accuracy (*p*s < .05), and this improvement was smaller for the high (*M* = 4.0%) than for low (*M* = 8.3%) probability colors, (*t*(17) = 4.48, *p* < .001). Bayesian testing (*BF* = 98.5 [± 0.01%]) suggests that this trend is of moderate to strong impact.

For comparison, the effect of spatial probability on final accuracy (*M* = 2.4%, *SD* = 3.9%) was significantly smaller than the effect of color probability on final accuracy (*M* = 8.4%, *SD* = 5.7%), (*t*(17) = 4.21 , *p* < .001). The increase in accuracy after vacillations was also lower in the spatial probability block (*M* = 0.5%) than in the color probability block (2.9%), (*t*(17) = 3.05, *p* = .007). Across the two experiments, discriminating of orientations (*Experiment 4.1-ori*) was not overall significantly more or less accurate than the discrimination of colors (*Experiment 4.1-col*), for either the spatial or the feature blocks (both *ps* > .05).

Post-experiment questionnaires

Consistent with the previous studies (Chapter 2 and 3), most of the participants did not realize that there were probability manipulations. Of the ones that did report something about the probability of a stimulus (7 out of 34 participants), only the spatial probability manipulation was accurate (e.g. ‘the top was more likely’). Comments on the orientation or color probability were limited to ‘things that look like ‘‘/’’ were more frequent’ or ‘blue felt more likely’. Neither of these were true, since the feature probability was spatially-contingent in all cases: If “/” was more likely when the grating appeared at the top location, “\” was more likely when the grating appeared at the bottom location. The same trends in the questionnaire responses were seen in the later experiments as well.

4.4. Experiment 4.1: Discussion

As with most studies on probability, participants were faster and more accurate when responding to probable objects than to improbable ones. Orientation, color and spatial probability appear to have affected RT in a similar manner *if* one ignores the breakdown into MTs and ITs. (*Figure 4.3*). While both spatial and feature probability affected ITs, only the latter had an impact on MTs, perhaps indicating increased confidence in responses. This logically should be tied down to the number of vacillations made or to response force, but those measures are likely too insensitive given a scenario where accuracies are high: Participants tend to initiate the correct trigger in the current experiment (see *Experiment 4.3* for comparison with a more difficult task).

In addition to MT, both forms of feature probability (orientation/color) also seem to have an exaggerated impact on discrimination accuracy as compared to spatial probability, by a factor of 3 or 4. One possible reason is that space in this context is uninformative to the response, whereas learning of the feature probabilities is informative. The other possibility is that feature probability is helping to shape perceptual precision. We return to this issue in *Experiment 4.3*.

Spatial and feature probability were non-equivalent manipulations in *Experiment 4.1*. The feature was directly informative to the response choice, but space was not. What happens when an *irrelevant* feature is made probable? Is the *color* of a grating with a probable *orientation* better discriminated than the *color* of a grating with an improbable *orientation*? Likewise, is the orientation of a grating with a probable color better detected than the orientation of a grating with an improbable color? A feature-tuning hypothesis might predict that only tuning the relevant feature is going to result in an effect. On the other hand, an *object-based* attentional hypothesis (e.g. Egly, Driver, & Rafal, 1994) might predict that instances of objects with the probable feature might be attended to more than if the feature is improbable.

Experiment 4.2 examined this idea. By looking at irrelevant feature probabilities and comparing the results to *Experiment 4.1* we can determine what the impact of feature-relevance (if any) might be on the probability effect, and to address whether feature probability is as domain-general as spatial probability appears to be.

4.5. Experiment 4.2: Methods

Experiment 4.2 was largely similar to *Experiment 4.1*. As before, both experiments consisted of a spatial block and a feature block, not necessarily in that order. However, for the feature blocks the discrimination task was on a feature that was not the feature that had its probability manipulated. *Experiment 4.2-col* examined the orientation probability manipulation from *Experiment 4.1-ori,* but with the color detection task of *Experiment 4.1-col*. *Experiment 4.2-ori* examined the color probability manipulation from *Experiment 4.1-col*, but with the orientation detection task of *Experiment 4.1-ori*. The spatial blocks were identical.

Participants

For *Experiment 4.2-col*, sixteen additional participants (*median* age = 21 years) were recruited (11 females, 5 males), in exchange for course credits. All reported themselves right-handed. For *Experiment 4.2-ori*, fifteen additional participants (*median* age = 21 years) were recruited (11 females, 4 males). Fourteen reported themselves right-handed.

Apart from flipping the task instructions, all aspects of *Experiment 4.2* were equivalent to *Experiment 4.1*.

4.6. Experiment 4.2: Results

Experiment 4.2-col (spatial + orientation probability, color detection)

Reaction Time

For *Experiment 4.2-col*, in the spatial block there was an effect of spatial probability on RT, with colors in the high-probability locations being detected faster (*M* = 482ms, *SD* = 71ms), than in the low-probability locations (*M* = 501ms, *SD* = 67ms), (*t*(15) = 5.14 , *p* < .001). ITs were significantly affected by spatial probability, with responses to colors in the high-probability locations being initiated faster (*M* = 392ms, *SD* = 71ms), than at the low-probability locations (*M* = 407ms, *SD* = 71ms), (*t*(15) = 4.61 , *p* < .001). However, there was no effect of spatial probability on MTs, with movement times associated with high-probability locations (*M* = 89ms, *SD* = 31ms) not being significantly different from low-probability locations (*M* = 95ms, *SD* = 30ms), (*t*(15) = 1.34 , *p* = .201). Force of the trigger pull was also not affected across high (*M* = 0.002) and low (*M* = 0.003) probability locations, (*t*(15) = 0.97 , *p* = .349).

For the feature (orientation) probability block there was no significant effect of probability on RT, with the color of high-probability orientations (*M* = 478ms, *SD* = 52ms), taking as long as the low-probability orientations (*M* = 478ms, *SD* = 58ms) to detect, (*t*(15) = 0.14 , *p* = .887). Unsurprisingly, neither IT nor MT nor force was affected by orientation probability (all *p*s > .05). The Bayesian tests of orientation probability on RT in *Experiment 4.1-ori* (*BF* = 798183 [± <0.01%]) and *Experiment 4.2-col* (*BF* = .258 [± <0.01%]) clearly support opposing hypotheses.

Accuracy and Vacillations

For *Experiment 4.2-col*, in the spatial block there was an effect of spatial probability on final accuracy, with colors in the high-probability locations being responded to more accurately (*M* = 95.7%, *SD* = 2.9%), than in the low-probability locations (*M* = 93.0%, *SD* = 5.2%), (*t*(15) = 3.57 , *p* = .003). There was no effect of spatial probability on the amount of vacillations, with colors in the high-probability locations being as directly responded to (*M* = 0.136, *SD* = 0.203), as in the low-probability locations (*M* = 0.153, *SD* = 0.214), (*t*(15) = 1.53 , *p* = .148). Vacillations improved initial to final accuracy (*p*s < .05), this improvement was not significantly different across high (*M* = 4.7%) and low (*M* = 5.2%) probability locations, (*t*(15) = .55 , *p* = .589).

In the feature block there was no effect of orientation probability on final color accuracy, with colors of high-probability orientations being responded to as accurately (*M* = 95.2%, *SD* = 2.6%), as the low-probability orientations (*M* = 95.5%, *SD* = 2.9%), (*t*(15) = 0.49 , *p* = .626). As with RT, the Bayesian tests of orientation probability on color accuracy in *Experiment 4.1-ori* (*BF* = 1141 [± <0.01%]) and *Experiment 4.2-col* (*BF* = .284 [± <0.01%]) clearly support opposing hypotheses. There was no significant effect of orientation probability on the amount of vacillations, or on the improvement of initial to final accuracy (all *p*s > .05).

Experiment 4.2-ori (spatial + color probability, orientation detection)

Reaction Time

For *Experiment 4.2-ori*, there was an effect of spatial probability on RT, with orientations in the high-probability locations being detected faster (*M* = 590ms, *SD* = 67ms), than in the low-probability locations (*M* = 655ms, *SD* = 82ms), (*t*(14) = 4.70 , *p* < .001). ITs were significantly affected by spatial probability, with responses to orientations at the high-probability locations being initiated to faster (*M* = 479ms, *SD* = 92ms), than at the low-probability locations (*M* = 538ms, *SD* = 105ms), (*t*(14) = 4.49, *p* < .001). However, there was no effect of spatial probability on MTs, with movement times associated with high-probability locations (*M* = 111ms, *SD* = 54ms) not being significantly different from low-probability locations (*M* = 117ms, *SD* = 54ms), (*t*(14) = 1.51 , *p* = .154). Force of the trigger pull was also not affected across high (*M* = 0.002) and low (*M* = 0.001) probability locations, (*t*(14) = 1.05, *p* = .308).

For the feature (color) probability block there was no significant effect of probability on RT, with the orientation of high-probability colors (*M* = 105ms, *SD* = 47ms), taking as long as the low-probability colors (*M* = 111ms, *SD* = 52ms) to detect, (*t*(14) = 1.71 , *p* = .109). Unsurprisingly, neither IT nor MT nor force was affected by color probability (all *p*s > .05). The Bayesian tests of color probability on RT in *Experiment 4.1-col* (*BF* = 9.06 [± <0.01%]) and *Experiment 4.2-col* (*BF* = 0.85 [± 0.01%]) clearly support opposing hypotheses. As with *Experiment 4.1*, it was again evident that discriminating of orientations (*Experiment 4.2-ori*) was in general associated with a larger RT than the discrimination of colors (*Experiment 4.2-col*), for the spatial and feature blocks (both *ps* < .001).

Accuracy and Vacillations

For *Experiment 4.2-ori*, there was an effect of spatial probability on final accuracy, with orientations at the high-probability location being responded to more accurately (*M* = 94.9%, *SD* = 3.3%), than at the low-probability location (*M* = 92.2%, *SD* = 6.5%), (*t*(14) = 2.28 , *p* = .039). There was no effect of spatial probability on the amount of vacillations, with orientations in the high-probability locations being as directly responded to (*M* = 0.202, *SD* = 0.295), as at the low-probability locations (*M* = 0.206, *SD* = 0.267), (*t*(14) = 0.27 , *p* = .790). Vacillations improved initial to final accuracy (*p*s < .05), this improvement was not significantly different across high (*M* = 5.7%) and low (*M* = 6.2%) probability locations, (*t*(14) = .77 , *p* = .453).

In the feature block there was no effect of color probability on final orientation accuracy, with orientations of high-probability color being responded to as accurately (*M* = 92.5%, *SD* = 6.6%), as the low-probability colors (*M* = 93.8%, *SD* = 5.1%), ($t(14) = 1.00$, $p = .336$). As with RT, the Bayesian tests of color probability on orientation accuracy in *Experiment 4.1-col* (*BF* = 2399 [± <0.01%]) and *Experiment 4.2-ori* (*BF* = .402 [± 0.01%]) clearly support opposing hypotheses. There was no significant effect of color probability on the amount of vacillations, or on the improvement of initial to final accuracy (all *p*s > .05). As with *Experiment 4.1*, discriminating of orientations (*Experiment 4.2-ori*) was overall not significantly more or less accurate than the discrimination of colors (*Experiment 4.2-col*), for either the spatial or the feature blocks (both *ps* > .05).

Post-experiment questionnaires

As with *Experiment 4.1*, most of the participants did not realize that there were probability manipulations. Of the ones that did (6 out of 31 participants), they only described the spatial probability manipulation accurately.

4.7. Experiment 4.2: Discussion

The spatial probability effects from *Experiment 4.1* were replicated. Even though the location was not the dimension that participants had to report, spatial probability effects on RT and accuracy were robust. The effects of spatial probability stem from modulations in IT rather than MT. The same does not hold true for either form of feature probability. When relevant for the task, the effects of orientation or color probability on RT and accuracy are robust (*Experiment 4.1*). When irrelevant to the task, their effects are completely eliminated (*Experiment 4.2*). This pattern of results is inconsistent with an object-based attention account (e.g. Egly, Driver, & Rafal, 1994). For example, in *Experiment 4.1*, the contingency of location and color has an apparent effect on color detection. If this is because blue objects, being the expected color for that location, are ‘attended’ to, orientations shown in blue should also be better detected as well. Instead, *Experiment 4.2* suggests that color probability creates no advantage in orientation detection of colored stimuli.

At first glance, the findings from *Experiment 4.2* seem counter to that demonstrated in the visual search literature. For example, colour probability affects search RT even if the response is about orientation (Sha, Remington, & Jiang, 2017). However, colour probability is known to affect search efficiency (Cort & Anderson, 2013): If the target is likely to possess a certain colour, the RT advantage might be due to participants prioritizing the stimuli possessing that colour in the search, rather than focusing on the stimuli that have non-probable colours. With the current task, stimuli are only presented one at a time, and therefore color probability is irrelevant to orientation discrimination.

Perhaps a feature-based attention account can explain the differences between *Experiments 4.1* and 4.*2*? If orientation probability is irrelevant to color discrimination, there is no need to attend to orientations, explaining the null result in *Experiment 4.2-col*. However, this argument cannot apply to *Experiment 4.2-ori*. The stimuli used were spatial gratings, and while color can be deciphered without having to process orientation information, the reverse is not true: Orientations can be deciphered only by processing (and attending to) the color-bands. This idea that orientation perception is more involved than color processing is supported by the finding that orientation discrimination consistently took longer than color discrimination. Yet, color probability did not affect orientation discrimination on any of the metrics.

While spatial probability creates object-general effects, feature probability creates feature-specific effects. This comports with the hypothesis that orientation probability shares some mechanisms with orientation training effects (such as that that has been demonstrated in monkeys). The tuning functions of orientation selective neurons in V1 that prefer trained orientations are ‘sharpened’ (Ringach, Hawken & Shapley, 1997; Schoups et al., 2001). A similar mechanism could also occur for other feature dimensions. Color-selective neurons also exist in V1 (Hanazawa, Komatsu, & Murakami, 2000; Johnson, Hawken, & Shapley, 2008), and the tuning process for color-probability could mechanistically work in the same way as for orientation-probability, only that a different pool of V1 neurons might be involved.

The non-selective benefits of spatial bias can be understood as a gain mechanism (e.g. Carrasco, 2011; Ling, Liu & Carrasco, 2009). Unlike with orientation probability (Ringach, Hawken & Shapley, 1997) where only the feature-relevant neurons are facilitated, spatial cues provide no information about the *features* of the upcoming target stimuli. These are the likely reasons why spatial probability improves detection, but not orientation precision (Chapter 3), particularly since it has been argued that tuning changes are required for perceptual precision to change (Yaeli & Meir, 2010).

If spatial probability affects detection rather than precision, an extension of the current task that emphasizes discrimination over detection should further differentiate spatial and feature probability effects. In *Experiment 4.3,* a restricted 41 degree range of orientations that bounded a 45 degree orientation was used. The task was to discriminate the more vertical and more horizontal orientations using the same triggers used in the earlier experiments. In addition, this manipulation allowed us to assess perceptual difficulty on a continuous scale by looking at accuracy rates for orientations nearer or further away from the 45 degree boundary. For example, in response to a 44 degree grating, spatial probability might increase the firing rate not just of neurons that prefer a 44 degree stimulus, but *all* neurons coding for that space. Given that the neurons preferring the 44 and 46 degree tilts likely have similar firings to begin with, non-selective gain process should not be expected to help the discriminate between the two tilts, since both might be equally gained, resulting in poor performance in this task. In contrast, neural tuning that result from perceptual learning is selective (Ringach, Hawken & Shapley, 1997; Schoups et al. 2001), and can lead to better perceptual discrimination (Gilbert, 1994; Yaeli & Meir, 2010). As a result, one might expect that orientation probability improves orientation discrimination even in cases where spatial probability cannot.

4.8. Experiment 4.3: Methods

Experiment 4.3 largely followed the design of *Experiment 4.1-ori*. Orientation and spatial probability were examined with an orientation discrimination task. The stimuli were shown in the same two locations as in the previous experiments, and in the same two colors. The main difference was that instead of having just two discrete possible orientations (45 or 135 degrees), orientations occupied the limited span of 25 to 65 degrees (any integer value, except for the middle 45 degree tilt was possible). Using the two triggers, participants were asked to discriminate clockwise and anti-clockwise from 45 degrees. Orientation probability was manipulated by making either the anti-clockwise (25-44 degrees) or clockwise (46-65 degrees) orientations more probable (80% occurrence). As before, this was counterbalanced by location.

Participants

Thirty-four additional participants (*median* age = 19 years) were recruited (26 females, 8 males), in exchange for course credits. Thirty-three reported themselves right-handed. The sample size was doubled for this experiment because we wanted to better examine the effect of probability across the different possible orientations. Data from one participant was dropped because of chance-level accuracy (beyond 2.5 *SDs* of the mean of other participants).

4.9. Experiment 4.3: Results

Reaction Time

In both the spatial and feature probability blocks, RTs and ITs were much longer (ps < .05) than in previous experiments, while MTs were not (*p*s > .05). The response force was also only about 10% of those from previous experiments (*p*s <.05).

There was an effect of spatial probability on RT, with responses to orientations at the high-probability locations being faster (*M* = 851ms, *SD* = 136ms), than at the low-probability location (*M* = 955ms, *SD* = 140ms), (*t*(32) = 9.45 , *p* < .001). ITs were significantly affected by spatial probability, with responses to orientations at the high-probability locations initiated faster (*M* = 740ms, *SD* = 133ms), than at the low-probability locations (*M* = 840ms, *SD* = 149ms), (*t*(32) = 9.25, *p* < .001). However, there was no effect of spatial probability on MTs, with movement times associated with high-probability locations (*M* = 110ms, *SD* = 36ms) not being significantly different from low-probability locations (*M* = 115ms, *SD* = 55ms), (*t*(32) = 0.81, *p* = .422). Force of the trigger pull was affected, with high probability locations being responded to more forcefully (*M* = 0.00018) than low probability locations (*M* = 0.00016), (*t*(32) = 6.63 , *p* < .001).

There was also an effect of feature (orientation) probability on RT, with high-probability orientations being detected faster (*M* = 839ms, *SD* = 133ms), than low-probability orientations (*M* = 881ms, *SD* = 140ms), (*t*(32) = 4.04 , *p* < .001). ITs were significantly affected by orientation probability, with responses to high-probability orientations being initiated faster (*M* = 730ms, *SD* = 121ms), than at the low-probability orientations (*M* = 767ms, *SD* = 140ms), (*t*(32) = 3.77, *p* < .001). However, there was no effect of orientation probability on MTs, with movement times associated with high-probability orientations (*M* = 110ms, *SD* = 44ms) not significantly different from low-probability orientations (*M* = 114ms, *SD* = 56ms), (*t*(32) = 0.89 , *p* = .379). Force of the trigger pull was also affected, with high probability orientations being responded to more forcefully (*M* = 0.00019) than low probability orientations (*M* = 0.00017), (*t*(32) = 4.09, *p* < .001).

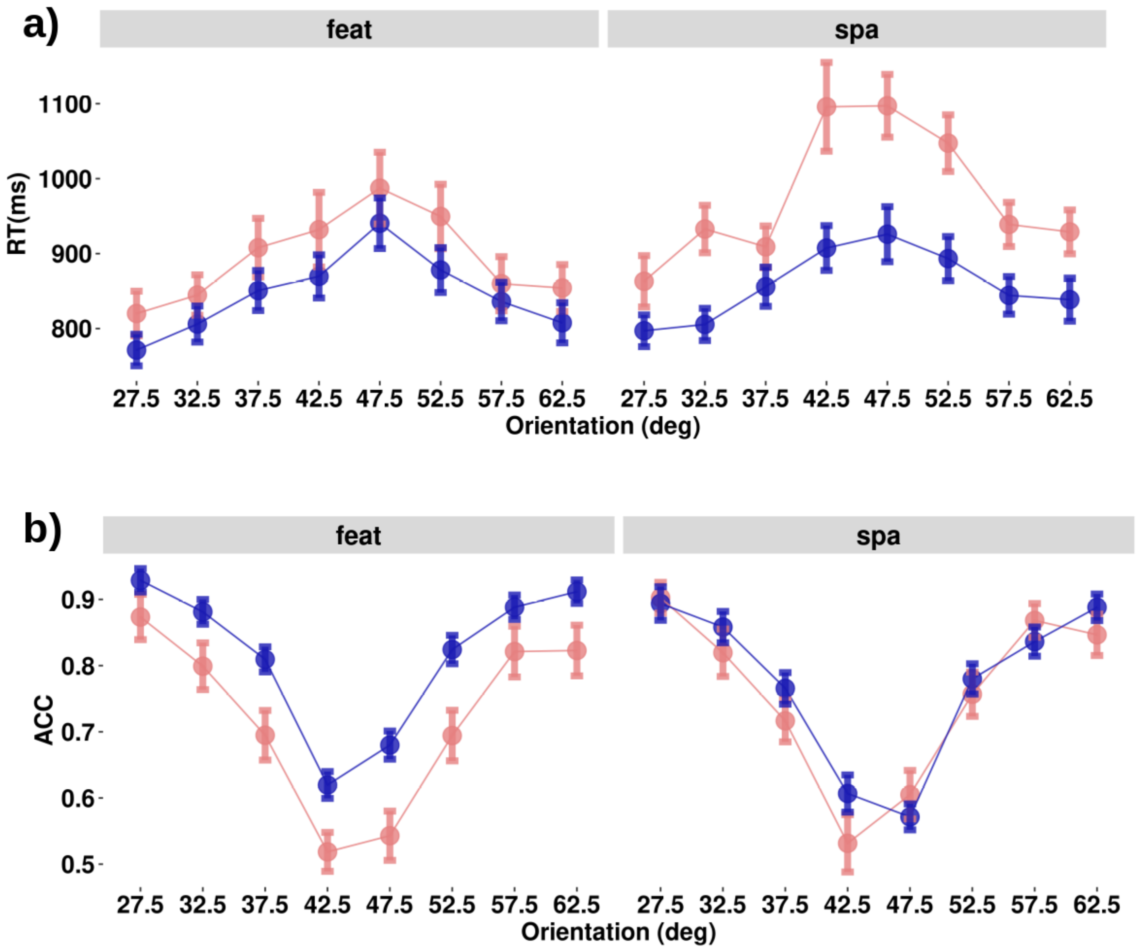

Figure 4.5. *Experiment 4.3* performance as a function of orientation and probability. a) Reaction Time and b) Accuracy across orientations in Experiment 4.3. Blue (dark) = High-probability; Red (light) = Low-probability. Error bars represent one standard error. The middle bins correspond to the orientations near the discrimination boundary (45 degrees).

Accuracy and Vacillations

In both the spatial and feature probability blocks, accuracy was reduced from previous experiments, as one would predict with a more difficult task (*p*s < .05). Both vacillations and the improvements from initial to final accuracy were also reduced (*p*s <.05).

There was a marginal effect of spatial probability on final accuracy, with orientations in the high-probability locations being responded to slightly more accurately (*M* = 77.6%, *SD* = 10.3%), than in the low-probability locations (*M* = 75.5%, *SD* = 12.5%), (*t*(32) = 1.80 , *p* = .082). There was no effect of spatial probability on the amount of vacillations (*M* (high) = 0.091, *SD* = 0.079; *M* (low) = 0.093, *SD* = 0.08), *t*(32) = 0.38 , *p* = .706). Vacillations improved initial to final accuracy (*p*s < .05), but this improvement was not significantly different across high (*M* = 1.3%) and low (*M* = 0.8%) probability locations, (*t*(32) = 1.48 , *p* = .148).

In contrast, there was a significant effect of feature (orientation) probability on final accuracy, high-probability orientations being responded to more accurately (*M* = 82.0%, *SD* = 7.5%), than the low-probability orientations (*M* = 72.6%, *SD* = 15.2%), (*t*(32) = 4.89 , *p* < .001). This accuracy increase was larger than in other experiments (*p*s < .05). There was no effect of orientation probability on the amount of vacillations, with high-probability orientations being as directly responded to (*M* = 0.096, *SD* = 0.078), as the low-probability orientations (*M* = 0.099, *SD* = 0.088), (*t*(32) = 0.38 , *p* = .705). Vacillations improved initial to final accuracy (*p*s < .05), but this improvement was not significantly different across high (*M* = 0.9%) and low (*M* = 1.2%) probability orientations, (*t*(32) = 0.81 , *p* = .423).

Orientation Binning

As a final set of analyses, the different sets of orientations were binned. In the RTs for the feature (orientation) block (*Figure 4.5a*), there was a significant main effect of probability ($F(1,30) = 17.4$, $MSE = 17766$, $p < .001$), a significant main effect of orientation bin ($F(7,223) = 9.25$, $MSE = 21124$, $p < .001$), and no significant interaction effect ($F(7,222) = 0.296$, $MSE = 13368$, $p = .955$). On the accuracy scores (*Figure 4.5b*) there was a significant main effect of probability ($F(1,32) = 26.9$, $MSE = 0.046$, $p < .001$), a significant main effect of orientation bin ($F(7,224) = 45.32$, $MSE = 0.022$, $p < .001$), and no significant interaction effect ($F(7,224) = 1.22$, $MSE = 0.011$, $p = .292$).

For the spatial block there was a significant main effect of probability ($F(1,31) = 91.5$, $MSE = 19640$, $p < .001$), significant main effect of orientation bin ($F(7,223) = 17.1$, $MSE = 17092$, $p < .001$), and a significant interaction effect ($F(7,222) = 2.2$, $MSE = 16079$, $p = .018$). To qualify this interaction, the maximum effect of bins was looked at for each of the four experimental conditions. This change in RT due to orientation bins was significantly larger for the low-probability locations (mean of 250ms), and opposed to the changes in RT due to orientation bins for the high-probability locations (120ms), for low probability orientations (140ms), and for high-probability orientations (130ms), (all ps $< .05$). Changes in RT due to orientation bins among the latter 3 conditions were not significantly different from each other (ps $> .05$).

On the accuracy scores while there was a significant main effect of orientation bin ($F(7,224) = 51.09$, $MSE = 0.021$, $p < .001$), there was no significant main effect of probability ($F(1,32) = 2.8$, $MSE = 0.017$, $p = .103$), and no significant interaction effect ($F(7,224) = 1.84$, $MSE = 0.014$, $p = .080$).

Post-experiment questionnaires

Again, most of the participants did not realize that there were probability manipulations. Of the ones that did (4 out of 33 participants), they only described the spatial probability manipulation accurately.

4.10. Experiment 4.3: Discussion

Experiment 4.3 was essentially a more difficult version of *Experiment 4.1-ori*. Accuracy scores confirmed the increased difficulty. RTs and ITs also increased. The increased difficulty may have also caused participants to be more deliberate as responses were made less forcefully, and with fewer vacillations. This is likely what caused differences in response force between high and low-probability locations/orientations to manifest. Binning the orientations suggests that the difficulty of the discrimination was maximal near the classification boundary, where RTs were longest and accuracy poorest.

As with *Experiment 4.1-ori*, RTs were faster for both the probable orientations and the probable locations. However, with the more difficult experiment, the effect on accuracy was wiped out for spatial probability, but magnified for orientation probability. Given the hypothesis that only orientation probability should improve perceptual precision when perceptual discrimination is difficult, it would seem puzzling why there would be an interaction between spatial probability and the orientation bins on RT. However, what should be noted is that the accuracy for the orientation discrimination approaches 50% (i.e. guessing) for the most difficult bin. Additionally, the effect of bins is much larger for the low-probability locations than for the high-probability locations or high/low orientation probability. Together these would suggest that there is a RT cost for guessing at low-probability locations rather than a benefit of perception for high-probability locations. These results match that of

Chapter 3, where orientation probability was found to affect perceptual precision, while spatial probability was only found to affect guessing.

4.11. Discussion

Probable objects are detected better (Miller & Pachella, 1973; Laberge & Tweedy, 1964; Hon, Yap & Jabar, 2013), as are objects in probable locations (Druker & Anderson, 2010; Fecteau, Korjoukov, & Roelfsema, 2009; Geng & Behrmann, 2005; Jiang et al., 2015; Rich et al., 2008; Vincent, 2011; Walthew & Gilchrist, 2006; Wolfe et al., 2007). However, using an orientation estimation task, it was previously demonstrated that manipulations of the orientation of a grating leads to effects distinct from spatial probability manipulations (Chapter 3). While both feature and location probabilities shortened reaction time, only feature (orientation) probability increased perceptual precision. Spatial probability affected detection/guess rates, with the precision of estimates of detected objects unchanged.

The current study provides additional support for the hypothesis that feature and spatial probability effects are distinct, similar to how feature and spatial attention effects are distinct (Carrasco, 2011). While the RT benefits as a whole seem superficially similar, breakdown of the RTs revealed that spatial probability consistently affected only the initiation times. In contrast, feature (orientation/color) probability also modulated the movement times associated with the trigger presses, at least when the detection task was easy (*Experiment 4.1*). These data highlight that participant ‘choices’ are often compressions of perceptual and response processes into binary (on/off) button responses. This can create problems when attempting to infer mechanisms from the shape of the RT distributions when a standard keyboard response is used, since it would be indecipherable whether the changes in RT derived from changes in perceptual or in response processes.

As a behavioral measure, initiation times might better capture the perceptual component of ‘detection’ time (Smeets, Wijdenes, & Brenner, 2016). On the other hand movement time may more closely track confidence (Jabar & Anderson, 2015). Shorter movement times suggest that feature probability affects both perception and confidence, the latter of which is diluted when the task is very difficult and confidence generally low (*Experiment 4.3*). Increased confidence is not due to pre-stimulus response preparation since probable features were spatially-contingent, and as spatial locations in the feature block was equiprobable, there was no opportunity to prepare a particular response prior *to a stimulus onset*. In addition, we found that there were no changes to response preparation prior to the point of response initiation, at which point the initial accuracy and IT already demonstrated effects of spatial and feature-probability.

Differences between the probability types extend beyond RT. In *Experiment 4.1*, feature probability had a pronounced effect on accuracy. This was amplified when the task was made more perceptually difficult. In *Experiment 4.3*, only feature probability affected accuracy. This is likely related to the idea that feature probability improves perceptual precision, while spatial probability does not (Chapter 3). While spatial probability might help with detection on a coarse level (*Experiments 4.1* and *4.2*), it does not aid in discriminating between similar orientations. This suggests separate mechanisms, particularly as spatial probability effects readily manifest when the response is not dependent on location. Feature probability only seems to manifest when the probable feature is the basis for the perceptual judgment.

What are these mechanisms? Space-based attentional manipulations are thought to invoke gain-mechanisms (Carrasco, 2011), such as increasing the input baseline levels of neural activity (Cutrone, Heeger, & Carrasco, 2014). Spatial probability has been found to modulate the posterior parietal cortex, a component of the dorsal attention network, as well as the frontal eye fields (FEF; Tseng et al., 2013). Spatial cues also activate the FEF, and the activity lingers even after the cue disappears, suggesting a role in spatial attention and the maintenance of spatial information (Armstrong, Chang, & Moore, 2009). Because spatial probability and spatial cuing share these higher-order regions, it is likely that spatial probability also modulates the perceptual processing regions like V1 in a 'top-down' and domain-general manner.

Feature probability, on the other hand, improves perceptual precision and is relevance-specific. In visual search, target color probability has been shown to affect the speed of judging the targets orientation (Sha, Remington & Jiang, 2017), presumably by helping in discriminating or prioritizing the target from distractors. However, in our *Experiment 4.2* color probability was not relevant to the orientation of the stimulus, since only the target was on-screen. Not only is this relevance-specificity opposing the trends seen with spatial probability, it is also in contradiction to an object-based attention account (e.g. Egly, Driver, & Rafal, 1994), since attending to the object with the probable feature should facilitate detection of its other features. An alternative to object-based attention is feature-based attention, which has been suggested to result in neural tuning (Carrasco, 2011; David et al., 2008; Paltoglou & Neri, 2012; Çukur et al., 2013; Ling, Jehee, & Pestilli, 2015), which could explain why perceptual precision is heightened in this (*Experiment 4.3*) and previous studies (Chapters 2 and 3). It has been suggested that neural tuning is necessary to increase perceptual precision (Yaeli & Meir, 2010). However, it has also been suggested that feature-based attention may additionally employ a gain mechanism, *in addition* to neural tuning (Ling, Liu & Carrasco, 2009). The contributions of top-down and bottom-up processes to tuning changes are presently an open question. Arguing that bottom-up changes may be sufficient are models of orientation training that do not have an ‘attentional’ component (e.g. Carandini & Ringach, 1997; Teich & Qian, 2003), and prior experiments where spatial attention and orientation probability had independent effects (Chapter 2).

Rather than an attentional account, the feature probability effects observed here can be parsimoniously explained as experience-dependent neural tuning. This would be similar to how orientation training in monkeys results in the ‘sharpening’ of the orientation selective neurons in V1 that prefer the trained orientations (Ringach, Hawken & Shapley, 1997; Schoups et al. 2001). This idea is also consistent with the suggestion that learnt likelihoods are reflected in the early phase of sensory processing (Summerfield & Egner, 2009), and that the site of plasticity must involve early cortical processing regions with narrow neural tuning (Gilbert, 1994).

As an alternative explanation, with feature probabilities in the current task, a participant could be reinforced over time to respond with the probable trigger given the location in which the stimulus was presented. While it is clear from the data that there was no *physical* or *motor* response preparation associated with feature probabilities, target prevalence has been suggested to affect a decision criterion (Wolfe & Van Wert, 2010), and perhaps this is true for probabilities associated with individual features as well. However this does not rule out a perceptual locus, since feature probability has also been shown to affect perceptual precision in non-2AFC tasks (e.g. Chapters 2 and 3). It could also be the case that probabilistic information for features might tune relevant perceptual channels, which *then* affect decision-making (Eckstein et al., 2009). Chapter 5 sought to provide electrophysiological evidence for the perceptual hypothesis, and provides a hint that there might be a relation between perceptual and decisional modulation.

Chapter 5: An Electrophysiological Investigation of Orientation Probability[9]

5.1. Introduction

The previous chapters have suggested that orientation probability effects are unlikely to occur through the same mechanisms as those involved in spatial cuing or spatial probability, likely gain-mechanisms (Carrasco, 2011), such as increasing the input baseline of neural responses (Cutrone, Heeger, & Carrasco, 2014). Instead, the suggested mechanism driving orientation-probability effects is the selective tuning of orientation-selective neurons in V1. This would be similar to how training monkeys with specific orientations affects the tuning width of V1 neurons preferring the trained orientation (Ringach, Hawken & Shapley, 1997; Schoups et al., 2001). Axiomatically, selective orientation training is a manipulation of orientation-probability. This idea is also consistent with the suggestion that learnt likelihoods are reflected in the early phase of sensory processing (Summerfield & Egner, 2009), and that the site of plasticity must involve early cortical processing regions with narrow neural tuning (Gilbert, 1994).

An alternative account is that orientation probability is due to feature-based attention, although this is also thought to result in neural tuning (David et al., 2008; Paltoglou & Neri, 2012; Çukur et al., 2013; Ling, Jehee, & Pestilli, 2015). For example, the pairing of an auditory cue with a particular orientation leads to both a decrease in fMRI activation and a concurrent increase in the sensory representation demonstrable by multi-voxel pattern analysis (Kok, Jehee & de Lange, 2012). Presumably, even in the absence of an explicit cue, orientation probability could yield a similar pattern if it also produced neural sharpening.

---

[9] A version of this chapter has been published as Jabar, S. B., Filipowicz, A. & Anderson, B. (2017). Tuned by experience: How orientation probability modulates early perceptual processing. *Vision Research*, *138*, 86-96. https://doi.org/10.1016/j.visres.2017.07.008

If orientation probability affects perception via sharpening the response profiles of neurons in early visual cortex, we should see both physiological changes in the early sensory systems tracking stimulus probability and in behavioural measures of perceptual precision. By employing spatial gratings in varying visual field locations we can study the effects of orientation probability on early visual processing via the ‘C1’ ERP component. The C1 is an early visual evoked component that has a unique visual field-dependent phase reversal. For this reason it has often been proposed as an ERP reflecting V1 activity, and Di Russo et al. (2002) used source localization methods to localize the C1 to the banks of the calcarine cortex. While a V1 source for the C1 has been traditional, some groups have more recently suggested that the C1 is less specific than V1 because forward-modelling suggests that either V2 and V3 might also result in the phase reversal property (Ales, Yates, & Norcia, 2010). However, the assumptions required for a V2/V3 C1 locus has been questioned, and also does not account for why there is a C1-V1 relation in primate neurophysiology (see Kelly, Schroeder, & Lalor, 2013).

Another interesting feature of the C1 component is the suggestion that it indicates feed-forward processing only. With a peak latency between 90-110ms the C1 has been felt to occur too early for significant top-down trial-specific modulation, unlike the P1, a later visual ERP (Di Russo, Martinez, & Hillyard, 2003). In addition, *endogenous* manipulations of *spatial* attention affect the P1 but not the C1 (Clark & Hillyard, 1996; Di Russo et al., 2003; Gonzalez, Clark, Fan, Luck, & Hillyard, 1994). However, the feed-forward only claim for the C1 is not universally accepted, and some recent data challenge this claim (for a review see Rauss, Schwartz, & Pourtois, 2011). For example, Kelly, Gomez-Ramirez, & Foxe (2008) found that endogenous manipulation of spatial attention *did* effect both the C1 and P1. Studies on the effects of *feature*-attention on these ERP components are considerably scarcer, although it seems that at least the P1 is affected. For example, when participants are instructed to attend to red instead of green stimuli, P1 is amplified, even if the stimuli appeared in an unattended *location* (Zhang & Luck, 2009).

There is also some disagreement about the *direction* of the effects different manipulations may have on the C1 potential. Kelly et al. (2008) reported that endogenous spatial attention *increased* C1 amplitude, consistent with a spatial gain mechanism (*Figure 1.1c*). Also consistent with a gain mechanism, Bao, Yang, Rios, He, & Engel (2010) and Zhang (2015) found increases in C1 amplitude for learned orientations in perceptual learning tasks. However, data from other perceptual learning experiments have shown decreased early visual activity to learned orientations. Gratton (1997) demonstrated this with optical imaging, and Pourtois, Rauss, Vuilleumier, & Schwartz (2008) found *reduced* C1 amplitudes. This direction is predicted by a neural sharpening mechanism similar to that seen in monkeys undergoing long term orientation discrimination training (*Figure 1.1b*, Ringach, Hawken & Shapley, 1997; Schoups et al., 2001). These mixed results emphasize the need to consider both the type of training and the experimental paradigm when interpreting effects (Hung & Seitz, 2014).

While the hypothesis that orientation probability effects are due to V1 sharpening is inspired by the perceptual learning studies done in monkeys (where direct neural recordings were possible; Ringach, Hawken & Shapley, 1997; Schoups, et al. 2001), it is important to note that the tasks explored in this thesis are probability manipulations, which is procedurally distinct in many ways from traditional perceptual learning paradigms. In typical perceptual learning paradigms either a single orientation is shown repeatedly in a predictable fashion (e.g. Ringach, Hawken & Shapley, 1997; Schoups et al., 2001) or only a pair of distinct orientations are possible, one of which has the higher exposure rate (Bao et al., 2010; Zhang et al., 2015). Perceptual learning studies also usually involve a training period that spans several sessions and may go on for days or even weeks. In the probability tasks explored in the previous chapters the orientations were not ‘trained’, but rather a subset of possible orientations were biased, and this was generally undetectable by participants. These probability effects develop quickly, within a single behavioral session, faster than the typical perceptual learning experience.

Probability effects also persist despite orientations being biased in a location-contingent manner. In contrast, retinotopic transfer has been noted with perceptual learning, suggesting a top-down influence on the visual cortices (Zhang, Xiao, Klein, Levi, & Yu, 2010). Therefore, while the hypotheses regarding the direction of the C1 effect were motivated by results from the perceptual learning literature, the specific interest of the present study was to examine the C1 effects associated with a probabilistic paradigm, a paradigm where orientations are biased rather than ‘trained’.

In summary, the behavioral results of manipulating orientation probability suggest perceptual effects. This raises the possibility of an early visual locus for the effect of orientation probability. While the principal electrophysiological focus of stimulus probability has most often been the decision-related P300 (Bledowski et al., 2004; Polich, 1990; Rohrbaugh, Donchin, & Eriksen, 1974), it is hypothesized that early visual cortical potentials could also show orientation probability effects and that based on orientation training experiments in monkeys that the nature of the effects would be decreases in waveform amplitudes.

5.2. Experiment 5.1: Methods

Since the ‘C1’ ERP component is the key variable of interest, the current study broadly replicated the design of Di Russo and colleagues (2002) while introducing the probability manipulation and estimation task outlined in Chapter 2.

Participants

Twenty paid participants were recruited from the University of Waterloo (10 females, 10 males). 18 were right-handed and 2 were left-handed. Participant ages ranged from 19 to 49 (median = 24). All participants had normal or corrected-to-normal vision, and did not declare any auditory deficits or any past neurological conditions/concussions. Informed consent was obtained. Participants came in for two sessions (first behavioral; second behavioral and EEG). Both sessions were always completed in the same week (Monday-Friday), with the majority having 2 or 3 days between sessions. Participants were asked to refrain from alcohol consumption for 24 hours prior to the 2nd (EEG) session, and to maintain their regular caffeine habits. Participants were paid 25 dollars for their involvement (one 30 minute session and one 2 hour session).

Stimuli

Stimulus locations and timings resembled those used in the Di Russo et al. (2002) study. Instead of checkerboard patterns, Gabors were used to allow for the orientation manipulation. The Gabors were oriented grayscale sine-wave gratings with a circular Gaussian mask (*Figure 5.1a*), at 50% contrast, with an average measured luminance of 22cd/m^2. They had a spatial frequency of 4 cycles per degree of visual angle, and were presented on a grey background with a similar luminance of 23cd/m^2. When viewed from a distance of 60cm, the Gabors subtended approximately 4 degrees of visual angle both vertically and horizontally. On any given trial, the centre of the Gabor was located 4 degrees away from the center of the display, which was marked by a black fixation symbol. As with Di Russo et al. (2002), the four stimulus positions were centered along an arc that was located at a polar angles of 25° above and 45° below the horizontal meridian (see *Figure 5.1c*). Lines, used as feedback and for participants to rotate to report their estimations, had a length of 4 visual degrees and always occurred in the same location as the Gabor for that trial.

Spatial Gabors were equally likely to appear in any one of the four locations surrounding the fixation symbol. *Collapsed across these four locations*, any orientation was equally likely. The critical manipulation was the occurrence-rate of the various *probability-location conjunctions*. Half the participants saw the conjunction depicted in *Figure 5.1c*: When a Gabor appeared on the top or bottom left, its orientation was more likely to be left-tilting, but this high-probability orientation was reversed if the Gabor appeared on the top or bottom right. Orientations were uniformly distributed across each quadrant, and were not shown at the exact cardinal orientations (0/90 degrees). High-probability orientations accounted for 80% of the trials. The lines in *Figure 5.1c* depicts an example distribution. The location-orientation conjunctions were counterbalanced across participants.

Probability distributions were counterbalanced across participants and maintained within participants for both experimental sessions. For example, in every set of 20 trials there were 5 Gabors that appeared in the top-left location. Of these, 4 might be left-tilting while one was right-tilting (or vice-versa). Participants were *not* informed about these probability distributions, and observed the same distribution across their two sessions. The first 20 trials of the first session, and the first 60 trials in the second session were practice trials which had uniformly distributed orientations.

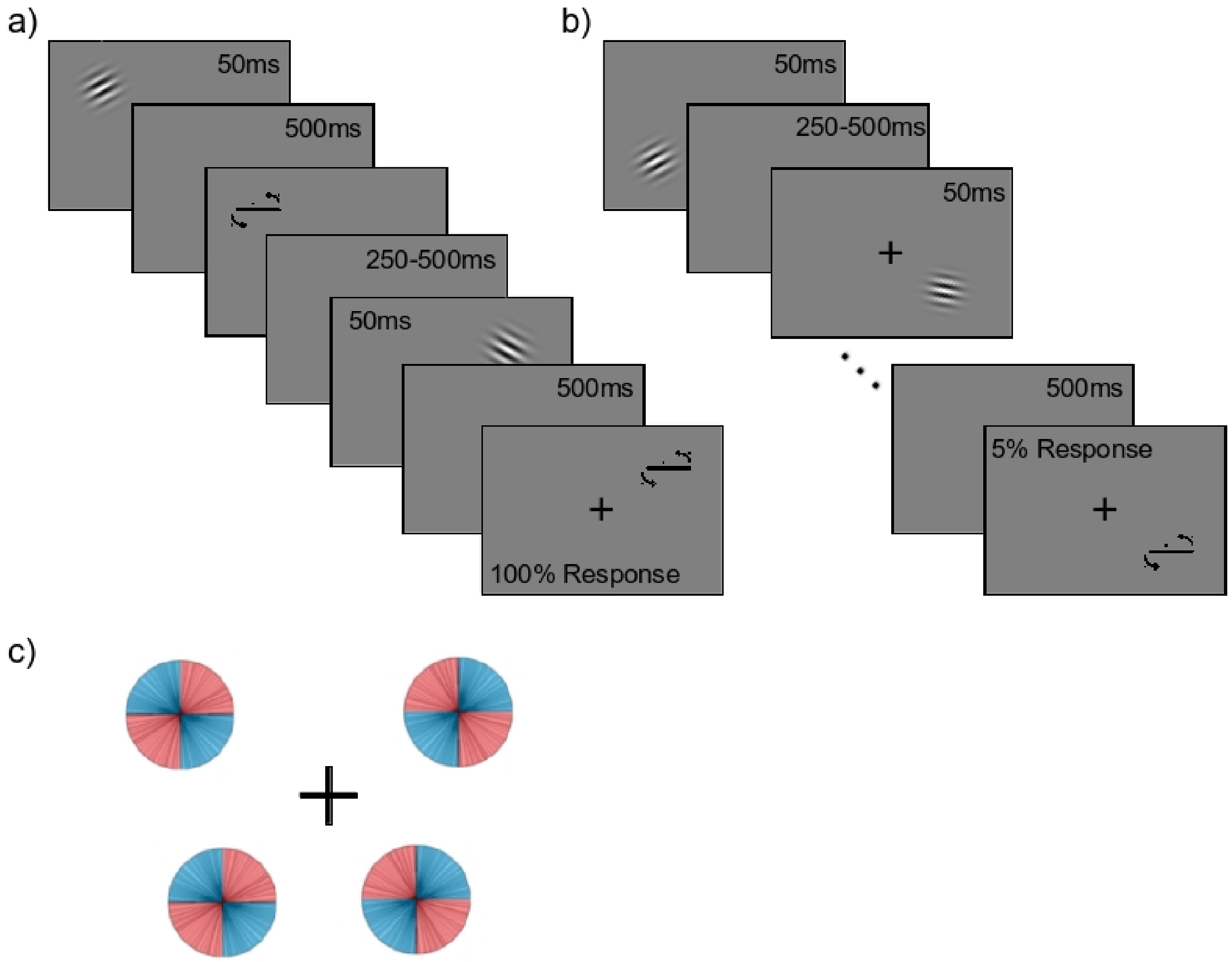

Figure 5.1. Experiment paradigm. a) Behavioural session. A Gabor stimulus was shown in one of four equiprobable locations for 50ms, followed by black screen for 500ms, followed by an estimation response. Training session trials (400 in total) always required this response. There was a variable ISI of 250-500ms to the next Gabor. b) EEG session was similar to the training session, except only 5% of the trials (3600 in total) required a response. Most of the time it was a 50ms Gabor presentation with an ISI of 250-500ms. c) Orientation probability distribution. In two locations orientations occurred in one quadrant four times more likely (high-probability: blue) than in the other (low-probability: red). The orientation probability was reversed in the other two locations. This location-orientation conjunction was counterbalanced across participants.

Auditory feedback was given after each response. A high pitched sound was played if the participant made an angular error less than 12 degrees. A lower pitch indicated an error greater than 12 degrees. Participants were not explicitly informed that 12 degrees was the error threshold.

Equipment

In both sessions, eye-tracking was done using an Eyelink1000 recording the dominant eye at 2000Hz and tracking both the pupil and corneal reflections. Participant head position was stabilized using a chin rest. Participants were instructed to fixate at the center of the screen and to avoid excessive blinking during stimulus presentations, but were free to blink or move their eyes prior to confirming the drift correction, during breaks or during responses. The task was gaze-contingent in that stimuli were not flashed until a central fixation (radius 0.5 visual angle) was detected. This was set-up so that all stimulus locations were always presented in roughly the same peripheral location retinotopically, and presumably involved a consistent pool of V1 neurons.

In the 2nd session, a BioSemi Active II system (64 channel) was used for the EEG recording. The electrode cap was placed according to the International 10-20 placement system. Additional electrodes were placed on the lateral side of each eye, inferior to each eye, and on each mastoid. EEG recording was done at 1024Hz.

Participants sat approximately 60 cm from a 32 cm x 24 cm gamma-corrected CRT monitor that refreshed at 85Hz (*mean* refresh time = 11.8ms, *SD* = 0.5ms). Prior to the task, participants were instructed to make their estimations of the Gabor orientations as accurately as they could. They were *not* told that they needed to be fast. Responses were made with a computer keyboard using their dominant hand.

Procedure

In the behavioral session (*Figure 5.1a*), every trial consisted of a Gabor flash and a response. Participants were first given 20 practice trials in which the orientations occurred uniformly in orientation. These data were not included in the analyses. The behavioral session had 400 probability-manipulated trials, which were sectioned into two blocks with a break in-between. Drift corrections were done every 20 trials (approx. 1-2 min), and an eye-tracker re-calibration was done after the break. This session took approximately 30 minutes.

In the EEG session (*Figure 5.1b*), there were 60 practice trials with uniform orientations. This was followed by 3600 actual trials, which were sectioned into three blocks with breaks in-between. For this session, responses were required on only 5% of the trials (i.e. randomly chosen out of every 20-trial segment). This was done to maximize the number of stimulus presentations (and corresponding evoked potentials), and also to maintain parity with the Di Russo et al. (2002) study, where infrequent detection responses were required. Drift corrections were done after every 6th response (approx. 2-3 min), and an eye-tracker re-calibration was done after each break. At the end of the main task, participants were given a questionnaire (same as the one used in Chapter 2). While the whole session lasted approximately two hours, the experimental task only took approximately 50-60 minutes, the rest of the time taken for electrode placement/removal.

Each spatial Gabor flashed for approximately 50ms (4 screen refreshes). If there was no response required, the next Gabor appeared between 250-500ms after (contingent on participants' fixation still being centred). On response trials, following a 500ms delay, a response line was drawn in the same spatial location as the last Gabor that had appeared. Participants made their estimations by rotating this line counter-clockwise or clockwise by pressing “Z” or “C” on a QWERTY keyboard. This rotation was at a maximum of 1 angular degree per frame refresh of the monitor. Participants pressed the “X” key to confirm their estimations. Once confirmed, the auditory feedback was given. On practice trials, a white feedback line with the actual orientation was displayed on top of the participant’s response. This visual feedback was not given during the main trials. This response procedure was the same as used in Chapter 2.

Analysis

Behavioral and eye data were analyzed using *R*. EEG data was processed using EEGLAB (Delorme & Makeig, 2004). As with Chapters 2 and 3, angular errors occupied an axial space, and the median absolute error was the primary measure. Reaction time (RT) for each trial was taken as the time from when the response line appeared to when the orientation was confirmed. Time taken to initiate movement (IT), time taken to make movements after initiation (MT), initial rotation direction and number of direction switches (vacillations) per trial were also recorded. The analyses reported used the median timings across trials (timings are positively skewed), although using the mean gave the same results.

EEG recordings used a linked-mastoid reference, and a 50ms window surrounding the onset of the stimulus (mean of -25ms to 25ms) was taken as the baseline voltage. Noisy electrodes, where the z-scored kurtosis across trials was 5 SD or greater (where most of the variance in the signal was the result of random large deviations), were removed and replaced by interpolation. Blink-contaminated trials were removed from analysis, as were trials where the electrode data had an average voltage that was greater than the mean +/- 2.5 SD of all trials. For the EEG analyses a low pass filter of 80Hz was used. There was some contamination at 85Hz, likely due to the refresh of the CRT monitor, and this was removed by an additional notch filter (a 1Hz high pass with a 30Hz low pass filter was also tried, led to the same results, and as such are not reported here). This pre-processing preserved 85% of the trials.

For the C1 ERP component, presumed to originate in V1, and therefore reflecting the anatomy of striate cortex, upper-field stimuli should show a negative deflection while lower-field stimuli should show a positive deflection (see *Figure 5.4a* here, and *Figure 5 and 6* in Di Russo et al., 2002). The electrode used for calculating C1 amplitude was the electrode showing the maximal (or minimal) amplitude of all the electrodes posterior to Cz [i.e., CP, P, PO and O], within the 50-150ms post-stimulus window. This was determined separately for each participant and for each of the four stimulus locations. This was done prior to the separation of high and low-orientation probability trials. The maximum amplitude of the trial-averaged waveform within that 50-150ms time window was used to determine the C1 amplitude for each trial-type. P300 amplitudes were determined similarly, but constrained to parieto-central electrodes [i.e. FC, C, CP, P and PO] with a positive deflection within the post-stimulus window of 250-500ms.

5.3. Experiment 5.1: Results

The first section reports the behavioural probability effects seen in the first session. Following that are the behavioral effects in the second (EEG) session, and then the effects of orientation probability on visual-evoked potentials.

Behavioral Session

There was a significant effect of probability on RT, with high-probability orientations (*M* = 2027ms, *SD* = 356ms) being confirmed faster than low-probability orientations (*M* = 2083ms, *SD* = 381ms), (*t*(19) = 2.27 , *p* = .035). This RT effect could either be due to the time required to initiate movements (IT) or the time required to complete moving the dial once movement had started (MT). ITs were significantly affected by probability (*Figure 5.2a*), with high-probability orientations (*M* = 746ms, *SD* = 107ms) initiated more quickly than low-probability orientations (*M* = 766ms, *SD* = 112ms), (*t*(19) = 4.57 , *p* < .001). Of the 20 participants, 17 showed this effect (*Figure 5.2a*). There was no effect on MT between high-probability (*M* = 1281ms, *SD* = 292ms) and low-probability orientations (*M* = 1317ms, *SD* = 318ms), (*t*(19) = 1.56 , *p* = .135). Vacillations, that is the average number of direction switches participants made in their response in a trial, was also not significantly different between high-probability (*M* = 0.11, *SD* = 0.08) and low-probability tilts (*M* = 0.13, *SD* = 0.10), (*t*(19) = 1.37 , *p* = .187).

For average (absolute) angular errors (*Figure 5.2b*) participants showed smaller errors in the behavioral session when compared to the behaviour measured in the EEG task. For the behavioral session alone, there was a marginal effect of orientation probability on precision (*Figure 5.2b*), with high-probability orientations (*M* = 8.90 deg, *SD* = 3.91 deg) being associated with smaller errors than low-probability orientations (*M* = 9.39 deg, *SD* = 4.19 deg), (*t*(19) = 2.04 , *p* = .055).

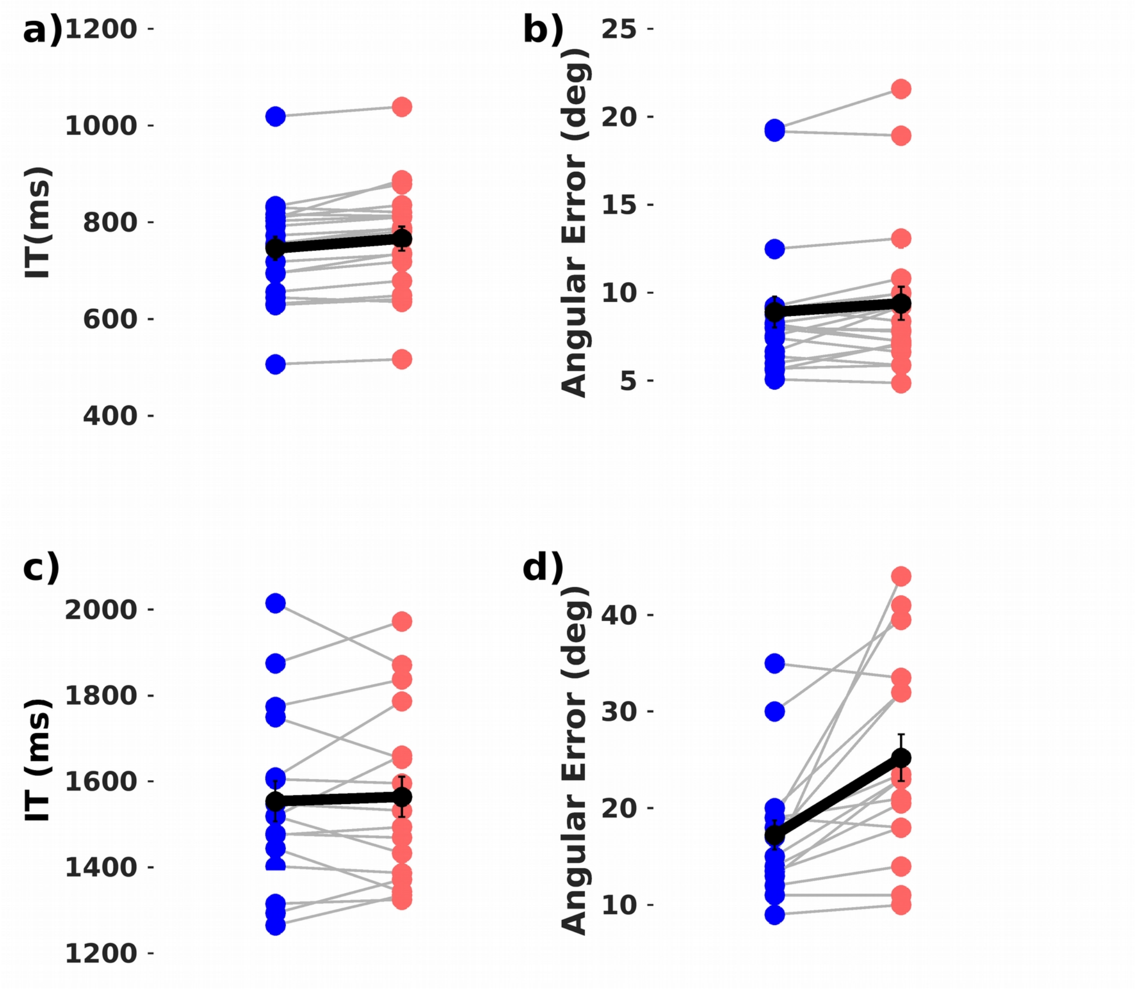

Figure 5.2. Behavioural results. a) Initiation times in the first session across subjects (separate lines). The solid black line represents the mean across subjects. b) Median absolute angular errors in the first session. c) Initiation times in the second session. d) Median absolute angular errors in the second session. High-probability orientations: blue; low-probability ones : red.

Behavioral effects in the EEG session

Two participants' data were removed because of noisy EEG signals. These participants' data were left out from the behavioural analysis to show that the trends observed were due to the remaining 18 participants that contributed to the EEG data. Unlike in the 1st session, there was no demonstrable effect of orientation probability on any aspect of reaction time (all *p*s >.05). Of interest is that the ITs (*Figure 5.2c*) for the high (*M* = 1554ms, *SD* = 199ms) and low-probability orientations (*M* = 1564ms, *SD* = 198ms) were significantly longer than in the behavioral session, (both *ps* < .001). This suggests that participants found the unpredictable, irregular behavioral assessments in the EEG session to be more difficult than the consistent behavioral probes used in the 1st session. This is supported by the precision differences (*Figure 5.2d*). Error magnitude was greater in the EEG session (both *ps* < .05) than in the 1st session. However, in the EEG session the benefit of probability on precision was significant with high-probability orientations (*M* = 17.2, *SD* = 6.40) being more precisely-estimated than low-probability orientations (*M* = 25.2, *SD* = 10.24), (*t*(17) = 4.13, *p* < .001), and 15/18 participants showing this direction of benefit.

Participants completed an open-ended post-experiment questionnaire after the second session (the one outlined in Chapter 2). Two participants mentioned being aware of a probability manipulation, neither of whom described it accurately. One participant mentioned that things that looked like '/' were more likely (which was incorrect, because they did not note the location-contingency). When participants were directly asked whether they noticed that some orientations were more likely in certain locations, they expressed either surprise or confusion. This lack of explicit awareness was also seen in the previous studies.

Effects of orientation probability on early visual-evoked potentials

An overview of the electrode data and the visual-evoked potentials at each stimulus location is given in *Figure 5.3*. The ‘C1’ component was identified based on a visual field location dependent phase reversal (*Figure 5.4a*) that matched the findings of Di Russo and colleagues (2002). Latencies were also comparable (*Figure 5.3; M* = 109ms) with that study, which is not surprising given the procedural similarities. Similar to the Di Russo et al. (2002) study, this component was maximally located at PO3/PO4, ipsilateral for the upper field stimuli and contralateral for the lower field stimuli (reflected in the topographies in *Figure 5.4a*). This was mostly consistent across participants. Two participants demonstrated a slight shift in how far lateralized the target electrodes were, PO7/PO8 were the maximal electrodes instead of PO3/PO4. Two other participants showed a slightly more posterior distribution, O1/O2 instead of PO3/PO4.

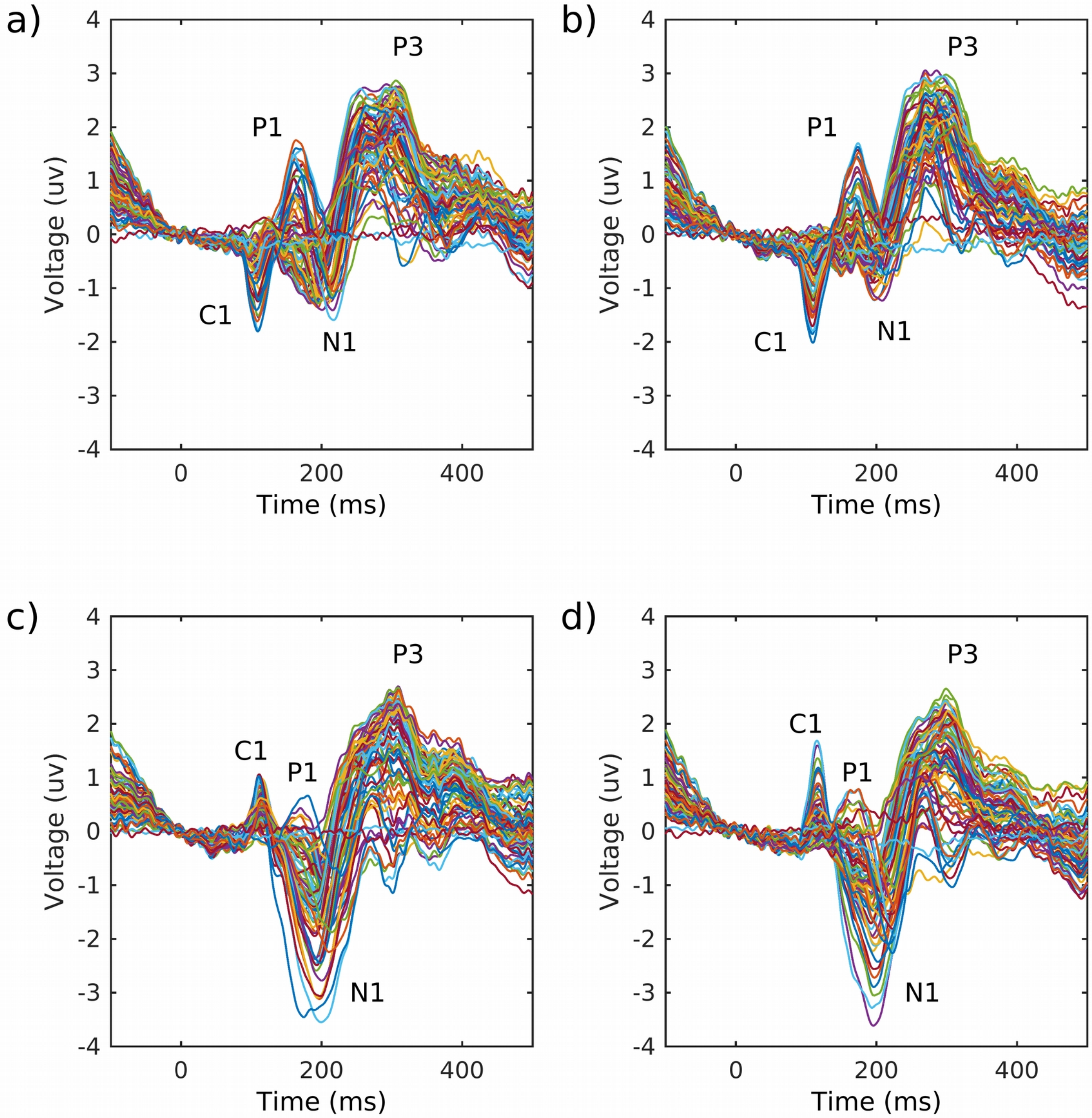

Figure 5.3. Averaged waveforms across trials for each visual field. Lines indicate different electrodes. The top-left panel represents the top-left visual field, etc. Labels indicate where the components of interest are.

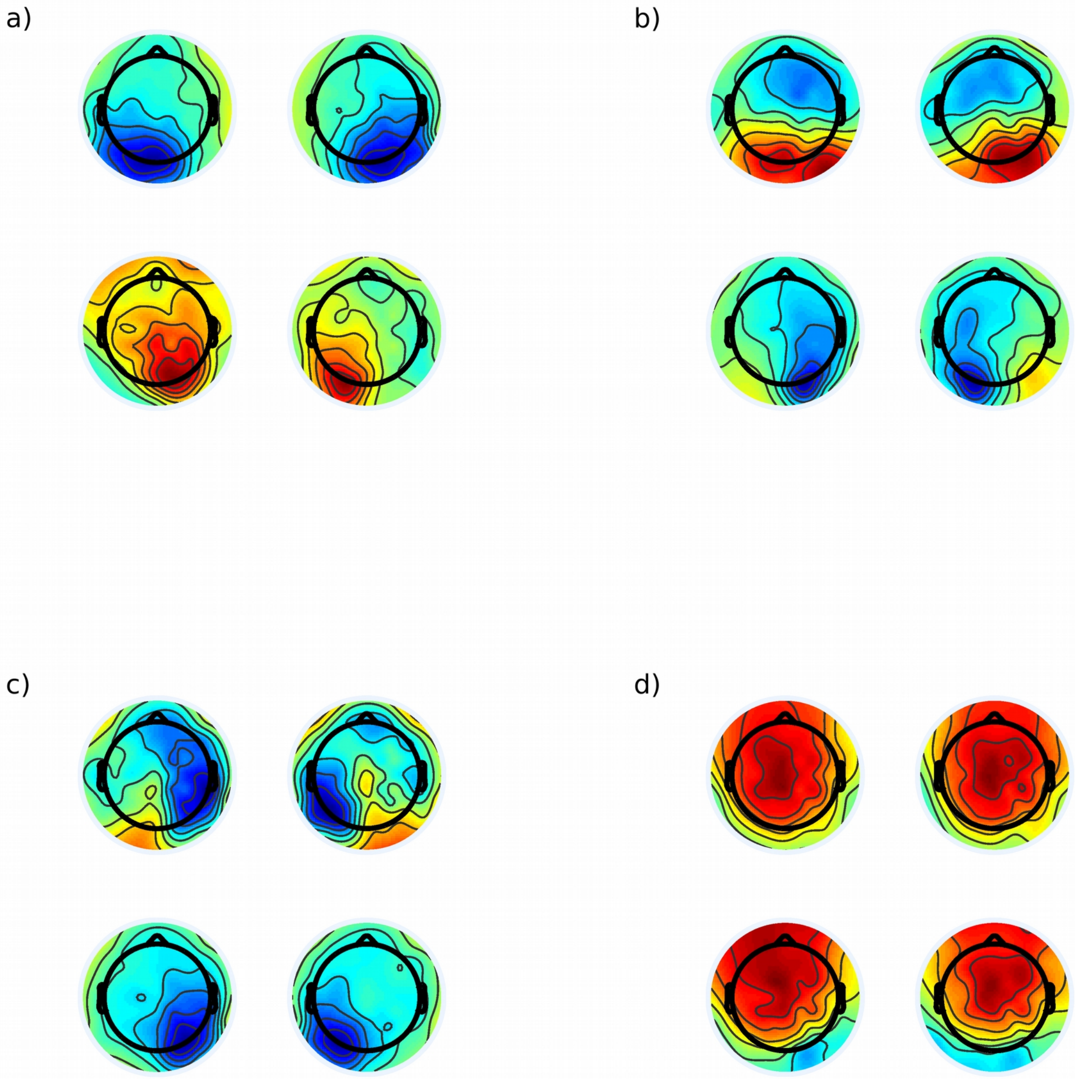

Figure 5.4. Scalp topographies. a) C1 distribution. Because the visual field in which the stimuli appears affects the direction of the C1, depicted is the average topography at 109ms post-stimulus at the corresponding location (top-left panel represents stimuli presented in the top-left, etc). b) P1 distribution at 165ms. Note that for the P1 in the lower fields, the positive peak was occluded by the concurrent N1, see *Figure 5.3*. c) N1 distribution at 200ms. d) P300 distribution at 295ms. Red = positive, blue = negative.

Having replicated the key behavioral effects of orientation probability, we turn to the key issue of the effects of orientation probability on early perceptual processing. There was a significant effect of orientation probability on the C1 amplitude that was consistent across participants (*Figure 5.5b/c*). The absolute amplitude of the C1 was significantly larger for the low-probability orientations than for the high probability orientations in *both* the upper-visual field stimuli (-2.95 vs. -2.35, *t*(17) = 3.89 , *p* = .001) and the lower-field stimuli (2.35 vs. 2.01, *t*(17) = 3.84 , *p* = .001).

Because there are four times as many high-probability trials as there were low-probability ones, sub-sampling was carried out to rule out the C1 effects as being simply due to differences in the volume of data. Because each participant only did 720 low probability trials across the four locations, in each run an equivalent number of high probability trials were randomly selected for comparison instead of the full 2880 trials. EEG analysis was otherwise the same as before (the same electrodes were used to measure the C1 amplitude, etc). This sampling was done 10000 times and in most cases, there was a significant (p<.05) difference in amplitudes between high and low probability trials. For the upper field, the mean *t*-value across these comparisons was 3.91 (*SD* = 0.72), with only 58 out of 10000 cases failing the p<.05 cutoff. For the lower field, the mean *t*-value across this sub-sampling was 3.80 (*SD* = 0.77), with only 82 out of 10000 cases failing the p<.05 cutoff.

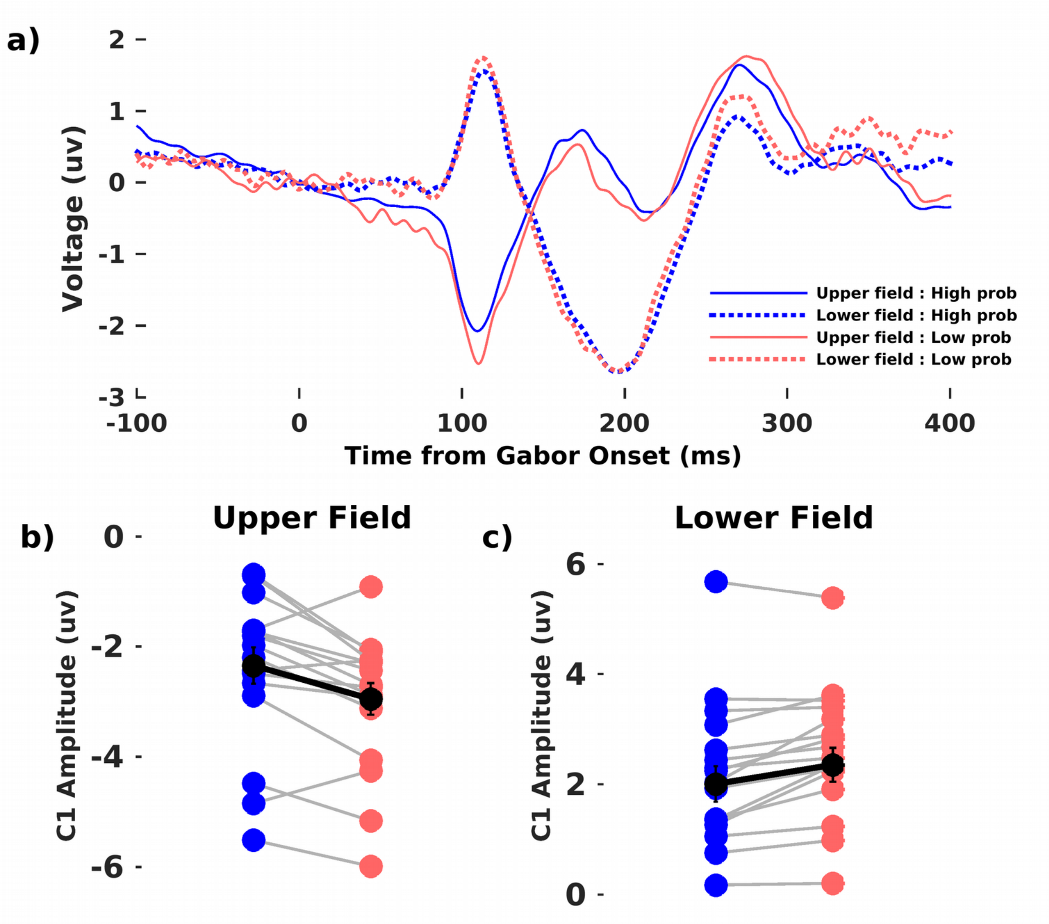

Figure 5.5. C1 analysis. C1s were calculated at the posterior electrodes that showed the maximum deflection at the 50-150ms time window, at each stimulus location. Left/right visual fields were collapsed. Defections (at about 100ms) are positive for upper-field stimuli (solid lines) and negative for lower-field stimuli (dotted lines), and split across high (blue) and low (red) probabilities. a) Average C1 wave across subjects. b) C1 amplitudes across subjects (separate lines) by probability, for upper-field stimuli. The solid line reflects the means across subjects. c) Same as previous panel, but for lower-field stimuli. C1 probability effects were consistent in both fields, with the amplitude of the low-probability trials being larger (direction is flipped due to the field-dependent nature of C1).

Repetition Suppression

To address the possibility of repetition suppression being a factor, repeats of high-probability orientations were examined. Trials of repeated high-probability orientations, but at different locations, had the same C1 amplitudes as when high probability orientations repeated at the same visual field location (*p*s > .05). Furthermore, when stimuli that repeated locations and were of similar orientations (within 20 degrees) were removed from the analyses (approximately 10% of trials), the C1 probability-based modulation remained, and was significant, in both the upper and lower visual fields (*p*s < .001).

Effects of orientation probability on P1 and N1

Because the P1 and N1 are modulated in spatial cuing tasks (e.g. Clark & Hillyard, 1996), it is of interest whether the probability manipulations affected these ‘attentional’ components in the same way.

There was a positive peak for both upper/lower visual fields in the 140-180ms time range (*Figure 5.3 & Figure 5.4b*). We looked at the electrodes showing maximal activity in this range, and found that this ‘P1’ was *not* significantly modulated by orientation probability in either the upper field (high: 1.17 vs. low: 2.19, *t*(17) = 1.50, *p* = .152) or the lower field (high: 1.38 vs. low: 1.65, *t*(17) = 1.61, *p* = .125). Removing ‘repeats’ did not change the outcome of the analyses. Also, note that the trend for the data is in the opposite direction to that found by Clark and Hillyard (1996) for attended stimuli.

A negative dip for both upper/lower visual fields occurred in the 170-230ms time range (*Figures 5.3* & 5.*4c*). We looked at the electrodes showing maximal (negative) activity in this range, and found that this ‘N1’ was *not* significantly modulated by orientation probability in either the upper field (high: -1.09 vs. low: -0.58, *t*(17) = 1.10 , *p* = .288) or the lower field (high: -0.93 vs. low: -1.02, *t*(17) = 1.09 , *p* = .293). For neither the N1 (*mean* = 200ms) nor the P1 (*mean* = 165ms) was the peak latency modulated by orientation probability (all *p*s >.05).

Time course of the C1 effect

Since the probability effect has to be acquired during the exposure to the trials, it was of interest to look at how quickly the C1 effect might occur. The 3600 trials were separated into 6 bins of 600 trials. Each bin therefore had 120 low-probability trials, with 30 in each of the four locations. Left/right amplitudes were averaged to obtain the mean amplitude for the upper and lower visual fields. Paired t-tests revealed a significant C1 probability effect (high - low probability) for the 2nd to 6th bins for lower visual field and all bins for the upper visual field (*p*s<.05, see *Figure 5.6*). The 1st bin in the lower field had a marginal probability effect, *t*(17) = 1.92, *p* = .071.

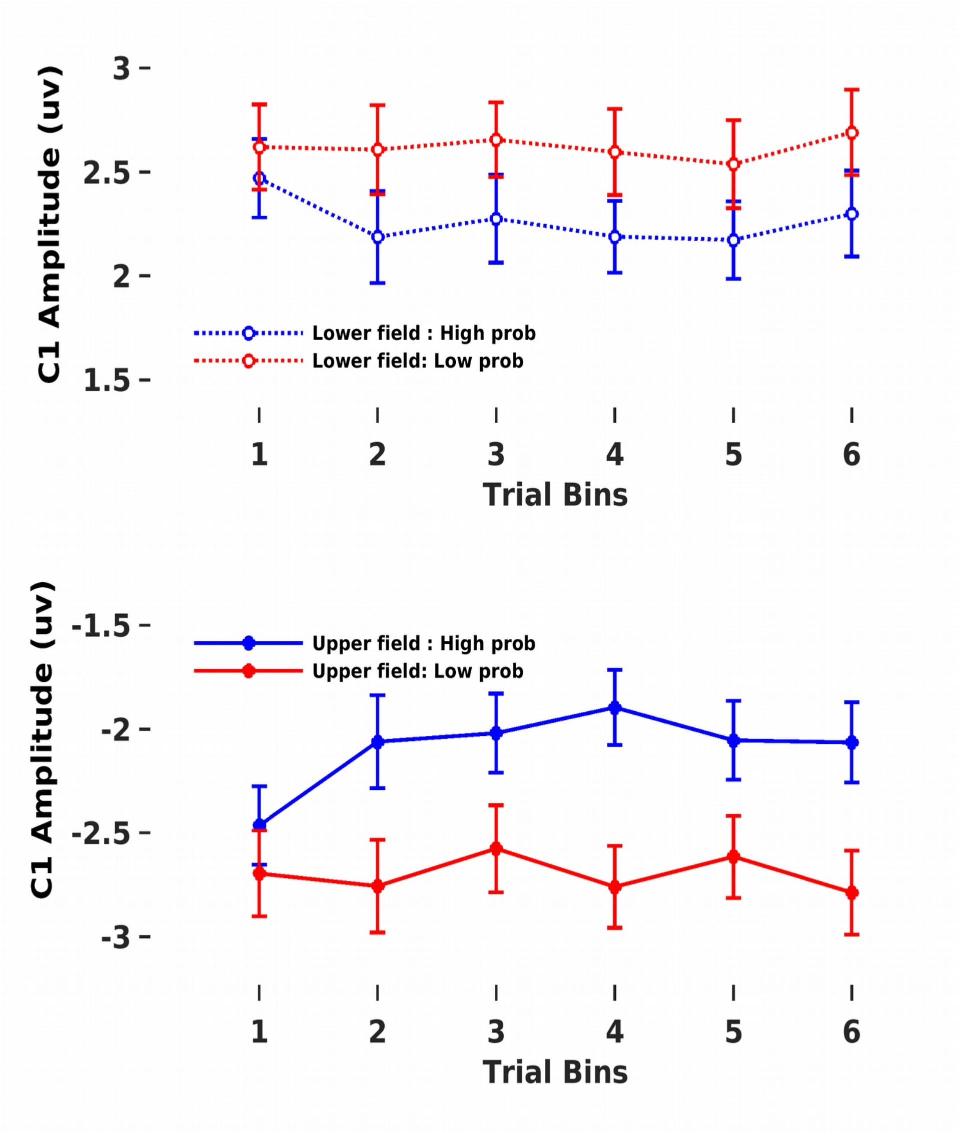

Figure 5.6. Time course analysis. The 3600 trials each participant went through was segmented into 6 bins, and the average C1 evoked potential per condition for each bin was calculated. The top panel depicts the lower visual field, and the bottom panel depicts the upper visual field. High-probability = blue, low-probability shown in red.

In the lower visual field, the C1 amplitudes for the low-probability trials was consistent and not significantly different across any bins (*p*s> .05). This was the same in the upper visual field as well, suggesting that C1 amplitudes relating to the low-probability events are unchanging. This was a different case for the high-probability trials. The C1 in the 1st bin for the lower field was significantly more positive ($M = 2.47$, $SD = 0.80$) than in the 2nd bin ($M = 2.19$, $SD = 0.94$) , $t(17) = 2.52$, $p = .022$. There was no difference across the 2nd to 6th bins (all pairwise comparisons, *ps* >.05). This trend was repeated in the upper field, where the 1st bin had a significantly more negative C1 ($M = -2.46$, $SD = 0.81$) than the 2nd bin ($M = -2.06$, $SD = 0.95$), $t(17) = 4.49$, $p < .001$. Again, there was no difference across the 2nd to 6th bins (all pairwise comparisons, *ps* >.05).

Effects of orientation probability on P300

The P300 is a late parieto-central component (*Figure 5.3* & *Figure 5.4d*) that shows a greater magnitude when targets are unexpected (Bledowski et al., 2004). I explored whether the orientation probability manipulation affected this component. This late positive component was maximal over the Pz electrode for most participants. The effect of orientation probability on P300 magnitude was not significant for the upper visual field ($t(17) = 0.46$, $p = .649$), but was significant (*Figure 5.7c*) for the lower visual field ($t(17) = 2.63$, $p = .018$), with low-probability orientations having the larger P300 (Low probability: $M = 3.29$, $SD = 1.59$; High probability: $M = 2.96$, $SD = 1.69$). There was no significant probability effect on (peak) latency ($M = 295$ms) in either field (*p*s >.05).

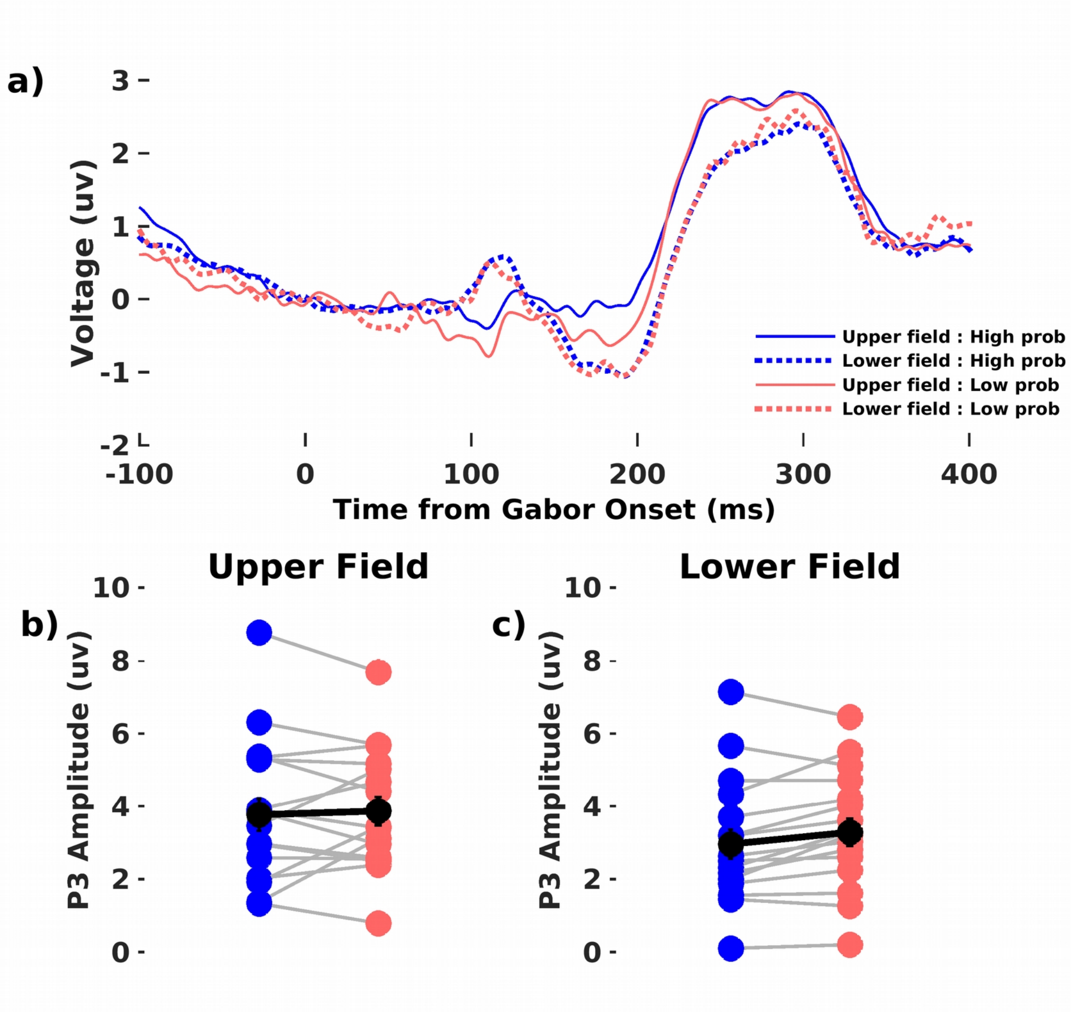

Figure 5.7. P300 analysis. P300s were calculated at the parieto-central electrodes showing maximum deflection at the 300-500ms time window, at each stimulus location. a) Average P300 wave across subjects for upper (solid) and lower-field (dotted) stimuli, across high (blue) and low (red) orientation probability. b) P300 amplitudes across subjects (separate lines) by probability, for upper-field stimuli. Sold line reflects the means across subjects. c) Same as previous panel, but for lower-field stimuli.

Lastly, it was examined whether the C1 and P300 amplitudes were related. The data presented in *Figures 5.5b/c and 5.7b/c* were recast into a separate analysis, which revealed that the C1-P300 correlation was significant for *both* the upper ($r = -.55$, $p < .001$) and lower visual fields ($r = .52$, $p = .001$). The direction of the correlation is reversed because the phase of the C1 reverses for lower and upper visual field stimulus presentations (*Figure 5.8*).

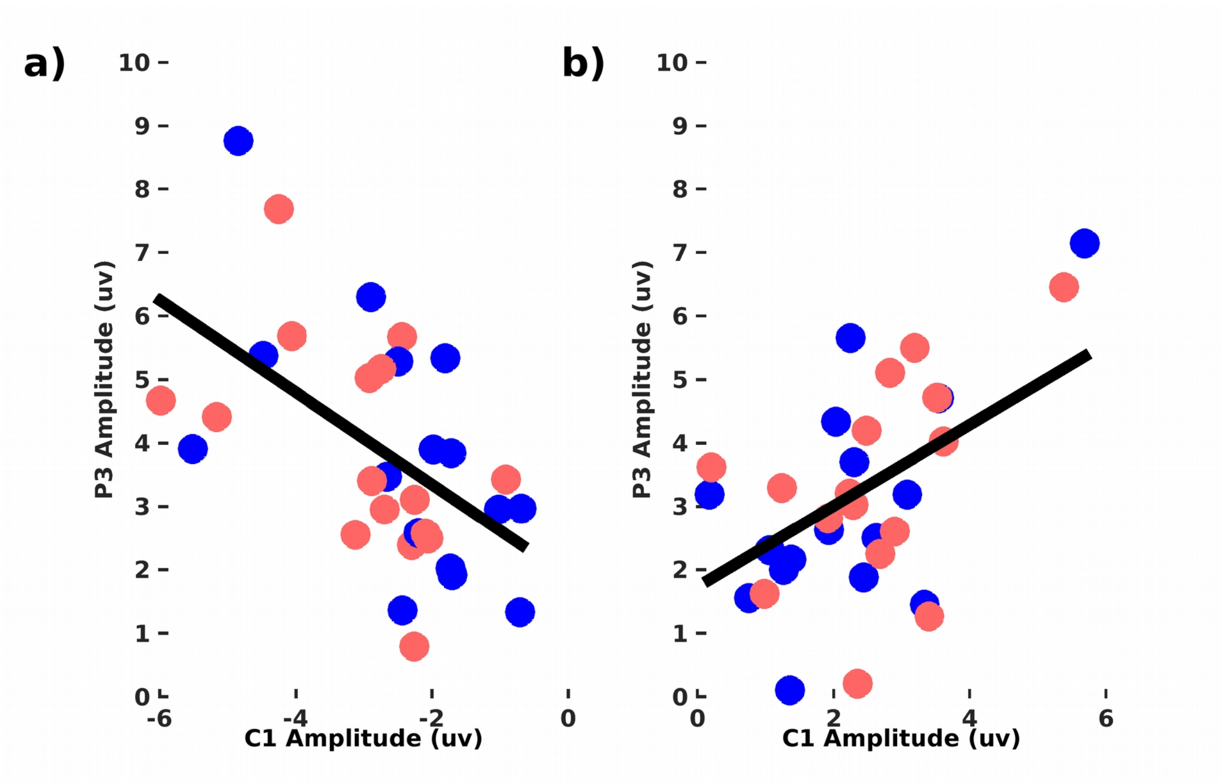

Figure 5.8. C1-P300 correlations. Amplitudes shown in *Figure 5.5b/c and Figure 5.7b/c* were recast as scatter plots. Significant correlations were found in both the a) upper and b) lower visual fields. Note that directions are reversed due to C1 directions depending on the field location. High-probability = blue, low-probability shown in red.

5.4. Experiment 5.1: Discussion

Probability effects were apparent in behaviour and were similar to the findings from the previous chapters. In the behavioral session, initiation times (IT) were consistently shorter for high-probability tilts. The probability effect on IT was probably not seen in the EEG session as that task was significantly harder with continuous rapid flashing of stimuli, and had unpredictable, irregular behavioral probes. Most likely the probability effect showed up more strongly in IT as compared to precision in the 1st session because, it being the easier task, participants managed to encode the stimulus fully in most instances. The time to fully encode was affected by probability in this case. Importantly though, there was a behavioural effect for trials in the EEG session with a clear probability effect on the angular errors: high-probability orientations were estimated more precisely.

To evaluate early visual processes as a locus of this benefit, the analyses focused on the ‘C1’ ERP component that has been source-localized to the primary visual cortex (Di Russo et al., 2002). In the current study, the early C1 component was topographically centred on the posterior parietal/occipital scalp area. In addition to replicating Di Russo and colleagues’ findings it was also shown that the C1 component was modulated by orientation probability: C1 amplitudes were *smaller* for the high-probability orientations as compared to the low-probability ones. Further, the time-course analysis (*Figure 5.6.*) would suggest that this effect is due to a quick reduction in the C1 amplitudes for the high probability trials, detectable within about 600 trial exposures (approximately 7-8 minutes in-task), and possibly plateauing within about 1200 trial exposures. This matches data from previous studies suggesting that perceptual effects due to probabilistic manipulations arise quickly.

The current study neither cued nor biased space. All locations were equiprobable but it was the *orientation-location conjunction* that was biased. Effects from this manipulation have been demonstrated to be distinct from spatial cuing (Chapter 2) and from spatial probability (Chapters 3 and 4). This is likely why the effects observed in this study are in contrast to those found with spatial attention, which either find no effect of spatial probability on C1 (Clark & Hillyard, 1996; Di Russo, Martinez, Hillyard, 2003), or that the cued space has a greater C1 (Kelly et al., 2008).

One account of our orientation probability effects would be to regard them as a form of expectation or top-down attention. But if this account is adopted, then it would imply that attentional effects for features are based on neural tuning (David et al., 2008; Paltoglou & Neri, 2012; Çukur et al., 2013; Ling, Jehee, & Pestilli, 2015), whereas feature-based attention has been suggested to involve *both* gain and tuning mechanisms (Ling, Liu, & Carrasco, 2009). While cuing of orientations do lead to reduced activity and sharper representations in V1 (Kok, Jehee & de Lange, 2012), attending to features has also been demonstrated to increase P1 ERPs (Zhang & Luck, 2009), which was not found in the current study. In contrast, *spatial* probability results in an *increase* of C1 and P1 ERP potentials (Slagter, Alilovic, Timmermans & Reteig, 2016). Fronto-parietal right brain damage also disproportionately affects the ability to benefit from spatial repetitions, as opposed to color repetitions (Shaqiri & Anderson, 2013) and this would make sense if spatial probability modulations rely on a 'top-down' mechanism more so than does feature probability modulations.

As an alternative to a top-down feature-based attention account of feature probability, the hypothesis in this thesis is that more frequent presentations of some orientations result in changes to the tuning widths of orientation-selective neurons in V1, similar to how orientation training in macaques leads to adaptive-tuning (e.g. Ringach, Hawken & Shapley, 1997), but on a much shorter time scale. While the source of the C1 potential could be V2/V3 (Ales et al., 2010), it is known that both cuing of orientations (Kok, Jehee & de Lange, 2012) and perceptual training (Ringach, Hawken & Shapley, 1997; Schoups et al., 2001) affect V1, making that the more likely locus in our opinion. Perhaps this modulation in V1 is mediated by recurrent local connections such as are the basis for the learning and adaptation effects seen in the Teich and Qian (2003) computational model of V1. Probable stimuli could “train” neurons with the result that tuning curves sharpen. Sensory neural tuning would offer a dynamic mechanism for our adaptations to uncertain and statistically non-stationary environments (Yaeli & Meir, 2010, also explored in Chapter 6). Such arguments have been also made before for the auditory domain (Dean et al., 2005) and visual (retina) domains (Hosoya et al., 2005).

Although the data from this study provides strong evidence for an early visual locus of the probability effect, it is possible that there might be other effects in addition to what was demonstrated here. Ghose, Yang, & Maunsell (2002), found that perceptual learning only produced weak reductions in V1 activity, and other studies, while confirming V1 effects for perceptual learning, have found V1 effects to be smaller than changes in V4 (Raiguel, Vogels, Mysore, & Orban, 2006; Yang & Maunsell, 2004). It could be that V4 modulations are downstream effects from initial changes in V1. Alternatively, V4 involvement could be taken as evidence for top-down control which trickles down to V1. The presence of retinotopic transfer in perceptual learning tasks (Zhang et al., 2010; Hung et al., 2014) supports a top-down account. Perhaps a combination of prolonged training and predictable stimuli enables participants to explicitly expect the stimuli, which does not occur here.

Probability effects could also go beyond visual processing. The other component examined was the late parieto-central positive component, the ‘P300’, which traditionally has been studied in the context of oddball paradigms. Although the component has been more studied in the auditory domain (Polich, 1990; Sutton, Braren, Zubin, & John, 1965), it has also been linked to the detection of rare visual events. Source-modelling suggests a source of either parietal and inferior temporal cortices or the insula/frontal cortices, depending on the specific sub-component (P3a/P3b; Bledowski et al. 2004). Although one might therefore expect the P300 to be modulated by our probability manipulations, this effect was weak compared to the C1. One potential reason is that P300s seem to be more susceptible to modulation when stimulus intervals are long (Strüber & Polich, 2002), and the ISIs in the current task were very short. It could also be that P300s are more clearly modulated by ‘oddballs’ which are not simply low-probability, but explicitly surprising events (Mars et al., 2008; Teigen & Keren, 2003). In the current case, low-probability orientations were less frequent, but not necessarily “surprising”.

Could repetition suppression or habituation be the explanation for probability effects on precision and C1 modulation? While it is generally true that more probable events will result in more repetitions, true repeats were rare in the current study because there were four equiprobable target locations, and the high-probability orientations was counterbalanced across them. Across locations, a high-probability tilt of 45 degrees is as likely to be followed by a high-probability tilt of 135 degrees as being followed by another 45 degree tilt. Comparing high-probability orientations that repeated location to those that those did not revealed no differences between C1 amplitudes. This may be because trials of repeated high probability are usually different angles as the range of possible orientations for the high probability quadrant spans 90 degrees. Even when ‘repeats’ of similar orientations were left out of the analysis, there was still a significant probability-based modulation of the C1 component: In prior studies, when repeats were left out of the analyses, effects on perceptual precision still remain (e.g. Jabar & Anderson, 2015). From an electrophysiological perspective, one would also think that if repetition-based habituation was the reason for the C1 differences, the same difference should be present in the P300 components (Ravden & Polich, 1998); however, the upper visual field P300s failed to show a difference between high and low probability orientations. The more likely explanation for orientation probability effect is the adaptive-tuning observed in primate training studies (e.g. Ringach, Hawken & Shapley, 1997; Schoups et al., 2001; Yaeli & Meir, 2010).

What has been demonstrated is that probability effects are accompanied by processing changes in early visual cortices. This is not to say that other processing levels are not also affected by stimulus probability. With conditions as complex as acquired badminton experience affecting C1 amplitudes (Jin et al., 2010), it seems likely that probability effects are mediated at multiple levels all working to link neural tuning to decisions. Probabilistic information for features might tune relevant perceptual channels, which then affects decision-making (Eckstein et al., 2009).

Assuming that the P300 also reflects decision-making processes (Rohrbaugh et al., 1974; Wang, Zheng, Huang, & Sun, 2015), the finding of a C1-P300 correlation can link early perceptual modulation to later decision-making processes. Alternatively, the causal direction might be reversed: The C1 might be the result of a top-down modulation that acts across the experimental block (e.g. Zhang et al. 2015). However, ERP amplitudes were variable across individuals, and while this C1-P300 relation leads to interesting hypotheses, the current study is not able to test this idea. The next chapter, however, will demonstrate that neither decision-making nor traditional attentional mechanisms are required to computationally model the neural and behavioral effects of probability manipulations.

Chapter 6: Probability Learning as a Passive V1 Circuit

6.1. Introduction

Chapters 2-4 provided behavioural support for the claim that orientation probability affects perception. Chapter 5, in showing that probable orientations dampen the amplitude of the C1 ERP component, provided neurophysiological support for the hypothesis that this perceptual effect occurs through the ‘sharpening’ of V1 neurons. This mechanism is similar to the one proposed for perceptual learning. Monkeys trained with an orientation demonstrate a sharpening of the V1 neuron preferring that orientation, relative to other neurons (Ringach et al., 1997; Schoups et al., 2001).

Whether sharpening or broadening would lead to encoding precision are contingent on the assumptions made about the nature of the neural population (Pouget, Deneve, Ducom, & Latham, 1999). For instance, the optimal direction depends on the dimensionality of the neuron. If a neuron is uni-dimensional (i.e., codes for one feature), the optimal would be to narrow in tuning, whereas the opposite might hold if the dimensionality is above two (Zhang, & Sejnowski, 1999). For orientation-selective V1 populations, this optimal half-width half-height in tuning has been estimated to be in the range of 12 degrees (Seung & Sompolinsky, 1993). V1 feed-forward layer (layer 4) are weakly selective and broad in tuning to begin with (Ringach et al., 1997), and the sharpening seen with orientation training improves discrimination performance (Schoups et al., 2001).

Models have been built that try to explain these V1 sharpening dynamics and how they lead to improved perceptual discrimination. V1 not only has the lateral geniculate nucleus (LGN) as its major input (Hubel & Wiesel, 1962), but also has intra-layer inhibitory and excitatory cortical interactions (Tsumoto, Eckart, & Creutzfeldt, 1979; Sillito, Kemp, Milson, & Berardi, 1980; Ahmed, Anderson, Douglas, Martin, & Nelson, 1994; Sato et al., 2016). It is only by taking both of these into account that crucial aspects of visual processing, e.g. that V1 tuning width is relatively independent of the contrast of the visual stimulus, is possible (Ben-Yishai, Bar-Or, & Sompolinsky, 1995). V1 sharpening in the Schoups et al. (2001) study was only observed in the supragranular (2/3) and infragranular (5/6) layers of the V1 cortex, and not in layer 4, where the thalamic input is received. This suggests that the changes were unlikely to have occurred pre-V1. Because V1 neurons are initially only weakly orientation-selective to feed-forward LGN input, recurrent intracortical inhibitory/excitatory feedback is required for the neurons to demonstrate their full orientation-selectivity (Teich & Qian, 2010). V1 neurons have a columnar organization, i.e. neurons preferring similar orientations are grouped together. It therefore is likely that cortical interactions also occur in an orientation-selective manner, i.e., a neuron preferring a 45-degree stimulus is more likely to interact with a neuron preferring a 44-degree stimulus, rather than one preferring a 10 degree stimulus (e.g. Jurjut, Georgieva, Busse & Katzner, 2017). By setting the neurons preferring the ‘trained’ orientation to have a a decreased excitation (or an increased inhibition) on neighbouring neurons, the Schoups et al. (2001) data could be modelled (Carandini & Ringach, 1997; Teich & Qian, 2003).

Given the hypothesis that orientation probability acts through neural tuning in V1, similar to how orientation training in monkeys occur (e.g., Ringach et al., 1997; Schoups et al., 2001), then a model that attempts to mechanistically describe the effects in those studies (i.e., Carandini & Ringach, 1997; Teich & Qian, 2003) should also be able to explain orientation probability effects.

A dynamic model

There are two issues with using Teich and Qian’s model for modelling probability effects. First, previous versions of the model compared two static states: baseline, and trained. There was no dynamic modulation that autonomously led *from* baseline *to* trained. For probability effects the question is not only whether decreased intracortical excitation replicates probability effects (as it did for perceptual learning), but *how* this change in intracortical excitation comes about. The second issue is time-span. Probability effects occur within a much shorter time-span than perceptual learning. Perceptual learning tasks in humans typically involve a limited set of stimuli and weeks of training (Zhang, Xiao, Klein, Levi, & Yu, 2010). In monkey studies (e.g. Ringach, Hawken & Shapley, 1997; Schoups et al., 2001) this process might take months. In contrast, for the probability studies described in Chapters 2-5, behavioural effects developed within a single session. That probability effects occur quickly was also demonstrated electrophysiologically, with C1 amplitudes for high probability orientations being dampened within the first 15 minutes or so of the estimation task (*Figure 5.6.*). Furthermore, instead of just having one ‘trained’ stimulus, the experiments presented in this thesis biased a range of orientations. Could the same mechanism of regulating intracortical excitatory connectivity explain both perceptual learning in monkeys and probability effects in humans?

By having the recurrent model adaptively modulate intracortical excitatory connectivity on a trial-by-trial basis both issues can be addressed. This is the idea explored in this chapter: Whether probability effects can be explained by a dynamic *perceptual-only* model where the lateral excitation of a neuron on its neighbours is dependent on the current rate of firing of that neuron and the similarity of their tuning preferences. This lateral excitation not only decreases if the neuron has been firing often, but also increases back to baseline if there has not been activity. This excitation can also change non-linearly: The smaller the current excitation is the more neuron firing is required to drive it down further. Conversely, the impetus for to increase connectivity back to baseline is greater the more connectivity has reduced from baseline (see *Model Instantiation* for details). This dynamic also allows for relatively recent information to be weighted more heavily than past information, which is a feature of both the visual and auditory cortices (Singer et al., 2017).

Sequential dependencies

A dynamic version of the Carandini & Ringach (1997) and Teich & Qian (2003) model will not only yield an ultimate sharpening of probable orientation tuning curves, it will also be sensitive to the sequence of stimuli presented. This also provides a greatly enhanced opportunity to test the model's ability to capture performance. If the effects from the stimulus sequences that the model is shown predict how participants respond to the same sequences, it provides strong evidence for the hypothesis that the probability effect operates through the modelled mechanisms. Sections *6.2* and *6.3* describe and provides test runs of the model to show that it accurately reflects the predicted probability effects. Sections *6.4* and *6.5* validate how participant performance tracks the model predictions.

6.2. Model Instantiation

This model is based on existing perceptual learning models (Carandini & Ringach, 1997; Teich & Qian, 2003) and therefore shares most of the same parameters and functions. Where possible, the same values (e.g. baseline J_e/J_i and intracortical connection exponents a_e/a_i) were used. The modifications and additional parameters added to the model allow for the model to respond to the stimulus sequence in a dynamic manner, without having to assume a Bayesian optimization process (e.g. Wei & Stocker, 2015). *Table 6.1* gives an overview of the parameters and their initial values. *Table 6.2* provides the equations that will be referred to in the next paragraphs. Appendix A provides the code implementation of the model.

Table 6.1. Model Parameters

N	Number of ‘V1 neurons’ (128)
τ	Membrane time constant (15ms)
α	Gain of spike encoder (10 spike/s/mV)
R	Firing rate of the neurons (Hz) = $\alpha \, max(V, 0)$
J_e°	Strength of intracortical excitation (baseline = 1.1 mV/spikes/s)
J_i°	Strength of intracortical inhibition (baseline = 1.1 mV/spikes/s)
J_f°	Strength of feed-forward excitation (0.5 mV/spikes/s)
σ_f	Gaussian width of feed-forward orientation bias (45°)
a_e	Exponent for the excitatory connection distribution (2.2)
a_i	Exponent for the inhibitory connection distribution (1.4)
A_e	Maximum fractional reduction of J_e° (0.010)
A_i	Maximum fractional reduction of J_i° (0.000, i.e. not reduced)
λ	Fractional decay rate of J_e°/J_i° towards baseline (0.010)
V	Membrane potential (units = mV)
θ	A neurons’ preferred orientation (range of -90° to +90°)
Φ	A stimulus orientation (range of -90° to +90°)

Note: Number in brackets are the values used in the presented simulations

Table 6.2. Model Computations

1 $V(\theta, t+\tau) = V_f(\theta,t) + V_e(\theta,t) - V_i(\theta,t) + \varepsilon$

2 $V_f(\theta, \Phi) = J_f {}^\circ \exp[-(\theta-\Phi)^2/(2\sigma_f{}^2)]$

3 $V_e(\theta, t) = J_e {}^\circ\ \mathrm{sum}[E(\theta - \theta')R(\theta',t)]$

4 $V_i(\theta, t) = J_i {}^\circ\ \mathrm{sum}[I(\theta - \theta')R(\theta',t)]$

5 $E(\theta) = c_e\ [\cos(2\theta) + 1]^{ae}$, *where* c_e *is a normalization factor*

6 $I(\theta) = c_i\ [\cos(2\theta) + 1]^{ai}$, *where* c_i *is a normalization factor*

7 $R(\theta)_{norm} = R(\theta)/\mathrm{sum}(R)$

8 $J(\theta)_{reduced} = J(\theta)\ [1 - (R(\theta)_{norm} * A)]$

9 $J(\theta)_{new} = J(\theta)_{reduced} + \lambda(J_{baseline} - J(\theta)_{reduced})$

Note: θ' indicates the preference of the other neurons that do not prefer θ

The implementation of the model is for a single receptive field location, and includes neurons with orientation preferences equally spaced across the possible range of orientations (-90 to 90 degrees). Note that because of the axial (half-circular) nature of orientations 90 degrees wraps back as -90 degrees. A ‘trial’ is defined as a single stimulus input. Following such a presentation the membrane potential of responsive neurons increase up until a plateau (*Figure 6.1a*). In all simulations, 1 second was the cutoff for determining firing rates. Computing neuronal activity uses equations 1 to 4. The membrane potential of a given neuron (V) at the next cycle ($t + \tau$) is the sum of the feed-forward membrane potential (V_f), the summed potentials of the intracortical excitatory connections (V_e), and the summed potentials of the intracortical inhibitory connections (V_i) at a given point in time (t). This is combined on each cycle by random noise (ε) which is scaled by the mean voltage of the feed-forward input across the neurons. The intracortical connections are governed by connection probabilities (E/I, equations 5/6) and the J_e/J_i parameters. The connection probabilities across V1 neurons, and J_i values are the same as for Teich & Qian (2003), and remain fixed throughout the simulations. However, the J_e values were permitted to change across trials (*Figure 6.1b*).

The J_e values start off at a baseline. What is novel in this adaptation of the model is that on each trial, J_e is reduced by a factor given by the normalized firing rate at the end of the trial (equation 8). Divisive normalization (equation 7) seems to be a computation that is made across the cortices (Carandini & Heeger, 2012), and it seems logical that the reductions in connection strengths would take this into account. Before the start of the next trial, the reduced J_e value is allowed to ‘decay’ or relax back to its baseline (equation 9). The net effect of this leaky integration is that neurons which fire often, e.g. those preferring probable orientations, get a gradual reduction in J_e, while neurons that prefer improbable orientations have J_e values that remain closer to the baseline. While J_i is also governed by equations 8/9, A_i was set to zero, which has the effect that J_i does not adapt. This is justified by noting that a decrease in J_e is computationally equivalent to an increase in J_i . One could also view our modulation of J_e as a modulation of the ratio of J_e to J_i .

It could be argued that a more biologically-realistic implementation of changes in each neuron's lateral connection strength (J_e) would have to be dependent on the current firing rate at that time step. Instead, *equations 8* and *9* were only adjusted at the end of each trial. This was primarily because this required less modifications to the Teich and Qian (2003) model. Lengths of stimulus presentation and inter-stimulus intervals were also not of particular concern for the current simulation as presentation times were fixed in the reported experiments, and orientation probability effects have been documented in studies with either long (Chapter 3) or short (Chapter 5) inter-stimulus intervals.

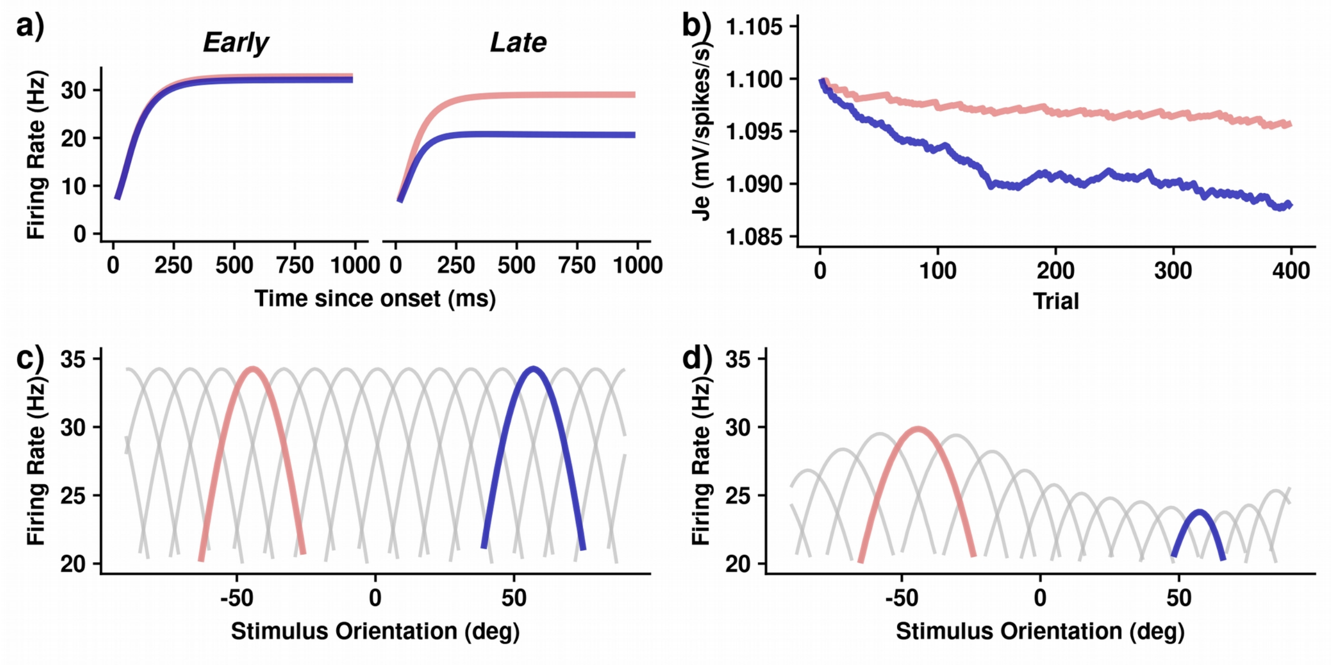

Figure 6.1. System evolution across a sample sequence. a) Example evolution of the firing rate of two neurons *within* a trial. For early trials, all neurons fire the same way. For later trials, maximum firing rate is attenuated as a function of neuron preference and stimulus probability. Neuron preferring high-probability orientations = blue, low-probability shown in red. b) Example change of the J_e parameter *across trials* for these two identified neurons. c) Neural tuning curves at the start of the sequence. d) Neural tuning curves for the same neurons at the end of the sequence. Note: Noise (ε in equation 1 of *Table 6.2)* was set to zero to produce these graphs. Noise was introduced for actual testing.

The model ‘estimate’ of the stimulus on each trial was taken as the population vector decoding of the firing rates (Georgopoulos, Schwartz, & Kettner, 1986; Pouget et al., 2000) at the end of the trial. For each trial the preferred orientation of each neuron was weighted by the firing rate, summed, and the reported angle was taken to the direction of this resultant vector. From this the angular error for that trial was taken as the difference between the input orientation and this estimate.

While the current implementation uses a standard decoding technique, there are optimal decoding methods such as Bayesian or Maximum-likelihood inference that in comparison would have smaller root-mean square errors between the actual and estimated stimuli (Dayan & Abbot, 2001). However, the advantage for these alternative decoding schemes is primarily in circumstances where neural preferences are unevenly distributed. In all the current simulations the distribution of orientation preferences was uniformly distributed. These optimal decoding techniques also assume the minimization of error or cost, and it is unclear what is the best cost function to apply to the visual cortex beyond accuracy of the classification. By having the decoded estimate be a linear sum of firing rates our model necessitated fewer assumptions.

6.3. Model Testing

Replicating our experiments with human participants, model simulations consisted of sequences of trials where 80% of the orientations were biased to one side (between 0 and 90 degrees) while 20% of stimuli consisted of the other orientation band (between -90 and 0 degrees). Also reproducing the psychophysical experiments, local probability was maintained in blocks of 20 trials (e.g. 16 left-tilting, and 4 right-tilting). This was counterbalanced across model runs. Each run had a randomized biased sequence that consisted of 400 orientations. One thousand runs were carried out. Every run began with the same initial parameters.

Since a neurons' firing rate (*R*) is directly proportional to its membrane potential (V), which in turn depends on J_e (equation 3), changes in J_e can modulate a neural response (compare *Figure 6.1a Early vs. Late*). Because J_e reduction is dependent on the stimulus, orientation probability causes the neurons that prefer probable orientations to get 'sharpened' (i.e., have narrower tuning), relative to other neurons. This similar to what happens with orientation training in monkeys (*Figure 1.1*). *Figures 6.1c* and *6.1d* depict the initial and final neural tuning for one example run. This is also the same effect that the Teich & Qian (2003) model enforced, but which is now observed emerging via dynamic adjustment. While the orientation biasing was binary (either left or right-tilting was probable), the J_e values showed a smooth sinusoidal pattern, which was seen in other runs as well. This smoothing is likely due to the columnar organization of V1 neurons. A neuron preferring the low-probability -1 degree orientation still fires to nearby high-probability orientations (e.g. a 1 degree stimulus), whereas a neuron preferring the low-probability -45 degree orientation fires little to any high-probability orientation. These differences in firing frequency led to smooth differences in tuning within each probability set.

Reproducing participant data, the median absolute error (*Figure 6.2b*) and the distribution of angular errors (*Figure 6.2a*) showed a significant difference in absolute (median) angular error between the high (*M* = 3.07 deg, *SD* = 0.26 deg) and low (*M* = 3.49 deg, *SD* = 0.50 deg) probability orientations, (*t*(999) = 21.7, *p* < .001). There was also a significant difference in kurtosis (refer to *Figure 2.2* for details on this metric) of the error distribution between the high (*M* = -0.48, *SD* = 0.19) and low (*M* = -0.75, *SD* = 0.28) probability orientations, (*t*(999) = 24.3, *p* < .001). Although this trend occurred in majority of the cases (approx. 76%), other cases showed either no probability effects, and some runs even showed a *reversed* probability effect.

While the size of the angular errors made by the model can be made similar to human data by parameter twiddling, the kurtosis of the error distribution cannot. This is likely because some of the errors in participant data are due to non-perceptual phenomena (e.g. motor-based errors, the participant missing the stimuli due to blinks, etc.), which adds to the ‘tails’ of the error distribution, inflating the kurtosis value. Therefore, further analyses of the model will focus on the median absolute errors.

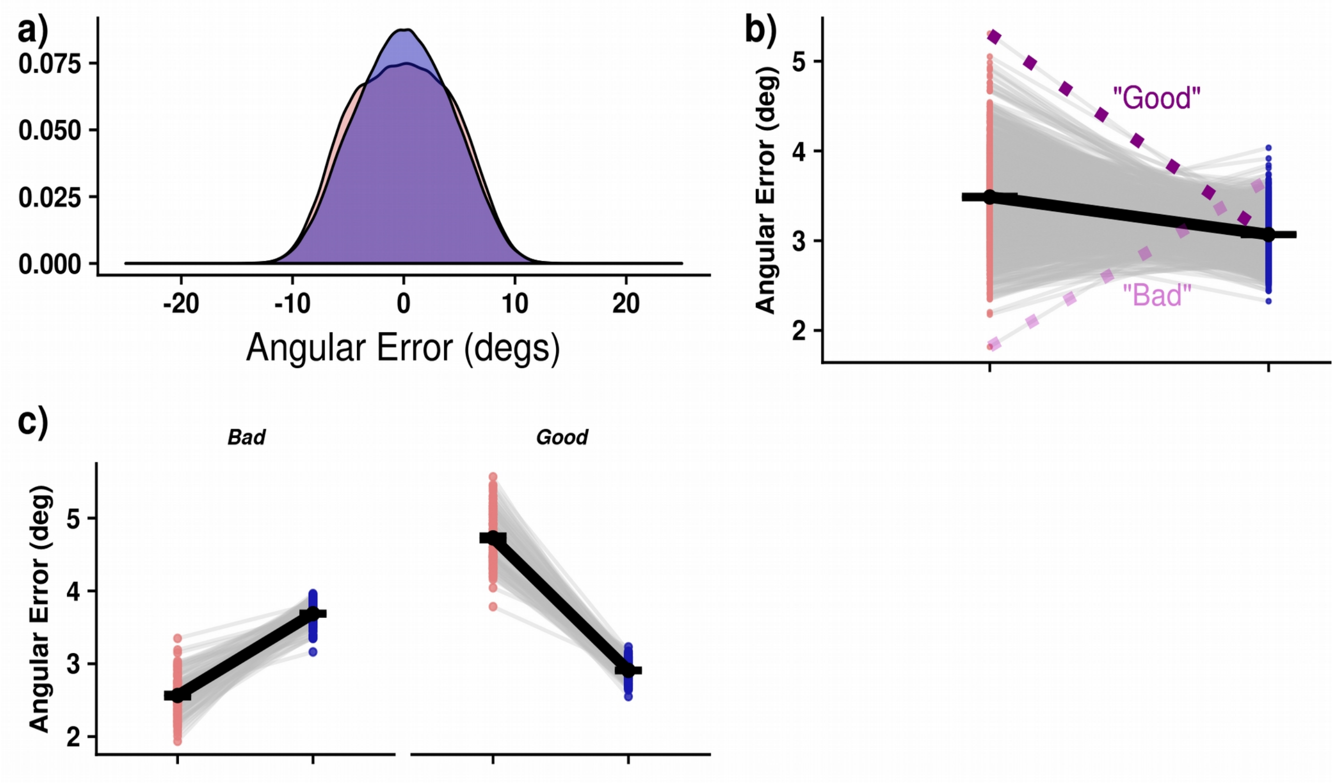

Figure 6.2. Modelled probability effects. a) Distribution of angular errors. b) Median absolute angular errors across random runs. The two runs highlighted were the runs with the most extreme results in each direction. c) Median absolute angular errors with just the two sequences highlighted in b). High-probability = blue, low-probability shown in red.

Reversed/Null Probability Effects

The cases where the expected probability effect had a reversed trend could have been due to random noise. But a more interesting possibility is that the learning differences were functions of the varying sequences of input stimuli. To examine the effect of noise on stimuli sequences, the model simulation sequences were fixed by selecting the run that led to greatest probability effect (i.e. a 'good' run), and the run that led to the maximally reversed probability effect (i.e. a 'bad' run[10]) and repeated these 100 times each with varying random noise inputs. The model parameters were otherwise identical. If the random noise was the reason behind the reversed probability effects, then the good and bad sequences should produce similar distributions of effects across runs. If the reversal was stimulus-driven, the 'bad' sequence should consistently produce reversed probability effects.

Under this repeated testing the 'bad' sequence consistently led to model estimates where the low-probability orientations were more precise (smaller error) than that of the high-probability orientations (*Figure 6.3c*). The 'good' run yielded a consistent probability effect (M = -1.82 deg, SD = 0.38 deg) that was significantly above zero, ($t(99) = 47.4$, $p < .001$). The 'bad' run produced a *reversed* probability effect (high – low, M = 1.12, SD = 0.33) that was significantly below zero, ($t(99) = 34.2$, $p < .001$). This suggests that differences in the probability effects are systematically affected by the sequence input into the model, above and beyond what can be explained by random noise.

---

[10] The labels used here simply relate to the fact that the occasional reversed/null probability effects were unanticipated by me.

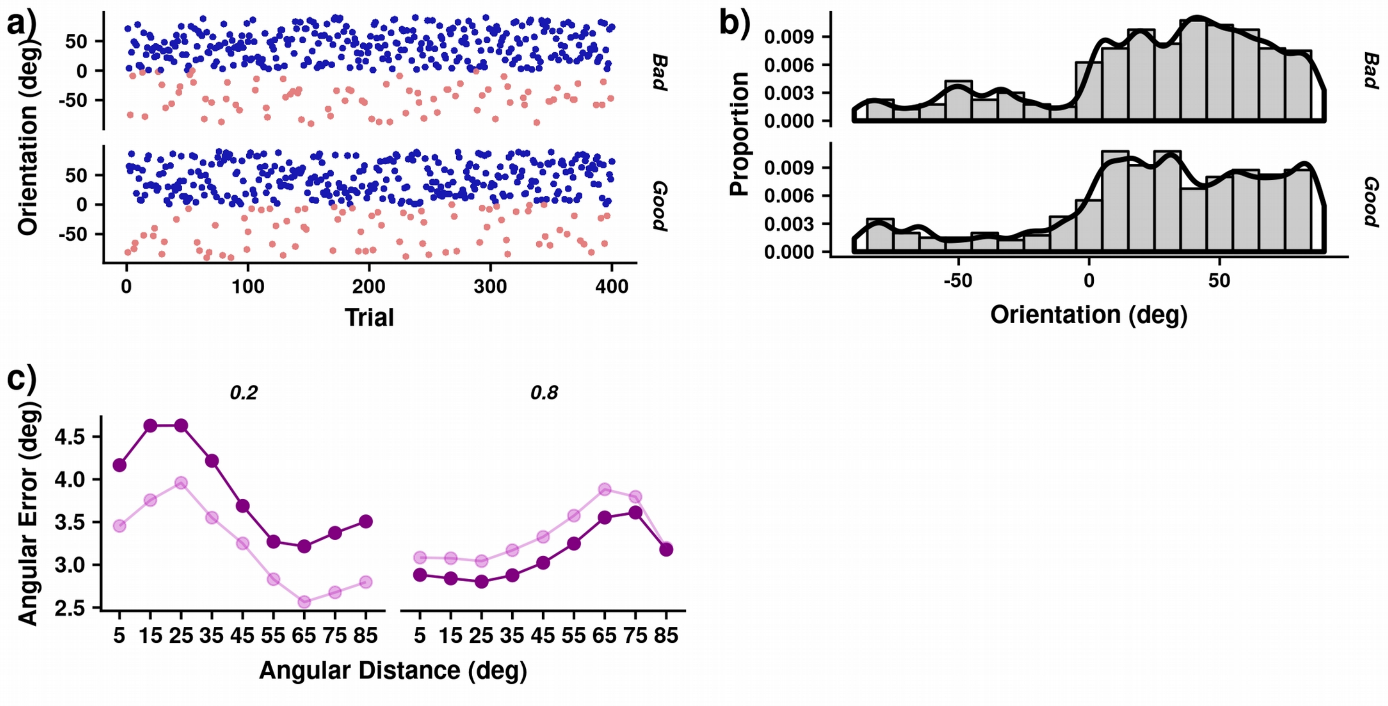

Figure 6.3. Probabilistic sequences. a) The example ‘good’ [probability effect in the predicted direction] and ‘bad’ [probability effect in the reversed direction] sequences. b) Overall trial distribution in the two example sequences. High-probability = blue, low-probability shown in red. c) Median absolute angular errors across the two sequences (light purple = ‘bad’; dark purple = ‘good’), as a function of what the change in stimulus orientation across trials was.

Figure 6.3a depicts the two sequences that were identified. A two-sample Kolmogorov-Smirnov (KS) test was used to compared orientation distributions (*Figure 6.3b*) of the two stimuli sequences, which suggested that they were drawn from a single distribution rather than from distinct ones (D = 0.045, p = .813). Rather than the global distribution of trials, it could be that the change between successive stimuli determined the pattern of results. The 1000 random stimuli sequences (the same sequences for the data in *Figure 6.2b*) were re-examined. For each stimulus sequence the orientation difference between consecutive trials was computed. For example, if trial 1 was 45 degrees and the next trial was 67 degrees, that would be an angular change of 22 degrees (the absolute change was taken here). The distributions of angular change for the ‘good’ and ‘bad’ sequences were compared with a KS test, which failed to reject the null hypothesis of different distributions (D = 0.003, p = .38).

In a further effort to probe for diagnostic differences in what constituted a ‘good’ or ‘bad’ sequence other than the end effect on probability learning, the angular error for bins of trials (*Figure 6.3c*) was assessed. While it is clear that angular change affects the extent of the angular errors, it also appeared that the ‘bad’ and ‘good’ sequences seem to systematically differ. For low-probability orientations, the errors made in the ‘bad’ sequences were consistently smaller than the ones made in the ‘good’ sequences. This reversed for the high-probability orientations, except for when the change in orientations was very large (85 degree bin). Given how little these surface characteristics of orientation sequences explained differences in model performance, more complex interactions are likely at play.

For example, take the case where the first two stimuli are a 45 degree grating followed by 70 degree grating. At the end of the first trial, the neurons preferring 45 degree stimuli will be the neurons with the lowest J_e value. After the next (70 degree) trial, the neurons with the lowest J_e value prefer orientations between 45 and 70 degrees, tending towards the latter. With more trials, the gradual accumulation of J_e reductions end up creating complex population tuning patterns. The precise population tuning would be difficult (if not impossible) to predict if the sequence of trials was disregarded.

Parameter space

Because the possibility of the complexity of the interactions between the sequence and the population representation, further steps were taken to identify the critical feature of good and bad sequences and how they depended on the specific parameter values chosen. The two previously identified sequences (see *Figure 6.3a/b*) were again fed into the model for another 500 simulated runs. Unlike prior simulations, A_e (fractional reduction of J_e) and λ (decay rate) were varied between 0.002 and 0.020 (at a resolution of 0.002).

Plotted in *Figure 6.4* are the means for each orientation probability (red = low prob, blue = high-prob), for each sequence (‘good’ = solid line, ‘bad’ = dotted line), for each value of A_e (across the panels lengthwise), for each value of λ (across the panels height-wise). Larger A_es and smaller λs lead to larger probability effects for the ‘good’ sequence. This makes sense given that bigger fractional reductions are going to drive more extreme shifts in J_e and will yield a larger effect on tuning. A smaller decay will slow the rate at which J_e relaxes to baseline. The ‘bad’ probability effects due to the parameters are negatively correlated with the size of the ‘good’ probability effects (Pearson’s $r = -.92$, $p < .001$; see *Figure 6.5*). Another noteworthy observation is that the change in the probability effect seemed to be due more to the low-probability orientations than the high-probability orientations.

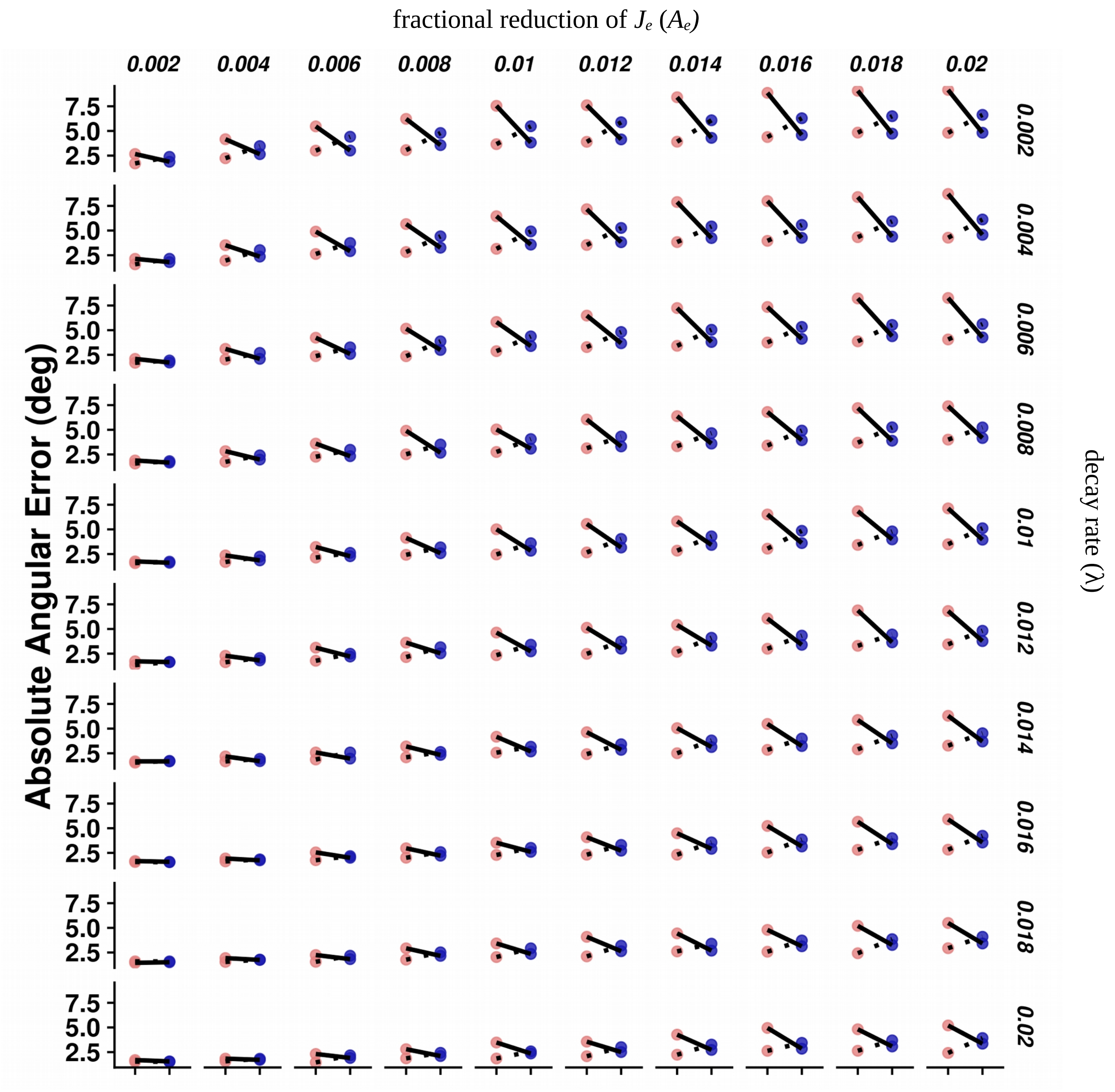

Figure 6.4. Effect of the model parameters on the probability effect, given the 'good' (solid lines) and 'bad' (dotted lined) sequences. Panels in the y-direction indicates the value of the decay rate (λ) and the x-direction indicates the fractional reduction of J_e (A_e). High-probability = blue, low-probability shown in red.

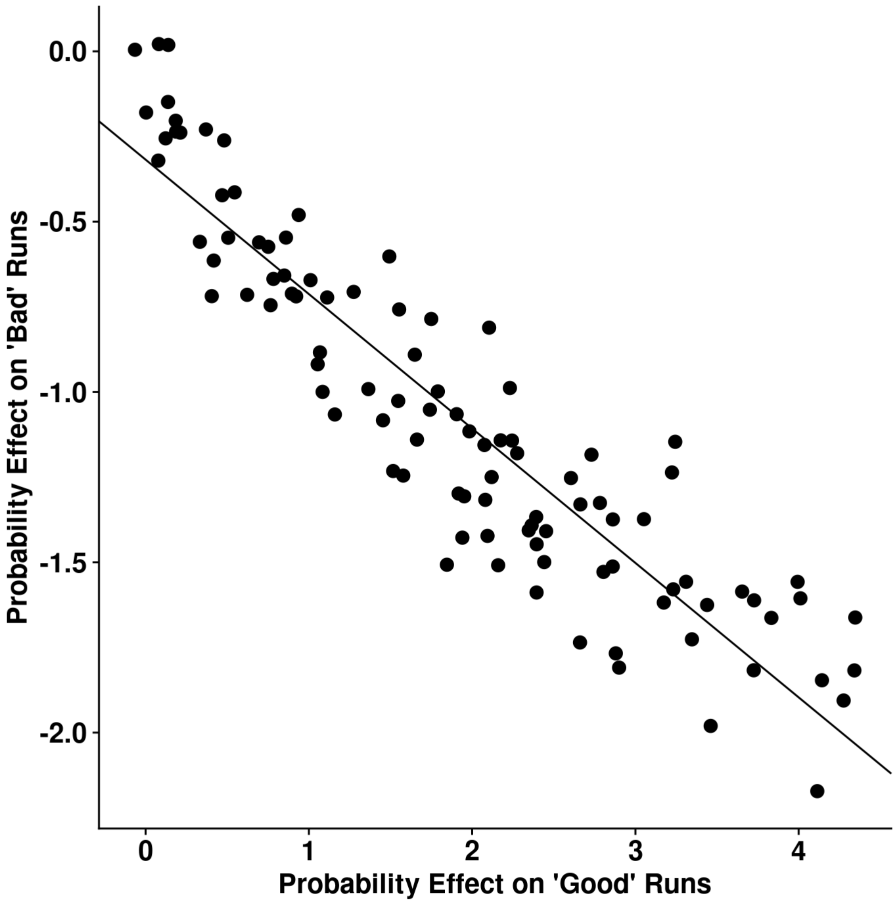

Figure 6.5. Relationship between the size (and direction) of the probability effect across the different runs with different parameters. Positive values indicate a bigger median (absolute) angular error for low-probability trials than high-probability trials.

While the model implementation is relatively simple, it makes very specific (and therefore testable) predictions about probability effects. If the orientation probability effect is indeed simply a low-level perceptual phenomena, we might expect participants to react to different sequences in a manner similar to the model. To test this prediction, a modified version of orientation estimation task outlined in Chapter 2 (*Figure 2.1*) was run. Using the exact same ‘good’ and ‘bad’ sequences tested with the model, participants were tested for orientation probability effects.

6.4. Experiment 6.1: Methods

Participants

22 participants were recruited (14 females, 8 males). 20 were right-handed and 2 were left-handed. Participant ages ranged from 17 to 28 (median = 21). All participants had normal or corrected-to-normal vision, and did not declare any auditory deficits.

Equipment & Stimuli

Participants sat approximately 60 cm from a 32 cm x 24 cm gamma-corrected CRT monitor that refreshed at 85Hz (*mean* refresh time = 11.8ms, *SD* = 0.5ms). Prior to the task, participants were instructed to make their estimations of the Gabor orientations as accurately as they could. They were *not* told that they needed to be fast. Responses were made with a computer keyboard using their dominant hand. As with the experiment in Chapter 2, stimuli were oriented grayscale sine-wave gratings with a circular Gaussian mask, at 50% contrast, with an average measured luminance of 40 cd/m^2. They had a spatial frequency of 4 cycles per degree of visual angle, and were presented on a grey background with a similar luminance of 40cd/m^2. When viewed from a distance of 60cm, the Gabors subtended approximately 4 degrees of visual angle both vertically and horizontally. Unlike the experiment in Chapter 2, no cues were used, and Gabors were always centrally located. Lines, for participants to rotate to report their estimations, had a length of 4 visual degrees and also always occurred at central fixation. As with Chapter 2, the response line started off horizontal (90 degrees) by default.

Procedure

Trials began with a 1000ms presentation of a black fixation cross, followed by a 500ms of a blank screen. A Gabor flashed for approximately 60ms, and after a 500ms blank screen, a response line was drawn in the same spatial location as where the Gabor had appeared. Participants made their estimations by rotating the response line counter-clockwise or clockwise by pressing “Z” or “C” on a QWERTY keyboard. This rotation was at a maximum of 1 angular degree per frame refresh of the monitor. Participants pressed the “X” key to confirm their estimations. Once confirmed, auditory feedback was given. On practice trials, a white feedback line with the actual orientation was displayed on top of the participant’s response. This visual feedback was not given during the main trials. This response procedure was the same as in Chapter 2.

Participants started with 20 practice trials of uniformly distributed orientations. The main task consisted of 800 trials. This was either the 400 trials of the ‘good’ or ‘bad’ run predicted by the model, followed by the other 400-trial sequence. Participants were given a break every 200 trials. After the experiment, participants were given a modified version of the previous questionnaires, which consisted of the following questions, in order:

1. Did anything about the experimental task stand out to you?
2. Please describe any strategies you may have used.
3. Did you feel that you perceived some stimuli better or differently than others, or in certain cases? Did you notice any change over time in your experience?

4. Do you think that some orientations were more likely? If yes, please elaborate.

5. Which of these do you think was more likely? Circle one of the choices.
[They were given a left and a right-tilting Gabor to choose from]

6. Was the first half of the experiment easier or more difficult, as compared to the second half?

7. Did you notice any differences in the first half of the experiment, as compared to the second half?

8. Did you think there were more repeats in one half of the experiment? E.g. (/) followed by another (/).

Questions 1-4 were standard questions used across all the experiments reported in the thesis. Questions 5-8 were new questions, and printed on the other side of the page. The idea here was that participants would be given the open-ended question first (Q4) before being forced to make a choice (Q5).

6.5. Experiment 6.1: Results

Data from the participants who did the ‘good’ sequence first, followed by the ‘bad’ ($n = 11$) was initially looked at separately from those that did the reverse. Because all the trends were identical across these two groups, the data were collapsed across the two sequence orders.

The median (absolute) angular error, kurtosis, initiation time (IT), and vacillation measures were computed in the same way as in Chapter 2. An overview of the data is depicted in *Figure 6.6*. For the angular error measure of precision (*Figure 6.6a*), a two-way (repeated measures) ANOVA revealed that there was a marginally significant main effect of probability, ($F(1, 21) = 4.31$, $MSE = 6.10$, $p = .050$). There was no main effect of sequence, ($F(1, 21) = 3.12$, $MSE = 1.46$, $p = .092$). Crucially, there was a significant interaction between these two factors, ($F(1, 21) = 6.77$, $MSE = 1.05$, $p = .017$). To analyze the interaction the probability effects in the ‘good’ and ‘bad’ sequences were looked at with separate t-tests. While there was the typical probability effect for the ‘good’ run (*t*(21) = 3.13, *p* = .005), with high-probability orientations being more precisely estimated (*M* = 5.16 deg, *SD* = 1.80 deg) than low-probability orientations (*M* = 6.32 deg, *SD* = 2.38 deg), the ‘bad’ run did not demonstrate a significant probability effect (*t*(21) = 0.07, *p* = .949), with high-probability orientations (*M* = 5.27 deg, *SD* = 1.88 deg) being estimated with similar precision to low-probability orientations (*M* = 5.29 deg, *SD* = 1.59 deg). To determine whether the stimulus sequence had a greater effect for high or low probability orientation, magnitude of angular error for high (or low) probability orientation trials across the two sequences was compared. There was no significant effect for high-probability orientations (*t*(21) = 0.45, *p* = .660), but there was an *increase* in the precision for low-probability orientations in the ‘bad’ sequence (*t*(21) = 2.53, *p* = .019)

The other three measures (kurtosis, IT, vacillations) did not show any interactions between sequence and probability (*p*s > .05). All revealed the established probability effects (*p*s < .05), but no sequence effects (*p*s > .05). While the low-probability precision (as measured by angular errors) was decreased when comparing the ‘bad’ to the ‘good’ run, this was not accompanied by a change in the kurtosis of the error distribution (*t*(21) = 0.97, *p* = .341). Instead, it was the standard deviation that was changed (*t*(21) = 2.99, *p* = .007), suggesting that the effect of the sequence on the low-probability trials was somewhat unlike the effect of probability.

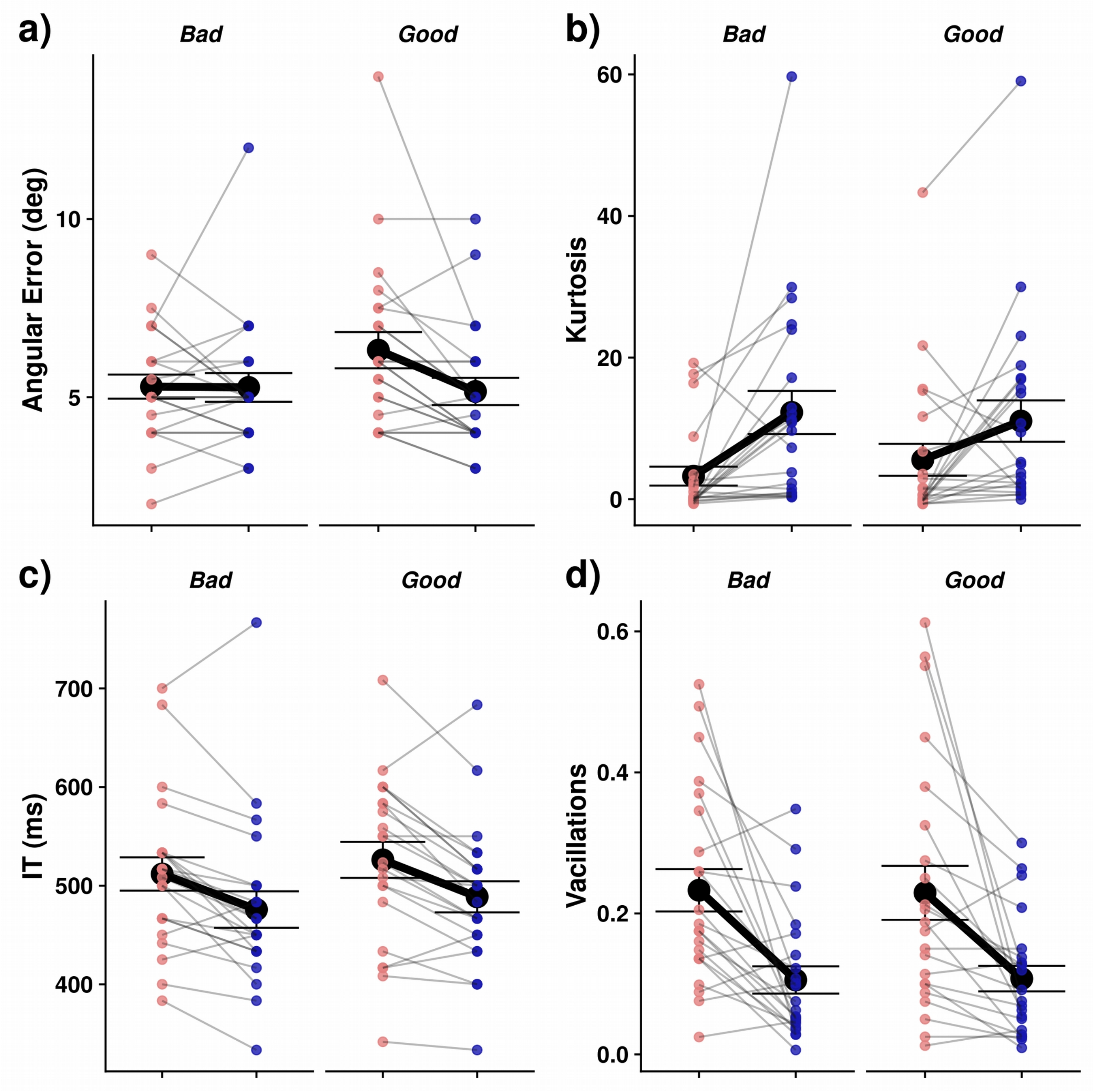

Figure 6.6. Probability effects in participants given the two sequences tested by the model. a) Median absolute angular errors. b) Kurtosis of the error distributions. c) Initiation time (ms). d) Number of vacillations. ‘Good’ and ‘bad’ sequences given to participants were identical to the ones depicted in *Figure 6.3a*. High-probability = blue, low-probability shown in red.

For the post-experiment questionnaire, the forced-choice post-experiment question (Q5) and the open-ended question that proceeded it was of particular interest. 16 participants chose the ‘45 degree’ option, which was correct. The others chose the wrong option. Of the 16 that did the forced-choice correctly, only 8 answered the previous open-ended question correctly. Given that only one stimulus location was used in this study, it is not surprising that some participants might explicitly learn the distribution used (as compared to previous chapters, which used a location-contingent orientation distribution). Even then, there was no clear pattern on whether the awareness of the probability manipulation was related to the sequence order, or to the size of the probability effects that the participant showed. It does not seem as if awareness (or attention) is a necessary condition of showing probability effects. This issue will be returned to in Chapter 7.

6.6. Discussion

While the model presented is a strictly low-level perceptual one, it nevertheless successfully replicates several facets of the orientation probability effect. High probability orientations are on average more precisely represented by the model, and this is the result of a sharpening process, which should result in less net activity in the system overall when the high-probability orientation is shown. This corroborates the direction of the effect demonstrated by the C1 ERP data in Chapter 5. This process can also occur quickly, within the time-scales of the experiments outlined in the previous chapters. The model is not only descriptive of probability effects, but also *predictive*.

The model suggested that sequences can affect the size, and perhaps the direction of the probability effect (*Figure 6.2*), even when there is nothing significantly different about the overall distribution of the trials that make up the sequences (*Figure 6.3*). Specifically, with the sequences that the participants were tested with, the model predicted that better performance for low-probability orientations in the ‘bad’ sequence, as compared to the ‘good’ sequence (*Figure 6.4*). When looking at the angular errors, participants showed exactly this trend (*Figure 6.6*). Because the model does not have a response component, angular errors are the important measure here, especially since the interest here is in the effects of probability on perception (via V1 tuning).

While it is interesting why there were probability effects in IT and in vacillations for the ‘bad’ sequence even while there was none for in angular error measure, this was likely because in an effort to make the task as close to the model data as possible, only a single stimulus location was used. Because one set of orientations was biased, this also created the possibility for response biases (previous experiments did not have this issue since probability was location-contingent). Future iterations of the model could introduce this spatial encoding, which would also allow us to model spatial probability. Note that this might require taking into account on/off-center neurons in the upstream LGN to determine how the orientation columns in V1 are related to spatial receptive fields (e.g. Miller, 1994).

While the angular error trends in the data were not reversed, they were different: A null effect for the “bad” sequence, and the standard probability effect for the “good” sequence. Further modifications of the model, for example, including orientation anisotropy - the common finding that cardinal (vertical/horizontal) orientations are more precisely represented than oblique orientations (Appelle, 1972) - could be modelled as differences in the tuning of cardinal-preferring neurons (Li et al., 2003), which can in-turn be implemented as pre-existing baseline differences in J_e or J_i values, and might lead to a better match between simulated and human data. Alternatively, orientation anisotropy could also be modelled as neuron stimulus preferences being unequally distributed across the possible range of orientations. Either way, this should have an effect on the model predictions, although on average the two sequences should show discernible differences in probability effects.

The perceptual model used to reproduce and predict human participant performance is simple, autonomous, and has no response or ‘attentional’ components. As a result the model offers a mechanistic account of how orientation probability effects can be acquired on-line through cortical interactions local to V1 cortex and as a function of environmental experience.

While the present model does a reasonable job of accounting for the data other models are available that also decrease in V1 activity and increase in perceptual precision (Chapter 5). Predictive coding models operate on recurrent prediction/error signal loops (Clark, 2013). Sensory signals from V1 could enable the downstream regions to make predictions about upcoming signals. The downstream regions would then send prediction signals back to V1 resulting in an error signal (the difference between incoming sensory input and the prediction signal). While computationally tractable, neural evidence for such mechanisms in the visual cortex is lacking. While studies have not shown clear evidence or reciprocal prediction signals nor error signals, they have shown that predictable stimuli decrease V1 responses (e.g. Alink et al., 2010). Kok, Jehee & de Lange (2012) demonstrated that predictive cuing of an orientation caused a decrease in V1 activation, and multi-voxel pattern analysis suggestied a strengthening/increasing coherence of the V1 signal in response to this predictive cue. If V1 is sending error signals rather than sensory signals then as predictions improve one would expect decoherence of the error signal. These data are arguably more simply accounted for by positing a V1 encoding of a sensory signal that gets sharpened with predictability.

While there remains ample room for expansions to the model or for comparisons with other models, what should be clear at this juncture is not only that the perceptual tuning hypothesis of orientation probability is supported by the available neurophysiological and behavioural data, but that the mechanisms that lead from one to the other can be computationally modelled without requiring abstract assumptions on the role of ‘attention’. The next experiment further examines the necessity for ‘attention’ in establishing (orientation) probability effects.

Chapter 7: Orientation Probability is Attention-free

7.1. Introduction

Chapter 6 suggested that orientation probability effects can be modelled through the use of a *local* recurrent V1 circuit, without having to make any assumptions about the role of top-down 'attention'. Attending to the feature of an object, e.g. its colour, increases the P1 evoked visual potential elicited by that object, even if that object appears in an unattended location (Zhang & Luck, 2009). This P1 increase was not observed with the electrophysiological experiment on orientation probability discussed in Chapter 5. Instead, the C1 potential, which is thought index V1 activity (Di Russo et al., 2002) *decreased* in amplitude. While feature-based cuing of orientations can modulate V1 activity (Kok, Jehee & de Lange, 2012), it is conceivable that it does so through feedback from a more downstream region, hence the modulation of P1, which has been source-localized to the extrastriate cortex (Clark & Hillyard, 1996; Di Russo, Martinez, Hillyard, 2003). More recent MEG studies on feature-based attention (to colour) further demonstrate that downstream, anterior visual processing areas have changes that *precedes* the changes in upstream, posterior sections, implying a top-down feedback process (Bartsch et al., 2017).

In contrast, the argument for orientation probability is that V1 activity is locally modulated as a result of stimulus exposure, rather than via top-down feedback. This predicts that tuning of the orientation selective V1 neurons would occur even if orientation probability is irrelevant to the task, e.g. if 'attention' was directed to colour rather than orientation. This chapter explicitly examines this by biasing orientation probability only on trials where orientation is irrelevant and colour is relevant. From the results of Chapter 4, orientation probability was not anticipated to impact colour detection performance.

However, if orientation-selective V1 neurons are tuned by the biased orientation distribution despite the task demands, then participants would have passively ‘learnt’ the orientation probability anyway. This notion of implicit learning can be assessed by intermixing occasional orientation estimation trials, where orientation probability was *not* biased. Finding a probability effect in such a situation would suggest that probability learning is a constant ongoing process driven by the distribution of sensory events, and could occur without ‘attention’ being directed to the probable feature.

7.2. Experiment 7.1: Methods

Participants

Twenty participants were recruited from the University of Waterloo (*median* age = 19 years; 15 females, 5 males), in exchange for course credit. 18 reported themselves right-handed. All participants had normal or corrected-to-normal vision and audition, and were not colour-blind.

Stimuli

Sine-wave gratings with a hard circular mask were used as the target stimuli in this experiment. These gratings were shown at 50% maximum contrast, and either appeared in grayscale, or were colored green, or colored blue. These colored stimuli had a similar appearance to the ones used in Chapter 4 (*Figure 4.1*), except the gratings were now sine waves instead of square waves. Regardless of color, the gratings always had a luminance of 40cd/m^2. These gratings had a spatial frequency of 4 cycles per degree of visual angle, and were presented on a gray background with a similar luminance of 40cd/m^2. When viewed from a distance of 60cm, the gratings subtended approximately 4 degrees of visual angle both vertically and horizontally. Gratings appeared either left or right of the fixation cross, with a distance of 4 degrees from the center of the cross to the center of the grating. The fixation cross either appeared in red or dark grey both having a luminance of 34cd/m^2. Lines, used for feedback and for participants to report their estimations, had a length of 4 visual degrees, and always occurred in the same location as the grating for that trial. Response lines always started off oriented horizontally.

Gratings were equally likely to appear at each of the two locations. Orientation probability distributions were maintained throughout the experiment. For every set of 24 trials, there were 20 color judgment trials, and 4 orientation judgment trials. The trials within each set were randomly shuffled to occur in any order. The rational for this uneven number of trials was to ensure that participants prioritised color over orientation information. On these 20 color trials, an orientation probability manipulation was done in a *location-contingent* manner. There were 8 left-tilting gratings on the left, 8 right-tilting grating on the right, etc. As in Chapter 6, orientations were uniformly distributed within each quadrant. *Collapsed across the two possible locations*, all orientations were equally likely. Color probability was always kept uniform. The final distribution of trials seen across participants is illustrated in *Figure 7.1*. There were two counterbalanced versions of the task (a/b), differing in their location-orientation mapping for the color discrimination trials (top panels). The orientation estimation trials (bottom panels) were rare, uniformly distributed in orientation, and mixed in with the biased trials.

The purpose of this design was to ensure that any orientation probability learning had to occur within the color discrimination trials, where ‘attention’ was directed to color rather than to orientation. To further support this, the fixation cross which preceded the stimulus explicitly (and accurately) cued participants to the type of upcoming stimulus (red for color trial, grey for orientation trial). While participants were explicitly informed what these cues meant, they were *not* informed about the probability distributions used.

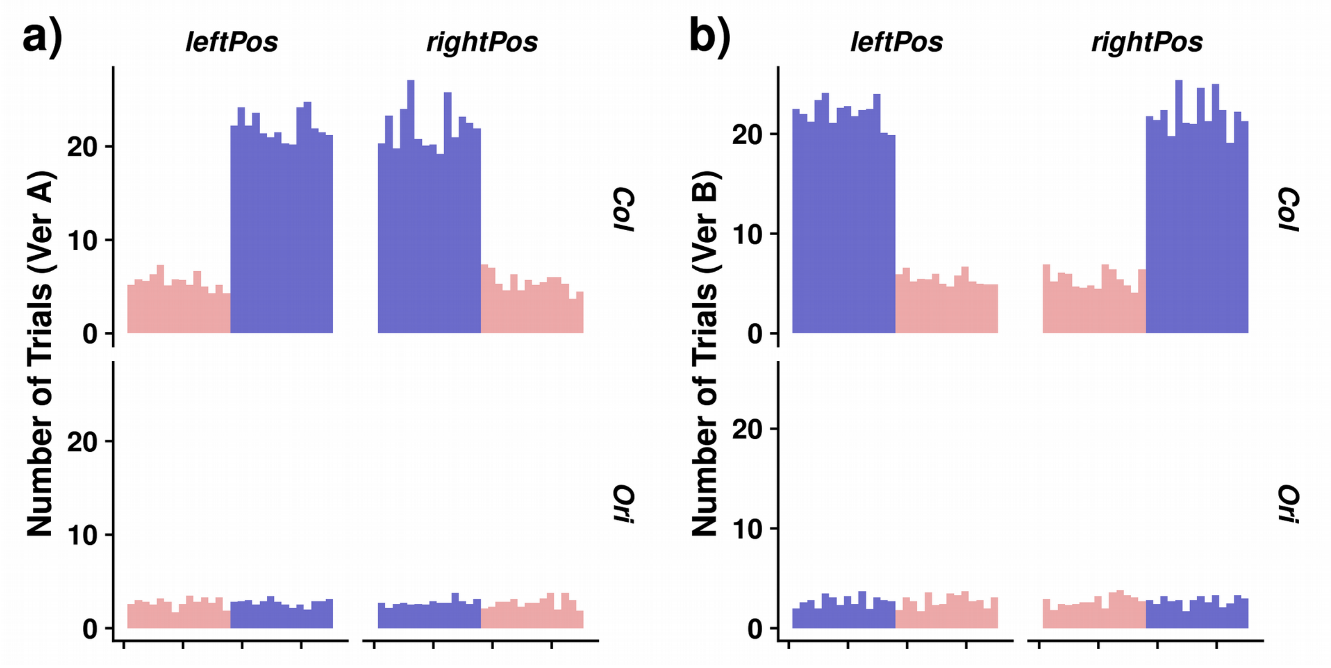

Figure 7.1. Stimulus distribution and performance in the color judgment trials. a) Distribution of stimulus orientations (x-axis) across the color and orientation trials, for one of the counterbalanced versions of the task. b) Distributions for the other version. Note: numbers refer to the *total* number of trials seen across *all* participants. High-probability = blue, low-probability shown in red.

Practice trials (40 in total) had uniformly distributed orientations and an equal split between color and orientation judgment trials. A gamma-corrected CRT monitor that refreshed at 89Hz was used, and stimulus timings were programmed as numbers of frames.

Auditory feedback was given after each response to maintain motivation. On orientation judgment trials, a high pitched sound indicated an error of less than 12 degrees, while a lower pitch indicated an error greater than 12 degrees. Color judgment was done in a binary way (either participants responded with the correct key, or they did not), and the same type of auditory feedback was given.

Procedure

Prior to the task, participants were instructed to be as accurate as they could be. They were *not* told that they needed to be fast. Responses were made with a QWERTY keyboard, specifically with the four direction keys (right hand), and the “X” key (left hand). Participants were also instructed to fixate on the central fixation symbol at the start of each trial. This fixation phase lasted 500ms, and the cross either appeared in red or grey to explicitly signal participants to the upcoming trial type. A grating then appeared for approximately 60ms (5 frames). After a delay of 500ms, either color patches or a response line appeared, depending on the task. The blue/green color patches were displayed as a reminder to participants as to the key assignments (the green patch was above the blue patch, because “UP” was the choice for green, and “DOWN“ was the choice for blue).

For the orientation estimation trials, participants made their orientation judgments by rotating the response line counter-clockwise or clockwise by pressing “LEFT” or “RIGHT” on the keyboard. This rotation was at a maximum of 1 angular degree per frame refresh of the monitor. Participants pressed the “X” key to confirm their estimations. The auditory feedback was then given. Following the 40 practice trials, there were 960 main trials (40 sets of 24 trials). These trials were separated into four blocks, with breaks in-between. A post-experiment open-ended questionnaire was given to each participant that probed with increasing specificity whether they had been aware of or could report the probability by location manipulation.

Post-experiment Questionnaire

The questionnaire consisting of the following seven questions:

1. Did anything about the experimental task stand out to you?

2. Please describe any strategies you may have used.

3. Did you feel that you perceived some stimuli better or differently than others, or in certain cases? Did you notice any change over time in your experience?

4. Do you think that some orientations were more likely? If yes, please elaborate.

5. Which of these do you think was more likely to occur on the left side? Circle one of the choices [They were given a left and a right-tilting grating to choose from].

6. Which of these do you think was more likely to occur on the right side? Circle one of the choices [They were given a left and a right-tilting grating to choose from].

7. Apart from the color, do you think there was a difference in the gratings used for the color task, as compared to the orientation task?

7.3. Experiment 7.1: Results

For the color discrimination trials, there was no difference between the blue and green gratings for either RT or accuracy (ps < .05, see *Figure 7.2*). The data also replicates the findings from Chapter 4, in that orientation probability had no effect on color discrimination. There was no difference between the time taken to judge the color of gratings with high (M = 663ms, SD = 135ms) or low (M = 660ms, SD = 138ms) probability orientations, ($t(39) = 0.85$, $p = .399$). There was also no difference between the accuracy in judging the color of gratings with high (*M* = .969, *SD* = 0.027) or low (*M* = .969, *SD* = 0.031) probability orientations, (*t*(39) = 0.13, *p* = .900).

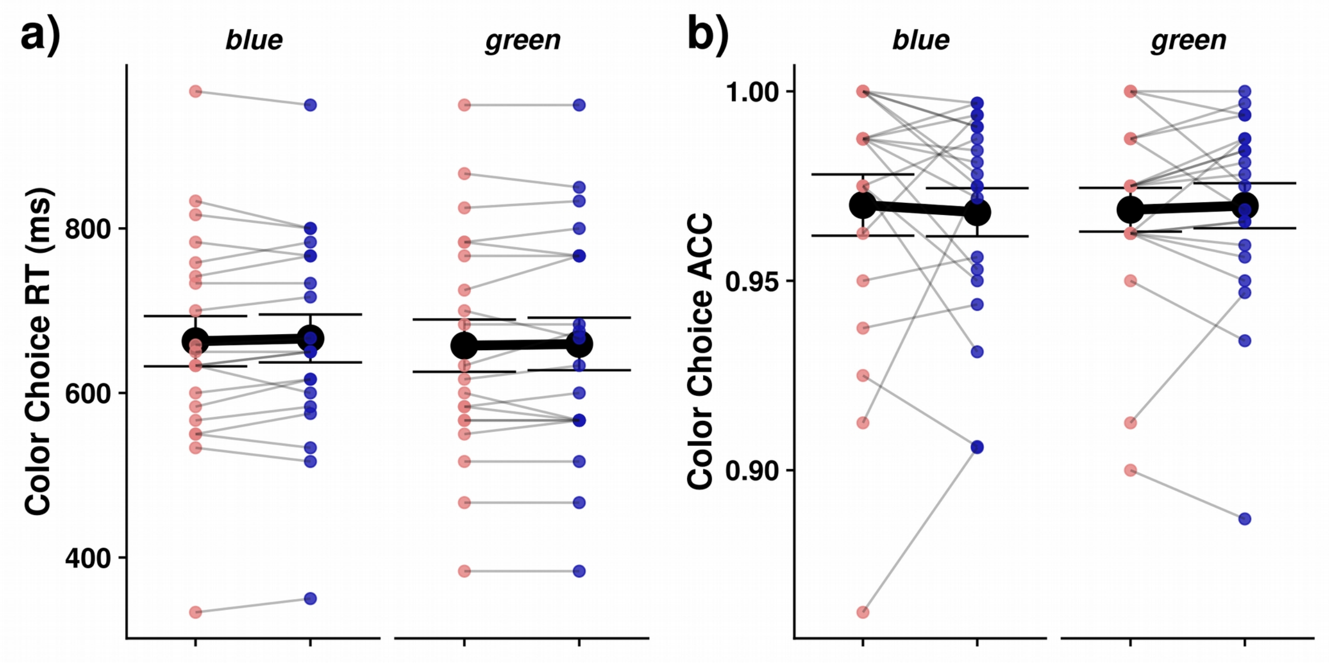

Figure 7.2. Color task data. a) Color choice reaction time (ms) associated with the stimulus color and orientation. b) Color choice accuracy associated with the stimulus color and orientation. Note, probability here is defined by the stimulus orientations shown during the color trials. High-probability = blue, low-probability shown in red.

Since orientation estimation trials had uniformly distributed orientations, these trials were divided based on the probability of those orientations from the color judgment trials. Comparisons between these ‘high’ and ‘low’ probability orientations yielded robust effects (*Figure 7.3*). Median absolute angular error for high-probability orientations were smaller (*M* = 6.68 deg, *SD* = 2.81 deg versus *M* = 7.88 deg, *SD* = 3.12 deg; *t*(19) = 4.66, *p* < .001). The same was true for the kurtosis of the angular error distribution. High-probability orientations had increased kurtosis (*M* = 3.85, *SD* = 4.93), as compared to low-probability orientations (*M* = 0.93, *SD* = 1.94), (*t*(19) = 2.88, *p* = .010).

Initiation times (IT; time taken for participant to start moving the response line) were unaffected by orientation probability. High-probability orientations (*M* = 726ms, *SD* = 124ms) were not significantly responded to faster than low-probability orientations (*M* = 720ms, *SD* = 126ms), (*t*(19) = 0.53, *p* = .600). This was the same when looking at movement times or reaction times as a whole (*p*s > .05). Vacillations in movement of the response line was marginally significantly predicted by probability, with high-probability orientations (*M* = .144, *SD* = 0.117) showing fewer vacillations compared to the low-probability orientations (*M* = .120, *SD* = 0.114), (*t*(19) = 2.00, *p* = .060).

Sometimes participants used the color response buttons for the orientation task (color trials were four times as likely). Less often they used the orientation buttons for the colored stimuli. Participants made these ‘false-initiations’ on a greater proportion of the orientation judgment trials (about 2.7% of the those trials) than the color judgment trials (about 0.4% of those trials), (*t*(19) = 3.14, *p* = .005). This suggests that participants were biased to responding towards the more common feature and indicates a prioritization of color over orientation.

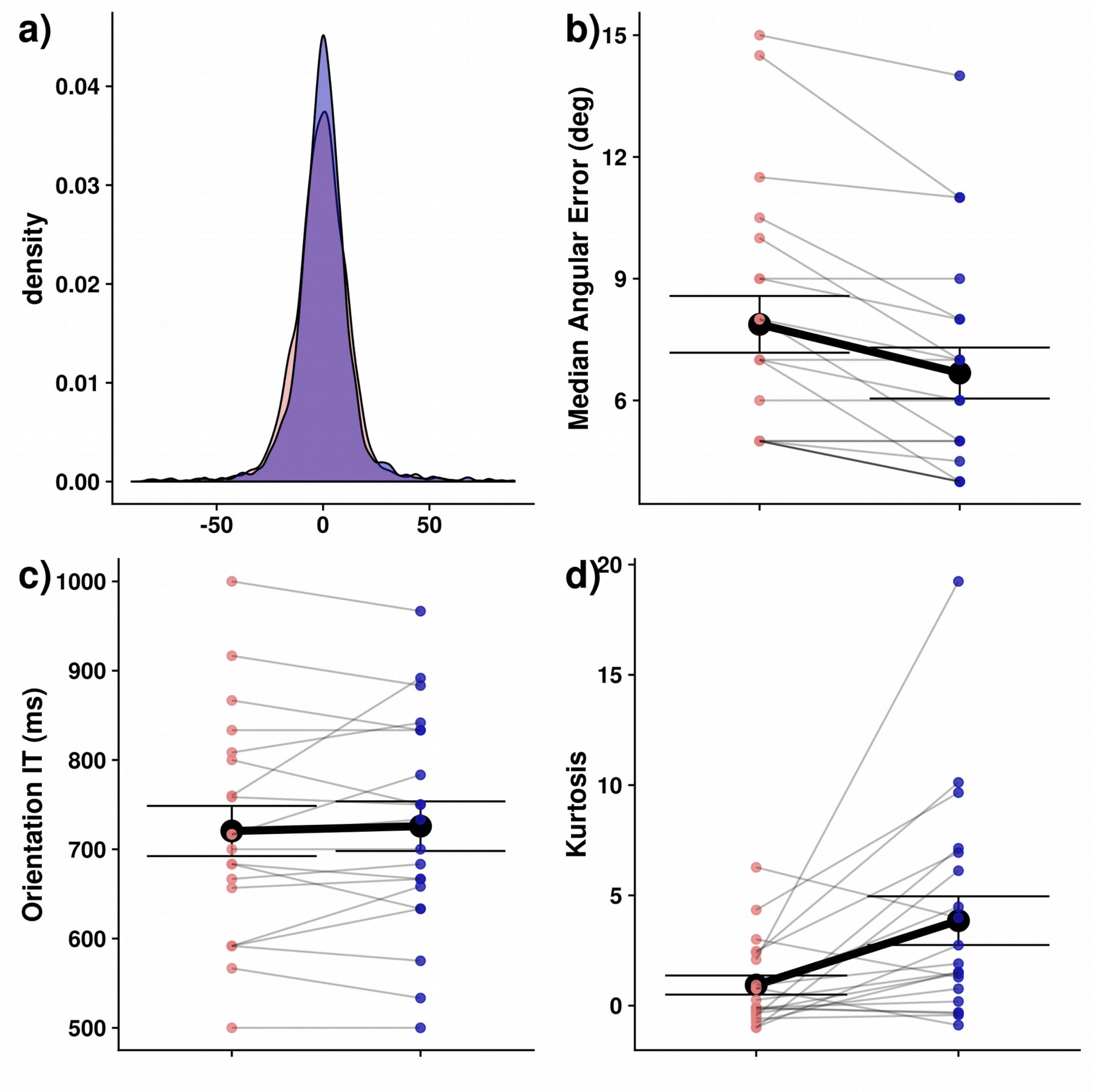

Figure 7.3. Performance metrics for the the orientation judgment trials. a) Distribution of angular errors. b) Median (absolute) angular errors. c) Initiation time (ms) associated with the orientation estimation. d) Kurtosis of the angular error distribution. Note, probability here is defined by the stimulus orientations shown during the color trials. High-probability = blue, low-probability shown in red.

For the post-experiment questionnaires, participants were at chance for the forced-choice questions (Q6/Q7) regarding orientation and color. Of the nine participants that circled the correct option in both questions, none were able to explicitly describe the distribution in the preceding open-ended question (Q5). Some participants spontaneously shared being aware that color trials were more likely, with some of them explicitly stating that using the fixation colors helped them to focus on the relevant feature.

7.4. Experiment 7.1: Discussion

Participants were *explicitly* clued to the type of judgment a trial would require. Because these cues were 100% valid, participants had every reason to attend to color for the color discrimination trials. This type of trial was also more probable (occurring 83.3% of the time). Any implicit attentional bias should have been towards color. This presumption is consistent with the ‘false initiation’ data where participants made color responses to orientation trials more often than color responses to orientation trials. Reaction times did not show orientation probability effects. But this task was procedurally more complex. Orientation responses were required infrequently and unpredictably, just as in the EEG session of Chapter 5, the other protocol where RT effects of probability were not found.

A clear orientation probability effect was observed median absolute angular error and kurtosis[11]. Participants were more precise for the probable orientation, and this orientation bias could only have been learnt during the color discrimination trials. This result aligns perfectly with the idea that orientation-selective V1 neurons are incrementally tuned by exposure to stimulus orientation, regardless of whether orientation estimations are explicitly required. This autonomous process obviates

---

[11] That even kurtosis was affected by probability in this experiment, where the two conditions have equal numbers of trials, further suggests that kurtosis changes seen in previous experiments were unlikely due to uneven sampling.

the need for any top-down attentional deployment and is consistent with modulation of *local* recurrent connectivity tested in the model of Chapter 6.

Complementary to the findings in Chapter 4, Chapter 7 demonstrated that orientation probability does not affect color discrimination performance. This is despite the same experiment also demonstrating that orientation information was encoded in the color discrimination trials, even though it was an irrelevant feature. This finding is problematic for object-based attention accounts (e.g. Egly, Driver, & Rafal, 1994). If the orientation probability was due to a 'whole object' selection, it would have implied better color judgment for gratings with the probable orientations. This was not the case. In Chapter 2, it was reported that spatial attention created effects that were independent from the effects of orientation probability. Combining all these findings, and the fact that a model that has no attentional component nevertheless successfully predicts when/how probability effects will manifest (Chapter 6), it is exceedingly difficult to posit any role of 'top-down attention' as the basis for probability learning. In fact, attentional accounts predict the *wrong* outcomes.

The next chapter ties in all the findings from this thesis, and explores why the V1 tuning mechanism posited is a better alternative explanation of probability effects, as compared with the use of 'attention' as a causal explanation.

Chapter 8: General Discussion

8.1. Probability Shapes Perception

Manipulations of probability typically cause one set of stimuli to be perceived more often than others. Training a monkey with the same orientations over an extended period of time selectively tunes the V1 neurons preferring that orientation (Ringach et al., 1997; Schoups et al., 2001). The question of interest here is whether probability can modulate perception in a similar way.

Chapter 2 demonstrated that while orientation probability manipulations cause effects on perceptual precision that is similar to the effects of an ‘attentional’ manipulation (peripheral spatial cuing), these effects were independent, suggesting that the underlying mechanisms are distinct. Chapter 3 examined how there could be different forms of probability with a different response method. Whereas orientation probability again affected perceptual precision, clustering analyses illustrated that spatial probability only affected detection/guess rates. This finding was corroborated with a novel approach to 2-alternative forced choice tasks in Chapter 4. An additional finding was that unlike spatial probability, feature-probability was domain-specific. Exposure to orientation probability was required for orientation discrimination to be affected, and exposure to color probability was required for color discrimination to be affected. This suggests a congruence with perceptual learning in monkeys, where selective tuning occurs as a consequence of exposure to the stimuli (Carandini & Ringach, 1997; Teich & Qian, 2003).

The tuning hypothesis was examined in Chapter 5 using electrophysiology. The C1 ERP component is thought to reflect early V1 activity (Di Russo et al., 2002). Consistent with a tuning-based account, high probability orientations dampened V1 activity while increasing perceptual precision. Chapter 6 demonstrated that simple modifications to computational models (Carandini &

Ringach, 1997; Teich & Qian, 2003) for perceptual learning (Ringach et al., 1997; Schoups et al., 2001) can be made to account for how V1 tuning occurs over time simply as a function of stimulus exposure. This model not only provided a mechanistic account of probability, but also made novel and *apriori* predictions about the impact of stimulus sequences, which was later verified by human data. This mechanism purports that probability learning is passive and automatic. Chapter 7 examined this by having orientation probability be exposed to participants while they were making color discriminations. Intermittent orientation estimation trials demonstrated that this probability information was acquired, even when the main color task did not benefit from this learning.

The evidence from psychophysical, electrophysiological and computational methods detailed in this thesis converge not only in indicating that probability shapes perception, but in indicating that the mechanism of feature probability is a selective tuning of the relevant feature processing units, as a result of exposure to that feature. While the ‘attentional’ (e.g. Hon & Tan, 2013) and ‘decisional’ (e.g. Menneer et al., 2010; Wolfe & Van Wert, 2010) accounts of probability are not necessarily incorrect, they are definitely incomplete since they do not describe *how* probability achieves its perceptual effects.

8.2. Why not attention?

The idea of ‘attention’ is associated with the idea of a ‘selection process of competing representations’ (Allport, 1993). While attentional manipulations might typically be thought of as involving the frontal and parietal cortices, these regions can modulate perceptual processing areas in a ‘top-down’ manner (e.g. Bressler et al., 2008). Cuing demonstrably modulates V1 in both human functional imaging (e.g. Brefczynski & DeYoe, 1999; Gandhi, Heeger & Boynton, 1999; Tootell et al., 1998), and in monkey single-cell recording (e.g. Sharma et al., 2015; Wang, Chen, Yan, Zhaoping, & Li, 2015) studies.

If ‘attention’ were a singular cause that links cuing and probability manipulations, the prediction for Chapter 2 would be that the two effects would interact, since they share the same stage of processing (Sternberg, 1969). Attention is not a unitary construct. The attention label can be (and has been) applied to a host of phenomena varying across multiple dimensions, e.g. spatial vs. feature-based cues, and these effects have distinctions which set them apart from each other (Carrasco, 2011). Is spatial probability equivalent to spatial attention? Chapters 3 and 4 suggest that, unlike orientation probability and spatial cuing, spatial probability does not affect perceptual precision. Further distinctions would have to be made between peripheral spatial cues and symbolic spatial cues, which are known to have differing temporal properties (Eimer, 2000). Peripheral cues can also affect the baseline firing of neurons (Cutrone et al., 2014; Wang et al., 2015), which would interact with the stimulus response. Explicit cues result in pre-frontal acetylcholine release (Parikh, Kozak, Martinez & Sarter, 2007), and cholinergic input affects the intracortical connections within the V1 cortex (Chen, Sugihara & Sur, 2015), which tunes V1 populations. Explicit spatial cuing might have different effects from (implicit) spatial probability due to any of these factors.

Is feature probability equivalent to feature attention? Similar to how orientation probability modulates V1 (Chapter 5), feature-based cuing of orientation modulates V1 activity (Kok, Jehee & de Lange, 2012). Feature-based attention modulates the P1 ERP component (Zhang & Luck, 2009) which has been source-localized to the extrastriate cortex (Clark & Hillyard, 1996; Di Russo, Martinez, Hillyard, 2003). Feature-based attention is characterized by changes to the anterior visual processing areas that *precede* the changes in upstream, posterior sections (Bartsch et al., 2017). P1 modulations were not observed in the electrophysiological investigation of orientation probability, and reflects a lack of top-down feedback processes. Chapter 7 demonstrates that orientation probability information is reliably acquired in a passive, automatic manner, with attention diverted to prioritize color instead of orientation information. It is unclear how top-down feature-based attention could account for this, since feature-based attention would predict a lack of probability effects.

Probability effects could fall under the umbrella of ‘attention effects’, in the sense that they all stem from some sort of a selection process of possible representations. However, appeals to attention as a *causal* explanation affords little-to-no predictive power, *if* the neural mechanisms involved are not explicitly specified.

8.3. Generalisability and Future Directions

This thesis has focused on the use of gratings as stimuli because they enable measurements of perceptual precision. Probability effects are known to affect other types of stimuli, and this section is meant to offer speculations on whether and how the mechanisms explored thus far might account for situations that do not involve gratings.

Probability effects obtained from manipulating the likelihood of certain glyphs (letters, digits, symbols, etc.) resemble probability effects obtained from manipulating gratings. Probable glyphs are detected faster and more accurately (Laberge & Tweedy, 1964; Miller & Pachella, 1973). These effects are also rapidly evoked (Hon et al., 2013) and can occur without explicit awareness (Hasher & Zacks, 1984). The experience-dependent neural tuning postulated for feature probability in this thesis could account for probability effects for such complex objects. Tuning can occur in the object-processing areas, e.g., in the inferior temporal (IT) or fusiform cortices. Similar to orientation columns in V1, the

IT cortex does have a columnar grouping of similar object-features (Tanaka, 2003; Tsunoda, Yamane, Nishizaki, & Tanifuji, 2001). Unsupervised visual experience rapidly affects the tuning of IT neurons in monkeys (e.g. Li & DiCarlo, 2010). Given the similarity in layout between V1 and IT cortices, the model presented in Chapter 6 can be adapted to other processing levels (e.g. object-representations in the IT cortex), to represent more complex probability effects. Experience-dependent neural tuning has also been postulated for other modalities, e.g. audition (Hanazawa et al., 2000), and changes in the primary auditory cortex could be modelled in a similar fashion to how V1 neurons change (e.g. Singer et al, 2017).

This thesis has focused on neural mechanisms relating to perception. This is not to exclude the possible role of decision-making (e.g. Menneer et al., 2010; Wolfe & Van Wert, 2010). Theoretical models demonstrate how early changes in perception, specifically with neural tuning, can affect later decision making stages (Eckstein et al., 2009; Summerfield & Egner, 2009; Yaeli & Meir, 2010). Exploratory analyses of evoked-related potentials in Chapter 5 revealed a C1-P300 correlation, providing support for this hypothesis. The idea that V1 changes *precede* behavioural changes was also recently demonstrated in mice (Jurjut, Georgieva, Busse & Katzner, 2017), which potentially explains why probability effects are hard to get rid of through instructions (Lau & Huang, 2010) or rewards (Wolfe et al., 2007): The underlying perceptual representations are still affected and not-countered.

In real-world searches or detection tasks, the target object is not only perceptually rare, but is also rarely responded to. High miss rates for rare targets might therefore not be strictly due to perception (Fleck & Mitroff, 2007). Given that normalization by synaptic inhibition/excitation is a computation that is not just apparent within the visual cortex (Sato et al.. 2016), but across other cortices as well (Carandini & Heeger, 2012), any manipulation that causes one subset of a population of neurons to fire more often than others would cause the type of modulation explored in Chapter 6. This suggests the possibility of local modulation in the motor cortices as a function of response probability. Alternatively, response probability, like spatial probability, could be due to an entirely distinct mechanism. While response probability was controlled for in this thesis, future research could examine what the relation between response and feature and/or spatial probability is.

8.4. Conclusion

There are facets of probability that remain to be explored (e.g. how perceptual changes relate to decision-making, what the role of response probability is, etc.). What is clear at this juncture is that when stimulus probability is manipulated, perception is affected. There are some forms of probability (e.g. spatial probability) that might recruit top-down mechanisms in a manner similar – but not identical – to ‘attentional’ cuing manipulations. However, certain probability effects can also be due to low-level changes occurring local to perceptual processing regions. For orientation probability (and perhaps color probability), this region is V1, but it is likely that similar synaptic mechanisms would be in play for cortical regions coding for other types of features/objects as well. Any theoretical account of probability going forward should consider the possibility for such mechanisms. Any practical use of probability manipulations should be mindful that probability is not synonymous with ‘attention’.

References

Ahmed, B., Anderson, J. C., Douglas, R. J., Martin, K. A. C., & Nelson, J. C. (1994). Polyneuronal innervation of spiny stellate neurons in cat visual cortex. *Journal of Comparative Neurology, 341*(1), 39–49. https://doi.org/10.1002/cne.903410105

Ales, J. M., Yates, J. L., & Norcia, A. M. (2010). V1 is not uniquely identified by polarity reversals of responses to upper and lower visual field stimuli. *NeuroImage*, *52*(4), 1401–1409. https://doi.org/10.1016/j.neuroimage.2010.05.016

Alink, A., Schwiedrzik, C. M., Kohler, A., Singer, W., & Muckli, L. (2010). Stimulus predictability reduces responses in primary visual cortex. *The Journal of Neuroscience*, *30*(8), 2960-2966. https://doi.org/10.1523/JNEUROSCI.3730-10.2010

Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. In *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 183–218).

Armstrong, K. M., Chang, M. H., & Moore, T. (2009). Selection and maintenance of spatial information by frontal eye field neurons. *The Journal of Neuroscience*, *29*(50), 15621-15629. https://doi.org/10.1523/JNEUROSCI.4465-09.2009

Anderson, B. (2011). There is no such thing as attention. *Frontiers in Psychology*, *2*, 1–8. https://doi.org/10.3389/fpsyg.2011.00246

Anderson, B. (2014). Probability and the changing shape of response distributions for orientation. *Journal of Vision*, *14*, 1–11. https://doi.org/10.1167/14.13.15.doi

Anderson, B., & Druker, M. (2013). Attention improves perceptual quality. *Psychonomic Bulletin & Review*, *20*(1), 120–127. https://doi.org/10.3758/s13423-012-0323-x

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: the “oblique effect” in man and animals. *Psychological Bulletin*, *78*(4), 266–278. https://doi.org/10.1037/h0033117

Bao, M., Yang, L., Rios, C., He, B., & Engel, S. A. (2010). Perceptual learning increases the strength of the earliest signals in visual cortex. *The Journal of Neuroscience*, *30*(45), 15080–15084. https://doi.org/10.1523/JNEUROSCI.5703-09.2010

Bartsch, M. V., Loewe, K., Merkel, C., Heinze, H. J., Schoenfeld, M. A., Tsotsos, J. K., & Hopf, J. M. (2017). Attention to color sharpens neural population tuning via feedback processing in the human visual cortex hierarchy. *The Journal of Neuroscience, 37*(43), 10346-10357. https://doi.org/10.1523/JNEUROSCI.0666-17.2017

Ben-Yishai, R., Bar-Or, R. L., & Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, *92*(9), 3844–3848. https://doi.org/10.1073/pnas.92.9.3844

Bledowski, C., Prvulovic, D., Hoechstetter, K., Scherg, M., Wibral, M., Goebel, R., & Linden, D. E. J. (2004). Localizing P300 generators in visual target and distractor processing: A combined event-related potential and functional magnetic resonance imaging study. *The Journal of Neuroscience*, *24*(42), 9353-9360. https://doi.org/10.1523/jneurosci.1897-04.2004

Brefczynski, J. A., & DeYoe, E. A. (1999). A physiological correlate of the “spotlight” of visual attention. *Nature Neuroscience*, *2*(4), 370–374. https://doi.org/10.1038/7280

Bressler, S. L., Tang, W., Sylvester, C. M., Shulman, G. L., & Corbetta, M. (2008). Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *The Journal of Neuroscience*, *28*(40), 10056-10061. https://doi.org/10.1523/JNEUROSCI.1776-08.2008

Carandini, M., & Heeger, D. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62. https://doi.org/10.1038/nrn3136

Carandini, M., & Ringach, D. L. (1997). Predictions of a recurrent model of orientation selectivity. *Vision Research*, *37*(21), 3061–3071. https://doi.org/10.1016/S0042-6989(97)00100-4

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*(13), 1484–1525. https://doi.org/10.1016/j.visres.2011.04.012

Carrasco, M., Giordano, A. M., & McElree, B. (2004). Temporal performance fields: Visual and attentional factors. *Vision Research*, *44*(12), 1351–1365. https://doi.org/10.1016/j.visres.2003.11.026

Carrasco, M., & McElree, B. (2001). Covert attention accelerates the rate of visual information processing. *Proceedings of the National Academy of Sciences*, *98*(9), 5363–7. https://doi.org/10.1073/pnas.081074098

Cheadle, S., Egner, T., Wyart, V., Wu, C., & Summerfield, C. (2015). Feature expectation heightens visual sensitivity during fine orientation discrimination. *Journal of Vision*, *15*(14), 14. https://doi.org/10.1167/15.14.14

Chen, N., Sugihara, H., & Sur, M. (2015). An acetylcholine-activated microcircuit drives temporal dynamics of cortical activity. *Nature Neuroscience*, *18*(6), 892-902. https://doi.org/10.1038/nn.4002

Chen, W. C., & Maitra, R. (2015). EMCluster: EM algorithm for model-based clustering of finite mixture gaussian distribution. Retrieved from http://cran.r-project.org/package=EMCluster

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204. https://doi.org/10.1017/S0140525X12000477

Clark, V. P., & Hillyard, S. A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *Journal of Cognitive Neuroscience*, *8*(5), 387–402. https://doi.org/10.1162/jocn.1996.8.5.387

Cort, B., & Anderson, B. (2013). Conditional probability modulates visual search efficiency. *Frontiers in Human Neuroscience*, *7*, 1–10. https://doi.org/10.3389/fnhum.2013.00683

Çukur, T., Nishimoto, S., Huth, A. G., & Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience*, *16*(6), 763–770. https://doi.org/10.1038/nn.3381

Cutrone, E. K., Heeger, D. J., & Carrasco, M. (2014). Attention enhances contrast appearance via increased input baseline of neural responses. *Journal of Vision*, *14*(14), 16. https://doi.org/10.1167/14.14.16

David, S. V., Hayden, B. Y., Mazer, J. A., & Gallant, J. L. (2008). Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron*, *59*(3), 509–521. https://doi.org/10.1016/j.neuron.2008.07.001

Dayan, P., & Abbott, L. F. (2001). *Theoretical Neuroscience*. Cambridge, MA: MIT Press.

Dean, I., Harper, N. S., & McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience*, *8*(12), 1684–1689. https://doi.org/10.1038/nn1541

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*(3), 292–307. https://doi.org/10.1037/1082-989X.2.3.292

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Di Russo, F., Martinez, A., & Hillyard, S. A. (2003). Source analysis of event-related cortical activity during visuo-spatial attention. *Cerebral Cortex*, *13*(5), 486–499. https://doi.org/10.1093/cercor/13.5.486

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, *15*(2), 95–111. https://doi.org/10.1002/hbm.10010

Druker, M., & Anderson, B. (2010). Spatial probability aids visual stimulus discrimination. *Frontiers in Human Neuroscience*, *4*, 1–10. https://doi.org/10.3389/fnhum.2010.00063

Dykes, J. R., & Pascal, V. (1981). The effect of stimulus probability on the perceptual processing of letters. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 528–537. https://doi.org/10.1037//0096-1523.7.3.528

Eckstein, M. P., Peterson, M. F., Pham, B. T., & Droll, J. A. (2009). Statistical decision theory to relate neurons to behavior in the study of covert visual attention. *Vision Research*, *49*(10), 1097–1128. https://doi.org/10.1016/j.visres.2008.12.008

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology. General*, *123*(2), 161–177. https://doi.org/10.1037/0096-3445.123.2.161

Eimer, M. (2000). The time course of spatial orienting elicited by central and peripheral cues: evidence from event-related brain potentials. *Biological Psychology*, *53*(2), 253-258. https://doi.org/10.1016/S0301-0511(00)00049-1

Fecteau, J. H., Korjoukov, I., & Roelfsema, P. R. (2009). Location and color biases have different influences on selective attention. *Vision Research*, *49*(9), 996–1005. https://doi.org/10.1016/j.visres.2009.03.013

Fink, G. R., Dolan, R. J., Halligan, P. W., Marshall, J. C., & Frith, C. D. (1997). Space-base and object-based visual attention: Shared and specific neural domains. *Brain*, *120*(11), 2013–2028. https://doi.org/10.1093/brain/120.11.2013

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738-743. https://doi.org/10.1038/nn.3689

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science*, *18*(11), 943–947. https://doi.org/10.1111/j.1467-9280.2007.02006.x

Gandhi, S. P., Heeger, D. J., & Boynton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences*, *96*(6), 3314–3319. https://doi.org/10.2307/47533

Gekas, N., Seitz, A. R., & Seriès, P. (2015). Expectations developed over multiple timescales facilitate visual search performance. *Journal of Vision*, *15*(9), 10. https://doi.org/10.1167/15.9.10

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics*, *67*(7), 1252–1268. https://doi.org/10.3758/BF03193557

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, *233*(4771), 1416-1419. http://dx.doi.org/10.1126/science.3749885

Ghose, G. M., Yang, T., & Maunsell, J. H. R. (2002). Physiological correlates of perceptual learning in monkey V1 and V2. *Journal of Neurophysiology*, *87*(4), 1867–88. https://doi.org/10.1152/jn.00690.2001

Gilbert, C. D. (1994). Early perceptual learning. *Proceedings of the National Academy of Sciences*, *91*(4), 1195-1197. http://dx.doi.org/10.1073/pnas.91.4.1195

Gonzalez, C. M. G., Clark, V. P., Fan, S., Luck, S. J., & Hillyard, S. A. (1994). Sources of attention-sensitive visual event-related potentials. *Brain Topography*, *7*(1), 41–51. https://doi.org/10.1007/BF01184836

Gorard, S. (2005). Revisiting A 90-year-old debate: The advantages of the mean deviation. *British Journal of Educational Studies*, *53*(4), 417–430. https://doi.org/10.1111/j.1467-8527.2005.00304.x

Gratton, G. (1997). Attention and probability effects in the human occipital cortex: an optical imaging study. *Neuroreport*, *8*(7), 1749–1753. https://doi.org/10.1097/00001756-199705060-00036

Hale, D. J. (1969). Repetition and probability effects in a serial choice reaction task. *Acta Psychologica*, *29*, 163–171. https://doi.org/10.1016/0001-6918(69)90011-0

Hanazawa, A., Komatsu, H., & Murakami, I. (2000). Neural selectivity for hue and saturation of colour in the primary visual cortex of the monkey. *European Journal of Neuroscience*, *12*(5), 1753–1763. https://doi.org/10.1046/j.1460-9568.2000.00041.x

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: the case of frequency of occurrence. *The American Psychologist*, *39*(12), 1372–1388. https://doi.org/10.1037/0003-066X.39.12.1372

Haskell, C., & Anderson, B. (2015). Attentional effects on orientation judgements are dependent on memory consolidation processes. *The Quarterly Journal of Experimental Psychology*, *69*(11), 2147–2165. https://doi.org/10.1080/17470218.2015.1105830

Hilgenstock, R., Weiss, T., & Witte, O. W. (2014). You'd better think twice: Post-decision perceptual confidence. *NeuroImage*, *99*, 323–331. https://doi.org/10.1016/j.neuroimage.2014.05.049

Hon, N., & Tan, C. H. (2013). Why rare targets are slow: evidence that the target probability effect has an attentional locus. *Attention, Perception & Psychophysics*, *75*(3), 388–393. https://doi.org/10.3758/s13414-013-0434-0

Hon, N., Yap, M. J., & Jabar, S. B. (2013). The trajectory of the target probability effect. *Attention, Perception & Psychophysics*, *75*(4), 661–666. https://doi.org/10.3758/s13414-013-0429-x

Hosoya, T., Baccus, S., & Meister, M. (2005). Dynamic predictive coding by the retina. *Nature*, *436*(7047), 71–77. https://doi.org/10.1038/nature03689

Hung, S.-C., & Seitz, A. R. (2014). Prolonged training at threshold promotes robust retinotopic specificity in perceptual learning. *The Journal of Neuroscience*, *34*(25), 8423–8431. https://doi.org/10.1523/jneurosci.0745-14.2014

Iarocci, G., Burack, J. A., Shore, D. I., Mottron, L., & Enns, J. T. (2006). Global–local visual processing in high functioning children with autism: structural vs. implicit task biases. *Journal of Autism and Developmental Disorders*, *36*(1), 117-129. https://doi.org/10.1007/s10803-005-0045-2

Jabar, S. B., & Anderson, B. (2015). Probability shapes perceptual precision: A study in orientation estimation. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(6), 1666–1679. https://doi.org/10.1037/xhp0000121

Jabar, S. B., & Anderson, B. (2017a). Not all probabilities are equivalent: Evidence from orientation versus spatial probability learning. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(5), 853–867. https://doi.org/10.1037/xhp0000354

Jabar, S. B., & Anderson, B. (2017b). Orientation probability and spatial exogenous cuing improve perceptual precision and response speed by different mechanisms. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.00183

Jabar, S. B., Filipowicz, A., & Anderson, B. (2017a). Tuned by experience: How orientation probability modulates early perceptual processing. *Vision Research*, *138*, 86–96. https://doi.org/10.1016/j.visres.2017.07.008

Jabar, S. B., Filipowicz, A., & Anderson, B. (2017b). Knowing where is different from knowing what. Distinct response time profiles and accuracy effects for target location, orientation, and color probability. *Attention, Perception & Psychophysics*, *79*(8), 2338-2353. https://doi.org/10.3758/s13414-017-1412-8

Jiang, Y. V., Sha, L. Z., & Remington, R. W. (2015). Modulation of spatial attention by goals, statistical learning, and monetary reward. *Attention, Perception, & Psychophysics*, *77*(7), 2189–2206. https://doi.org/10.3758/s13414-015-0952-z

Jin, H., Xu, G., Zhang, J. X., Ye, Z., Wang, S., Zhao, L., Lin, C. D., & Mo, L. (2010). Athletic training in badminton players modulates the early C1 component of visual evoked potentials: A preliminary investigation. *International Journal of Psychophysiology*, *78*(3), 308–314. https://doi.org/10.1016/j.ijpsycho.2010.09.005

Johnson, E. N., Hawken, M. J., & Shapley, R. (2008). The orientation selectivity of color-responsive neurons in macaque V1. *The Journal of Neuroscience*, *28*(32), 8096–8106. https://doi.org/10.1523/jneurosci.1404-08.2008

Kastner, S., & Ungerleider, S. (2000). Mechanisms of visual attention in the human cortex. *Annual Review of Neuroscience*, *23*(1), 315–341. https://doi.org/10.1146/annurev.neuro.23.1.315

Kelly, S. P., Gomez-Ramirez, M., & Foxe, J. J. (2008). Spatial attention modulates initial afferent activity in human primary visual cortex. *Cerebral Cortex*, *18*(11), 2629–2636. https://doi.org/10.1093/cercor/bhn022

Kelly, S. P., Schroeder, C. E., & Lalor, E. C. (2013). What does polarity inversion of extrastriate activity tell us about striate contributions to the early VEP? A comment on Ales et al. (2010). *NeuroImage*, *76*, 442–445. https://doi.org/10.1016/j.neuroimage.2012.03.081

Kok, P., Jehee, J. F. M., de Lange. (2012). Less is more: expectation sharpens representations in the primary visual cortex. *Neuron*, *75*(2), 265–70. https://doi.org/10.1016/j.neuron.2012.04.034

Laberge, D., & Tweedy, J. R. (1964). Presentation probability and choice time. *Journal of Experimental Psychology*, *68*(5), 477–481. https://doi.org/10.1037/h0048691

Lau, J. S. H., & Huang, L. (2010). The prevalence effect is determined by past experience, not future prospects. *Vision Research*, *50*(15), 1469–1474. https://doi.org/10.1016/j.visres.2010.04.020

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*(8), 1279–92. https://doi.org/10.3758/BF03194543

Li, B., Peterson, M. R., & Freeman, R. D. (2003). Oblique effect : A neural basis in the visual cortex oblique. *Journal of Neurophysiology*, *90*, 204–217. https://doi.org/10.1152/jn.00954.2002

Li, N., & DiCarlo, J. J. (2010). Unsupervised natural visual experience rapidly reshapes size-invariant object representation in inferior temporal cortex. *Neuron*, *67*(6), 1062–1075. https://doi.org/10.1016/j.neuron.2010.08.029

Ling, S., Jehee, J. F. M., & Pestilli, F. (2015). A review of the mechanisms by which attentional feedback shapes visual selectivity. *Brain Structure and Function*, *220*(3), 1237–1250. https://doi.org/10.1007/s00429-014-0818-5

Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Research*, *49*(10), 1194–1204. https://doi.org/10.1016/j.visres.2008.05.025

Liu, T., & Becker, M. W. (2013). Serial consolidation of orientation information into visual short-term memory. *Psychological Science*, *24*(6), 1044–1050. https://doi.org/10.1177/0956797612464381

Liu, T., Stevens, S. T., & Carrasco, M. (2007). Comparing the time course and efficacy of spatial and feature-based attention. *Vision Research*, *47*(1), 108–113. https://doi.org/10.1016/j.visres.2006.09.017

MacLeod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology*, *111*(1), 107. http://dx.doi.org/10.1037/0021-843X.111.1.107

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *The Journal of Neuroscience*, *28*(47), 12539–12545. https://doi.org/10.1523/JNEUROSCI.2925-08.2008

Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, *14*(9), 744–751. https://doi.org/10.1016/j.cub.2004.04.028

Mattes, S., Ulrich, R., & Miller, J. (2002). Response force in RT tasks: Isolating effects of stimulus probability and response probability. *Visual Cognition*, *9*(7), 477–501. https://doi.org/10.1080/13506280143000548

Menneer, T., Donnelly, N., Godwin, H. J., & Cave, K. R. (2010). High or low target prevalence increases the dual-target cost in visual search. *Journal of Experimental Psychology. Applied*, *16*(2), 133–44. https://doi.org/10.1037/a0019569

Miller, J. O., & Pachella, R. G. (1973). Locus of the stimulus probability effect. *Journal of Experimental Psychology*, *101*(2), 227–231. https://doi.org/10.1037/h0035214

Miller, K. D. (1994). A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON-and OFF-center inputs. *The Journal of Neuroscience*, *14*(1), 409-441. http://www.jneurosci.org/content/14/1/409.short

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. Retrieved from http://cran.r-project.org/package=BayesFactor

Mounts, J. R., & Tomaselli, R. G. (2005). Competition for representation is mediated by relative attentional salience. *Acta Psychologica*, *118*(3), 261-275. https://doi.org/10.1016/j.actpsy.2004.09.001

Jurjut, O., Georgieva, P., Busse, L., & Katzner, S. (2017). Learning enhances sensory processing in mouse V1 before improving behavior. *The Journal of Neuroscience*, *37*(27), 6460-6474. https://doi.org/10.1523/JNEUROSCI.3485-16.2017

Paltoglou, A. E., & Neri, P. (2012). Attentional control of sensory tuning in human visual perception. *Journal of Neurophysiology*, *107*(5), 1260–74. https://doi.org/10.1152/jn.00776.2011

Parikh, V., Kozak, R., Martinez, V., & Sarter, M. (2007). Prefrontal acetylcholine release controls cue detection on multiple timescales. *Neuron*, *56*(1), 141-154. https://doi.org/10.1016/j.neuron.2007.08.025

Peirce, J. W. (2008). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*, 10. https://doi.org/10.3389/neuro.11.010.2008

Petrusic, W. M., & Baranski, J. V. (2009). Probability assessment with response times and confidence in perception and knowledge. *Acta Psychologica*, *130*(2), 103–114. https://doi.org/10.1016/j.actpsy.2008.10.008

Polich, J. (1990). P300, probability, and interstimulus interval. *Psychophysiology*, *27*(4), 396–403. https://doi.org/10.1111/j.1469-8986.1990.tb02333.x

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In *Attention and performance X: Control of language processes* (pp. 531–556).

Pouget, A., Deneve, S., Ducom, J. C., & Latham, P. E. (1999). Narrow versus wide tuning curves: What's best for a population code?. *Neural Computation*, *11*(1), 85-90. https://doi.org/10.1162/089976699300016818

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience*, *1*(2), 125-132. https://doi.org/10.1038/35039062

Pourtois, G., Rauss, K. S., Vuilleumier, P., & Schwartz, S. (2008). Effects of perceptual learning on primary visual cortex activity in humans. *Vision Research*, *48*(1), 55–62. https://doi.org/10.1016/j.visres.2007.10.027

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: 1. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 261–282. https://doi.org/10.1037/0096-1523.24.1.261

Raiguel, S., Vogels, R., Mysore, S. G., & Orban, G. A. (2006). Learning to see the difference specifically alters the most informative V4 neurons. *The Journal of Neuroscience*, *26*(24), 6589–602. https://doi.org/10.1523/JNEUROSCI.0457-06.2006

Rauss, K., Schwartz, S., & Pourtois, G. (2011). Top-down effects on early visual processing in humans: A predictive coding framework. *Neuroscience and Biobehavioral Reviews*, *35*(5), 1237–1253. https://doi.org/10.1016/j.neubiorev.2010.12.011

Ravden, D., & Polich, J. (1998). Habituation of P300 from visual stimuli. *International Journal of Psychophysiology*, *30*(3), 359–365. https://doi.org/10.1016/S0167-8760(98)00039-7

Rich, A. N., Kunar, M. A., Van Wert, M. J., Hidalgo-Sotelo, B., Horowitz, T. S., & Wolfe, J. M. (2008). Why do we miss rare targets? Exploring the boundaries of the low prevalence effect. *Journal of Vision*, *8*(15), 1–17. https://doi.org/10.1167/8.15.15

Ringach, D. L., Hawken, M. J., & Shapley, R. (1997). The dynamics of orientation tuning in macaque primary visual cortex. *Nature*, *387*(6630), 281–284. https://doi.org/10.1038/387281a0

Rohrbaugh, J. W., Donchin, E., & Eriksen, C. W. (1974). Decision making and the P300 component of the cortical evoked response. *Perception & Psychophysics*, *15*(2), 368–374. https://doi.org/10.3758/BF03213960

Samaha, J., Barrett, J. J., Sheldon, A. D., LaRocque, J. J., & Postle, B. R. (2016). Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00851

Sato, T. K., Haider, B., Häusser, M., & Carandini, M. (2016). An excitatory basis for divisive normalization in visual cortex. *Nature Neuroscience*, *19*(4), 568-570. https://doi.org/10.1038/nn.4249

Schoups, A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, *412*(6846), 549–553. https://doi.org/10.1038/35087601

Schwartz, S., & Kirsner, K. (1982). Laterality effects in visual information processing: Hemispheric specialisation or the orienting of attention?. *The Quarterly Journal of Experimental Psychology*, *34*(1), 61-77. https://doi.org/10.1080/14640748208400858

Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences*, *90*(22), 10749-10753. https://doi.org/10.1073/pnas.90.22.10749

Sha, L. Z., Remington, R. W., & Jiang, Y. V. (2017). Short-term and long-term attentional biases to frequently encountered target features. *Attention, Perception, & Psychophysics*, *79*(5), 1311–1322. https://doi.org/10.3758/s13414-017-1317-6

Shaqiri, A., & Anderson, B. (2013). Priming and statistical learning in right brain damaged patients. *Neuropsychologia*, *51*(13), 2526–2533. https://doi.org/10.1016/j.neuropsychologia.2013.09.024

Sharma, J., Sugihara, H., Katz, Y., Schummers, J., Tenenbaum, J., & Sur, M. (2015). Spatial attention and temporal expectation under timed uncertainty predictably modulate neuronal responses in Monkey V1. *Cerebral Cortex*, *25*(9), 2894–2906. https://doi.org/10.1093/cercor/bhu086

Sillito, A. M., Kemp, J. A., Milson, J. A., & Berardi, N. (1980). A re-evaluation of the mechanisms underlying simple cell orientation selectivity. *Brain Research*, *194*(2), 517–520. https://doi.org/10.1016/0006-8993(80)91234-2

Singer, Y., Teramoto, Y., Willmore, B., Schnupp, J. W., King, A. J., & Harper, N. S. (2017). Sensory cortex is optimised for prediction of future input. *BioRxiv*, 224758. https://doi.org/10.1101/224758

Slagter, H. A., Alilovic, J., Timmermans, B., & Reteig, L. C. (2016). Predictions, not attention, modulate the first feedforward sweep of cortical information processing. In *Annual meeting of the Society for Neuroscience*. San Diego, CA. http://www.abstractsonline.com/pp8/#!/4071/presentation/26117

Smeets, J. B. J., Wijdenes, L. O., & Brenner, E. (2016). Reacting with or without detecting. *Motor Control*, *20*(2), 200–205. https://doi.org/10.1123/mc.2015-0081

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, *30*, 276–315. https://doi.org/10.1016/0001-6918(69)90055-9

Strüber, D., & Polich, J. (2002). P300 and slow wave from oddball and single-stimulus visual tasks: Inter-stimulus interval effects. *International Journal of Psychophysiology*, *45*(3), 187–196. https://doi.org/10.1016/S0167-8760(02)00071-5

Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*(9), 403–409. https://doi.org/10.1016/j.tics.2009.06.003

Sun, P., & Landy, M. S. (2016). A two-stage process model of sensory discrimination: An alternative to drift-diffusion. *The Journal of Neuroscience*, *36*(44), 11259–11274. https://doi.org/10.1523/jneurosci.1367-16.2016

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*(3700), 1187-1188. https://doi.org/10.1126/science.150.3700.1187

Tanaka, K. (2003). Columns for complex visual object features in the inferotemporal cortex: Clustering of cells with similar but slightly different stimulus selectivities. *Cerebral Cortex*, *13*(1), 90–99. https://doi.org/10.1093/cercor/13.1.90

Teich, A. F., & Qian, N. (2003). Learning and adaptation in a recurrent model of V1 orientation selectivity. *Journal of Neurophysiology*, *89*(4), 2086–2100. https://doi.org/10.1152/jn.00970.2002

Teich, A. F., & Qian, N. (2010). V1 orientation plasticity is explained by broadly tuned feedforward inputs and intracortical sharpening. *Visual Neuroscience*, *27*(2), 57-73. https://doi.org/10.1017/S0952523810000039

Teigen, K. H., & Keren, G. (2003). Surprises: Low probabilities or high contrasts? *Cognition*, *87*(2), 55–71. https://doi.org/10.1016/s0010-0277(02)00201-9

Tootell, R. B. H., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., & Dale, A. M. (1998). The retinotopy of visual spatial attention. *Neuron*, *21*(6), 1409–1422. https://doi.org/10.1016/S0896-6273(00)80659-5

Tseng, P., Chang, C. F., Chiau, H. Y., Liang, W. K., Liu, C. L., Hsu, T. Y., Hung, D. L., Tzeng, O. J. L., Juan, C. H. (2013). The dorsal attentional system in oculomotor learning of predictive information. *Frontiers in Human Neuroscience*, *7*, 404. https://doi.org/10.3389/fnhum.2013.00404

Tsumoto, T., Eckart, W., & Creutzfeldt, O. D. (1979). Modification of orientation sensitivity of cat visual cortex neurons by removal of GABA-mediated inhibition. *Experimental Brain Research*, *34*(2), 351–363. https://doi.org/10.1007/BF00235678

Tsunoda, K., Yamane, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience*, *4*(8), 832–838. https://doi.org/10.1038/90547

Vincent, B. (2011). Covert visual search: Prior beliefs are optimally combined with sensory evidence. *Journal of Vision*, *11*, 1–15. https://doi.org/10.1167/11.13.25

Walthew, C., & Gilchrist, I. D. (2006). Target location probability effects in visual search: an effect of sequential dependencies. *Journal of Experimental Psychology. Human Perception and Performance*, *32*(5), 1294–1301. https://doi.org/10.1167/6.6.842

Wang, F., Chen, M., Yan, Y., Zhaoping, L., & Li, W. (2015). Modulation of neuronal responses by exogenous attention in macaque primary visual cortex. *The Journal of Neuroscience*, *35*(39), 13419–13429. https://doi.org/10.1523/JNEUROSCI.0527-15.2015

Wang, L., Zheng, J., Huang, S., & Sun, H. (2015). P300 and decision making under risk and ambiguity. *Computational Intelligence and Neuroscience*, *2015*, 1–7. https://doi.org/10.1155/2015/108417

Wei, X. X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain'anti-Bayesian'percepts. *Nature Neuroscience*, *18*(10), 1509-1517. https://doi.org/10.1038/nn.4105

Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental Psychology. General*, *136*(4), 623–38. https://doi.org/10.1037/0096-3445.136.4.623

Wolfe, J. M., & Van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology*, *20*(2), 121–124. https://doi.org/10.1016/j.cub.2009.11.066

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*(8), 2031-2046. https://doi.org/10.3758/APP.72.8.2031

Wyart, V., Nobre, A. C., & Summerfield, C. (2012). Dissociable prior influences of signal probability and relevance on visual contrast sensitivity. *Proceedings of the National Academy of Sciences*, *109*(9), 3593–3598. https://doi.org/10.1073/pnas.1120118109

Yaeli, S., & Meir, R. (2010). Error-based analysis of optimal tuning functions explains phenomena observed in sensory neurons. *Frontiers in Computational Neuroscience*, *4*, 130. https://doi.org/10.3389/fncom.2010.00130

Yang, T., & Maunsell, J. H. (2004). The effect of perceptual learning on neuronal responses in monkey visual area V4. *The Journal of Neuroscience*, *24*(7), 1617–1626. https://doi.org/10.1523/jneurosci.4442-03.2004

Yantis, S., & Serences, J. T. (2003). Cortical mechanisms of space-based and object-based attentional control. *Current Opinion in Neurobiology*, *13*(2), 187–193. https://doi.org/10.1016/S0959-4388(03)00033-3

Zhang, G. (2015). ERP C1 is top-down modulated by orientation perceptual learning. *Journal of Vision*, *15*, 1–11. https://doi.org/10.1167/15.10.8

Zhang, K., & Sejnowski, T. J. (1999). Neuronal tuning: To sharpen or broaden?. *Neural Computation*, *11*(1), 75-84. https://doi.org/10.1162/089976699300016809

Zhang, T., Xiao, L. Q., Klein, S. A., Levi, D. M., & Yu, C. (2010). Decoupling location specificity from perceptual learning of orientation discrimination. *Vision Research*, *50*(4), 368–374. https://doi.org/10.1016/j.visres.2009.08.024

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233–235. https://doi.org/10.1038/nature06860

Zhang, W., & Luck, S. J. (2009). Feature-based attention modulates feedforward visual processing. *Nature Neuroscience*, *12*(1), 24–25. https://doi.org/10.1038/nn.2223

Appendix A: Python implementation of sequence-driven recurrent V1 tuning

Note: This code uses Python version 2. The code is also accessible at https://github.com/syaheed/public-code/blob/master/V1model.py

```
### Import libraries
import numpy as np
import random

### Build functions
def readSeq(filename):  #function read in a sequence (e.g. to test a fixed sequence)
    f = open(filename)
    content = f.readlines()
    num_stim = len(content)-1
    stims = np.zeros(num_stim)
    for t in range(0,num_stim):
        line = content[t+1]
        trial,stim = line.split(',')
        stims[t] = stim.strip()
    return(stims)

def degDiff(a,b): #function to calculate distance in circular space
    diff =  a - b
    diff[diff > 90.0] = diff[diff > 90.0] - 180
    diff[diff < -90.0] = 180 + diff[diff < -90.0]
    return(diff)

def memPotential(Jf,pref,stim,sigf): # formulae from Teich and Qian (2003)
    diffs = degDiff(pref,stim)
    return(Jf * np.exp( (-(diffs)**2)/(2*sigf**2)    )  )

def rand(items): # to generate a noise vector
    return( np.array([(random.random()-0.5)*2 for r in range(items)]) )

def cosd(d): # just a degree form of cosine for convenience
    return np.cos(np.radians(d))

def conn(theta,a): # connection formulae (for E and I)
    c  = (cosd(2*(theta))+1)**a
    return(c/sum(c))

def even_spread(N): # function to evenly spread stimulus/neurons across the possible orientations
    interval = 180.0/N
    return(np.linspace(interval/2.0 - 90.0,90.0 - interval/2.0,num = N))
```

```
def V_calc1(c,J,R): # calculate Ve/Vi for one neuron
    return(np.dot(c*J,R))

def V_calc2(clist,Jlist,R): # calculate Ve/Vi across the whole set of neurons
    V = np.empty(N, dtype=float)
    for n in range(N):
        V[n] = V_calc1(clist[n,:],Jlist[n],R)
    return(V)

def R_calc1(V_init,e_list,i_list,Je_list, Ji_list, R): # calculate R for one time point
        Ve = V_calc2(e_list,Je_list,R)
        Vi = V_calc2(i_list,Ji_list,R)
        V =  Ve - Vi + V_init
        noise = rand(N) * np.mean(V) * noise_level
        R = (V+noise) * alpha
        R[R<0.0] = 0.0
        return(R)

def R_calc2(V_init,e_list,i_list,Je_list, Ji_list, R): # calculate R across time points
    for t in range(0,1000,tau):
        R = R_calc1(V_init,e_list,i_list,Je_list, Ji_list, R)
    return(R)

def make_stim(num_trials,prob): # make a probability sequence
    high_N = int(round(prob*num_trials))
    low_N = num_trials - high_N
    high_prob = np.round(np.random.uniform(low=-90, high=0, size=high_N))
    low_prob = np.round(np.random.uniform(low=0, high=90, size=low_N))
    stim = np.concatenate((high_prob,low_prob),axis = 0)
    return(np.random.permutation(stim))

def decode(fire): # for vector decoding
    fireX = np.dot(xvect,fire)/N
    fireY = np.dot(yvect,fire)/N
    return( np.degrees(np.arctan2(fireX, fireY))/2 ) # convert doubled (circular) space to axial space


def normalised_firing(R): #Normalization response of an individual neuron given summed activity
    return(R/sum(R))

def J_update(J_orig,J,A,R,decay_rate):
    reduction = normalised_firing(R) * A # excitation reduction is based on normalised firing on trial
    J = J * (1- reduction) # excitation reduction is also scaled on what current excitation is
    decay = (J_orig-J) * decay_rate # Assuming that there is also an x% loss of reduction per trial
    J = J + decay
    J[J<0] = 0
    return(J)
```

```
# Parameters to change
num_runs  = 1000 # number of times to run the model
Ai = 0.000 # reduction of Ji
Ae = 0.010 # maximum reduction of Je
decay_rate = 0.010  # proportion decay of Je towards its original value
noise_level = 2.50 # multiplier for the noise added to the membrane potential at each time step

# Parameters that are fixed (inherited from Teich & Qian,2003)
N = 128 # number of neurons
Je_orig = 1.1 # strength of intracortical excitation
Ji_orig = 1.1 # strength of intracortical inhibition
Jf = 0.5 # feedforward strength
sigf = 45.0 # feedforward tuning
ae = 2.2
ai = 1.4
tau = 15 # membrane time constant (ms)
alpha = 10.0 # spikes/s/mV

# initialize vectors
pref_list = even_spread(N) # neurons are evenly distributed
zero_index = np.argmin(abs(pref_list-(0)))
zero_pref = pref_list[zero_index]

diffs = degDiff(pref_list,zero_pref)
e  = conn(diffs,ae)
i  = conn(diffs,ai)

e_list = np.zeros(shape=(N,N))
i_list = np.zeros(shape=(N,N))
for n in range(N):
    e_list[n,:] = np.roll(e,n-zero_index) # to circle shift the numbers
    i_list[n,:] = np.roll(i,n-zero_index)

yvect = cosd(2*pref_list) # yaxis in double (circular) space
xvect = cosd(2*pref_list-90)  # xaxis in double (circular) space
```

```
# build a set of input stimuli with a probability distribution
for run in range(1,num_runs+1):
    side_type = random.choice([1,-1]) # decide which set of orientations is probable
    Je = [Je_orig]*N
    Ji = [Ji_orig]*N

    stims = side_type * make_stim(20,0.8) # to constrain probability in 20 trial chunks
    for i in range(20-1): # to get 400 trials in total (20 sets of 20)
        temp = side_type * make_stim(20,0.8)
        stims = np.concatenate((stims,temp),axis = 0)

    num_trials = len(stims)
    final_Je = np.zeros(shape=(N,len(stims)))
    final_R = np.zeros(shape=(N,len(stims)))
    decoded = np.zeros(len(stims))
    error = np.zeros(len(stims))

    ### This is the main loop where the tuning occurs
    for s in range(num_trials):
        stim = stims[s]
        V_init = memPotential(Jf,pref_list,stim,sigf)
        noise = rand(N) * np.mean(V_init) * noise_level
        R = (V_init + noise) * alpha
        R[R<0.0] = 0.0 # floor all negative firing rates due to noise
        R = R_calc2(V_init,e_list,i_list,Je, Ji, R)
        Je = J_update(Je_orig,Je,Ae,R,decay_rate) # update Je for next trial
        Ji = J_update(Ji_orig,Ji,Ai,R,decay_rate) # no need to update Ji
        final_R[:,s] = R
        final_Je[:,s] = Je
        decoded[s] = decode(R)
        error[s] =  decoded[s] - stim

    ### Save the data of each run
    cwd = '/home/syaheed/RecurV1' # change to intended path of data output
    fname = cwd + '/output/firingRate_' + str(run) + '_' + '.csv' # this is the main output
    fname2 = cwd + '/output/Je_' + str(run) + '_' + '.csv'  # Je value for every neuron at each trial

    run_c = np.ones(len(stims)) * run
    side_c = np.ones(len(stims)) * side_type
    Ae_c = np.ones(len(stims)) * Ae
    decay_c = np.ones(len(stims)) * decay_rate

    data = np.array([run_c, side_c, Ae_c, decay_c, stims, decoded, error])
    np.savetxt(fname, np.transpose(data), delimiter=',')
    np.savetxt(fname2, np.transpose(final_Je), delimiter=',')
```