A common framework for visual crowding in typical and amblyopic vision

A thesis submitted for the degree of Doctor of Philosophy at University College London (UCL)

by

Alexandra V Kalpadakis-Smith

B.Sc in Psychology, MRes in Cognitive Neuroscience

Department of Experimental Psychology

Division of Psychology and Language Sciences

January 11th 2019

Declaration of Authorship

I, Alexandra Kalpadakis-Smith, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Amblyopia is a developmental disorder of vision characterised by reduced acuity in one eye despite optical correction. When associated with strabismus, foveal vision is impaired by crowding: objects that are readily recognised in isolation become indistinguishable in clutter. In typical vision, crowding is minimal in the fovea and increases in the visual periphery. According to pooling accounts, the increase of crowding in the periphery arises from the integration of adjacent objects to promote perceptual homogeneity where sampling is insufficient and neurons have large receptive fields. It is unclear whether amblyopic crowding represents the same process. In this thesis I characterise amblyopic crowding, and investigate whether it can be understood within the same pooling framework as peripheral crowding. First, I show that amblyopic crowding systematically shifts the appearance of crowded objects to promote perceptual homogeneity, matching the perceptual effects in the periphery. A model simulating pooled responses of populations of visual neurons accurately characterises these effects in amblyopic and peripheral crowding, suggesting a common underlying mechanism. Second, I investigate the pattern of amblyopic crowding across the visual field and its neural correlates. I show that amblyopic crowding is elevated relative to typical vision in both fovea and periphery. At a group level, the increase of crowding in amblyopia and the typical periphery matches the increase of fMRI population receptive field (pRF) estimates in V1, V2, and V3, but at an individual level there is no correlation. Finally, I investigate the effects of higher-level grouping processes by examining whether uniformity in global clutter configuration modulates amblyopic crowding. I find that in most cases, clutter disrupts recognition in amblyopia regardless of global configuration, suggesting that contrary to the periphery, amblyopic crowding is largely unaffected by higher-level grouping processes. Therefore, on the whole pooling provides a successful framework for both amblyopic and peripheral crowding.

Impact Statement

Amblyopia is one of the most common developmental disorders of vision, with a prevalence of 2-5% of children in the UK. It is clinically defined as a two-line difference in acuity between the two eyes. When amblyopia is associated with strabismus (a misalignment of the visual axes), in addition to the acuity deficit, the vision in the amblyopic eye is further affected by an elevation in visual crowding, a disruption of object recognition in clutter. However, crowding is rarely highlighted as a specific therapeutic target, and the crowding deficit persists post-childhood treatment into adulthood. It is likely that the significant crowding deficits in amblyopia are going untreated due to a lack of understanding of its effects.

The aim of this thesis was to investigate visual crowding in strabismic amblyopia. In order to achieve this, insight was drawn from the literature on crowding in the typical visual system. In typical vision, crowding is minimal in central vision, and rises with increasing eccentricity (i.e. distance from the centre of gaze). It is thought that crowding in typical vision represents the compulsory pooling of visual input, promoting perceptual similarity in the peripheral visual field. By this theory, crowding is a mechanism that integrates multiple objects when the visual field is insufficiently sampled, and neurons with large receptive fields are required for adequate coverage.

Here, I examined whether crowding in amblyopia can be considered within the same framework. Increasing the understanding of the mechanism underlying crowding and its neural basis in the amblyopic visual system will aid in characterising behavioural and neural markers of crowding that could inform treatment. Importantly, by determining whether the mechanism underlying crowding in strabismic amblyopia matches that in the typical visual system, this opens the door for other clinical instances of crowding, such as crowding in posterior cortical atrophy, nystagmus, and dyslexia, to be considered within the same framework.

This thesis is also relevant for the screening of amblyopia. There is a lack of appropriate tests of crowding in the clinic, and clinical tests of crowding typically rely on the measurement of the smallest letter that can be recognised within a line of other letters. However, the literature shows that crowding in amblyopia does not depend on the size of the object to be identified, but the spacing between it and the surrounding clutter. In this thesis, the Vac-Man battery of visual tests for children was used. These video-game based tests have been developed with the aim of examining young children's visual function, and particularly measuring acuity and

crowding. For the aims of this thesis, these tests were adapted, contributing to their refinement and validation as tools for clinical use.

Acknowledgements

My hope with this section is that it will achieve a balance between being a dry list of the important people in my life during the PhD period and a soppy dedication piece. So here goes.

First of all, I would like thank my supervisor, John Greenwood, for taking me on as his first PhD student, training and teaching me, allowing me to pursue my ideas, and pushing me to constantly improve my work.

Then, I would like to thank Sam Schwarzkopf, for his invaluable help and advice on fMRI population receptive field mapping. Annegret Dalhmann-Noor, for opening the door for me to do clinical research. Alan Johnston, for his important insight into the project when it was still at its early stages, and for making it possible for me to do this PhD. Vijay Tailor, not only for his help in the clinics, but for being a great office and lab mate, and a very good friend. Nonie Finlayson, for the help with the delineations. Tessa Dekker and her lab for sharing research space with me.

I am thankful to the friends and colleagues (and their friends and children) who sat through thousands of presentations of the letter C on a grey screen. I am also especially grateful to those with amblyopia who participated in my experiments, giving me the precious gift of highly variable but also highly interesting clinical data.

When it comes to friends and family, I'd like to first thank my Dad, for being the voice of reason in moments of crisis, and supporting me. My close friends in London, Nat, Pan, Hermes, Haris, Zois, Marita, for being there for me. Josephine, for quietly taking care of me during the writing period. Daniele, for being my family, making the tiny flat our home. Athene, I cannot fathom how much more difficult this would have been had we not gone through it together – you gave me strength. My Olympia, for her support and love, especially during the torturous final months. Very special thanks to Esme, my partner in everyday life from afar, who was there for me not only when things got tough, but for every single day the past four years.

Mum, I do not have the words to say how grateful I am, for everything. You won't understand a word of it, but this is for you.

Table of Contents

List of Figures

List of Tables

1 **Chapter 1:** Introduction

1.1 General Introduction

Recognising our friend who just entered the room or our favourite coat hanging behind the door may appear easy, everyday tasks. Our visual system detects and classifies objects among hundreds of thousands of possibilities with great apparent ease and within a fraction of a second. From an evolutionary perspective, our ability to recognise objects so readily is not a surprise as our everyday activities, and therefore our survival, depend on the accurate and rapid extraction of object identity from the patterns of photons that reach our retinae. However, when looking for our friend in a crowded airport arrivals lounge or trying to find our favourite coat in a big pile of clothes, this previously efficient and rapid process can be compromised by the presence of clutter (as in Figure 1.1[1]). The disruption of object recognition by clutter is called visual crowding.

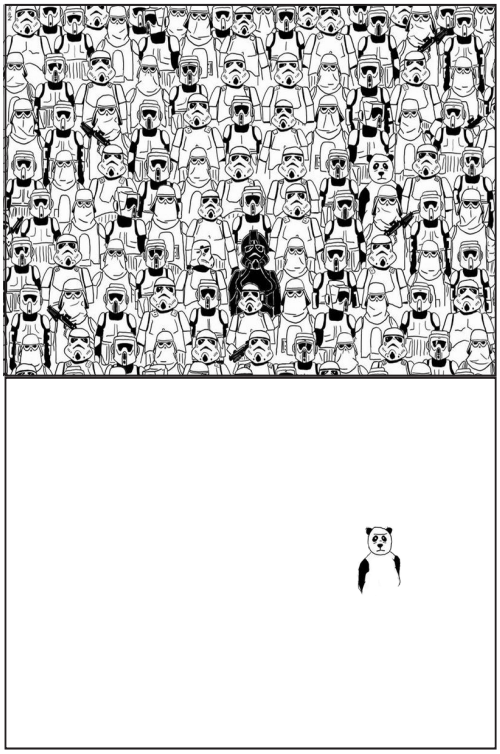

Figure 1-1 Effect of clutter

The reader can experience the disruptive effect of clutter by trying to find the panda in a group of stormtroopers (top panel). When the panda is on its own it is immediately identified (bottom panel) [1]

In the typical visual system of adults, the effect of clutter is typically minimal in central vision (the fovea) (Flom, Heath, & Takahashi, 1963; Toet & Levi, 1992; L. Liu & Arditi, 2000; Siderov, Waugh, & Bedell, 2013) but greatly affects the visual periphery (Bouma, 1970; Toet & Levi, 1992). However, we are usually oblivious to the limitations of our peripheral vision. By regularly moving our eyes to scan the scene before us, we create the illusion of a high-resolution and relatively uncluttered representation of our environment by using our fovea to look directly at the objects we are interested in. This illusion is easily dispelled when we try to describe a face or read a word in our peripheral visual field, and the limitations of our peripheral vision become apparent. Part of this difficulty to recognise or scrutinise objects in our periphery is due to the decline in visual acuity (the ability to resolve fine details)

[1] Original image retrieved from https://technabob.com/blog/2016/01/04/find-panda-stormtroopers

in peripheral vision. As Wertheim (1894) first reported in what has since become a classical demonstration, acuity drops rapidly within 5° degrees from the centre of the fovea, and continues to drop, albeit at a slower rate, out to the far periphery. One might thus think that the effect of clutter is greater in peripheral vision just because peripheral vision is like foveal vision, but with poorer acuity and thus lower resolution.

However, acuity limitations can be easily abolished in the visual periphery by enlarging an object. In contrast, even if the object is larger, when clutter is placed around it, it remains very difficult to recognise. This effect does not result from poor resolution (Lettvin, 1976), and thus a reduction in acuity, but is due to visual crowding. In fact, the reduced acuity in the visual periphery is only modest when compared with the disruptive effect of crowding (Rosenholtz, 2016). As such, crowding is considered the greatest detriment to object recognition in peripheral vision (Whitney & Levi, 2011; Rosenholtz, 2016). The reader can experience the effect of crowding in their visual periphery in Figure 1.2. By fixating on the red cross in the centre of the figure, it should be easy to recognise the identity of the letter (the *target*) in isolation (left), as it is large enough to overcome peripheral acuity limitations. However, when surrounding letters (*flankers*) are placed next to the target (right) this task becomes significantly harder – this is the effect of crowding.

Figure 1-2 Demonstration of the effect of crowding on letter recognition

The study of crowding is not only relevant for the understanding of object recognition. Crowding sets the reading speed thus limiting reading (Levi, Song, & Pelli, 2007a; Pelli et al., 2007), and impairs visual search (Vlaskamp & Hooge, 2006; de Vries, Hooge, Wiering, & Verstraten, 2011; Moores, Cassim, & Talcott, 2011). Additionally, there is an association between crowding and saccadic eye movements, suggesting a link with oculomotor signals (Nandy & Tjan, 2012; Harrison, Mattingley, & Remington, 2013; Harrison, Retell, Remington, & Mattingley, 2013; Yildirim, Meyer, & Cornelissen, 2015), For example, during eye-movement preparation, the presence of flankers at the end position of a saccade interferes with the ability to recognise an isolated target at the same location once the saccade is

completed (Harrison, Retell, et al., 2013). As such, investigating crowding can lead to a better understanding of visual function in general.

Additionally, the study of crowding is not only relevant to understanding visual function in the peripheral visual field. During development, the effect of clutter has been found to be more disruptive in the fovea of typically developing children relative to adults up until the age of 11 (Atkinson, Anker, Evans, & McIntyre, 1987; Jeon, Hamid, Mauer, & Lewis, 2010). Crucially, crowding also affects central vision in several clinical disorders. Particularly, it affects the central vision in strabismic amblyopia (Levi & Klein, 1985; Greenwood et al., 2012), as well as the central vision of individuals with dyslexia (Geiger & Lettvin, 1987; Martelli, Di Filippo, Spinelli, & Zoccolotti, 2009; Moores et al., 2011), nystagmus (Chung & Bedell, 1995; Pascal & Abadi, 1995), and posterior cortical atrophy (Crutch & Warrington, 2007, 2009). These instances of crowding are not as well researched and understood as crowding in the typical visual periphery.

In this thesis, I focus on one such clinical instance of crowding – crowding in strabismic amblyopia. Amblyopia is a developmental disorder of vision, characterised by reduced acuity in one eye despite optical correction. When amblyopia is associated with a misalignment of the visual axes (i.e. one eye may turn in, out, up or down), vision is further affected by crowding (Levi & Klein, 1985; Greenwood et al., 2012). The purpose of this thesis is to investigate visual crowding in amblyopia, and examine whether it shares common characteristics with crowding in typical vision. Hence, this chapter begins with a review of the literature on peripheral crowding which has been the primary focus of the field in the past two decades. The chapter then proceeds with a general description of amblyopia as a disorder of vision, and finally considers what is known about amblyopic crowding and what remains unclear.

1.2 Crowding in peripheral vision

1.2.1 When does crowding occur?

In this section, I consider what determines whether crowding occurs. Tasks measuring crowding typically involve an object (e.g. a letter) that the observer is asked to identify, and clutter in the form of other objects surrounding or "flanking" the object-to-be-identified. In this thesis, the object that the observer has to identify will be referred to as the *target*, and the flanking objects will be referred to as the

flankers. Here I consider the characteristics of the flankers that determine how disruptive they are on target recognition.

The two most important factors determining how disruptive nearby flankers are on the recognition of a target are the spacing between them and the target, and the target eccentricity (i.e. the distance of the target from fixation). In his highly influential report, Bouma (1970) demonstrated that for complete isolation of a target letter at a peripheral visual field location, there is a "critical spacing" between the target and flankers: for a target to be accurately identified as if there was no crowding, no other letters should be present approximately within half of the target eccentricity. For example, if a letter is positioned at 10° degrees eccentricity, flankers as far as 5° degrees away can disrupt its recognition. In a review of the critical spacing values reported in the relevant literature, Pelli, Palomares, and Majaj (2004) found that they ranged between 0.1 and 0.7 times the target eccentricity, with a median of 0.5, nicely confirming Bouma's (1970) early findings. From here forward in this thesis, the separation between the target and the flanker elements over which the flankers disrupt target identification will be referred to as the *spatial extent* of crowding.

The dependency of crowding on target-flanker separation demonstrates that by increasing the separation above the critical spacing, crowding can be relieved. A factor that needs to be considered in the relationship between critical spacing, and thus the extent of crowding, is target size. Are the large crowding extents found in the visual periphery simply a consequence of the large target sizes used to measure peripheral crowding (due to poor peripheral acuity)? If that is the case, then the extent of crowding should be proportional to the target size. Levi, Hariharan, and Klein (2002b) measured the extent of crowding at 5° and 10° eccentricity using a large range of target sizes. They showed that the large extent of crowding found in the periphery is not simply a consequence of target size, as it was not proportional to the size of the target. Small targets did not require smaller critical spacings, and thus the extent of crowding did not scale to target size, but was disproportionally large. Rather, the smallest critical spacing depended on eccentricity, being larger at 10° degrees eccentricity compared to 5°. At both 5° and 10° eccentricity, the smallest critical spacing was 10% of the target eccentricity (i.e. 0.5° at 5° and 1° at 10° eccentricity). This independence of peripheral crowding from target size has been confirmed by other studies (Tripathy & Cavanagh, 2002; Pelli et al., 2004; Pelli et al., 2007).

Importantly, the large extent and size invariance of peripheral crowding differentiates it from foveal crowding. In the fovea, the spatial extent of crowding is typically extremely small (Siderov et al., 2013). This poses a methodological challenge for researchers, who in order to obtain a measurable disruption in target identification with the addition of flankers, have required additional manipulations including reduced stimulus luminance (Bedell et al., 2013) and contrast (Siderov et al., 2013), and shortened presentation times to below 100ms (Lev, Yehezkel, & Polat, 2014). With such stimulus manipulations, estimates of the foveal extent of crowding range between 4-5 minutes of arc (Bedell et al., 2013; Siderov et al., 2013; Lev et al., 2014). Recently, Coates, Levi, Touch, and Sabesan (2018) used a custom-made setup utilising adaptive optics that allowed them to measure target-flanker interactions in the fovea below the magnitude of one minute of arc. They showed that the edge-to-edge critical spacing of foveal crowding was substantially smaller than what had been previously measured, extending between 0.75 to 1.3 minutes of arc.

Regarding the dependence of foveal crowding on target size, a number of studies have shown that the extent crowding is dependent on target size. Particularly, it has been shown that crowding scales with the size of the target (Levi, 2000; Chung, Levi, & Legge, 2001), and is proportional to target size over more than a 50-fold range of sizes (Levi, Klein, & Hariharan, 2002; Hariharan, Levi, & Klein, 2005). On the other hand, a study by Danilova and Bondarko (2007) has shown that when the size of the target is increased slightly above acuity limitations, the extent of crowding does not increase with target size, but is reduced. However, measurements of the extent of crowding with a target at the resolution limit may simply represent floor effects, and are thus not an appropriate benchmark to establish size independence. Together, findings on the extent of foveal crowding point to differences between crowding in the periphery and the fovea: the extent of peripheral crowding is large, dependent on eccentricity and independent of target size, whereas foveal crowding is tiny and depends on target size.

In the periphery, in addition to the dependence on eccentricity and target-flanker spacing, the extent of crowding also varies depending on the positioning and location of the flankers in the visual field. Toet and Levi (1992) measured the extent of crowding at the vertical, horizontal, and diagonal meridian in the lower visual field. They positioned flankers around the target horizontally, vertically, and diagonally in order to measure the shape of the crowding zone. They found that radially

positioned flankers (i.e. flankers positioned along the axis connecting the target with the fovea; see Figure 1.3A) were more disruptive on target recognition than tangentially positioned ones (i.e. flankers positioned orthogonally to the axis connecting the target with the fovea; see Figure 1.3B). This difference in the extent of crowding depending on the position of the flankers relative to fixation creates an elongated crowding zone diagonally with an elliptical shape, which has been replicated by others (Pelli et al., 2007; Petrov & Popple, 2007). This radial-tangential anisotropy has also been found in gap resolution, bisection, and saccades, suggesting that this bias is not specific to crowding but rather reflects a more general "topology of spatial vision" (Greenwood, Szinte, Sayim, & Cavanagh, 2017). Greenwood et al. (2017) have proposed that these variations in spatial vision may reflect idiosyncrasies in visual field retinotopy (e.g. in cell density or receptive field size).

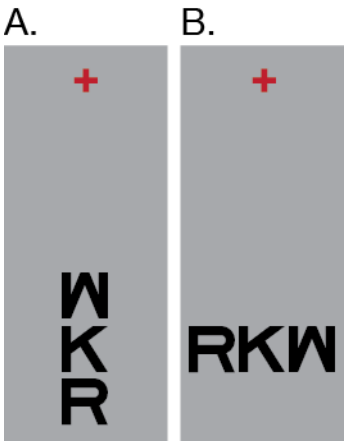

Figure 1-3 Radially and tangentially positioned flankers.

A. Flankers are positioned along the radial axis connecting the target with the fovea (red fixation cross).
B. Flankers are positioned along the tangential axis orthogonal to the radial.

In addition to this radial-tangential anisotropy of peripheral crowding, the extent of crowding depends on the position of the target in the visual field. He, Cavanagh, and Intriligator (1996) showed that the extent of crowding is greater in the upper relative to the lower visual field. More recently, it has been shown that this effect can be accounted for the shape of the visual field. The visual field is naturally asymmetric, and although there are some individual differences, typically the upper visual field is smaller (extends to ~50° eccentricity) than the lower visual field (extends to ~70°) (Niederhauser & Mojon, 2002). The asymmetrical shape of the visual field indicates that a target presented at 25° eccentricity in the upper visual field would be mid-way between fixation and the visual field edge, whereas in the lower visual field it would only be one-third of the way. Fortenbaugh, Silver, and Robertson (2015) showed that when distance from fixation was expressed in the percentage of the visual field extent (i.e. relative to the visual field boundary), this upper-lower asymmetry of crowding was eliminated.

Up to this point I have considered how crowding in the periphery differs according to the combined location of both flankers, one on each side of the target. Bouma (1970) reported that although two flankers, one placed on each side of the

target, were much more disruptive than one, their effect was not equally strong. He showed that crowding was stronger with a single flanking letter placed at an eccentric location greater than the target, than with a single flanking letter at a location closer to the fovea. This effect has since been replicated numerous times with letters (Banks, Larson, & Prinzmetal, 1979; Chastain, 1982; Bex, Dakin, & Simmers, 2003) and shown to occur with other types of stimuli, such as faces (Farzin, Rivera, & Whitney, 2009). The inner-outer asymmetry of crowding might at first appear counterintuitive, since one would expect the inner flanker to be more visible, and thus be more readily confused with the target. One explanation for this asymmetry relates to the cortical mapping of the stimulus: although in visual space the angular separation from the target for inner and outer flankers is the same, in cortical space the outer flanker is closer to the target (Motter & Simoni, 2007). It has also been shown that if attention is biased towards the flanker closest to the fovea this asymmetry can be reversed, making the inner flanker more disruptive (Petrov & Meleshkevich, 2011b). The disruptive effect of the outer flanker can also be magnified if the attentional demands of the task are increased (Petrov & Meleshkevich, 2011a). Based on these results Petrov and Meleshkevich (2011b, 2011a) speculated that the inner-outer asymmetry could also arise due to a mislocalisation of attention, with attention typically being directed towards the inner flanker. However, the outer flanker may be more disruptive due to increased positional uncertainty, that can be relieved if cues to its position are provided. Regardless, these results suggest that an explanation of the inner-outer asymmetry based solely on cortical mapping may not be sufficient.

When considering the characteristics of the flankers that determine the extent of crowding, another important question is to what extent do the target and flankers need to be similar in order for crowding to occur. Crowding has been found to be selective for differences between the target and flankers along fundamental visual dimensions including orientation (Wilkinson, Wilson, & Ellemberg, 1997), colour (Põder & Wagemans, 2007; Kennedy & Whitaker, 2010), contrast polarity (Chakravarthi & Cavanagh, 2007; Chung & Mansfield, 2009; Rosen & Pelli, 2015), and motion (Bex & Dakin, 2005). This selectivity means that crowding is strong (i.e. flankers are most disruptive and the spatial extent is large) when features in the target and flankers are similar along these dimensions, and weak (i.e. flankers are least disruptive and the spatial extent is small) when they differ. For example, a red target surrounded by green flankers is more easy to recognise than a green target

surrounded by green flankers (Kooi, Toet, Tripathy, & Levi, 1994). Selectivity for target-flanker similarity is typically symmetric: a red target surrounded by green flankers will release crowding as much as a green target surrounded by red flankers. Neurons selective for such dimensions have been found in early visual cortex. For example, neurons in the primary visual cortex (V1) respond selectively to orientation (Hubel & Wiesel, 1959; Schiller, Finlay, & Volman, 1976) and to differences in the direction and speed of motion (Gur & Snodderly, 2007). This selectivity of crowding for target-flanker similarity may arise from visual neurons in the earliest (i.e. lower) visual areas along the visual processing hierarchy, and thus can be termed selectivity for *low-level* similarity between the target and flankers.

However, there is also evidence suggesting that crowding is determined by similarities in higher-level properties of the target and flankers. Particularly, crowding has also been found to be selective for "holistic similarity" in faces (Louie, Bressler, & Whitney, 2007). Louie et al. (2007) showed that the selectivity of crowding for faces is asymmetric: with an upright target face (processed holistically) crowding was strong with upright flankers and weak with inverted flankers; with an inverted target face (processed featurally) crowding was strong with both upright and inverted flankers. In a follow-up study, instead of using photographic faces like Louie et al. (2007), Farzin et al. (2009) used Mooney faces (Mooney, 1957), which are claimed to lack individual features and thus only be perceived as faces when upright and processed holistically. They found that on an upright Mooney face, crowding was strong with upright Mooney flankers and weak with inverted. Although they did not test a case where the target was inverted, similarly to Louie et al. (2007) they concluded that their results were indicative of crowding being selective for holistic similarity in faces. The processing of holistic identity information has been shown to involve the fusiform face area (FFA) (Caldara & Seghier, 2009; J. Liu, Harris, & Kanwisher, 2010; Zhen, Fang, & Liu, 2013), a region in the higher-levels of the ventral stream of visual processing. As such, these findings have been taken as evidence that crowding can be determined by more complex "higher-level" properties of the target and flankers. As crowding has been shown to occur between face parts (i.e. the eyes, nose, mouth) (Martelli, Majaj, & Pelli, 2005), and whole faces as in Louie et al. (2007), these findings suggest that crowding occurs over multiple stages of visual processing (Chaney, Fisher, & Whitney, 2014; Manassi & Whitney, 2018).

Recently it was demonstrated that the asymmetric selectivity found in Louie et al. (2007) could be explained based on increased difficulty in recognising identity in inverted faces (Kalpadakis-Smith, Goffaux, & Greenwood, 2018). Louie et al. (2007) used an identity-matching task, asking observers to determine whether the identity of the target was the same or difference from a reference face. In identity discrimination tasks difficulty is not matched between upright and inverted conditions. For isolated (uncrowded) targets, recognising identity is more difficult in inverted than upright faces – also referred to as the face inversion effect (Tanaka & Farah, 1993; Rossion, 2008). Kalpadakis-Smith et al. (2018) found that when a task with matched task difficulty between upright and inverted target conditions was used involving the discrimination of differences in the eye positioning between faces (Goffaux & Rossion, 2007), then crowding was symmetric: crowding was strong with similarly oriented flankers and weak when the flanker orientation was reversed for *both* upright and inverted targets. They argued that in identity-matching tasks, the poor performance with uncrowded inverted target faces does not provide the dynamic range for the release from crowding to occur with dissimilar flankers, as performance is already low. These findings are contrary to crowding being selective for higher-level holistic similarity, and demonstrate that crowding in the visual periphery depends on task difficulty.

In this first section on peripheral crowding, research on the factors that modulate the spatial extent of crowding have been reviewed. Overall, this body of research has demonstrated that the in the periphery, the spatial extent of crowding depends on the separation between the target and flankers, scales with eccentricity (Bouma, 1970; Toet & Levi, 1992), and does not depend on target size (Levi, Hariharan, et al., 2002b; Tripathy & Cavanagh, 2002; Pelli et al., 2004; Pelli et al., 2007). The extent of peripheral crowding also shows a radial-tangential anisotropy (Toet & Levi, 1992), an inner-outer asymmetry (Bouma, 1970; Banks et al., 1979; Chastain, 1982; Bex et al., 2003; Petrov & Meleshkevich, 2011b, 2011a), and is greater in the upper compared to the lower visual field (He et al., 1996). Importantly, the extent of crowding varies according to the similarity between the target and flankers, extending over larger distances when the target and flankers are similar (Kooi et al., 1994; Chakravarthi & Cavanagh, 2007; Põder & Wagemans, 2007; Chung & Mansfield, 2009; Kennedy & Whitaker, 2010; Rosen & Pelli, 2015). In the next section, I consider the models that have been put forward to account for peripheral crowding.

1.2.2 Models of crowding

What mechanism could be underlying crowding and the effects described above? When an object is crowded, it does not simply disappear. This is evident from studies showing that there is little or no effect of crowding on the detection of a target object (Andriessen & Bouma, 1976; Levi, Hariharan, et al., 2002b; Pelli et al., 2004; Livne & Sagi, 2007). Rather, crowding affects the discrimination of the target. This distinguishes crowding from ordinary masking, where the detection of a target is disrupted because of the presentation of a second stimulus, the "mask" (in crowding studies, the flankers). Additionally, unlike masking, the extent of peripheral crowding is dependent on eccentricity and is size invariant (Levi, Hariharan, et al., 2002b; Tripathy & Cavanagh, 2002; Pelli et al., 2004; Pelli et al., 2007), and the strength and extent of crowding are much greater than those of masking (Andriessen & Bouma, 1976; Levi, Hariharan, et al., 2002b). Finally, masking does not show the characteristic inner-outer asymmetry of crowding, as the more eccentric mask is as disruptive as the mask positioned on the target side closer to the fovea (Petrov, Popple, & McKee, 2007).

Based on the findings that crowding affects object discrimination but not detection (Andriessen & Bouma, 1976; Levi, Hariharan, et al., 2002b; Pelli et al., 2004; Livne & Sagi, 2007), crowding is thought to affect the second stage of object recognition, feature combination. The first stage of object recognition involves the detection of object features, that are typically simple and non-overlapping (Pelli, Burns, Farell, & Moore-Page, 2006). This first stage could occur in V1, for example when a feature matches the selectivity and receptive field of a V1 neuron (i.e. the region in the visual field that neuron responds to). In the second stage of object recognition, the detected features are assumed to be combined, in order for the object to be then recognized at a later stage (Treisman & Gelade, 1980; Riesenhuber & Poggio, 2000). Crowding is assumed to be a breakdown or a disruption of this second stage of object recognition, setting a bottleneck and limit for recognising objects in the peripheral visual field (Levi, 2008; Whitney & Levi, 2011).

To better understand the underlying mechanism of crowding, the effects crowding has on the appearance of the target have been studied. These effects will be referred to as the *perceptual effects* of crowding, and have been integral in informing theories on the underlying crowding mechanism. When tasked with reporting the target, the information available to the observer about the crowded target is taken as indicative of the perceptual outcome of crowding. In this section,

first the evidence on the perceptual effects of crowding will be considered, and then the mechanisms proposed to account for these perceptual effects and other characteristics of crowding will be discussed.

To investigate the perceptual effects of crowding, Greenwood, Bex, and Dakin (2010) used a change-detection paradigm, where observers had to respond when they noticed a change in the appearance of the target. They found that when a target was a patch of visual noise surrounded by oriented Gabor flankers, observers did not notice when the target was changed to an oriented Gabor that was identical to the flankers. However, when the noise was switched for a Gabor with a different orientation to the target, this was readily detected by the observers. Crucially, rotations of the flankers whilst keeping the noise target the same induced illusory target rotations: target patches of noise, as well as blank targets, became perceptually oriented, adopting the orientation of the flankers. These findings demonstrated that crowding shifts the appearance of the target to resemble the flankers. Greenwood et al. (2010) argued that crowding is a regularization process that simplifies the appearance of clutter in the peripheral visual field to promote consistency in the appearance of adjacent objects.

In line with evidence showing that crowding shifts the target appearance to make it more like the flankers are studies showing that when tasked to report the identity of the target, observers report average identities between the target and flankers. In an influential study, Parkes, Lund, Angelluci, Solomon, and Morgan (2001) found that crowding did not prevent observers from being able to report the average orientation of a cluster of oriented elements, while simultaneously being unable to report the orientation of a single target patch. Greenwood, Bex, and Dakin (2009) used a cross-like target, in which the position of the horizontal line was varied, and asked observers to report whether the horizontal line was positioned above or below the stimulus midpoint. They found that when the target was crowded with flanker crosses, the perceived position of the horizontal line in the target appeared shifted towards the direction of the horizontal line in the flankers. They showed that while observers also made random reports of the horizontal line position (i.e. reported positions that did not correspond to either the target or flanker line position), the observers' responses were indicative of reports of the average position of the horizontal line, consistent with the findings by Parkes et al. (2001).

Other studies have shown that the errors observers make when reporting the identity of a crowded target are not just indicative of an intermediate or partial shift

towards the identity of the flankers, but are reports of the flanker identity. Strasburger, Harvey, and Rentschler (1991) found that the majority of incorrect responses observers made when tasked to identify a crowded number digit consisted of a flanker response. Similarly, Ester, Zilber, and Serences (2015) showed that when observers were asked to adjust the orientation of a clock-like reference stimulus to match the perceived orientation of a crowded target in their visual periphery, they either correctly reported the target orientation or reported the orientation of the flanker. Substituting the flanker identity for the target has been reported in numerous of studies in the crowding literature (Krumhansl & Thomas, 1977; Chastain, 1982; Huckauf & Heller, 2002; Strasburger, 2005; Põder & Wagemans, 2007; Vul, Hanus, & Kanwisher, 2009).

Taken together, these results indicate that observers make two types of errors when tasked to report the identity of the target: averaging or assimilation errors, where they report intermediate identities between the target and flankers, and substitution errors, where they substitute the target with the flanker and report the flanker identity. In order to account for these effects, a model of the underlying crowding mechanism must receive the target and flanker stimuli and produce the corresponding type of observer response. Here, the following accounts of crowding will be discussed: pooling, substitution, attentional accounts of crowding, population response models, and texture models.

Traditional pooling accounts portrayed crowding as a process of compulsory integration of target and flanker features (Parkes et al., 2001). Based on the premise that crowding represents a breakdown of the second stage of object recognition, feature combination, Pelli et al. (2004) introduced the idea of an “integration field”. They proposed that the visual system may have many integration fields of various sizes distributed across the visual field and overlapping one-another. According to this theory, when available the visual system uses an integration field of matching size and location as the target object to be identified. This is what normally happens in the fovea, where crowding, and thus feature integration, is minimal (Flom, Weymouth, & Kahneman, 1963; Toet & Levi, 1992; L. Liu & Arditi, 2000; Siderov et al., 2013) or absent (Strasburger et al., 1991; Levi, Klein, et al., 2002; Pelli et al., 2004). In the periphery, where the extent of crowding is large (Bouma, 1970; Toet & Levi, 1992), there is a lack of small integration fields. The visual system thus has to do with the integration fields it has available, which can be inappropriately large. Within those large peripheral integration fields, both the target and the flanking

objects are integrated, and thus crowding occurs. The concept of an integration field is consistent with a receptive field that scales with eccentricity (Zeki, 1978; A. T. Smith, Singh, Williams, & Greenlee, 2001; Rosa & Tweedale, 2005; Dumoulin & Wandell, 2008), and integrates the features that are detected by neurons at an early stage of visual processing, such as V1, into an object (Levi, 2008).

Simple pooling accounts have proposed that the computation of this pooling mechanism is an averaging of the target and flanker features (Parkes et al., 2001; Greenwood et al., 2009; Dakin, Cass, Greenwood, & Bex, 2010). Greenwood et al. (2009) put forward a weighted averaging account of crowding, and incorporated different weights for the target and flankers at the stage of combination. These weights determined the exact contribution of the target and the flankers in the averaging process. The advantage of using weights is that they can be incorporated to account for the dependency of crowding on the separation between the target and flankers (Bouma, 1970), and the selectivity of crowding for target-flanker similarity (e.g. Kooi et al., 1994). Flankers that are more closely positioned to the target and/or are more similar to it can be assigned a high weight resulting in averaging of the target and flanker identities. In contrast, flankers that are more widely separated from the target and/or dissimilar to it can be assigned a lower weight, resulting primarily in target responses. A prediction of the averaging accounts of crowding is that as more flankers are placed within the integration field the amount of irrelevant activity is increased and target signal is diluted, thus making it harder to identify the identity of target and increasing the strength of crowding. Although some studies are consistent with this prediction (Strasburger et al., 1991; Parkes et al., 2001; Põder & Wagemans, 2007), others have not found an effect of increased number of flankers on crowding strength (Pelli et al., 2004).

On the other hand, simple substitution accounts argue that crowding arises from the features of the target being substituted onto the target. The basic difference between substitution accounts and pooling approaches is that simple substitution accounts do not assume any integration of the target and flanker elements. According to these accounts, observers simply confuse the locations of the target and flankers, resulting in an incorrect flanker report based on a single item, from the wrong place (Wolford, 1975; Estes, Allmeyer, & Reder, 1976; Krumhansl & Thomas, 1977; Chastain, 1982; Ester et al., 2015). Substitution accounts propose that observers have access to both the target and flanker identities, but due to the increased positional uncertainty associated with peripheral

vision (Westheimer, 1975; Levi, Klein, & Yap, 1987; Levi & Klein, 1990), they are unable to bind these identities to their corresponding positions in the visual field. However, as simple substitution is based on a single object from the wrong place, it would fail to account for the selectivity of crowding for target-flanker similarity (e.g. Kooi et al., 1994), as well as findings showing that a flanker is more likely to be substituted if it is similar to the target (J. Freeman, Chakravarthi, & Pelli, 2012). More complex substitution models account for influences of target-flanker similarity by incorporating weights, with a substitution error being more likely to occur if the flanker is similar to the target (Põder & Wagemans, 2007).

The attentional account of crowding by (Strasburger et al., 1991; Strasburger, 2005) proposes a concept similar to the attentional spotlight (Eriksen & Yeh, 1985; Posner & Petersen, 1990), a “zoom lens” of a limited size that aids the processing of stimuli within its beam. Although in the fovea this lens is very precise and able to select the right object (the target) resulting in minimal crowding, in the periphery it is imprecise, often zooming in and selecting the wrong object (a flanker instead of the target). Strasburger et al. (1991) argue that errors reflect the inability to precisely focus spatial attention at peripheral visual field locations in the face of an accompanying loss of positional information (Westheimer, 1975; Levi et al., 1987; Levi & Klein, 1990). As such, this account is similar to simple substitution accounts in that it does not assume any integration or pooling between the features of the target and flankers, and predicts only substitution errors due to source confusion because of the positional uncertainty that characterises the periphery (Westheimer, 1975; Levi et al., 1987; Levi & Klein, 1990).

A different attentional account of crowding has been put forward by He et al. (1996), Intriligator and Cavanagh (2001), and Tripathy and Cavanagh (2002). Similarly to (Strasburger et al., 1991; Strasburger, 2005), which instead argues that crowding occurs due to the limitations of attentional resolution. That is, when multiple objects are close to each other, instead of focusing on the wrong item, attention fails to individuate the target from the flankers. This failure consequently may result in the binding of the target and the flankers, and observer being unable to select and identify the target. Although it is possible that the breakdown in individuating the target from the flankers could potentially result in assimilation errors, this attentional account does not specify the perceptual effects of crowding or make predictions about the resulting errors. However, within the framework of this account, the selectivity of crowding to target-flanker similarity (e.g. Kooi et al.,

1994), may be driven by attentional tuning to the properties of the target, that prevents interference and facilitates target recognition when the flankers are dissimilar to the target (Kravitz & Behrmann, 2011).

More recently, it has been argued that averaging and substitution are merely descriptors of the errors observers make under crowded conditions, not distinct mechanisms underlying crowding. Population response pooling models, a variant of the pooling approach, have provided a very successful framework for crowding (van den Berg, Roerdink, & Cornelissen, 2010; Harrison & Bex, 2015) in that they account for both averaging/assimilation and substitution errors. Population response pooling models are based on the principles of population coding (Pouget, Dayan, & Zemel, 2000), an approach which mathematically applies the idea that information encoding in the brain occurs in populations of neurons rather than single neurons. Similar to the selectivity of V1 neurons for orientation (Hubel & Wiesel, 1959; Schiller et al., 1976; Celebrini, Thorpe, Trotter, & Imbert, 1993; Mazer, Vinje, McDermott, Schiller, & Gallant, 2002; Ringach, Shapley, & Hawken, 2002; Gur, Kagan, & Snodderly, 2005), each neuron-like detector within a population is assumed to have a peak sensitivity for particular orientation and some lesser sensitivity to nearby orientations. Crowding is simulated as the combined ("pooled") population response to the target and flankers. The perceptual effects of crowding could thus result from the pooling of the responses of populations of neurons in visual areas whose size, like the spatial extent of crowding, increases with eccentricity (Dow, Snyder, Vautin, & Bauer, 1981; Van Essen, Newsome, & Maunsell, 1984).

Harrison and Bex (2015) put forward a population response model to account for crowded orientation signals. They showed that an idealised population code can successfully produce both assimilation and substitution errors in a probabilistic fashion. They thus argued that these types of errors do not represent an underlying mechanism of either averaging or substitution, but rather the perceptual reports that are drawn from the population code to the target and flanker stimuli. By varying the weight of the flankers and thus their contribution to the combined population response to target and the flankers, Harrison and Bex (2015) simulated the differential effect of the flankers depending on target-flanker separation. As with weighted averaging models (Greenwood et al., 2009), weights can be incorporated into these models to also simulate the selectivity of crowding for target-flanker similarity (e.g. Kooi et al., 1994). Weights could also be applied differentially to the response to the flankers depending on their position with regards

to the target to simulate the radial-tangential anisotropy (Toet & Levi, 1992; Pelli et al., 2007; Petrov & Popple, 2007; Greenwood et al., 2017) and the inner-outer asymmetry (Banks et al., 1979; Chastain, 1982; Bex et al., 2003; Farzin et al., 2009).

A different approach to investigating the perceptual effects of crowding from examining whether observers' reports indicated averaging or substitution errors was followed by Balas, Nakano, and Rosenholtz (2009). They grounded their study on the notion that under crowded conditions, the region over which crowding occurs may perceptually appear like texture. Using a texture analysis and synthesis routine (Portilla & Simoncelli, 2000), they created mongrels: new synthesized patches of texture that appear to have the same texture as the sample patches. These mongrels evoke the jumbled percept of a mixture of target and flanker features typically produced by crowding. To test their account empirically, they had a first group of observers perform a letter recognition task with crowded arrays in the visual periphery, and a second group of observers doing the same task while viewing mongrels of the original crowded arrays foveally. They found comparable recognition accuracies between the two groups, and reasoned that since the performance of the latter group could predict the performance of the former, then the information encoded by the summary statistics in the mongrels foveally must be the same as the information that is available in the visual periphery.

Based on these findings, a different approach to pooling was put forward by Balas et al. (2009), viewing crowding as a mechanism that statistically summarises the target and flankers. They argued that the visual system computes summary statistics of the visual input over some local pooling field or region that grows with eccentricity. As such, instead of taking as the unit of the crowding mechanism the target and flanker features (e.g. orientation) like pooling models (Greenwood et al., 2009; Dakin et al., 2010; van den Berg et al., 2010; Harrison & Bex, 2015), or the entire target and flanker objects like substitution models (Strasburger et al., 1991; Strasburger, 2005; Ester et al., 2015), they take the image statistics. Balas et al. (2009) argue that crowding can be seen as a strategy the visual system employs to deal with a bottleneck in visual processing by representing the visual input in summary statistics in order to reduce the information passing through the bottleneck, whilst still encoding a large amount of information about the visual scene. This account suggests that the observer might be treating the crowded signal as texture, with statistical descriptors such as the average orientation signal being available for report. As such, summary statistic models of crowding do not

contradict population response pooling models, but are rather based on different descriptions of the perceptual effects of crowding.

In summary, this section has reviewed evidence on the perceptual effects of crowding and the models attempting to account for these effects have been considered. Together, the evidence suggests that when reporting the identity of an object in clutter, observers make systematic assimilation and substitution errors. Although averaging (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010), and substitution-based accounts (Estes et al., 1976; Krumhansl & Thomas, 1977; Chastain, 1982; Strasburger et al., 1991; Strasburger, 2005; Ester et al., 2015) primarily account for only one type of error, population response pooling models (van den Berg et al., 2010; Harrison & Bex, 2015) can account for both assimilation and substitution errors and with the use of weights, incorporate spatial characteristics of crowding described in section 1.2.1. Texture-based models (Balas et al., 2009) could also potentially account for both types of errors, depending on the resulting summary statistics available to the observer. Therefore, population response pooling models provide an attractive framework for crowding. In the following section, I consider findings that cannot be accounted for by the models of crowding described above that suggest that crowding can be modulated by grouping.

1.2.3 Grouping effects on crowding

In section 1.2.1, spatial aspects of crowding considered. A hallmark of crowding, Bouma's (1970) critical spacing, states that flankers disrupt performance only when placed within a restricted region around the target that is equal to half the target eccentricity. However, studies measuring the critical spacing of crowding have typically used one flanker on each side of the target (Toet & Levi, 1992). Curiously, when more flankers are placed next to the local flankers neighbouring the target, forming a multi-element flanker array, Bouma's (1970) critical spacing no longer holds.

In an early study measuring crowding in the periphery and varying the number of flankers, Banks et al. (1979) showed that when a target letter was flanked by one flanker letter, target recognition deteriorated relative to when the target letter was presented in isolation. However, when the flankers were increased to five, the target was more readily recognised relative to when only one flanker was presented, relieving the effects of crowding. More recently, Manassi, Sayim, and Herzog (2012)

presented two vertically positioned lines with a small offset (i.e. a vernier) and asked observers to discriminate the direction of the offset (left or right). The vernier target was flanked by 2,4,8 or 16 flankers of either equal length to the target, shorter length, or greater length. Adding more identical flankers of same length to the target kept their disruptive effect on target recognition constant. However, with dissimilar flankers (i.e. of shorter or longer length to the target), the target offset was more easily identified, and thus crowding was relieved. Importantly, the more dissimilar flankers were added, the easier it was to identify the vernier offset. This relief from crowding with additional flankers dissimilar to the target has also been found in the fovea (Malania, Herzog, & Westheimer, 2007; Sayim, Westheimer, & Herzog, 2008). Taken together, these results demonstrate that flankers outside Bouma's (1970) critical spacing can modulate crowding. These findings are thus difficult to account for with models that assume a restricted region over which crowding occurs, either due to an integration field (Pelli et al., 2004) or a pooling zone (van den Berg et al., 2010; Harrison & Bex, 2015), or limited attentional resolution (Strasburger et al., 1991; He et al., 1996; Intriligator & Cavanagh, 2001; Tripathy & Cavanagh, 2002; Strasburger, 2005).

Unlike the local flankers neighbouring the target, the disruptive effect of multi-element flanker arrays is not determined by low-level selectivity for target-flanker similarity along fundamental visual dimensions. Manassi et al. (2012) presented a red target vernier in the visual periphery, and asked observers to identify the side of the offset. When the vernier was surrounded by two red lines, one each side (Figure 1.4A), performance deteriorated relative to when the vernier

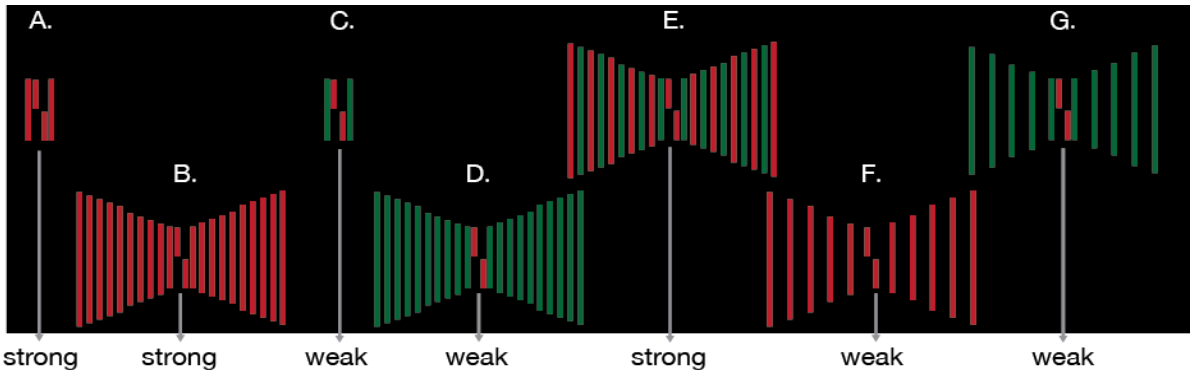

Figure 1-4 Illustration of stimuli and findings by Manassi, Sayim, & Herzog (2012), JoV

The target was a red vernier (two horizontal lines with a small offset), and the observers' task was to discriminate the target offset (left or right). Each flanker condition is assigned a letter (A-G) for ease of refence in text. Under each flanker condition, it is indicated whether crowding was "strong" (i.e. significantly elevated offset discrimination thresholds relative to when the target was presented in isolation) or weak (i.e. did not have a marked effect on raising thresholds relative to when the target was presented in isolation).

was presented in isolation, and crowding was strong. Similarly, when red ten flankers were on each side of the target (Figure 1.4B) that extended outside Bouma's (1970) critical spacing, vernier offset discrimination deteriorated relative to isolated targets and crowding was strong. When two flanker lines were green (Figure 1.4C), there was very little deterioration in performance, and crowding was weak. This was also the case with ten green flankers on each side of the target (Figure 1.4D). Up to this point, these results could be explained based on crowding being selective for similarity in colour between the target and the local flankers neighbouring it, as these remained constant in cases of strong crowding (Fig. 1.4, A & B) and weak crowding (Fig. 1.4, C & D). However, when the flankers were alternating red and green (Figure 1.4E), crowding was strong. In fact, threshold elevation matched the condition with flankers identical to the target. This cannot be explained by the similarity in colour between the target and the neighbouring flanker, as they were of different colour. It also cannot be explained by the inclusion of similar red flankers in the multi-element flanker array, as removing the green flankers lowered thresholds (Figure 1.4F), releasing the disruptive effect of the flankers. Note that this was also the case when the similar red flanker flankers were removed (Figure 1.4G). The disruptive effect of multi-element alternating flankers has also been found with polarity (Rosen & Pelli, 2015), and is not specific to the visual periphery, as it has also been found in the fovea (Sayim et al., 2008).

If not target-flanker separation and selectivity for target-flanker similarity along fundamental visual dimensions, what then determines these curious instances of crowding with multi-element flanker arrays in both foveal and peripheral vision? It has been proposed that the effects of these additional flankers on crowding depend the global stimulus configuration, and particularly grouping (Herzog & Manassi, 2015; Herzog, Sayim, Chicherov, & Manassi, 2015; Francis, Manassi, & Herzog, 2017). The notion that the overall configuration of a number of stimuli has effects on perception has been described within the principles of Gestalt (Wertheimer, 1923; Koffka, 1935). Wertheimer (1923) argued that given any number of stimuli, the visual system tends to group these stimuli and perceive greater wholes instead of independent, individual objects or parts. The way in which the visual system groups objects together was described by Wertheimer (1923) as several laws or principles. For example, the good Gestalt principle states that elements are likely to be grouped together if they are parts of a pattern that creates an orderly, regular, and balanced configuration.

These Gestalt principles of grouping (Wertheimer, 1923) have been applied to explain the effects of multi-element flanker arrays on crowding (Herzog & Manassi, 2015; Herzog et al., 2015; Francis et al., 2017). When flankers are added that "fit" *only* with the previous ones neighbouring the target but not with the target, they are segregated from the target, forming a coherent configuration or group with each other but not with the target. In this case, *flanker-flanker grouping* occurs, and crowding is weak. For example, when additional shorter or longer flankers from the target were added in Manassi et al. (2012), the disruptive effect of the flankers became increasingly weaker – this is taken to indicate flanker-flanker grouping. On the other hand, when additional flankers are added that match *both* the local flankers neighbouring the target *and* the target, a coherent configuration is formed complimenting the percept. In this case, *target-flanker grouping* occurs, and crowding is strong. For example, when flankers form a regular alternating pattern of colour or polarity with the target (Sayim et al., 2008; Manassi et al., 2012; Rosen & Pelli, 2015), they also form a coherent configuration following the principle of good Gestalt, and there is target-flanker grouping.

Grouping accounts of crowding (Herzog & Manassi, 2015; Herzog et al., 2015; Francis et al., 2017) argue that whether or not crowding occurs is determined at a higher-level stage of visual processing, when information about the objects is combined across the entire visual field to obtain the perceptual organization of the whole stimulus. Grouping therefore precedes crowding (Manassi, Sayim, & Herzog, 2013), and the effects of flankers on the target are determined by how objects look and whether or not they group (Herzog et al., 2015). Effects that have been attributed to selectivity for low-level target-flanker similarity using two local flankers can be also accounted for by the grouping principle of similarity (Wertheimer, 1923): similar flankers to the target are more disruptive on its recognition because there is target-flanker grouping based on similarity, whereas dissimilar flankers are less disruptive because they do not form a coherent configuration with the target, and no such target-flanker grouping occurs. As such, it has been argued that grouping determines the strength of crowding (i.e. how disruptive the flankers are) regardless of the number of flankers and whether or not they extend beyond Bouma's (1970) critical spacing.

Francis et al. (2017) proposed a model of grouping aiming to explain how grouping may operate on crowding. At the first stage of the model, they assume that a pooling process occurs that creates oriented boundaries for the visual input

consisting of the target and flankers. However, contrary to pooling models of crowding (van den Berg et al., 2010; Harrison & Bex, 2015), this pooling stage does not determine the effect flankers have on target recognition. At a second stage, a segmentation process occurs on the oriented boundaries that utilises the boundary groupings in order to change the representation of the visual information and thus modulate crowding. At the final decision-making stage, it is assumed that the process by which observers respond to the target is template matching. For example, if the target is a crowded vernier, observers have one template for a vernier shifted to the left and a different template for a vernier shifted to the right. When crowding occurs, boundaries from the flanking objects contribute to the template calculation, thus making it harder for the observer to match a template to the visual input. Francis et al. (2017) find that this model successfully predicts many flanker configuration effects, such as the reduction of crowding when the flankers are shorter or longer and thus ungroup from the vernier (Malania et al., 2007; Manassi et al., 2012), and the reduction of crowding by the addition of more flankers (Manassi et al., 2012).

However, the grouping model by Francis et al. (2017) has only been tested with experimental data in which the target is a vernier or a Gabor grating, and the flankers consist either vertical and/or horizontal lines or Gabors that create very clear boundaries. It thus remains unclear how it would perform with more complex stimuli such as letters, which rely less heavily on positional information and create more complex boundaries. The relief from crowding obtained when flankers that match the flankers but not target are added may derive from a better representation of the target, due to the boundaries providing a cue to the target location (Rosenholtz, Yu, & Keshvari, 2017; Yu & Rosenholtz, 2018). Such cues would alleviate positional uncertainty with regards to the vernier offset, especially in the visual periphery where such uncertainty is high (Levi et al., 1987{Westheimer, 1975 #261)}. Studies minimising positional uncertainty are thus essential to lending support that a higher-level grouping stage is necessary to account for the findings with multi-element flanker arrays. A relief from crowding using stimuli other than vernier targets in the fovea, where positional accuracy is high (Westheimer, 1975, 1981), would lend support to these effects occurring due to flanker-flanker grouping and not positional cuing.

In this section, studies were reviewed using multi-element flankers showing that flankers extending beyond Bouma's (1970) critical spacing can modulate

crowding, while also not adhering to low-level selectivity for target-flanker similarity. These findings have been taken as evidence of grouping modulating crowding in both in the fovea and the visual periphery (Herzog & Manassi, 2015; Herzog et al., 2015; Francis et al., 2017), suggesting higher-level influences on crowding. As discussed in section 1.2.2, studies on crowding and face recognition (Louie et al., 2007; Farzin et al., 2009) have also suggested that crowding could involve higher stages of visual processing. In the next section, I discuss findings on the neural locus of crowding, and whether they are in support of crowding occurring over one or multiple stages of visual processing.

1.2.4 Neural basis of crowding

An important question regarding crowding is the underlying neural locus. A number of studies have shown that peripheral crowding occurs when the target and flankers are presented to different eyes (Flom, Weymouth, et al., 1963; Westheimer & Hauske, 1975; Levi, Klein, & Aitsebaomo, 1985; Kooi et al., 1994). This finding has the important implication that crowding must occur at or beyond the point of binocular fusion, the site where information from the two eyes is combined, which is considered to be V1 (A. Smith, 2015). As such, it is fair to assume that crowding occurs in the cortex, not in the eyes.

Psychophysical studies using adaptation have implicated both V1 and later visual areas in crowding. He et al. (1996) provided the first direct evidence that crowding takes place beyond V1. They found that even though flankers disrupted the identification of a crowded target as observers did not perform above chance for discriminating the target orientation, strong adaptation after-effects were still observed. As adaptation is thought to occur in V1 (Movshon & Lennie, 1979), this provided evidence that crowding occurs beyond V1. However, these findings were questioned more recently by a study by Blake, Tadin, Sobel, Raissian, and Chong (2006). Blake et al. (2006) showed that the threshold-elevation aftereffect was significantly reduced during crowding. They argued that the strong after-effect found by He et al. (1996) could be accounted for by response saturation because the adapters they used were at a very high contrast level. Blake et al. (2006) thus argued that their findings indicate that the neural events that underlie crowding rest at an early stage of visual processing, because the threshold elevation after-effect arises at least partially from adaptation in V1. As such, based on adaptation studies, V1 cannot be excluded as a neural site involved in crowding.

Studies using texture synthesis models trying to simulate visual scene perception have also provided mixed results with regards to involvement of V1 in the perceptual effects of crowding. J. Freeman and Simoncelli (2011) used the same texture synthesis model (Portilla & Simoncelli, 2000) used by Balas et al. (2009) in their summary statistics model of crowding. They developed a model that first decomposed an image based on a population of oriented V1-like receptive fields, and then computed local averages within the image over pooling regions that scaled in size with eccentricity. They assumed that if this model captured the appearance of the visual scene, then two images that are identical to the model, should also appear identical to an observer – what they called "metamers". J. Freeman and Simoncelli (2011) found that when images were statistically matched within small pooling regions, observers' performance was at chance, and thus the images were indistinguishable. When however the pooling regions were larger, the performance was at ceiling, and the images were clearly perceived as different. They showed that the pooling regions over which images were indistinguishable to observers were consistent with the eccentricity-dependence of crowding and Bouma's (1970) critical spacing of 0.5 the target eccentricity, and matched the receptive field size of V2 neurons in macaques. These results were interpreted as providing a link between receptive field scaling in V2, crowding, and the rich experience of visual scene perception (Movshon & Simoncelli, 2014; Seth, 2014; Cohen, Dennett, & Kanwisher, 2016). They are also consistent with pooling accounts of crowding, arguing that the perceptual experience of crowding results from the pooling of target and flanker features within receptive fields that increase with eccentricity.

However, there was an important limitation in this study. Observers in J. Freeman and Simoncelli (2011) were not shown original images to compare them to the ones synthesised by the model. Rather, they were shown two synthesised images generated by the model and were required to compare them with each other. Using natural images and model-generated images, Wallis, Bethge, and Wichmann (2016) and Wallis et al. (2018) found that observers could very easily discriminate between them when the pooling regions used in J. Freeman and Simoncelli (2011) were applied. Therefore, Wallis et al. (2018) argued that lower scaling factors (i.e. pooling zones) than even V1 receptive fields may be required to generate model-synthesised images that are perceptually indistinguishable from the original real natural image. This suggests that scene appearance, and the perceptual

experience of crowding, must involve significantly smaller pooling regions than the size of receptive fields in V2.

However, V1 neurons do not follow Bouma's (1970) critical spacing of 0.5 the target eccentricity, as they scale with eccentricity at a much lower rate. Motter (2002, 2018) showed that in primates, receptive field size in V1 scales about 0.1 times the eccentricity, and that cortical receptive fields that show capabilities to integrate across a range matching Bouma's (1970) critical spacing are first found in primate V4. Motter (2006) also showed that the responses of V4 neurons are consistent with the effects of temporal crowding, as measuring in a serial presentation sequence. V4 receptive fields are also suitable for the crowding locus because they show an anisotropy in the representation of the visual field (Pinon, Gattass, & Sousa, 1998), similarly to the radial-tangential anisotropy of crowding resulting in the region over which crowding being elliptical (Toet & Levi, 1992; Pelli et al., 2007; Petrov & Popple, 2007; Greenwood et al., 2017). Additionally, consistent with crowding occurring for multiple stimulus types, from Gabors (Hariharan et al., 2005; Greenwood et al., 2010) to letters (Bouma, 1970; Greenwood et al., 2009), and affecting multiple stimulus dimensions (Kooi et al., 1994; Kennedy & Whitaker, 2010), physiological studies suggest that V4 combines signals from different stimuli (Logothetis & Charles, 1990; Ferrera, Nealy, & Maunsell, 1994). However, V4 lesions have little effect of crowding, challenging the idea of V4 as the neural locus of crowding (Merigan, 2000).

There have also been a few neuroimaging studies attempting to shed light on the cortical locus of crowding that in turn point to the involvement of not one, but multiple visual areas in crowding. Functional magnetic resonance imaging (fMRI) is based on the increase in blood flow to the local vasculature that accompanies neural activity (Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001), and measures changes in Blood Oxygen Level Dependent (BOLD) signal. Millin, Arman, Chung, and Tjan (2014) compared BOLD response in the visual cortex to crowded and uncrowded stimuli. They found that crowding was associated with a decrease in BOLD signal in V1 to V4. Interestingly, they found that this suppression was correlated with crowding strength. The strongest suppression in BOLD was induced when the target-flanker separation was the smallest, and thus the effects of crowding the strongest. The crowding-related suppression in BOLD was found regardless of whether attention was directed to the stimuli or away from the stimuli

due to the observers also having to perform a task at fixation, suggesting that attention did not modulate these effects.

However, other studies have failed to find any effect of crowding on BOLD signal as early as V1 when attention was directed away from the stimuli (Fang & He, 2008; T. Bi, Cai, Zhou, & Fang, 2009; J. Freeman, Donner, & Heeger, 2011a). T. Bi et al. (2009) utilised the adaptation paradigm used by Blake et al. (2006) in the MRI scanner. They measured the threshold elevation after-effect both when an oriented Gabor target was presented in isolation (uncrowded) and with flankers (crowded). When the orientation of the test grating changed from parallel to orthogonal to the target (the adapter), if this was accompanied by a large positive amplitude difference in BOLD, it was taken as an indication of adaptation. They found that orientation-selective adaptation in V1 was not influenced by crowding. However, in V2 and V3, they showed that crowding weakened the adaptation effect. These results involve V2 and V3, but not V1 in the cortical locus of crowding, and point to a role of attention in crowding, in line with attentional accounts (Strasburger et al., 1991; He et al., 1996; Intriligator & Cavanagh, 2001; Tripathy & Cavanagh, 2002; Strasburger, 2005).

J. Freeman, Donner, and Heeger (2011b) used fMRI in a different way than described above. Instead of focusing on the reduction in BOLD activation due to crowding, they examined the effects of crowding on correlations between responses in different visual areas. They had observers view target letters at different locations in the visual periphery. Targets were either presented in isolation in the uncrowded condition or surrounded by flankers in the crowding condition. Their aim was to examine whether crowding (i.e. the addition of flankers) affected the dynamic interactions between visual areas, as indicated by the correlation between their fMRI time series. They found that correlations between the responses in early visual areas and visual word form area (VWFA) were lower when the target was crowded compared to when it was uncrowded. These differences in correlation between uncrowded and crowded targets were retinotopically specific to the peripheral targets, and thus were not caused by other confounding factors such as arousal. These differences were found even when attention was diverted away from the target letter, suggesting that they were not caused by any modulations in attention. J. Freeman et al. (2011b) argued that if feature integration involves a cascade of transformations along the ventral visual pathway (Riesenhuber & Poggio, 1999), then crowding may disrupt these transformations from one visual processing

stage to the next, leading to disruptions in the correlations in activity from one visual area to the next. According to their argument, the transformations during feature integration from early to higher visual areas would become unstable due to crowding, resulting in lower correlations between the visual areas involved in this process.

The discrepancy between findings from neuroimaging studies with regards to the involvement of V1 could be explained by differences in the techniques used to measure crowding. A common principle in these studies is that in order to identify the cortical locus of crowding, they have induced a physical change to the stimulus, such as the introduction of flankers, making it hard to tease apart what differences in BOLD activation relate to the change in stimulus and what differences relate to the crowding percept. To overcome this issue, Anderson, Dakin, Schwarzkopf, Rees, and Greenwood (2012) had observers adapt to a peripheral patch of noise surrounded by oriented Gabor flankers, and used the same change-detection paradigm used in Greenwood et al. (2010) to study crowding-induced changes in target appearance on BOLD activation. Note that the behavioural study showed that when a noise target was substituted with a Gabor identical to the flanker orientation (change-same), this change went undetected, whereas when the noise was substituted with an orthogonally-oriented Gabor (change-different), the change was easily noticed (Greenwood et al., 2010). They predicted that the visual areas that represent the physical properties of the stimulus would show repetition suppression in trials in which the target remained the same (i.e. remain a noise target), but would be released from adaptation when the target was changed to the orientation of the flankers, or to an orthogonal orientation to the flankers. In contrast, areas that represent the crowded percept should show repetition suppression not only in trials in which there is no change to the target, but also in trials in which the target is switched to match the orientation of the flankers (change-same), as this change is not perceived. In contrast, they should show a release from adaptation only when the target was changed to an orientation orthogonal to the flankers (change-different). This pattern of brain activation was observed throughout V1 to V4, with strength increasing from early to late areas, suggesting that crowding is a multi-stage process.

Collective work on the neural basis of peripheral crowding clearly shows that one single crowding stage or cortical locus is unlikely, as the findings discussed above implicate areas V1 to V4. This is consistent with crowding affecting many

stimulus dimensions, from orientation to colour and motion (Wilkinson et al., 1997; Bex & Dakin, 2005; Põder & Wagemans, 2007; Kennedy & Whitaker, 2010), and being influenced by higher-level grouping processes (Manassi et al., 2012, 2013; Herzog & Manassi, 2015; Herzog et al., 2015). If the perceptual effects of crowding are a consequence of pooling within large receptive fields, as pooling accounts argue (Parkes et al., 2001), then the increase in the size of receptive fields at corresponding eccentricities from V1 to V4 could underlie the progressive increase in the crowding-related pattern of activation observed in Anderson et al. (2012). Advances in fMRI have allowed for the measurement of the aggregate receptive field of populations of neurons (pRF) within an fMRI voxel (Dumoulin & Wandell, 2008; Amano, Wandell, & Dumoulin, 2009; Harvey & Dumoulin, 2011) in humans, and can thus shed light to whether there is a relationship between receptive field size and crowding.

1.3 Amblyopia

Amblyopia, often referred to as “lazy eye”, is a developmental disorder that affects the spatial vision of one (unilateral) or both eyes (bilateral) in the absence of an obvious organic defect. Clinically, the generally accepted definition of amblyopia is reduced visual acuity despite optical correction when measured with an optotype chart, such as the LogMAR chart. These charts display several rows of optotypes, which are standardised symbols for measuring acuity in the clinic. When tested monocularly, if a child performs two or more rows worse with one eye than the other (i.e. has a two-line difference in acuity), this is typically taken as evidence of amblyopia. Amblyopia is the most common cause of unilateral visual deficit during development (Webber & Wood, 2005; Gunton, 2013), with a prevalence in children estimated between 0.2 and 5.4% (Preslan & Novak, 1996, 1998; Lim et al., 2004; Grönlund, Andersson, Aring, Hård, & Hellström, 2006; Matsuo, Matsuo, Matsuoka, & Kio, 2007; Robaei et al., 2008). It also persists in adulthood, remaining one of the most common causes of unilateral visual impairment with a prevalence between 0.35 and 3.6%, including in populations in which advanced medical care is available. Bilateral amblyopia is much less common, affecting just 0.1-0.45% of the population (Robaei et al., 2005; McKean-Cowdin et al., 2013).

Amblyopia is considered to derive from the degradation of the retinal image associated with abnormal visual experience during infancy and early childhood. As such, children with conditions that disrupt equal binocular vision are at risk of

developing amblyopia. Amblyopia is classified based on the type of underlying pathology causing abnormal binocular vision and/or form vision deprivation, and is typically divided into three categories. First, strabismic amblyopia, in which a misalignment in the visual axes of the eyes (i.e. one or both eyes turn in, out, up or down) causes decorrelated images to be received by the visual cortex. The deviating eye can turn inward (esotropia) or outward (exotropia), and the deviation can be intermittent or constant, with the angle of deviation being either stable or variable (B. T. Barrett, A. Bradley, & P. V. McGraw, 2004). Second, anisometropic amblyopia, in which a significant difference in the refractive error (i.e. blur) between the two eyes creates dissimilar images. In anisometropia, the eyes may be myopic or hyperopic to different extents. Third, when there is a physical obstruction to vision in one eye (for example due to the presence of a cataract or drooping of the upper eyelid), amblyopia can develop due to form deprivation. It is thought that amblyopia arises from the mismatch in the images between the two eyes, either due to misalignment (strabismus), differences in blur (anisometropia), or obstruction (form deprivation). In order to prevent diplopia (double vision) or confusion, the image from one eye becomes favoured, while the image from the other eye is suppressed (Harrard, 1996).

Amblyopia is not a single abnormality that can be characterised merely by the difference in optotype acuity between the eyes. McKee, Levi, and Movshon (2003) recruited a large number of observers (N= 427) and measured four visual functions that are known to be impaired in amblyopia: grating acuity, vernier acuity, binocularity, and contrast sensitivity (i.e. the lowest contrast at which a grating of a particular spatial frequency can be identified). Using factor analysis, they determined that two explanatory variables were needed to characterise the underlying functional losses in their sample: acuity and sensitivity. The acuity factor relied heavily on acuity measures (i.e. optotype, vernier, and grating acuity), and the sensitivity factor relied on contrast sensitivity measures, that were taken as edge contrast and the Pelli-Robson contrast thresholds (Pelli, Robson, & Wilkins, 1988). The “amblyopia map” representing the scores of the individuals on each of the two factors revealed that there are marked differences in patterns of visual loss among the clinically defined categories. Those with strabismic amblyopia showed moderate acuity losses but normal or better-than-normal contrast sensitivity at low spatial frequencies. On the other hand, observers with anisometropic amblyopia showed moderate losses in acuity but worse-than-normal contrast sensitivity. Those with

deprivational amblyopia had an indistinguishable pattern of visual deficits from anisometropes. These results suggest that each type of amblyopia is characterised by different losses in visual function.

McKee et al. (2003) argued that a developmental factor that played a role on visual function was the development of binocular vision in the central visual field. Under normal every day viewing, when both eyes are open, the vision for most individuals with amblyopia is dominated by one eye, and is thus monocular. This indicates that impaired stereoscopic depth perception is a key deficit in amblyopia (Webber & Wood, 2005). McKee et al. (2003) found that individuals with a complete loss of binocularity had a particular pattern of visual loss: better than normal contrast sensitivity, and moderate to severe acuity losses. This clearly corresponded to the pattern of visual loss for individuals with strabismus, and indeed most of the observers in their study that did not have binocular vision also had a misalignment in the visual axes. However, the few anisometropic observers with aligned eyes who did not show binocular vision also had this pattern of visual loss, whereas the anisometropic observers with binocular function were distinct from observers with strabismus. This suggests that the loss of binocular function during development plays an integral role in the pattern of visual deficits in amblyopia.

Taking into account these marked differences between amblyopia types and evidence suggesting that crowding primarily affects the strabismic type (Polat, Bonneh, Ma-Naim, Belkin, & Sagi, 2005; Song, Levi, & Pelli, 2014), this introductory chapter will focus entirely on research on unilateral strabismic amblyopia. Therefore, unless otherwise specified, the words *amblyopia* and *amblyopes* in this thesis will refer to the strabismic amblyopia. Additionally, I will refer to the deviating eye with the reduced optotype acuity as the amblyopic eye, and the eye used for preferred seeing as the fellow fixating eye.

Crowding has been found to be elevated in the central vision of the amblyopic eye, and is considered an important characteristic of vision in strabismic amblyopia. It has been known for more than half a century that individuals with amblyopia have better acuity for a letter presented in isolation than when a letter appears in a line with other letters (Stuart & Burian, 1962). Flom, Weymouth, et al. (1963) measured the extent of crowding in the typical and amblyopic fovea by placing flanking bars at various distances from a target Landolt-C. They found that the extent of crowding in the affected eye of observers with amblyopia was significantly greater than in the fovea of observers with typical vision. Since this

observation, many studies have found that the extent of crowding is greater in the fovea of the amblyopic eye relative to the fellow fixating eye, in both children (Greenwood et al., 2012) and adults (Flom, Weymouth, et al., 1963; Hess & Jacobs, 1979; Simmers, Gray, McGraw, & Winn, 1999; Levi, Hariharan, & Klein, 2002a; Bonneh, Sagi, & Polat, 2004, 2007), suggesting that amblyopic crowding persists after childhood treatment.

As indicated by the clinical definition of amblyopia that emphasises reduced visual acuity, amblyopia is primarily thought of as a disorder of spatial vision and is associated with a variety of deficits in spatial vision, not only crowding. Although as reported by McKee et al. (2003) contrast sensitivity for the amblyopic eye is normal or better-than-normal at low spatial frequencies, amblyopia shows a contrast sensitivity deficit for high spatial frequencies (Hess & Howell, 1977; Levi & Harwerth, 1978; Sjöstrand, 1981; Howell, Mitchell, & Keith, 1983). Impaired vernier acuity has been also found in the central vision of the amblyopic eye (Levi & Klein, 1982, 1985). Strabismic amblyopes also show difficulties in tasks involving spatial localisation, such as vertically aligning targets and positioning a line so that it is located in the middle of two others (Bedell & Flom, 1981; Bedell & Flom, 1983), and are said to exhibit positional uncertainty (Fronius & Sireteanu, 1989; Demanins & Hess, 1996; Wang, Levi, & Klein, 1998). Perceptual distortions have also been described in strabismic amblyopia. These distortions are idiosyncratic, and may manifest as straight gratings appearing wavy, “pointy” abrupt positional shifts orthogonal to the orientation of a grating, and stripes appearing fragmented (Pugh, 1958; Sireteanu, Wolf-Dietrich, & Constantinescu, 1993; Barrett, Pacey, Bradley, Thibos, & Morrill, 2003; Sireteanu, Thiel, Fikus, & Iftime, 2008; Piano, Bex, & Simmers, 2015).

In addition to the deficits in spatial vision, higher-level visual functions have also been shown to be affected by the disorder, such as global motion processing and temporal integration. The amblyopic eye shows deficits in the detection of global motion, independently of the contrast sensitivity deficit (Simmers, Ledgeway, Hess, & McGraw, 2003; Simmers, Ledgeway, Mansouri, Hutchinson, & Hess, 2006; Aaen-Stockdale, Ledgeway, & Hess, 2007). This deficit in global motion processing is not reliant on the spatial properties of the stimuli (Aaen-Stockdale & Hess, 2008), and is thus independent from the spatial vision deficits described earlier. In terms of temporal integration, it has been shown that amblyopic observers require elements to be presented closer in time when required to detect a light target among two noise elements (Altmann & Singer, 1986). When presentation times were short, their

performance was close to those of controls, thereby ruling out poor acuity as the source of disrupted performance. It should be noted that other higher-level visual functions, such as face detection and biological motion perception have been found to be intact in amblyopia (Neri, Luu, & Levi, 2007; Cattaneo et al., 2013).

Overall, this section suggests that vision in the amblyopic eye of individuals with strabismic amblyopia is characterised by a variety of deficits, including an increased extent of crowding in the fovea (Levi & Klein, 1985; Greenwood et al., 2012). Based on the variety of deficits, one would predict that on a cortical level, amblyopia would affect multiple visual areas, as psychophysical studies have demonstrated both low-level acuity deficits (Levi & Klein, 1985; McKee et al., 2003{Levi, 1982 #264)} and higher-level deficits in global motion processing (Simmers et al., 2003; Simmers et al., 2006; Aaen-Stockdale et al., 2007). An overview of the evidence on the neural basis of amblyopia is given in the section below.

1.3.1 Neural Basis of Amblyopia

A long-standing question concerns the site of damage in amblyopia. Extensive research has been undertaken on both animal models and more recently on humans to determine the neural correlates of the amblyopic deficit. Exhaustive anatomical and physiological experiments have failed to find any abnormalities in the retina of monkeys reared with experimentally induced amblyopia (Hendrickson et al., 1987). Animal studies have demonstrated a lack of abnormalities in how the visual input from both eyes is relayed through the retina to the thalamus and LGN, suggesting that the inputs to the visual cortex must be normal (Movshon et al., 1987; Kiorpes, Kiper, O'Keefe, Cavanaugh, & Movshon, 1998). As such, the earliest functional physiological abnormalities have been placed in V1, which is where information from the two eyes is first combined (A. Smith, 2015).

In animal models of amblyopia, a misalignment of the eyes during development has been shown to disrupt the binocular connections of cortical neurons in V1 (Hubel & Wiesel, 1965; Harwerth, Smith, Boltz, Crawford, & von Noorden, 1983; Kiorpes et al., 1998) and reduce the number of neurons responding binocularly (Baker, Grigg, & von Noorden, 1974; Crawford & von Noorden, 1979). There is still some controversy on whether the number of cells responding to the amblyopic eye in V1 is smaller than the number of cells responding to the fellow fixating eye. Electrophysiological recordings on cats with artificial strabismus have

shown that fewer neurons are driven by the amblyopic eye relative to the fellow fixating eye (Berman & Murphy, 1981; R. D. Freeman & Tsumoto, 1983; Sireteanu & Best, 1992). Electrophysiological recordings on non-human primates have yielded mixed results, with some showing an equal number of neurons responding to both eyes and thus balanced ocular dominance in V1 (E. L. Smith, Chino, Cheng, Crawford, & Harwerth, 1997; Kiorpes et al., 1998), whereas others have shown an overall shift of eye dominance towards the fellow fixating eye (Crawford & von Noorden, 1979).

Studies in non-human primates have found that single neurons in V1 show lower contrast sensitivity and spatial frequency tuning for stimuli presented to the amblyopic eye relative to the fellow fixating eye (Kiorpes et al., 1998). However, even in severe cases of amblyopia, a relatively small proportion of V1 neurons show reductions in contrast sensitivity at high spatial frequencies (Kiorpes et al., 1998; Kiorpes, Movshon, Chalupa, & Werner, 2003). Additionally, the magnitude of these contrast sensitivity losses, when present, is often too small to account for the severity of the behavioural losses in the same animals (Kiorpes et al., 1998). The emerging view on the neural basis of amblyopia is thus that it involves cortical alterations beyond V1 that result in the variety of behavioural deficits described in the previous section of this Introduction (Kiorpes & McKee, 2006; Kiorpes & Daw, 2018).

The involvement of V2 in the amblyopic deficit has recently been supported by recent animal studies on primates. H. Bi et al. (2011) analysed the receptive field properties of V2 neurons of macaque monkeys raised with strabismic amblyopia, and compared them to V1 neurons in the same animals. They found that the behavioural loss of visual sensitivity was associated with a reduction in the functional connections from V1 to V2 that were severely reduced for the amblyopic eye. They also showed that the spatial resolution and orientation selectivity of V2, but not V1 neurons were abnormal for the amblyopic eye. In both V1 and V2 binocular suppression was robust and the magnitude of suppression was related with the severity of amblyopia. Further supporting a V2 deficit, Shooner et al. (2015) recorded from populations of neurons in V1 and V2, and found a shift in ocular dominance with more neurons responding to the fellow fixating eye. Crucially, they also showed that the magnitude of this shift correlated with the severity of the behavioural visual deficits found in the amblyopic eye.

Neuroimaging studies on human observers with amblyopia have the potential to provide the additional information needed on the site of the deficit. Studies using fMRI have generally confirmed losses at the level of V1, showing a reduction and delay in BOLD response in V1 for the amblyopic eye (Barnes, Hess, Dumoulin, Achtman, & Pike, 2001; Algaze, Roberts, Leguire, Schmalbrock, & Rogers, 2002; Farivar, Thompson, Mansouri, & Hess, 2011). Barnes et al. (2001) presented large high-contrast sinuisoidal stimuli to ensure visibility to both the amblyopic and fellow fixating eye of observers with amblyopia. In comparison to responses driven by the fellow fixating eye, BOLD activation driven by the amblyopic eye was reduced not only in V1, but also in V2. These findings suggested that the cortical deficit in amblyopia extended beyond V1.

Further confirming that the cortical deficit in amblyopia can be seen in areas beyond V1, Li, Dumoulin, Mansouri, and Hess (2007) used fMRI and compared BOLD responses in V1, V2, and V3 to wedge and annulus checkerboard stimuli in observers with amblyopia and typical vision. They found a consistent reduction in activation in area V1, V2, and V3 when driven by the amblyopic eye compared to the fellow eye. This reduction in V2 and V3 activation correlated with V1, indicating that the extrastriate losses could follow as a simple consequence of the V1 loss. Interestingly, Muckli et al. (2006) found a progressive reduction of BOLD response to stimulation of the amblyopic eye in areas V4, V8 and LO as compared to lower visual areas (V1/V2), suggesting that transmission of activity is increasingly impaired while it is relayed towards higher-level visual areas. However, it is hard to resolve from these findings what underlies this reduction in activation in later visual areas, and whether it is simply a result of V1 losses carried across the visual hierarchy.

Recently, Clavagnier, Dumoulin, and Hess (2015) used fMRI to obtain estimates of population receptive field (pRF) size for human observers with typical vision and amblyopia. They showed that pRF sizes were enlarged for 1-6° eccentricity in visual areas V1, V2, and V3 relative to the fellow fixating eye and the eyes of observers with typical vision. Clavagnier et al. (2015) showed that this enlargement in pRF size cannot simply be explained due to anomalies in V1 being reflected in later areas, but rather that additional processing deficits occur in V2 and V3. Although these differences could be underlying the reduced visual function, such as the acuity losses and the increased extent of crowding for the amblyopic eye, the relationship between fMRI estimates of pRF size and measures of visual function in strabismic amblyopia has not been investigated.

Very little is known about the areas in the dorsal visual pathway in amblyopia including middle temporal (MT) and medial superior temporal (MST) area, which are involved in motion processing (Zeki et al., 1991; Tootel et al., 1995). Psychophysical studies point to a global motion processing deficit in amblyopia, implying neural deficits in the dorsal pathway (Simmers et al., 2003; Simmers, Ledgeway, & Hess, 2005; Simmers et al., 2006; Aaen-Stockdale et al., 2007; Aaen-Stockdale & Hess, 2008). Secen, Culham, Ho, and Giaschi (2011) asked observers with typical vision and amblyopia to track one, two or four moving objects. They compared BOLD activation during this attentive tracking condition to passive viewing during which the observers were asked to look at the moving stimuli without tracking them, and to a condition where no moving stimuli were presented. They found that activation in MT was reduced relative to observers with typical vision for both the amblyopic and the fellow fixating eye of observers with amblyopia in both passive viewing and the conditions in which the observers were required to track the moving objects, suggesting that the deficit in amblyopia extends to the dorsal stream.

From the above brief review of evidence on the neural basis of strabismic amblyopia, it is clear that the disorder affects V1 (Hubel & Wiesel, 1965; Harwerth et al., 1983; Kiorpes et al., 1998; Barnes et al., 2001; Algaze et al., 2002; Farivar et al., 2011; Li, Mullen, Thoompson, & Hess, 2011; Clavagnier et al., 2015), with recent studies implicating V2 (H. Bi et al., 2011; Li et al., 2011; Clavagnier et al., 2015; Shooner et al., 2015), as well as higher visual areas (Li et al., 2011; Secen et al., 2011; Clavagnier et al., 2015). If multiple visual processing stages are affected in amblyopia, then the neural basis of crowding in amblyopia could also involve multiple visual areas, similarly to what the evidence on peripheral crowding reviewed in section 1.2.4 suggests. However, due to the multiple of deficits present in the vision of the amblyopic eye, it is not clear whether crowding in strabismic amblyopia is the same as crowding in typical visual periphery. The variability in the deficits of spatial vision, and particularly the acuity reduction, could lead to impaired object recognition in clutter. In the following section, findings in amblyopic crowding are reviewed in order to examine whether the evidence suggests commonalities with crowding in the visual periphery.

1.3.2 Is amblyopic crowding the same as crowding in typical vision?

In the typical visual periphery, the extent of crowding is greater than what would be predicted based on reduced peripheral acuity and as such independent

from target size (Levi, Hariharan, et al., 2002b; Tripathy & Cavanagh, 2002; Pelli et al., 2004; Pelli et al., 2007), but dependent on the separation between the target and the flankers (Bouma, 1970). Peripheral crowding also shows a radial-tangential anisotropy, where radially positioned flankers are more disruptive on target recognition than tangentially positioned flankers, making the shape of the spatial zone over which crowding occurs elliptical (Toet & Levi, 1992; Pelli et al., 2007; Petrov & Popple, 2007; Greenwood et al., 2017). Peripheral crowding is also selective for the similarity between the target and flankers in dimensions such as orientation or colour (Kooi et al., 1994; Kennedy & Whitaker, 2010), and can be determined by whether or not the target and flankers group (Manassi et al., 2012, 2013; Herzog & Manassi, 2015; Herzog et al., 2015). In this section, the available evidence on amblyopic crowding is reviewed to examine whether it also shows these characteristics.

A question that has concerned researchers is whether the extent of amblyopic crowding could be accounted for by the reduced acuity of the amblyopic eye. In the typical fovea, it has been shown that crowding is proportional to acuity (Flom, Weymouth, et al., 1963; Song et al., 2014; Yehezklel, Sterkin, Lev, & Polat, 2015). As such, if the extent of crowding in the amblyopic fovea were proportional to acuity for an isolated pattern, then this would suggest that crowding in amblyopia is essentially “normal” (i.e. a scaled up version of crowding in the typical fovea). In the typical periphery however, crowding is not proportional to acuity, but extends over larger distances (e.g. Toet & Levi, 1992). Levi and Klein (1985) measured the extent of crowding in amblyopia with a vernier target and flanking lines placed at different separations from the target. They compared the measurements from amblyopic observers with those from observers with typical vision tested at 5° eccentricity in the visual periphery. They found that the amblyopic eye showed elevated vernier acuity and crowding compared to the unaffected eye of the same amblyopic observers, and the eyes of observers with typical vision. Interestingly, they found that the extent of crowding was proportional to vernier acuity. This relationship of crowding to acuity has also been found in other studies (Flom, Weymouth, et al., 1963; Hess & Jacobs, 1979; Simmers et al., 1999), and suggests that the extent of amblyopic crowding is essentially “normal” when the elevation in acuity is taken into account.

However, more recent studies have shown that crowding in amblyopia extends over greater distances than what would be expected based on the acuity

deficit. Hess, Dakin, Tewfik, and Brown (2001) found that the extent of crowding for Landolt-Cs with flanking bars was much larger in the amblyopic eye than what would be predicted when expressed relative to acuity. Levi, Hariharan, et al. (2002a) extended these findings using a tumbling E target and flanking lines comprised by smaller Gabor or Gaussian patches. They found that crowding extended over larger distances than what would be predicted by the reduction in acuity, even when tested with low spatial frequency to match performance levels to the typical fovea. In a study that had a significantly larger sample than the previous studies reported above (N=50), Bonneh et al. (2004) showed that the greater extent of crowding in the affected eye of observers with amblyopia was largely independent from the acuity deficit. Similarly, Song et al. (2014) tested the typical periphery and the amblyopic fovea, and showed that the extent of crowding was disproportionately large relative to acuity in both cases, with flankers positioned at separation up to 3 times the target size interfering with target identification. Taken together, these findings are contrary to earlier studies and suggest that amblyopic crowding extends over larger distances than what would be predicted by the reduction in acuity, and point to similarities between crowding in the typical periphery and the amblyopic fovea.

A potential explanation for the discrepancy in the findings with regards to the relationship between acuity and crowding in amblyopia can be provided by the results of Greenwood et al. (2012). Greenwood et al. (2012) tested children with unaffected vision and amblyopia, and found that although acuity was correlated with the extent of crowding for both groups of children, some children with strabismic and mixed strabismic and anisometropic amblyopia showed crowding of a magnitude that exceeded acuity predictions. The cases that showed this uncorrelated crowding were the ones with the greatest interocular differences in acuity. This suggests that the discrepancies in the literature might stem from variations in the clinical characteristics of the patients tested, with studies showing that amblyopic crowding is disproportional to the acuity deficit including patients with a greater depth of amblyopia.

When considering the dependence of amblyopic crowding on target size, Flom, Weymouth, et al. (1963) found that in the amblyopic fovea, the disruption caused by flanking bars on the recognition of the target was reduced when the separation between them increased. Similar results have been obtained by Polat et al. (2005) and Levi and Carney (2011), demonstrating that there is a critical spacing

for observers with amblyopia over which crowding occurs in the fovea of the amblyopic eye. Levi, Hariharan, et al. (2002a) and Hariharan et al. (2005) not only showed that increasing centre-to-centre separation alleviated crowding in amblyopic observers, but also demonstrated this critical spacing does not depend on target size. These results demonstrate that similarly to crowding in the typical visual periphery (Bouma, 1970; Levi, Hariharan, et al., 2002b; Tripathy & Cavanagh, 2002; Pelli et al., 2004; Pelli et al., 2007), the extent of amblyopic crowding depends on the separation between the target and flankers and does not scale to the size of the target (e.g. decreasing with small targets), but remains disproportionately large.

A characteristic of the extent of peripheral crowding is the radial-tangential anisotropy (Toet & Levi, 1992; Levi & Carney, 2009). Levi and Carney (2011) investigated the existence of such an anisotropy in amblyopic crowding. With a Gabor stimulus as the target, they placed two flankers horizontally (one on each side) or vertically to the target (one on the top and the second on the bottom). They found that there was no difference in performance in a contrast discrimination task between the two flanker conditions, indicating that the crowding zone in the fovea of the amblyopic eye is not elliptical, similarly to the typical fovea. The lack of differences between the horizontal and vertical positioning of the flankers suggests that this elliptical shape of the crowding zone is specific to peripheral crowding (Toet & Levi, 1992).

An additional consideration is whether amblyopic crowding affects the same visual dimensions as peripheral crowding, and if it does, whether it shows selectivity along these dimensions. Studies on amblyopic crowding have primarily used simple stimuli such as Gabor patches (Levi & Carney, 2011), verniers (Levi & Klein, 1985), Gabor and Gaussian patches forming letter characters (Levi, Klein, et al., 2002; Bonneh et al., 2004; Hariharan et al., 2005), letters (Flom, Weymouth, et al., 1963; Giaschi, Regan, Kraft, & Kothe, 1993; Hess, Dakin, et al., 2001; Greenwood et al., 2012; Song et al., 2014), and numbers (Bonneh et al., 2007). With the use of these stimuli, it has been shown that crowding affects the discrimination of orientation (Flom, Weymouth, et al., 1963; Hess, Dakin, et al., 2001; Levi, Hariharan, et al., 2002a; Bonneh et al., 2004; Hariharan et al., 2005; Levi & Carney, 2011; Greenwood et al., 2012), position (Levi & Klein, 1985), and contrast (Levi & Carney, 2011), as well as the recognition of letters (Giaschi et al., 1993). Evidence on dimensions such as colour and motion is lacking, and complex stimuli such as faces have not been used to investigate amblyopic crowding. Interestingly however, amblyopic crowding has

also been found to extend in the temporal as well as the spatial domain, with observers with amblyopia having difficulty identifying the target letter in rapid serial presentations (Bonneh et al., 2007)

There are very few studies that have attempted to shed light on whether amblyopic crowding is selective for target-flanker similarity and their results have been mixed. In the visual periphery, crowding is modulated by the contrast polarity of the flankers: on a black target, crowding is strong and extends over large distances with black flankers, but weak and extends over smaller distances when the polarity of the flankers is reversed to white (Kooi et al., 1994; Chakravarthi & Cavanagh, 2007; Chung & Mansfield, 2009; Rosen & Pelli, 2015). Hess, Dakin, et al. (2001) used a Landolt-C target with flanking bars of the same or opposite polarity to the target to investigate whether amblyopic crowding is modulated by flanker polarity. They found that for some observers with amblyopia the extent of crowding depended on the polarity of the flanking bars, with bars of opposite polarity having a reduced disruptive effect on target identification compared to same-polarity flankers. However, in other amblyopic observers, flanking bars of opposite polarity disrupted target recognition. Extending these findings, Hariharan et al. (2005) found that the extent of crowding was similar with same- and opposite- polarity flankers in observers with amblyopia. These results suggest that for some individuals with amblyopia, crowding is sensitive to differences in polarity between the target and flankers, but for others it does not show such selectivity.

Levi, Hariharan, et al. (2002a) showed that crowding in the amblyopic fovea occurs regardless of the similarity in contrast between the target and the flankers. In the typical visual periphery, when the target has a higher contrast than the flankers, the flankers are less disruptive than when their contrast is matched. However, when the target has a lower contrast than the flankers, the flankers are more disruptive than when their contrast is the same as the target (Kooi et al., 1994; Chung et al., 2001). Levi, Hariharan, et al. (2002a) used Gaussian patches comprising a tumbling E target, and four flanker lines, one on each side of the target E. They kept the target contrast constant, and varied the contrast of the flankers. They found that for both amblyopic observers tested, target recognition was disrupted regardless of flanker contrast. Importantly, the target was crowded even when the flanker contrast was below the observers' individual flanker detection threshold, making the flankers invisible. These findings suggest that amblyopic crowding is not modulated by the similarity in contrast between the target and flankers.

In order to test the effect of orientation, for the same observers Levi, Hariharan, et al. (2002a) used the same stimulus configuration but used Gabor patches to comprise the tumbling E and flanker lines. They varied the orientation of the Gabor patches making up the flankers whilst keeping the contrast constant. They found that flankers with an orthogonal carrier orientation to the target substantially disrupted target identification, but to a lesser degree than flankers that had the same carrier orientation to the target, indicative of some selectivity for target-flanker orientation similarity. This finding however has not been replicated for other amblyopic observers, who have shown similar performance on crowding tasks with both similarly- and orthogonally- oriented flankers (Hariharan et al., 2005). This is contrary to crowding in the visual periphery measured with the same stimulus configurations, where crowding is orientation specific, with flankers of orthogonal orientation to the target showing little or no disruptive effect on target recognition (Levi, Hariharan, et al., 2002b; Hariharan et al., 2005).

As discussed in the section 1.2.3 on grouping effects in crowding in typical vision, whether or not the target forms a uniform configuration with the flankers can determine how disruptive crowding is on target recognition. Levi and Carney (2009) varied the number of flankers and tested their effect on orientation discrimination of a Gabor patch at 5° eccentricity in the typical visual periphery. The flankers were segments of an annular surround, and eight such segments formed the full surround. In the periphery, they found that performance dropped and the flankers became increasingly more disruptive from one to four flanker segments. However, with eight flankers forming a full annular surround, performance improved and crowding was less disruptive than with four segment flankers. Levi and Carney (2011) used the exact same stimuli and paradigm to investigate grouping effects in the amblyopic fovea. Interestingly, they found the same pattern of results: up to four flankers, crowding was strong, but with eight flankers crowding was significantly reduced. These findings could be interpreted based on grouping principles: four segment flankers group with the target resulting in strong crowding, whereas in the eight-segment flanker condition, there is flanker-flanker grouping, as the flankers form a uniform annular configuration without the target and crowding is released. However, with no other studies investigating the effects of grouping in amblyopia, it is unclear whether target-flanker grouping and flanker-flanker grouping can modulate amblyopic crowding. As such, more research is needed to determine

whether amblyopic crowding can be determined by grouping like crowding in the typical visual system.

From the above, it can be concluded that amblyopic crowding extends over large distances (Flom, Weymouth, et al., 1963; Yehezklel et al., 2015), in some cases over what would be predicted by acuity limitations (Hess, Dakin, et al., 2001; Levi, Hariharan, et al., 2002a; Bonneh et al., 2004; Song et al., 2014), is dependent on target-flanker separation but independent of target size (Flom, Weymouth, et al., 1963; Hariharan et al., 2005; Polat et al., 2005; Levi & Carney, 2011), shows little evidence of selectivity for target-flanker similarity (Hess, Dakin, et al., 2001; Levi, Hariharan, et al., 2002a; Hariharan et al., 2005) and might be modulated by grouping effects (Levi & Carney, 2011). However, it is unclear what effects amblyopic crowding has on the appearance of crowded objects. As seen in section 1.2.3, the investigation on the perceptual effects of peripheral crowding have been integral in informing the underlying mechanism, and would be similarly informative in determining the mechanism underlying amblyopic crowding. Additionally, it is not clear whether amblyopic crowding can be determined by higher-level grouping modulations, like crowding in the typical visual periphery. Nonetheless, it is important to consider what the current findings can tell us about the underlying mechanism of amblyopic crowding. Models of amblyopic crowding are considered in the next section.

1.3.3 Models of amblyopic crowding

Amblyopic crowding has been argued to be a consequence of a shift to the spatial scale of analysis, and thus a mere consequence of the acuity deficit. Flom, Weymouth, et al. (1963) argued that crowding in amblyopia is related to the size of the receptive fields that respond to the target. According to this theory, amblyopic crowding resembles crowding in the typical fovea that is proportional to acuity (Flom, Weymouth, et al., 1963; Yehezklel et al., 2015) and is thus “normal” with the only difference lying on the scale factor. As discussed above, although for some individuals with amblyopia the extent of crowding can be predicted by the reduced acuity (Flom, Weymouth, et al., 1963; Hess & Jacobs, 1979; Levi & Klein, 1985; Simmers et al., 1999), others show crowding zones which far exceed what would be predicted by acuity (Hess, Dakin, et al., 2001; Levi, Hariharan, et al., 2002a; Bonneh et al., 2004; Greenwood et al., 2012; Song et al., 2014). This casts doubt on the notion that crowding in amblyopia is essentially “normal”, with the only difference

lying in the scale factor (i.e. the acuity). Additionally, the independence of the extent of amblyopic crowding on target size (Levi, Hariharan, et al., 2002a; Hariharan et al., 2005) and its dependence on target-flanker separation (Polat et al., 2005; Levi & Carney, 2011) suggest some fixed region within which the crowding process occurs, independent of the acuity deficit. The findings discussed above suggest that a scale-shift account of amblyopic crowding is unlikely.

Further questioning the scale-shift hypothesis is the evidence that amblyopic crowding does not affect the detection of an object: even objects that are strongly crowded and difficult to recognise can be detected by amblyopic eyes (Levi, Hariharan, et al., 2002a). This is similar to crowding in the visual periphery, where discrimination but not detection is affected (Andriessen & Bouma, 1976; Levi, Hariharan, et al., 2002b; Pelli et al., 2004; Livne & Sagi, 2007). As such, it has been argued that similarly to the periphery, crowding in amblyopia must occur after the features of the target object are detected. It has thus been suggested that a similar mechanism resulting from inappropriate feature integration could be underlying both peripheral and amblyopic crowding (Levi, Hariharan, et al., 2002a; Hariharan et al., 2005). Such integration or pooling could be resulting from an engagement of large receptive fields leading to inappropriate feature or object integration beyond V1. As discussed in the previous subsection of this Introduction, there is evidence of an enlargement of population receptive size in humans with strabismic amblyopia (Clavagnier et al., 2015), although this is not necessarily indicative of an enlargement in the receptive field size of individual neurons comprising the pRF. As a first step to test a pooling account of amblyopic crowding, evidence on the perceptual outcomes would shed light on whether these effects could be a result of pooling of the target and flanker features.

Alternatively, there are some empirical indications that amblyopic crowding could be a result of increased positional uncertainty and unfocused attention. Similarly to the periphery, amblyopic vision is characterised by high degrees of positional uncertainty (Levi et al., 1987; Hess & Field, 1994; Wang et al., 1998). For example, high-contrast remote flankers positioned at a distance from the target facilitate the detection of a single patch in the typical fovea (Levi, Klein, et al., 2002), by providing positional cues to the location of the target. However, such facilitation is not evident in the amblyopic fovea (Polat, Sagi, & Norcia, 1997; Levi, Hariharan, et al., 2002a). It has been suggested that weaker facilitation in amblyopia could be a

consequence of increased positional uncertainty, so that even the location of high contrast flankers is uncertain (Levi, Hariharan, et al., 2002a).

Alternatively, this positional uncertainty may arise due to unfocused visual attention. Although there has been very little research on higher-level attentional deficits in amblyopia, the impairment in identifying targets in rapid serial visual presentations of stimuli (Bonneh et al., 2007; Popple & Levi, 2008) suggests that attention in amblyopia might be unfocused. Additionally, Sharma, Levi, and Klein (2000) asked observers with strabismic amblyopia to count briefly presented features using only their amblyopic eye. They found that observers with amblyopia significantly undercounted the number of features presented. These errors were not reduced by increasing the stimulus presentation duration or by increasing feature visibility. Cueing to the target locations however reduced errors made by observers with amblyopia. Sharma et al. (2000) also showed that observers with amblyopia underestimated the number of features missing from a uniform grid. They argued that this cannot be accounted for by perceptual distortions, as adding more noise (i.e. positional jitter) on the grid for observers with typical vision did not result in underestimations of missing features, but rather overestimations, as the noise resulted in apparent 'holes'. It is however still likely that perceptual distortions in strabismic amblyopia are idiosyncratic (Barrett et al., 2003; Piano et al., 2015) and do not manifest as simulated in this study. Nonetheless, the authors concluded that of observers with amblyopia have a difficulty in individuating objects due to a deficit in the ability of attentional mechanisms to isolate each object. With the limited evidence on attentional deficits in amblyopia, and the multitude of evidence on deficits in spatial vision (McKee et al., 2003; Levi, 2006), it is more likely that crowding in amblyopia is a deficit of spatial vision.

From the above review on research in amblyopic crowding, it is clear that many aspects of the phenomenon still remain to be investigated. Although it appears unlikely that amblyopic crowding is merely a result of the acuity deficit or a consequence of unfocused visual attention, it still cannot be determined from the research available what the underlying mechanism is.

1.4 Thesis Outline

With the evidence available, it is unclear what the mechanism underling amblyopic crowding is, and whether crowding in amblyopia is the same as crowding in the typical visual system. In this thesis, a pooling framework of crowding was

adopted. Pooling accounts propose that crowding in typical vision arises due to the integration features of adjacent objects (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010; van den Berg et al., 2010; Harrison & Bex, 2015), in order to promote perceptual homogeneity where sampling is insufficient and neurons have large receptive fields (Parkes et al., 2001). In this thesis, I explored whether crowding in strabismic amblyopia can be accounted for by a pooling framework of crowding. In order to achieve this, I examined three different components of amblyopic crowding: first, the perceptual effects (i.e. the effects of amblyopic crowding on the appearance of target objects), second, the neural basis of amblyopic crowding, and third, whether grouping can determine crowding in amblyopia.

The first empirical chapter (Chapter 2) investigated whether amblyopic crowding shares the same perceptual effects as crowding in the visual periphery. In the periphery, crowding has been shown to have systematic effects on the appearance of the target: it shifts the identity of the target towards that of the flankers (Greenwood et al., 2010). Adults are reported to make assimilation errors, reporting intermediate or average orientations between the target and flankers (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010), and in other instances substitution errors (Strasburger et al., 1991; Ester, Klee, & Awh, 2014; Ester et al., 2015). I used an orientation-matching task to investigate the perceptual effects of crowding in the fovea of children with amblyopia and typical vision, and in the adult periphery. I reasoned that if the errors of children with amblyopia are systematic and resemble those made by adults in the periphery, this would demonstrate common perceptual effects. It would also suggest that these effects can be modelled with a common mechanism, such as a population response pooling model that can account for both assimilation and substitution errors (van den Berg et al., 2010; Harrison & Bex, 2015, 2016). Alternatively, crowding could be the result of the perceptual distortions that characterise amblyopic vision (Bedell & Flom, 1981; Bedell & Flom, 1983; Flom & Bedell, 1985; Barrett et al., 2003; Sireteanu, Thiel, et al., 2008). Such distortions would alter the appearance of the crowded target non-systematically, resulting in random errors across the amblyopic sample. These errors would indicate that a distinct mechanism underlies crowding in amblyopia than the typical visual periphery.

The second empirical chapter (Chapter 3) investigated the neural basis of crowding in amblyopia. In unaffected vision, crowding is minimal in the fovea but

rises in the periphery (Bouma, 1970; Toet & Levi, 1992). This effect has been attributed by pooling models (Parkes et al., 2001) to the cortical undersampling of the peripheral visual field, with the extent of crowding increasing because receptive field size increases with eccentricity (Van Essen et al., 1984). However, the relationship between receptive field size and crowding in typical vision is still untested. In this chapter, I investigated whether there is a relationship between crowding and receptive field size in typical vision and amblyopia. In order to characterise variations in visual function across the visual field, I measured acuity and crowding in the fovea and the periphery of observers with amblyopia and typical vision. Second, using functional magnetic resonance imaging (fMRI), measures of population receptive field size (pRF) (Dumoulin & Wandell, 2008) were obtained. I reasoned that if the pattern of variation in the extent of crowding across the visual field follows the same pattern as pRF size in both typical and amblyopic vision, then this would indicate that an increase of receptive field size could be underlying crowding in amblyopia and typical vision.

The third empirical chapter (Chapter 4) investigated whether grouping can determine crowding in strabismic amblyopia. In the typical visual system, it has been shown that crowding can be determined by grouping of the target and flankers (e.g. Herzog & Manassi, 2015). When the target and flankers form a uniform configuration, there is target-flanker grouping and crowding is strong. In contrast, when the flankers form a uniform configuration with each other but not the target, there is flanker-flanker grouping and crowding is weak (Banks et al., 1979; Malania et al., 2007; Sayim et al., 2008; Sayim, Westheimer, & Herzog, 2010; Manassi et al., 2012, 2013; Manassi, Hermens, Francis, & Herzog, 2015; Manassi, Lonchampt, Clarke, & Herzog, 2016). Here, I investigated whether crowding in strabismic amblyopia shows evidence for target-flanker grouping and flanker-flanker grouping with multi-element flanker arrays. I reasoned that if crowding in amblyopia were strong (i.e. the spatial extent is large) in cases of target-flanker grouping, and weak (i.e. the spatial extent is smaller) in cases of flanker-flanker grouping, this would demonstrate that amblyopic crowding is determined by grouping. Alternatively, if there are no variations in the extent of crowding between conditions in which there is target-flanker grouping and flanker-flanker grouping, then this would indicate that amblyopic crowding is not determined by grouping.

2 **Chapter 2:** The perceptual effects of crowding in amblyopic, developing, and peripheral vision

2.1 **Introduction**

As discussed in Chapter 1, crowding in the fovea of observers with strabismic amblyopia shares certain characteristics with crowding in peripheral vision. In the visual periphery, the extent of crowding is determined by the separation between the target and the flankers, extending up to half the target eccentricity (Bouma, 1970), and is largely independent from target size (Levi, Hariharan, et al., 2002b; Tripathy & Cavanagh, 2002; Pelli et al., 2007). The extent of crowding in amblyopia also depends on the separation between the target and flankers, and does not scale with the size of the target (Flom, Weymouth, et al., 1963; Levi, Hariharan, et al., 2002a; Polat et al., 2005; Levi & Carney, 2011). However, there are also differences between amblyopic and peripheral crowding. Peripheral crowding shows selectivity for target-flanker similarity. For example, similarly-oriented flankers to the target are more disruptive than flankers with a different orientation to the target (Andriessen & Bouma, 1976; Hariharan et al., 2005). In contrast, for most observers with amblyopia, flankers strongly disrupt target recognition regardless of whether they are similar or dissimilar to the target in dimensions such as orientation (Levi, Hariharan, et al., 2002a; Hariharan et al., 2005). Therefore, it is unclear whether amblyopic and peripheral crowding are instances of the same underlying mechanism.

In the typical visual periphery, insights into the mechanisms underlying crowding have come from research into the errors observers make when tasked with reporting the identity of a crowded target. In an early study, Parkes et al. (2001) found that although observers' performance was poor when asked to identify the orientation of a crowded target, they could accurately report the average orientation of the target and flankers. A number of other studies have shown that the errors observers make are best predicted by a weighted average of the target and flankers features, such as their orientation (Greenwood et al., 2009; Dakin et al., 2010). However, there is also evidence to suggest that observers make a disproportionate amount of flanker reports instead of reporting the target (Strasburger et al., 1991; Strasburger, 2005; Nandy & Tjan, 2007; Ester et al., 2014; Ester et al., 2015). Together, these results suggest a shift in the appearance of crowded targets, either partial or complete, towards the flanker identity. As changes in the flanker can induce illusory changes in the identity of the target even when it is blank

(Greenwood et al., 2010), this shift towards the flankers has been taken as a *perceptual* effect of crowding, and not merely a decisional bias.

Of the many theories put forward to explain peripheral crowding, two types of models with contrasting predictions have primarily been used to explain the perceptual effects that arise due to crowding. First, averaging models (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010) that are based on the principles of pooling accounts of crowding. These models posit that crowding is compulsory pre-attentive averaging, resulting in observers perceiving and thus reporting an average feature (e.g. orientation) of the target and flankers. In contrast, substitution models (Ester et al., 2014; Ester et al., 2015) are based on the principle that crowding emerges due to the substitution of a flanker onto the target, resulting in observers reporting the flanker identity. This substitution is either attributed to the increased positional uncertainty that characterises peripheral vision (Wolford, 1975; Krumhansl & Thomas, 1977; Strasburger et al., 1991), or unfocused spatial attention (Strasburger et al., 1991; Strasburger, 2005). However, the issue with these two models of the perceptual effects of peripheral crowding is that they can only predict one type of error.

More recently, population response models of peripheral crowding have been put forward (van den Berg et al., 2010; Harrison & Bex, 2015), that can account for both averaging and substitution errors. Harrison and Bex (2015) used an orientation matching psychophysical paradigm, and asked adult observers to match the orientation of a reference Landolt-C to a crowded Landolt-C target presented in the visual periphery. They found that observers made a range of errors: they reported average or intermediate orientations between the target and flankers, which will be referred to as *assimilation errors*, as well as reporting the orientation of the flankers, which will be referred to as *substitution errors*. Importantly, similarly to an earlier model by van den Berg et al. (2010), Harrison and Bex (2015) showed that both these types of errors can be accounted for by a population response model of crowding.

The model by Harrison and Bex (2015) is based on the principles of population coding (Pouget et al., 2000), and as such assumes a population of detectors that are selective across fundamental visual dimensions, such as V1 neurons for orientation (Schiller et al., 1976). According to this model, crowding results from the pooling of the population responses to the target and flankers. After pooling occurs, the perceived identity of a crowded target is then read-out as the

peak of this pooled response. Using weights, the differential contribution of the population response to the flanker orientation and the target orientation in the pooled response can be determined. An equal weighting of the response to the target and flankers results in a pooled population response with a peak on an average orientation, and thus an assimilation error. A high flanker weight results in a greater contribution of the response to the flanker orientation in the pooled population response, and thus the peak would be closer to the flanker orientation resulting in a substitution error. Hence, population response models of crowding can account for the different types of systematic errors seen for peripheral crowding, reconciling compulsory averaging and substitution models.

It is unclear whether such a population response pooling model could account for the perceptual errors made by observers with amblyopia, as the perceptual effects of amblyopic crowding have not been investigated. Based on previous studies measuring different visual functions, assumptions can be drawn on the perceptual effects of amblyopic crowding. Similarly to the periphery, vision in the fovea of observers with strabismic amblyopia is characterised by increased positional uncertainty. Positional accuracy is compromised both when targets are presented at the acuity limit (Levi & Klein, 1982, 1985) and at significantly larger sizes (Demanins & Hess, 1996). Based on this increased positional uncertainty, it could be expected that observers with amblyopia would confuse the flanker for the target, and similarly to the periphery make substitution errors when tasked with reporting the identity of a crowded target. It is however unclear whether they would also make assimilation errors.

Alternatively, the perceptual effects of amblyopic crowding may arise due to perceptual distortions affecting the crowded target. When tasked to align a light stimulus with two vertically arranged reference marks at different locations in the peripheral visual field using their amblyopic eye, observers with strabismic amblyopia not only showed marked positional uncertainty, but also consistent displacements with relation to the reference marks (Fronius & Sireteanu, 1989). Additionally, when asked to memorise and reconstruct circles of different radii, strabismic amblyopes showed considerable distortions, including shrinkage, expansion, and torsion of specific regions of the visual field when using their amblyopic eye (Sireteanu, Lagreze, & Constantinescu, 1993). Based on these and other findings (Lagreze & Sireteanu, 1991; Sireteanu, Baumer, & Iftime, 2008), amblyopic vision is said to be characterised by perceptual distortions. Although

these distortions are reported to be consistent over time within an observer, they are idiosyncratic and thus random between observers and across space (Barrett et al., 2003). As such, their effect on a crowded target would result in an interaction between the distortion and the target, but it would not depend on the similarity between the target and the flankers. Therefore, if these distortions underlie the perceptual effects of amblyopic crowding, then perceptual errors would be random, suggestive of a distinct mechanism than the typical periphery.

In this chapter, I investigated the perceptual effects of crowding in children with strabismic amblyopia. As a comparison, I also investigated the perceptual effects of crowding in the fovea of children with typical vision. The extent of crowding in the fovea of typically developing children has been found to be greater relative to the adult fovea (Atkinson & Braddick, 1983; Atkinson et al., 1987; Greenwood et al., 2012), with this elevation reported to persist up until 11 years of age (Jeon et al., 2010). Similarly to strabismic amblyopia, it is unknown what effects developing crowding has on the appearance of a target in clutter. However, studies have demonstrated that children make a disproportionate amount of random errors in psychophysical tasks relative to adults (Witton, Talcott, & Henning, 2017; Manning, Jones, Dekker, & Pellicano, 2018). These errors are made in “catch” trials (Treutwein, 1995) that are above children’s threshold and are considered easy. Because children are expected to respond correctly in these catch trials, these errors can be considered random, and are typically attributed to attentional lapses and poorer short-term memory than adults (Witton et al., 2017; Manning et al., 2018). As such, these random errors could also dominate responses when children are tasked to report the identity of a target both when the target is isolated and when it is crowded. This would make it difficult to determine the perceptual effects of crowding in children with typical vision, as well as children with amblyopia.

The aim in this chapter was to investigate whether the perceptual effects of crowding in amblyopic and typically developing vision are the same as in the adult visual periphery. In order to achieve this, I tested children aged 3-9 with strabismic amblyopia and typical vision in the fovea, and adults with typical vision in the periphery using an orientation-matching task similar to the one used by Harrison and Bex (2015). If children with amblyopia and typical vision make both assimilation and substitution errors, this would demonstrate that crowding in these instances has the same systematic effects on the appearance of the target as in the adult periphery. In contrast, random errors would suggest that crowding affects the

appearance of the target in a non-systematic manner. Common perceptual effects would be indicative of a common underlying mechanism for amblyopic, developing, and peripheral crowding, and it should be possible to simulate them using the same population response pooling model. Alternatively, random errors would require a crowding mechanism exclusive to strabismic amblyopia and/or developing vision.

2.2 **Methods**

2.2.1 Observers

2.2.1.1 Children

40 children were tested, between 3 and 9 years of age, divided into two groups: a control group with typical vision (*n*= 20, mean= 73.2 months), and a group with strabismic amblyopia (*n*= 20, mean= 72.1 months). All children were tested at the Richard Desmond Children's Eye Centre at Moorfields Eye Hospital in London, UK.

Prior to taking part in the study, children underwent a full orthoptic assessment to ensure they met our inclusion and exclusion criteria. In the control group, children were selected to have a best-corrected acuity of 0.1 logMAR (logarithm of the minimum angle of resolution) or better, as measured by orthoptic charts (Thompson V200), in the absence of any pre-existing visual or neurological deficits. Note that 0.0 logMAR corresponds to about 1 minute of arc (arcmin).

In the group with strabismic amblyopia, inclusion was made based on the presence of amblyopia as indicated by a two-line difference in best-corrected logMAR acuity between the eyes, as well as on manifest heterotropia (deviation of the visual axes). This heterotropia could be either esotropia (inward deviation of the visual axes) or exotropia (outward deviation of the visual axes). Children with additional visual deficits (e.g. macular dystrophies) and developmental or neurological deficits (e.g. autism) were excluded. We did not exclude cases of joint anisometropia and strabismus, as these mixed cases show a very similar pattern of visual deficits to those seen in pure strabismic amblyopes (McKee et al., 2003). Crucially for the purposes of this chapter, both strabismic amblyopes and mixed cases have large foveal extents of crowding (Greenwood et al., 2012; Song et al., 2014).

Three children did not complete all experimental tasks and were excluded from the analysis. They are not included in the tallies above. Clinical details of the

children tested from both groups can be found in Appendix A section 6.1. The experimental procedures were performed with the informed consent of the observers and were approved by the East of England – Cambridge South Research Ethics Committee of the National Health System (NHS) Health Research Authority.

2.2.1.2 Adults

10 adults were tested (4 males, M= 28.7 years, range 24-35 years). All had normal or corrected-to-normal visual acuity. As indicated by their own report, none had amblyopia or strabismus, or any history of binocular dysfunction.

2.2.2 Apparatus

2.2.2.1 Children

Each child completed three tasks: acuity, crowding extent, and an orientation-matching task. Experiments were programmed using Matlab (The Mathworks, Ltd., Cambridge, UK) on a Dell PC running PsychToolBox (Brainard, 1997; Pelli, 1997). Stimuli were presented on an ASUS VG278HE LCD monitor, with 1920 × 1080 resolution and 120 Hz refresh rate. The monitor was calibrated using a Minolta photometer and linearised in software, to give a maximum luminance of 150 cd/m^2. A second Dell UltraSharp 2208WFP monitor, with 1680 × 1050 resolution and 75 Hz refresh rate, was positioned above the first. In the acuity and crowding tasks, this monitor was used to display a running tally of the points children received by playing the games. In the orientation-matching task, it displayed the response stimulus.

Figure 2.1A shows the experimental setup for the children. Children wore stereo-shutter glasses (nVidia Corp., Santa Clara, CA) alternating at 120 Hz. These glasses were used to present the stimuli monocularly. The glasses were custom-fitted in a ski mask frame in order for the children to be able to wear them comfortably above their optical correction. Children were seated 3 meters away from the screen. For the acuity and crowding-extent tasks large pictures of the ghosts were placed at the monitor edges as aid. The experimenter recorded the children's' responses in the tasks using the keyboard. For the orientation-matching task a Griffin Powermate response dial was used to register responses.

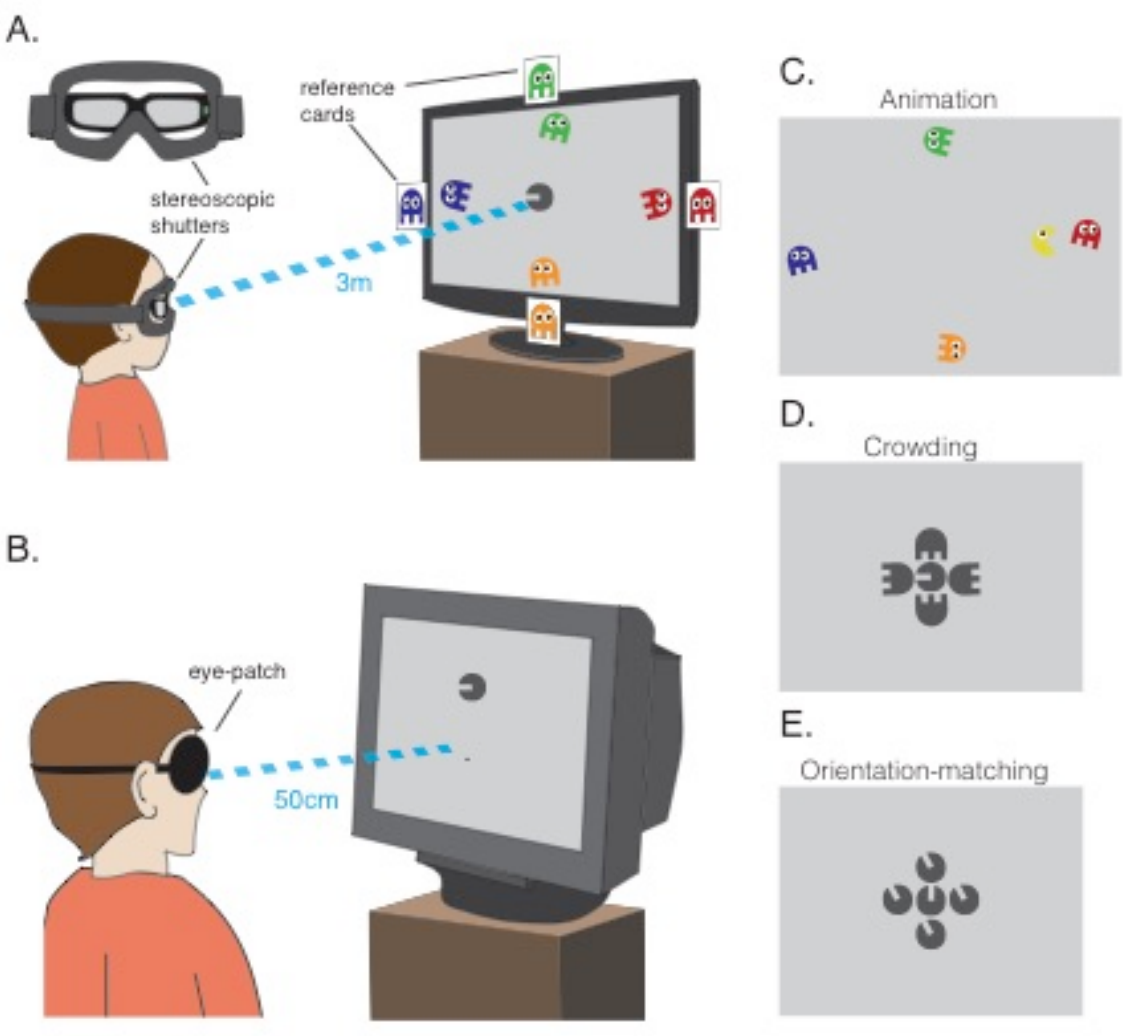

Figure 2-1 Apparatus and stimuli

A. For children, the stimuli were viewed through stereoscopic shutter glasses fitted in a children's ski mask and presented on a monitor that was 3D-compatible at a distance of 3 m. An example trial of the acuity task is depicted on the screen. Children were tasked to report the colour of the ghost that VacMan was facing. Coloured cards of the ghosts were placed on the monitor edges to aid the children in selecting the ghost.
B. An example frame from the "reward animation" that was presented every three correct trials.
C. Illustration of the stimuli in the crowding extent task. Ghost flankers were presented at random orientations at a fixed relative separation (1.1 × stimulus diameter) while their absolute separation was varied by QUEST. Note that these stimuli were used to measure crowding for both children and adults.
D. Illustration of the stimuli in the orientation matching tasks. Flankers identical to the target (filled-in Landolt-Cs) were presented in the same orientation at a fixed separation. Note these stimuli were used in both children and adults' version of this task.

2.2.2.2 Adults

Adults completed the same three tasks as children: acuity, crowding, and an orientation-matching task. Experiments were run using Matlab on a Viglen Genie PC running PsychToolBox (Brainard, 1997; Pelli, 1997). Stimuli were presented on a Sony GDM-FW900 cathode ray tube (CRT) monitor with 2304 × 1440 resolution and 80Hz refresh rate. The monitor was calibrated and linearised to give a maximum

luminance of 122 cd/m^2. The use of a CRT ensured uniform luminance values across the periphery, which is difficult with LCD monitors due to bleed-through. For the acuity and crowding extent tasks, observers registered their response using a keyboard. For the orientation-matching task, they used a Griffin Powermate dial to register their response.

Figure 2.1B shows the experimental setup for the adults. Observers were seated 50 cm from the monitor. Stimuli were presented monocularly to the dominant eye, with observers wearing an eye-patch covering their non-dominant eye. Eye-dominance was established using the Miles test (Miles, 1928).

2.2.3 Stimuli and Procedures

2.2.3.1 Children

As previously mentioned, each child completed three different tasks. Acuity was measured to determine the minimum target size each child could detect, and crowding was measured to determine the size of the spatial zone of crowding in the amblyopic and typically developing fovea. The orientation-matching task was the central task to the aims of this chapter, as it was used to investigate the perceptual effects of crowding. The acuity and crowding measures were used to determine the size of the target in the orientation-matching task, in order to ensure that it was above any limits of acuity but also within the spatial region over which crowding occurs.

The three tasks involved five video-game characters: Vac-Man (Visual Acuity Man) and four ghosts. Note that our acuity and crowding tasks were taken and adapted from Greenwood et al. (2012). Vac-Man was a circle with a horizontal gap for a “mouth” in its centre, resembling a filled-in Landolt-C. The size of the mouth was equal to one-fifth of the stimulus diameter, similarly to Sloan letters (Sloan, 1959). Vac-Man was the centrally located target stimulus in all three tasks, but also served as flanker and response stimuli in the matching task. The ghost characters acted either as colour aids for the identification of Vac-Man’s orientation in the acuity task, or achromatic flanker stimuli in the extent task (as in Figure 2.1A). The gap for each of the ghosts’ “legs” was also one-fifth of the stimulus diameter.

All children began with the acuity task, followed by the crowding extent. In these tasks, children were asked to report which of the ghosts Vac-Man was facing (four-alternative forced choice, 4AFC). Feedback was given after each trial through brief animations, with Vac-Man smiling with correct responses and frowning with

incorrect. Every three correct responses a longer animation was presented, in which Vac-Man ate a ghost (see Figure 2.1C for illustration of the animation). Children had unlimited time to respond.

Vac-Man was presented at the centre of the screen, and rendered in black at 90% Weber contrast against a mid-grey (45cd/m2) background. In the acuity task, the ghosts, each with a distinct colour (green above, red to the right, orange in the bottom, and green to the left) moved slowly along the monitor edges (as in Figure 2.1A). The ghosts were presented at the edges of the screen, at a large separation from the target, minimising the chance of any interference with the target and thus crowding occurring.

In the crowding extent task, the four ghosts surrounding Vac-Man were abutting and achromatic. As amblyopic crowding depends on target-flanker separation, the close separation between the target and flankers ensured the target was crowded. Although it is unclear whether amblyopic and developing crowding are selective for target-flanker similarity in colour, by making both the target and flankers achromatic it was ensured that if they did, the target would be crowded (Kooi et al., 1994). The flanker ghosts were located above, below, left, and right of Vac-Man, with each ghost randomly oriented either 0°, 90°, 180°, or 270°. Children could look at the picture of the ghosts that were stuck around the screen and either report the colour of the ghost, its location (e.g. top, bottom, left, right), or point to the ghost they thought Vac-Man was facing. Normal colour naming abilities were checked using the stimuli prior to participation.

Acuity thresholds were measured by varying the overall size of Vac-Man, (and thus the visibility of the mouth gap that indicated his orientation) using a QUEST algorithm staircase procedure (Watson & Pelli, 1983) converging at 62.5% correct performance. Correct responses resulted in a decrease in Vac-Man's size and incorrect responses in an increase in size. The spatial extent of crowding was also measured by varying the size of the Vac-Man and ghosts with QUEST converging at 80% correct performance. The higher convergence point in the crowding extent task was chosen in order for the resulting Vac-Man gap threshold to be used as the upper bound when setting the target gap size in the orientation-matching task.

The relative centre-to-centre separation between the Vac-Man target and a ghost was 1.1× target diameter, but by varying the size of the target and flankers, their absolute centre-to-centre separation was also varied simultaneously. Varying

the size of the target and flankers with QUEST in order to measure the spatial extent of crowding differs from the method used by Greenwood et al. (2012), where the size of the target and flankers was fixed and the centre-to-centre separation between them was varied. Varying size while keeping the relative centre-to-centre separation fixed has been proposed as the most efficient method to measure crowding extent (Levi, Song, & Pelli, 2007b; Song et al., 2014). In amblyopic and peripheral vision where crowding is largely not limited by acuity (Hess, Dakin, et al., 2001; Levi, Hariharan, et al., 2002b; Bonneh et al., 2004), this method allows to disentangle the extent of crowding from any limitations in acuity, as it does not depend on a fixed target size. Although this method confounds size and absolute separation by varying both simultaneously, the extent of amblyopic and peripheral crowding are limited by centre-to-centre separation and not target size (Hariharan et al., 2005). As such, only variations in absolute centre-to-centre separation should affect the measurements of the extent of crowding, not the variations in size. Note that the spacing value of 1.1× target diameter was chosen as it been shown to be the optimal value for to measure the extent of crowding without fixing the target size in both typical vision (fovea and periphery) and the amblyopic fovea (Song et al., 2014).

The QUEST routine for both acuity and crowding extent tasks was tailored to suit testing children in three ways. First, in the beginning of the task children were given 3 practice trials with a target mouth-size of twice the acuity guess threshold, which was the LogMAR acuity value measured during orthoptic testing. Second, easier trials with a mouth size twice the current threshold estimate were presented every fifth trial. This minimised the frustration arising from the presentation of numerous trials near threshold. Third, an exit criterion was added to reduce the time taken for threshold estimation: if the standard deviation of the estimated threshold for the last 8 trials was below 0.03 log units, the experimenter was given the option to exit the task. Otherwise, the QUEST terminated after 30 trials were completed for each eye. The average number of trials needed to estimate threshold, excluding practice trials, was 44 for acuity and 46 for crowding extent. Both eyes were tested in one experimental run, with two QUEST staircases, one for each eye, running simultaneously. The output of each QUEST staircase gave the size of Vac-Man's mouth in degrees of visual angle at the predefined threshold.

With the final orientation-matching task the perceptual effects of crowding were investigated. Four achromatic "imposter" Vac-Men were positioned on the

cardinal directions around the “real” target Vac-Man, with both target and flanker VacMen at 90% Weber contrast. A second response Vac-Man was presented on the response screen, twice the size of the target to ensure visibility. The size of the stimuli in this task was determined individually for each child. It was crucial to ensure that Vac-Man was both visible (i.e. above the acuity limit) and crowded (i.e. fell within the interference zone of crowding). A multiple of the mouth acuity threshold was thus used. If 3× the acuity exceeded the maximum centre-to-centre separation for crowding for that child, lower values were used (2.5×, 2×, and 1.5× the acuity threshold). In the group with amblyopia, three children were tested an acuity multiple of 3, seven with 2.5×, five with 2×, and five with 1.5×. In the group with typical vision, one child was tested with an acuity multiple of 2.5, six with 2×, and thirteen with 1.5×.

Children were asked to make the Vac-Man presented on the response screen the “same” as the real Vac-Man presented on the central monitor by adjusting its orientation using the dial. They had unlimited time to respond. The orientation of the target varied randomly between ±45° from vertical and the orientation of the flankers differed from the target either 30° or 90°. The orientation difference from the target could either be positive (i.e. +30° from the target orientation, counter-clockwise rotation), or negative (i.e. -30° from the target orientation, clockwise rotation). This resulted in five flanker conditions: uncrowded (targets presented in isolation), flankers with a +30° difference from the target, flankers with a -30° difference, flankers with a +90° difference, and flankers with a -90° difference. 12 trials were tested for each condition, resulting in 60 trials in total for the orientation matching task. When children’s responses deviated from the orientation of the target by more than ±35°, they received feedback in the form of a frowning Vac-Man, whereas when they responded within that range, Vac-Man smiled. This was done to maintain children’s engagement in the task and reward them for participating.

Only two target-flanker orientation differences were selected to avoid fatigue. These two orientation differences were deemed most informative for the following reasons: first, orientation differences bellow ~90° might not be informative in distinguishing between averaging and substitution errors (Harrison & Bex, 2015), so 90° flanker differences were chosen to help distinguish between the two error types; second, 90° orientation differences have been found to lead to a relief from crowding in the typical visual periphery (Hariharan et al., 2005). By also using 30°

flanker differences that make flankers more similar to the target, selectivity for target-flanker similarity in orientation for amblyopic and developing crowding could be examined by comparing the two flanker conditions.

2.2.3.2 Adults

Adults completed the same three tasks: acuity, crowding extent, and the orientation-matching task. For all tasks, a filled-in Landolt-C identical to Vac-Man was used as the target (as in Figure 2.1B). For the crowding extent task, achromatic ghosts were used as flankers. Stimuli were presented monocularly to the dominant eye at four eccentricities: 2.5°, 5°, 10°, and 15° in the upper visual field.

For the acuity and extent tasks, observers completed two runs of the QUEST staircase for each eccentricity. On each trial, a fixation dot first appeared in the bottom of the screen for 500ms. This was followed by the target, either presented in isolation (acuity task) or surrounded by the ghost flankers (crowding extent task) for 500ms. Then, a 1/f mask of elliptical shape was presented for 250ms to avoid adaptation effects at the target location. A different 1/f mask was presented on each trial. The size of the mask was fixed at 1/3 the target eccentricity. After the presentation of the mask, adult observers had unlimited time to make a response on the target orientation (identically to the kids, 4AFC). Following their response, the inter-trial interval was 500ms during which the fixation dot was on screen. Each staircase consisted of 45 trials, resulting in a total of 90 trials per eccentricity. For each subject, the average threshold value of the gap size from the two thresholds estimated by QUEST for each eccentricity was taken as the acuity and extent values.

For the orientation-matching task, flankers were filled-in Landolt-Cs identical to the target, matching the stimulus configuration in the children's version (Figure 2.1E). A multiple of 3 times the acuity threshold was used as the stimulus size. This ensured that stimuli fell within the crowding zone for all observers (i.e. the target was greater than the acuity threshold but smaller than the crowding extent threshold).

The trial presentation sequence in the orientation-matching task was identical to the acuity and crowding tasks. On each trial, a fixation dot first appeared near the bottom of the screen for 500ms, followed by the target for 500ms. The target was either presented in isolation (uncrowded condition) or surrounded by flankers of a 30° or 90° orientation difference. Then, a 1/f mask of elliptical shape was presented for 250ms. After the presentation of the mask, a reference stimulus

identical to the filled-in Landolt-C target appeared at fixation at a random orientation. The size of the reference target was equal to the target. Adult observers had unlimited time to adjust the reference stimulus to match the orientation of the target previously presented. Following their response, the inter-trial interval was 500ms during which the fixation dot was on screen.

Adults completed 5 blocks of 100 trials per eccentricity, resulting in a total of 500 trials per eccentricity. In each block, 20 trials were included per flanker condition (uncrowded, flankers with -30° or +30° difference from the target, flankers with a -90° or +90° difference from the target). Blocks from different eccentricities were interleaved to counter any practice effects. Observers received auditory feedback in the form of a beep when their estimate of the target orientation was offset by more than ±35°. All other parameters were identical to the children's version of the tasks described above.

2.3 **Results**

In the Results section I first present findings from the acuity and crowding tasks, followed by the orientation-matching task. Note that a different order is followed from the Methods section with regards to the observer groups. Here, I start by presenting the results from the adult observers, followed by the children. I follow this order in order to relate results from the adult periphery from the orientation-matching task to previous findings in the literature on perceptual errors and examine whether they match. Then by presenting the results from the children with typical vision and amblyopia, I can examine whether they make systematic or random errors, and subsequently whether their errors match those made by adults in the periphery.

2.3.1 Acuity and Crowding Extent

The acuity and crowding extent tasks each gave a measure of the gap size of the filled-in Landolt-C (VacMan) once performance reached a particular point (62.5% for acuity; 80% for crowding). For acuity, this was the value of interest. For the crowding extent task, I was interested in the spatial extent of the crowding zone (i.e. the radius from the centre of the target to the centre of one flanker); this corresponded to the absolute centre-to-centre separation between the target and a flanker, equal to the target diameter which was five times the gap size, multiplied by

the relative separation fixed at 1.1 the target size (GapSize × 5 × 1.1). Analyses on these two measures are discussed below.

2.3.1.1 Adults

Acuity values for the four eccentricities tested can be seen in Figure 2.2A. In the adult periphery, acuity thresholds increased with eccentricity, averaging 2.3 arcmins at 2.5° eccentricity, 3.7 arcmins at 5°, 7 arcmins at 10°, and 11.4 arcmins at 15°. A one-way ANOVA revealed a significant effect of eccentricity, ($F[1.47, 13.20] = 110.34$, $P < .0001$, Greenhouse-Geisser corrected), demonstrating the well-known reduction of acuity in the visual periphery.

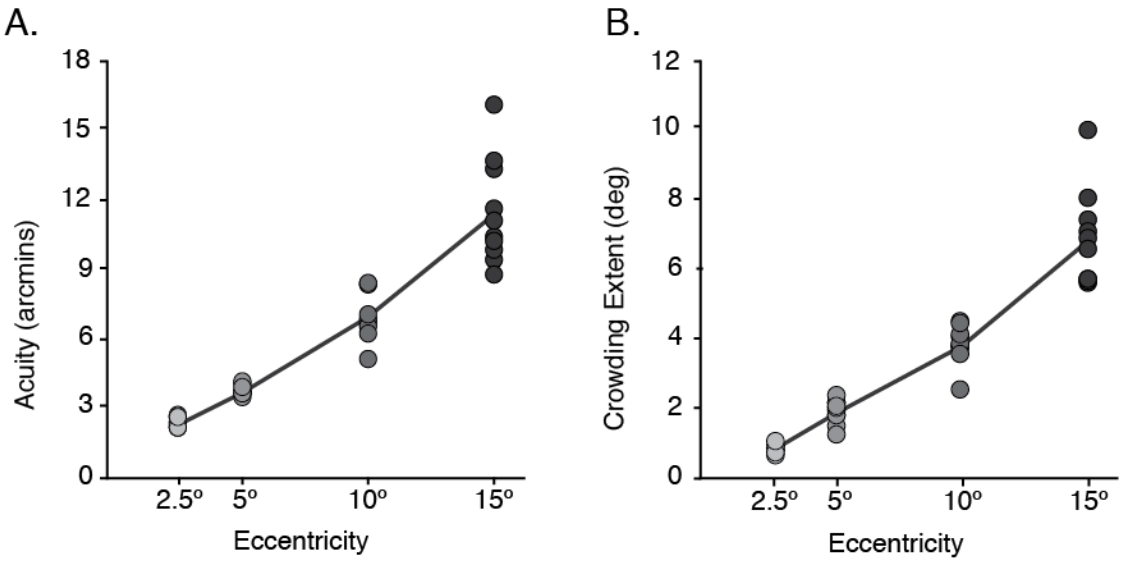

Figure 2-2 Acuity and crowding in the adult periphery

A. Acuity measured as the gap of the filled-in Landolt-C target at 62.5% performance on the QUEST for adult observers in the visual periphery (2.5°, 5°, 10° and 15° eccentricity). Dots indicate the values for each observer and the line shows the mean of the adult sample (N=10). Values are presented in minutes of arc.
B. Extent of crowding measured as the centre-to-centre separation between the target and a flanker at 80% performance on the QUEST for adult observers in the visual periphery (2.5°, 5°, 10° and 15° eccentricity). Plotted with conventions as in A. Values are in degrees of visual angle.

The extent of crowding at each eccentricity is presented in Figure 2.2B. Similarly to the pattern found for acuity, the extent of crowding increases from an average of 0.82° at 2.5°, 1.88° at 5°, 3.81° at 10°, to 6.83° at 15° eccentricity. Note the vast difference in scale between the acuity and crowding extent measures, as acuity is presented in minutes of arc and the extent of crowding in degrees of visual angle. A one-way ANOVA revealed a significant effect of eccentricity, ($F[1.23, 11.07]$

= 146.20, *P*< .0001, Greenhouse-Geisser corrected), showing that the extent of crowding scales with eccentricity in the adult periphery.

2.3.1.2 Children

Acuity values for the children with typical vision and the children with amblyopia can be seen in Figure 2.3A. Acuity thresholds for the left and right eyes of the group with typical vision averaged 1.1 and 0.9 arcmins respectively, both equivalent to Snellen acuity of 6/6. There was no significant difference between these values (paired samples t-test: t[19]=2.37, *P*= .5), indicating no interocular differences in acuity for the group with typical vision. Reduced acuity levels were evident in the affected eye of the amblyopic group, with an average of 4 arcmins, compared with an average acuity of 1.1 arcmin for the unaffected fellow fixating eye (equivalent to Snellen acuities of 6/24 and 6/6). This resulted in a significant difference in acuity between the two eyes (paired samples t-test: t[19] =4.13, *P*< .001), characteristic of amblyopia. Acuity in the unaffected eye did not differ from the acuity of the children with typical vision (unpaired t-test between unaffected eye and both eyes of children with typical vision: t[58]= -1.06, *P*= .29).

Values for the extent of crowding for the group with typical vision and the group with amblyopia can be seen in Figure 2.3B. For the group with typical vision, the extent of crowding averaged 0.17° for the left eye and 0.15° for the right eye, with no significant difference between the eyes (paired samples t-test: t[19]=1.58, *P*= .13). For the amblyopic group, the extent of crowding was greater in the affected eye, averaging 0.79° compared to 0.16° for the fellow fixating eye, resulting in a significant difference between the two eyes (paired samples t-test: t[19]=485, *P*< .001). I found no difference in the extent of crowding between the two eyes of children with typical vision and the fellow fixating eye of the children with amblyopia (unpaired t-test between the FFE and both eyes of children with typical vision: t[58]=-0.18, *P*= .86).

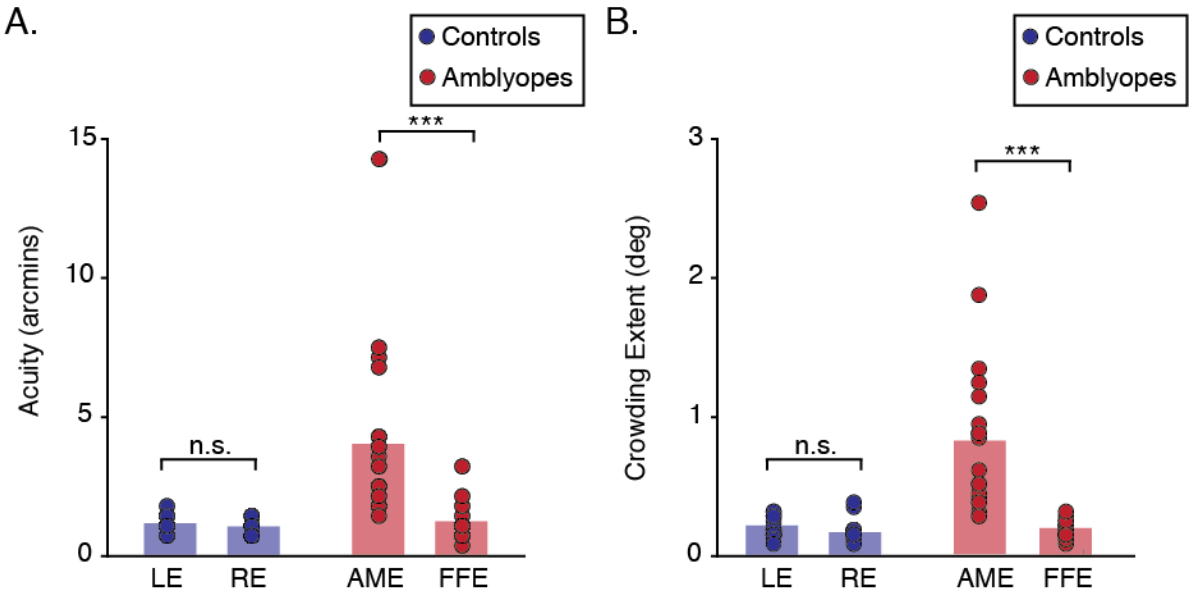

Figure 2-3 Acuity and crowding in the typically developing and amblyopic fovea

A. Acuity measured as the gap of the filled-in Landolt-C target (i.e. VacMan) at 62.5% performance on the QUEST for children with typical vision (N=20) and amblyopia (N=20). Dots indicate the values for each eye and bars indicate the mean. Values are presented in minutes of arc. LE = left eye, RE= right eye, AME= amblyopic eye, FFE = fellow fixating eye. n.s. = no significant difference; ****P*<.001.
B. Extent of crowding measured as the centre-to-centre separation between the target and flankers at 80% performance on the QUEST for children with typical vision (N=20) and amblyopia (N=20). Plotted with conventions as in A. Values are in degrees of visual angle.

When considering which eccentricity in the adult periphery is most similar to vision in the amblyopic eye of children, there appears to a difference between acuity and crowding. Acuity for the amblyopic eye on average (4 arcmins) was most similar to acuity at 5° eccentricity in the adult periphery (3.7 arcmins). On the other hand, the extent of crowding in the amblyopic eye on average (0.79°) was most similar to the extent of crowding at 2.5° eccentricity in the adult visual periphery (0.82°). Therefore, in the context of the typical adult periphery, acuity in the amblyopic fovea would have a greater equivalent eccentricity than crowding, suggestive of a greater disruption.

2.3.2 Orientation-matching

In this section I discuss the results from the orientation-matching task. For this task, responses were recorded as the perceived orientation of the target on each trial. For the adults, responses were amalgamated across the five blocks collected for each eccentricity. For the children this was not required, as only one block of responses was obtained in the orientation-matching task.

Each response was converted into a value of error from the target, with 0° indicating no error. Frequency histograms were constructed to tally the error

number within a range of ±180°, in 10° bins (37 bins in total). The response error was plotted separately for each flanker condition. This resulted in five distributions of response errors per observer. I determined that the pattern of errors in conditions with equivalent target-flanker differences of opposite sign (i.e. -30°and 30°, -90° and 90°) did not differ significantly for each observer. Therefore, I reversed the sign of the response errors in the conditions with negative orientation differences in order to sum the distributions. This resulted in three distributions of response error per observer (and per eccentricity for the adult observers): uncrowded, 30° target-flanker difference, and 90° target-flanker difference.

After visual inspection of the distributions, I determined that their shape approximated a Gaussian distribution. Therefore, after smoothing the distributions by applying a boxcar filter (width 1/3), I fitted unimodal and bimodal Gaussian functions to the uncrowded condition and the 30° and 90° target-flanker orientation difference conditions. The fitted unimodal Gaussian distribution had four parameters: a mean (indicative of the location of the peak of the distribution), a standard deviation (indicative of the bandwidth), a scale (indicative of the height of the distribution), and an offset (indicative of the base height). The fitted bimodal Gaussian had two means and two scales, one for each of the two modes. Below I discuss the fitted parameters that best characterise the pattern of response error in each crowding condition, starting from the adult periphery. Note that only the distribution that best fitted the data for each condition, either the unimodal or the bimodal Gaussian, is presented.

2.3.2.1 Adults

2.3.2.1.1 Group Distributions

Figure 2.4 shows the distributions of the group response error for the adults at the four eccentricities tested in the visual periphery. For uncrowded targets presented at 2.5° eccentricity (Figure 2.4A), the distribution of response errors was unimodal, with the response “error” reported with the highest frequency being 0°, indicating that the majority of the responses had no error. There was also little variability in the spread of the response errors. The pattern of response error for uncrowded targets at the higher eccentricities was identical (Figures 2.4 B-D). The fitted Gaussian unimodal distribution was centred on 0° and it was narrow at 2.5° (M=-0.94°, SD= 15.07°), 5° (M= -0.85°, SD= 14.69°), 10° (M= -0.11°, SD= 14.28°), and 15° eccentricity (M= -0.24°, SD= 14.28°). This indicates that when the target

was presented in isolation, observers reported the orientation of the target with little error and good precision, with increasing eccentricity having no effect on estimates of the orientation of uncrowded targets.

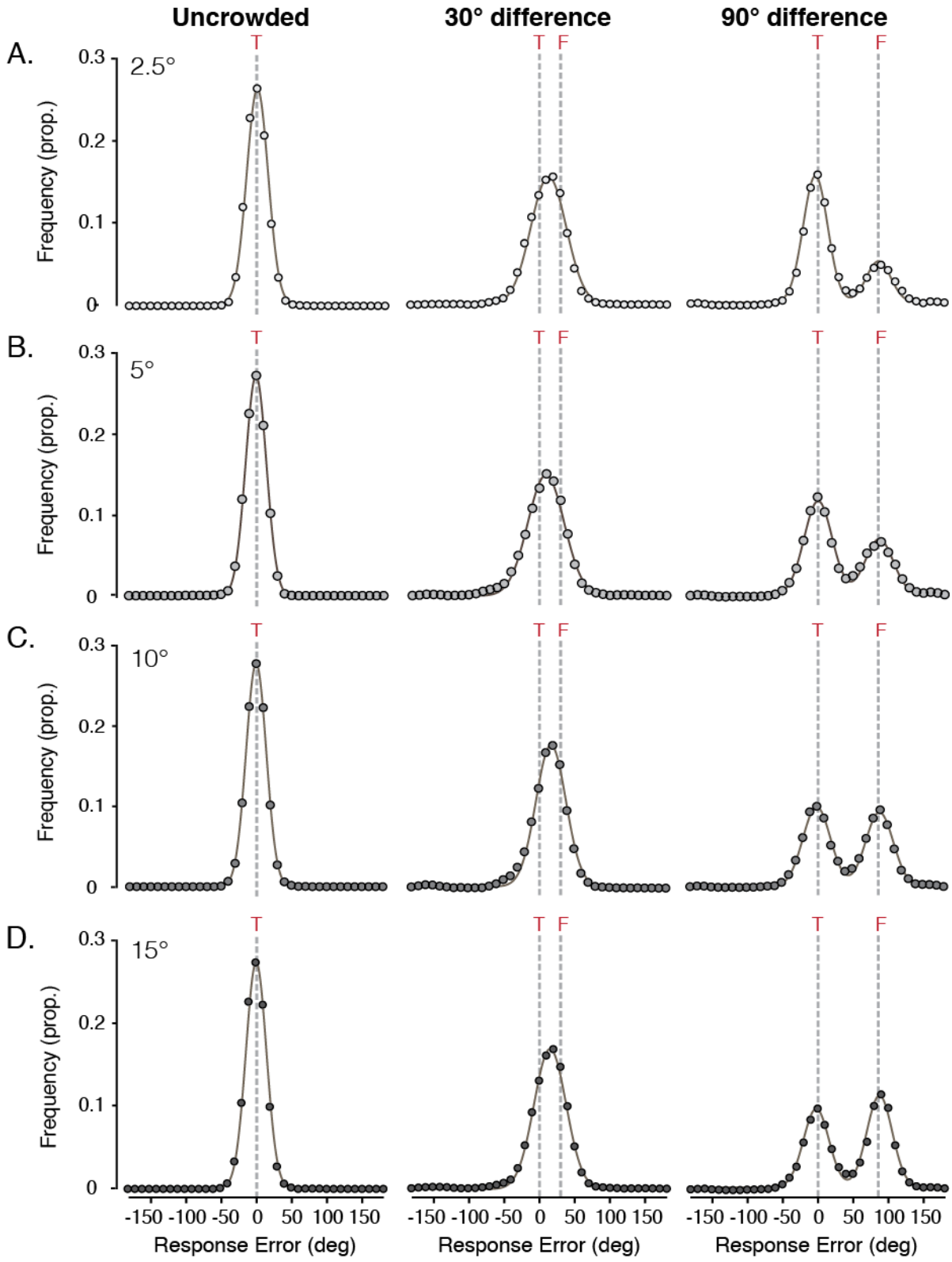

Figure 2-4 Group response error distributions in the adult periphery

A. Group response error distributions for adults (N=10) at 2.5° eccentricity. Dots indicate the proportion of responses for each of the 37 bins of the histogram. Grey lines indicate the target location (T), and when present, the flankers (F).
B.C.D. Group response error distributions for adults (N=10) at 5°, 10°, and 15° eccentricity, plotted with conventions as in A.

As can be seen in Figure 2.4, when flankers of a 30° difference surrounded the target, the distribution of response error was also unimodal at all four eccentricities. However, relative to the uncrowded condition where there was hardly any error on average, here there was an increase in error. The response errors with the greatest frequency were to orientations between the target and flankers, indicating that observers primarily reported intermediate orientations between the target and flankers. This was captured as a shift towards the flankers in the peak of the unimodal Gaussian distribution. This shift was smaller at 2.5° (M= 12.42°) and 5° (M= 9.86°) eccentricity, compared to 10° (M= 16.64°) and 15° (M= 15.8°) eccentricity. In addition to this shift, there was also an increase in the spread of response errors, indicating more variable responses. The addition of the 30° difference flankers thus also led to an increase in the bandwidth of the unimodal distribution relative to the uncrowded condition at 2.5° (SD=25.06°), 5° (SD= 25.65°), 10° (SD= 21.22°), and 15° (M= 22.71°) eccentricity. This increase in response error variability indicates that crowding has a disruptive effect on response precision.

When flankers that differed by 90° surrounded the target, the distribution of response errors was bimodal. The first peak of response errors was concentrated at 0° indicating either no error and thus target responses, whereas the second peak was centred at 90°, corresponding to responses of the flanker orientation. Increasing eccentricity did not have an effect on the location of the two peaks, indicating the observers reported both the target and the flanker orientations at 2.5° (M_1= -1.64°, M_2= 88.38°), 5° (M_1= -0.22°, M_2=86.08°), 10° (M_1= -1.19°, M_2= 87.47°), and 15° eccentricity (M_1= -0.97°, M_2= 89.01°). However, as can be seen in Figure 2.4, the height of the second peak located near 90° gradually increased with eccentricity; an effect clearly demonstrated by the change in the scale of the two peaks with eccentricity. At 2.5° eccentricity (Figure 2.4A), the scale of the first peak centred on the target orientation (S_1= 0.15) was larger than that of the second peak centred on the flanker orientation (S_2= 0.05), indicating more responses with 0° error and thus of the target orientation. This was also the case at 5° eccentricity (Figure 2.4B), although the difference between the two scale values decreased (S_1= 0.12, S_2= 0.06). At 10° eccentricity (Figure 2.4C), the height of the two peaks was near equal as indicated by the similar scale values (S_1= 0.10, S_2= 0.09), suggesting an almost equal number of target and flanker responses. The scale of the second peak (S_2= 0.11) was slightly larger than the first (S_1= 0.09) at 15° eccentricity, and Figure 2.4D shows that flanker responses were frequent than target responses. This

suggests that increasing eccentricity had an effect on the proportion of target to flanker responses. At lower eccentricities, the frequency of target responses was greater, as indicated by the highest of the two peaks being centred on 0°, whereas at higher eccentricities, responses of the flanker orientation became more frequent. Thus, there was an increase of flanker responses with eccentricity. Finally, I note that the bandwidth of the bimodal distribution increased compared to uncrowded for all eccentricities in this condition (2.5°: SD= 17.48°; 5°: SD= 19.79°; 10°: SD= 19.52°; 15°: SD= 18.06°), showing a decrease in response precision. This shows that flankers of a 90° orientation difference increased response variability compared to the uncrowded condition, although not as much as flankers with a 30° difference.

From the group data of the adult observers, a number of conclusions can be drawn on the pattern of response errors, and thus on the perceptual effects of crowding. First, when the target was uncrowded, observers reported the orientation of the target with high accuracy and good precision. Increasing the target eccentricity thus had no effect on the ability to discriminate the orientation of uncrowded targets. Second, with flankers of a 30° difference from the target, the perceptual error increased relative to the uncrowded condition, and observers primarily reported orientations between the target and flankers, with a small shift of the reported orientations towards the flankers with eccentricity. Adding flankers also led to an increase in response variability and precision suffered. Finally, with flankers of a 90° difference, error increased relative to the uncrowded condition, and observers reported either the target or the orientation of the flankers. Response variability increased, but not to the same extent as in the 30° flanker difference condition. Increasing eccentricity reduced accurate target responses and increased flanker responses. These results show that with 30° target-flanker differences, crowding leads primarily to responses of intermediate orientations between the target and flankers, that can be classified as assimilation errors. Crowding with 90° target-flanker differences leads primarily to target responses and responses of the flanker orientation, that can be classified as substitution errors that increased with eccentricity.

2.3.2.1.2 Individual Distributions

In order to examine whether the pattern of response errors on a group level is representative of the underlying individual distributions, I consider two cases of adult observers. Although the types of errors individual observers made with

uncrowded targets matched the group data observed above, the pattern of response errors in the flanker conditions showed some individual variation across eccentricities. Particularly, the rate by which observers shifted to reporting the flanker orientation with increasing eccentricity differed. Here, I present the individual response errors for observer P02 that showed a gradual shift towards flanker responses with eccentricity, and for observer P10 who represents an extreme case, as they showed the most rapid shift towards flanker responses with eccentricity.

Figure 2.5 shows response error distributions from observers P02 and P10. For observer P02 (Figure 2.5A), the fitted Gaussian distribution of response errors to uncrowded targets was unimodal and centred near 0° at 2.5° (M= 1.53°), 5° (M= 0.65°), 10° (M= -0.26°) and 15° eccentricity (M= -0.86°), indicating responses of the target orientation. The bandwidth of this distribution was narrow for all eccentricities (2.5°: SD= 15.01°; 5°: SD= 15.75°; 10°: SD= 15.66°; and 15°: SD= 16.50°), demonstrating that observer P02 showed little variability in their responses to uncrowded targets. As can be seen in Figure 2.5B, observer P10 shows an identical pattern of response errors with uncrowded targets: at all eccentricities, they accurately reported the target orientation with good precision. Therefore, for uncrowded targets increasing eccentricity did not affect response error, similarly to what is observed in the group data in Figure 2.4.

When flankers of a 30° orientation difference surrounded the target, for observer P02 the distribution of response errors was unimodal at all eccentricities. At 2.5° eccentricity (Figure 2.5Ai), the peak was located between 0° and 30° (M= 15.3°), indicating that reports consisted primarily of intermediate orientations between the target and flankers, similarly to the group data for that eccentricity. The peak shifted gradually towards the flankers at 5° (M= 19.93°) and 10° eccentricity (M= 22.29°). At 15° eccentricity (Figure 2.5A iv), the shift of the peak to the flankers was complete (M= 29.09°), indicative of observer P02 primarily reporting the flanker orientation. The bandwidth of the distribution increased compared to the uncrowded condition, but did not differ across eccentricities (2.5°: SD= 22.13°; 5°: SD= 19.26°; 10°: SD= 20.57°; 15°: SD= 18.42°). Therefore, with increasing eccentricity observer P02 shifted their responses from intermediate orientations between the target and flankers to the flanker orientation. Interestingly, for observer P10 this shift from intermediate orientations to the flanker orientation occurred at an earlier eccentricity. At 2.5° eccentricity the peak was centred between the target and flanker orientations (M= 15.60°), but already at 5° eccentricity the peak shifted to 30° (M=

26.64°), similarly to 10° (M= 27.59°) and 15° (M= 27.40°). At 2.5° where the observer was reporting primarily intermediate orientations between the target and flankers, the bandwidth of the distribution was notably larger (SD= 24.01°) than in the higher eccentricities where responses consisted of flanker reports (5°: SD= 17.40°; 10°: SD= 15.98°; 15°: SD= 16.15°).

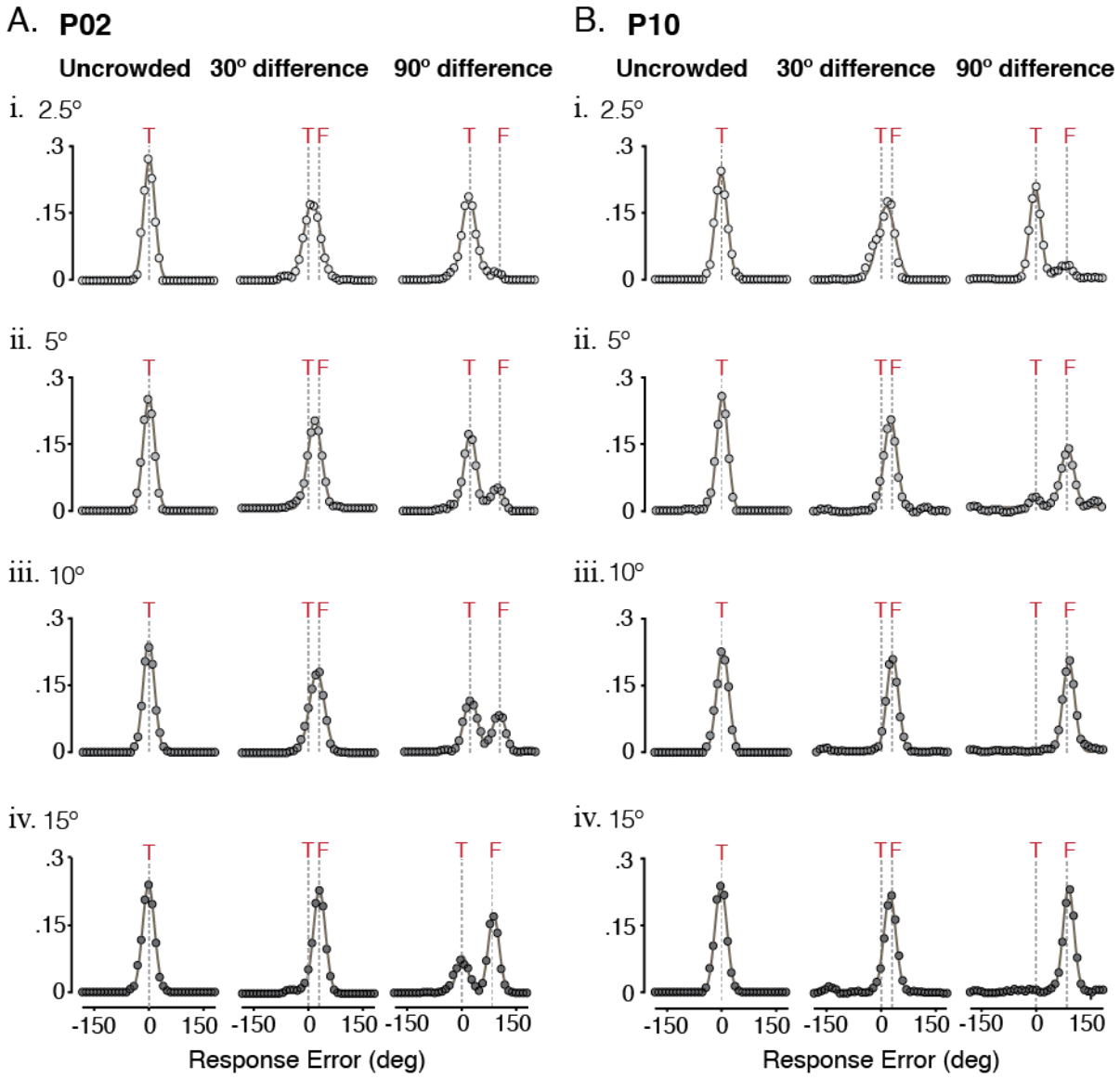

Figure 2-5 Individual error response distributions in the adult periphery

A. Individual response error distributions for adult observer P02 at 2.5°, 5°, 10° and 15° eccentricity depicted at subplots i, ii, iii, and iv respectively. Dots indicate the proportion of responses for each of the 37 bins of the histogram. Each bin has a width of 10°. Grey lines indicate the target location (T), and when present, the flankers (F).
B. Individual response error distributions for adult observer P10, plotted with conventions as in A.

When the flankers differed by 90° from the target, for observer P02 the distribution of response errors was bimodal at all eccentricities. The first peak was

centred near 0° whereas the second peak was located near 90° at 2.5° (M_1=0.34°; M_2=67.43°), 5° (M_1=2.48°; M_2=77.07°), 10° (M_1=1.54°; M_2=79.57°), and 15° (M_1=3.38°; M_2=88.64°) eccentricity. As can be seen in Figure 2.5A, the height of the second peak centred near the flankers gradually became larger as the eccentricity increased, similarly to what was observed in the group data. Interestingly, response variability showed only a small increase compared to uncrowded for 2.5° (SD= 19.29°) and 10° eccentricity (SD = 18.64°), and no increase at 5° (SD= 16.75°) and 15° (SD= 14.91°). As shown in Figure 2.5B, observer P10 showed a similar pattern of response errors to observer P02 at 2.5° and 5° eccentricity. However, similarly to the condition in which the flankers differed by 30°, the rate of increase of flanker responses with eccentricity was greater for this observer. This was particularly evident at 10° and 15° eccentricity, where the distribution that best fitted the data was unimodal. The peak was centred near 90° for 10° (M= 87.65°) and 15° (M= 87.92°), suggesting that observer P10 made exclusively flanker responses. Despite the rapid shift to flanker responses, the bandwidth of the distributions did not change across eccentricities (2.5°: SD= 15.29°; 5°: SD= 18.13°; 10°: SD= 16.23°; 15°: SD= 14.84°).

Overall, the following conclusions can be made based on the examination of the individual data from observers P02 and P10. Similarly to the group data, with uncrowded targets reports of the target orientation were made with high accuracy and good precision, and increasing eccentricity had no effect on response error. With flankers of a 30° difference from the target, reports of the target orientation decreased relative to uncrowded, and as in the group data, the distribution of response errors shifted towards the flanker orientation. However, instead of an intermediate shift to reports between the target and flanker orientations with increasing eccentricity (as in the group data), there was a complete shift to flanker responses. The rate at which this shift occurred was different in the two observers, occurring at 15° for observer P02, and at 5° eccentricity for the extreme case of observer P10. With flankers of a 90° difference, observer P02 matched the response error pattern observed in the group data: they reported either the target or the flanker orientation. Their responses of the flanker increased with eccentricity and only overtook target responses at 15° eccentricity. On the other hand, observer P10 showed an extreme case of this pattern of increasing flanker reports with eccentricity observed in the group data: flanker responses overtook target responses at 5° eccentricity, and at greater eccentricities they primarily reported the

flankers, with no target responses. Taken together, these individual response error distributions indicate that similarly to the group data, flanker responses, and thus substitution errors, increase with eccentricity, but the rate of increase differs across observers.

2.3.2.2 Children

Here I consider the group results from the orientation-matching task for the children with typical vision and strabismic amblyopia, in order to assess whether they made the same types of errors as adults in the peripheral visual field.

2.3.2.2.1 Group Distributions

Figure 2.6 shows the distributions of response errors with their fitted Gaussian distributions for the children with typical vision and amblyopia. For the children with typical vision (Figure 2.6A), when the target was uncrowded the unimodal distribution was centred near 0° (M=1.34°) and was relatively narrow (SD=20.18°), as compared with the adults in the periphery whose distributions were very narrow (SD ~15° across eccentricities). The group with amblyopia (Figure 2.6B) showed an almost identical pattern of response errors, with the peak centred on 0° (M=-0.56°) and a relatively narrow bandwidth (SD=19.38°). This suggests that the two groups did not differ in their responses to uncrowded targets, and both accurately reported the target orientation with good precision.

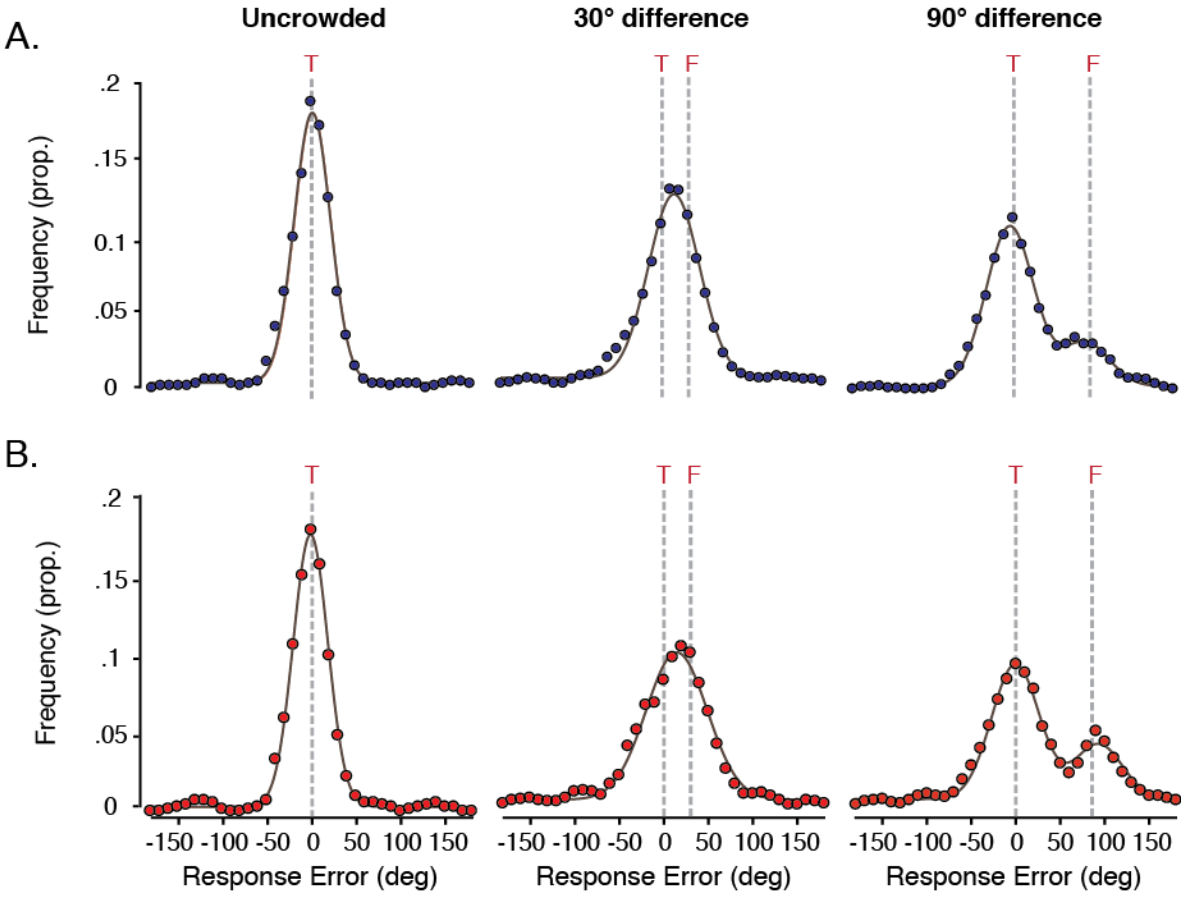

Figure 2-6 Group response error distributions in the typically developing and amblyopic fovea

A. Group response error distributions for the group with children with typical vision (N=20). Dots indicate the proportion of responses for each of the 37 bins of the histogram. Each bin has a width of 10°. Grey lines indicate the target location (T), and when present, the flankers (F).
B. Group response error distributions for the group with amblyopia (N=20), plotted with conventions as in A.

When the flankers differed from the target by 30°, the peak was shifted towards the flanker orientation for both groups. For the group with children with typical vision, the peak of the distribution was centred between 0° and 30° (M= 15.43°). As is evident from the wider distributions in Figure 2.6A, there was an increase in response error variability compared to the uncrowded condition (SD=28.06°). The amblyopic group also showed a shift of the distribution from 0° towards the flankers (M=14.73°), and a larger increase in bandwidth compared to the typical group (SD=34.25°). Therefore, with flankers of a 30° difference children from both groups primarily reported intermediate orientations between the target and flankers and their response precision was reduced. This response pattern is very similar to what was observed for this condition in the near adult periphery, particularly at 2.5° eccentricity (see Figure 2.4A).

When the flankers differed from the target by 90°, the distribution of response errors for both groups of children was bimodal. Similarly to the adults in the near periphery (Figure 2.4 A&B), children in this condition primarily made responses with little error indicative of reports of the target orientation, and with a lesser frequency responses with a 90° error indicative of flanker reports. For the group of children with typical vision, the first peak of the bimodal distribution was near 0° (M= -3.76°), whereas the second peak was near 90° (M= 81.38°). The height of the first peak (S_1=0.10) was greater than the second (S_2= 0.03), indicating that the target orientation was reported more often than the flankers. As in the 30° flanker difference condition, the bandwidth of the distribution increased compared to the uncrowded condition (SD = 27.48°). The pattern of response error for the children with amblyopia was very similar. The first peak of the distribution centred on 0° (M= -0.57°) and the second peak near 90° (M= 92.70°), and similarly to the children with unaffected vision, the height of the first peak (S_1=0.09) was greater than the second (S_2=0.04). There was also an increase in the bandwidth of the distribution (SD= 26.90°), but to a lesser degree than when the flankers differed by 30°. Note that this was the case in the adult periphery too, with 90° flankers having less of an effect on response precision than flankers with a 30° difference. Hence, with flankers of a 90° difference, the groups of children with unaffected vision and amblyopia primarily reported the orientation of the target, and with less frequency the orientation of the flankers, also showing a reduction in response precision compared to the uncrowded condition.

From the group data of children with typical vision and amblyopia, a number of conclusions can be drawn. First, when the target was uncrowded, children with typical vision and amblyopia reported the orientation of the target with relatively high precision, when compared to responses of adult observers. When flankers differed by 30° from the target, children from both groups primarily reported orientations between the target and flankers. Response variability increased compared to the uncrowded condition, and as such responses were less precise. When flankers differed by 90°, the majority of children's responses were of the target orientation, and they also reported the flanker orientation, but with less frequency. Response variability increased relative to the uncrowded condition, but for the group with amblyopia not to the same extent as in the 30° flanker difference condition. Taken together, with 30° target-flanker differences, the errors children made were reports of intermediate orientations between the target and flankers, and as such can be

classified as assimilation errors, similarly to the group responses of adults in this condition. With 90° target-flanker differences, children primarily reported the target, and less often the flanker, in responses that can be taken as substitution errors, matching the group response error pattern of adults at 2.5° eccentricity in this condition. Therefore, these results suggest that on a group level children with unaffected vision and amblyopia make the same systematic errors as adults in the peripheral visual field.

2.3.2.2.2 Individual Distributions

I now consider individual response error distributions in order to determine whether the systematic effects of crowding observed at a group level can also be seen in individuals. Figure 2.7 shows the response errors in the three conditions from example observers in the groups with typical vision and amblyopia. For example, observer C06 from the group with typical vision (Figure 2.7A) shows a very similar pattern of response error to what was observed at a group level, albeit with a complete shift of the distribution to the flankers instead of a partial shift towards intermediate orientations in the 30° flanker difference condition. This is also the case for observer A10 from the group with amblyopia (Figure 2.7B). Note that the pattern of response error for these two children is similar to that of observer P02 at 10° eccentricity (see Figure 2.5iii). Similarly to these two examples, most children with unaffected vision (n= 12/20) and amblyopia (n= 9/20) make the same systematic errors observed at the group level. Below I discuss individual cases that deviate from the group pattern of response error. These can be classified in the following categories: flanker responses, target responses, and responses with increased variability.

In the first response category, children who primarily reported the orientation of the flankers instead of the target are included. From the subset of children with typical vision (n= 2/20) I take the example observer C07 (Figure 2.7C). With uncrowded targets, the distribution was unimodal, centred near 0° (M =-2.15°), and had a very narrow bandwidth (SD = 15.13°), indicating precise reports of the target orientation. With flankers of a 30° difference, the shift of the peak of the unimodal distribution towards 30° was greater than in the group data (M = 24.57°), and the bandwidth substantially increased compared to uncrowded (SD = 33.31°). When the flankers differed by 90°, the distribution was bimodal, with the first peak indicating reports of orientations proximal to the target (M= 21.9°), and the second indicating responses near the flanker orientation (M = 92.08°). The second peak corresponding

to the flanker orientation was greater (S_2= 0.11) than the first corresponding to the target (S_1= 0.05). This shows that the majority of responses were of the flanker orientation instead of target responses as in the group data. The bandwidth of the distribution increased compared to uncrowded (SD = 21.56°). Although noisier, this response pattern is similar to what is observed at 15° eccentricity in the periphery (Figure 2.4D).

Reporting the flanker orientation instead of the target was also evident in a subset of children with amblyopia (n= 3/20). Figure 2.7 D shows the response errors for observer A16. With uncrowded targets, observer A16 reported the orientations near the target (M = 8.36°) with relatively good precision (SD = 24.25°). With flankers of a 30° difference, the distribution was bimodal with the first larger peak located near 30° (M= 27.98°), and a second peak that was not associated with either the target or the flanker orientations (M = 107.52°), suggesting a small number of random responses. The height of the first peak (S_1=0.13) was greater than the second (S_2= 0.03), indicating that observer A16 reported primarily reported orientations near the flankers. With flankers of a 90° difference, the distribution was unimodal with its peak centred near 90° (M= 103.18°). Interestingly, the bandwidth remained comparable to that for uncrowded targets for both the 30° (SD = 22.03°) and 90° (SD= 21.94°) flanker difference conditions. However, responses were noisier in all conditions, with errors at the tails of the distributions that did not correspond to the target, the flanker, or intermediate orientations. This response pattern is very similar to that of adult observer P10 at 10° and 15° eccentricity (see Figure 2.5 B iii and iv).

In the second category, I include children that make errors indicative of responses near the target orientation in all conditions. For the subset of children with typical vision that can show this response pattern (n= 5/20), observer C12 is taken as an example. As can be seen in the left panel of Figure 2.7E, in the uncrowded condition average error was with little variability, as indicated by a peak of the distribution near to 0° (M= 7.29°) and the narrow bandwidth (SD = 17.99°). With flankers that differed by 30° from the target, the distribution did not shift towards the flankers as in the group data (M= 7.17°), but the bandwidth increased compared to the uncrowded condition (SD = 21.58°). With flankers of a 90° difference, the peak was centred on 0° (M =0.73°), with no secondary peak at 90°, but a bandwidth similar to the uncrowded distribution (SD=17.29°). In these cases, despite the addition of flankers, children primarily reported orientations near the

target. Flankers with a 30° difference from the target disrupted precision but 90° flankers did not.

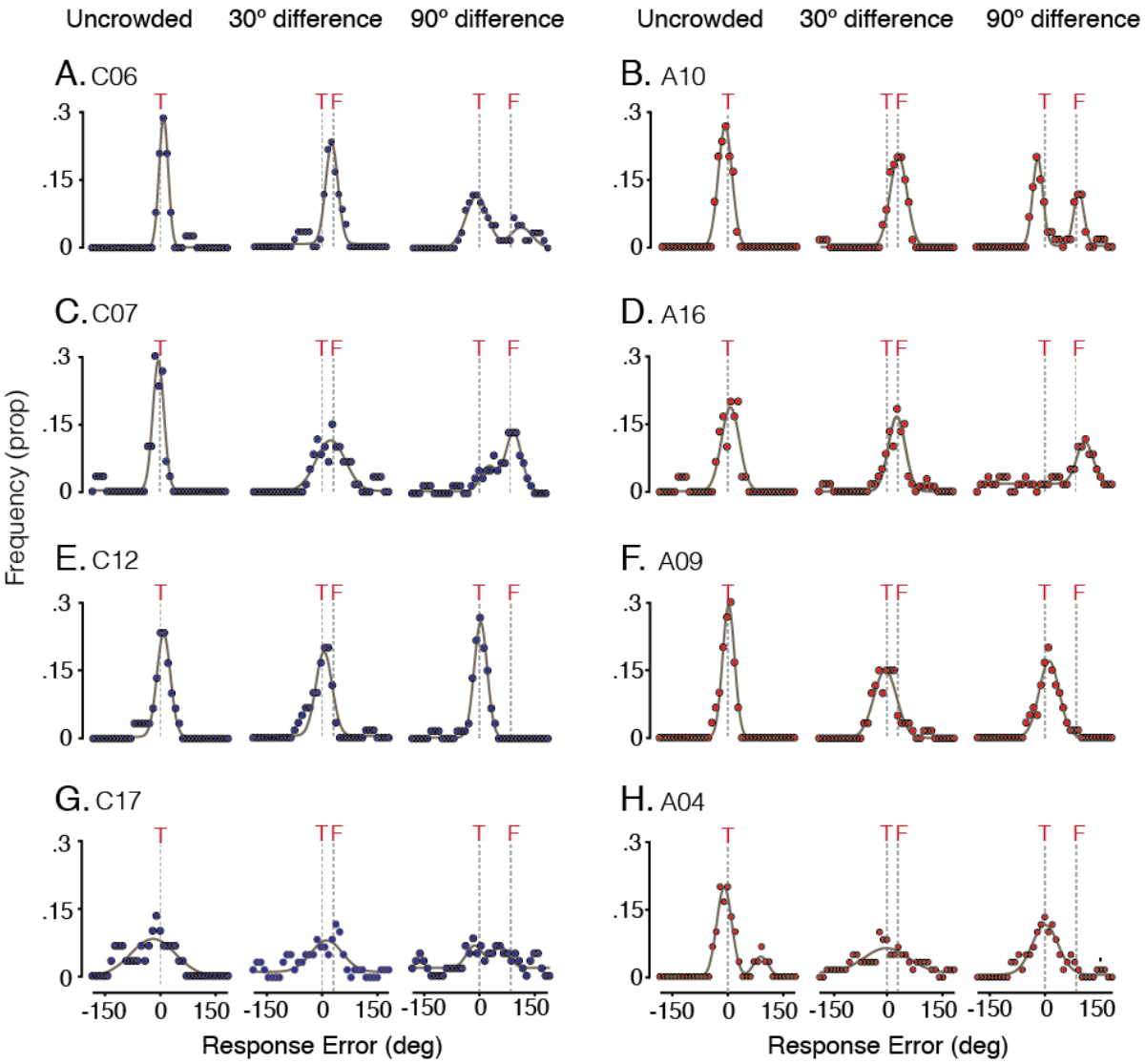

Figure 2-7 Individual response error distributions in the typically developing and amblyopic fovea

A. Individual response error distributions for one child with typical vision (C06) in the three flanker conditions (uncrowded, flankers with a 30° orientation difference, and flankers with a 90° difference). Dots indicate the proportion of responses for each of the 37 bins of the histogram. Each bin has a width of 10°. Grey lines indicate the target location (T), and when present, the flankers (F).
B – H. Individual response error distributions for children in the group with typical vision (blue) and group with amblyopia (red). Plotted with the same conventions as in A.

In the group with amblyopia, there was also a subset of children who made errors indicative of reports of the target orientation under crowded conditions (n= 4/20). However, they showed an important difference from the individuals in the control group: the addition of flankers made their responses significantly more variable in both crowding conditions. This is evident in the response error

distributions of observer A09 (Figure 2.7F). For uncrowded targets, performance showed little error and variability, with the peak of the distribution near 0° (M= 3.42°) and a very narrow bandwidth (SD = 16.11°). However, with flankers of a 30° difference, although the peak remained near 0° (M= -4.44°), the bandwidth of the distribution nearly doubled compared to uncrowded (SD = 30.36°). This was also the case with flankers of a 90° difference, where the peak remained close to 0° (M= 11.17°) and the bandwidth increased substantially (SD = 26.86°). As such, in these cases of children with amblyopia, crowding might not shift responses towards the flankers, but it significantly affects the precision by which the target orientation is reported by increasing response variability.

In the third category, children with substantial response variability are included. Only one child with typical vision showed this response pattern – observer C17, whose response distributions can be seen in Figure 2.7G. For uncrowded targets, response error was highly variable. This is clear from the bandwidth of the distribution that was very wide (SD= 58.96°), and centred relatively near the target (M= -17.59). This was also the case with flankers of a 30° difference from the target, with the unimodal distribution centred between the target and flankers (M= 12.49) but with a very broad bandwidth (SD= 43.34°). With flankers of a 90° difference the distribution was also very broad (SD = 54.32). Due to this variability, and the lack a substantial proportion of responses at the peaks, the mean fitted values are not necessarily indicative of the response pattern. In the case of observer C17, the increased response variability cannot be attributed to the addition of the flankers, and thus to crowding, as it is evident in the uncrowded condition too.

For the children with amblyopia in the third category (n= 4/20), this increased response variability was evident in the conditions with flankers, but not for uncrowded targets. Figure 2.7H shows the response error distributions for observer A04, who shows this response pattern. With uncrowded targets, the distribution that best fit the data was bimodal. The first peak was centred near 0° (M_1= -7.47°), indicating target responses, and the second peak did not correspond to either the target or the flanker orientation (M_2= 123.42°). The non-target responses were substantially less frequent, as the scales indicate (S_1= 0.05, S_2=0.0007). Response variability in this condition was good, and generally similar to that of other children previously discussed. With flankers of a 30° orientation difference, the peak of the unimodal distribution was centred close to 0° (M= -2.47°), suggesting target responses on average. Characteristically for this response category, there was a

significant increase in response variability (SD= 59.37°), to an extent that was not found for children in the other two categories, or the group data. With flankers of a 90° difference, on average participant A04 reported the target (M= 0.12°). Response variability significantly increased compared to uncrowded targets (SD= 38.01°), although as was characteristic for this condition, not to the same extent as with the 30° flanker difference. Therefore, in this subset of children with amblyopia, the addition of flankers resulted in very noisy distributions of target responses.

Overall, most children in both the group with typical vision and amblyopia made response errors similar to the pattern observed in the group response distributions, and thus also the near adult periphery. However, there were also children that deviated from that pattern, that were classified in three categories. In the first category, children made primarily flanker responses with both 30° and 90° flanker orientation differences. This response pattern was similar to the group data of adults at 15° eccentricity and the adult observer P10, who primarily substituted the target orientation with the flanker orientation. In the remaining two categories, the response pattern was distinct to children and was not observed in the adult periphery. In the second category, children reported orientations near the target under conditions with flankers, indicative of target responses, but their response variability increased. In the third response category, one child with typical vision made highly variable responses for both uncrowded and crowded targets, indicative of a general difficulty with response precision. For children with amblyopia in this category, increased response variability was observed only when flankers surrounded the target, suggesting an increase in noise due to crowding. Therefore, although the majority of children with typical vision and amblyopia make the same perceptual errors as adults in the typical periphery, there is also a subset of children in both groups with distinct perceptual errors.

2.3.3 Modelling

In the above section, the errors children with typical vision and amblyopia make were examined, and for the majority of children were found to systematic and match the errors adult observers in the typical periphery made. Therefore, I devised a computational model to simulate the perceptual errors in typically developing, amblyopic and peripheral vision. Note that two alternative models with a distinct crowding stage were also tested, but did not capture the data as well as the final

model presented here. Details on these alternative models are presented in Appendix A (sections 6.1.2 and 6.1.3).

Following traditional pooling accounts (Parkes et al., 2001; Greenwood et al., 2009), I assumed that crowding results from the integration of the target and flanker features. Pooling accounts argue that after a first stage where the features in the target and flankers are detected, these features are pooled (Pelli et al., 2004). The extent of crowding is taken as indicative of the area of visual space over which the features in the target and flankers are pooled, and thus the size of the "integration fields" (Pelli & Tillman, 2008). In the adult typical visual system the extent of crowding, and thus the size of the integration fields, increases with eccentricity (Bouma, 1970; Toet & Levi, 1992). This increase has been argued to arise due to cortical undersampling of the visual periphery (Parkes et al., 2001), with neurons having large receptive fields (Van Essen et al., 1984) in order to ensure adequate coverage of the visual field. Although the model in this chapter is agnostic with regards to the neural site of the pooling process, I nonetheless sought to implement pooling in a neurophysiologically plausible way.

My approach to modelling pooling was inspired by the models by Harrison and Bex (2015) and van den Berg et al. (2010). I first simulated a population of detectors selective for orientation, based on the well-documented orientation selectivity of neurons in V1 (Schiller et al., 1976). I assumed that each detector is sensitive to a range of orientations, with a Gaussian tuning function and a peak sensitivity centred on a particular orientation, and lesser sensitivity to nearby ones. Based on the principles of population coding (Pouget et al., 2000), the population activity distribution is a Gaussian function centred on the orientation of the Landolt-C stimulus, with a bandwidth equivalent to the underlying sensitivity bandwidth of the detectors. The perceived orientation is then read out from the peak of the response distribution.

For the ease of modelling, instead of simulating the absolute perceived orientation and then subtracting it from the target orientation to obtain the orientation difference from the target, I directly simulated this difference. As in the results from the orientation matching task presented above in 2.3.2, 0° was indicative of no error. In order to cover the entire range of orientation differences from the target in the experiment (-180° to 180°), I included one detector with a peak sensitivity for each integer orientation within this range, resulting in 361 detectors. Given the relationship between the sensitivity of the underlying detectors and the

response of the population, I generated the population response as a Gaussian function with a base value of 0 and a peak of 1.

The model involved three distinct stages that are summarised in Figure 2.8. Section A shows example stimuli for each flanker condition that was inputted in the model. In order to fit the sensitivity bandwidth of the underlying detectors, I used one free parameter for the standard deviation. At the first stage of the model (Figure 8 section B), I generated the response of the population of detectors to the stimuli:

$$y = \theta + \alpha\sigma_e$$

Where θ represented the target orientation θ_t, σ_e was the early noise and α was the magnitude of this noise, and the second free parameter of the model. Early Gaussian noise was applied to the population response to simulate the known response variability of cortical neurons (Tolhurst, Movshon, & Dean, 1983). On a single trial when the input to the model was a target presented in isolation (uncrowded), the population response to the orientation of the target was centred near 0°. An example uncrowded trial is shown in section B of Figure 2.8 (top panel). For illustration purposes, in this example trial the bandwidth of the underlying distribution is set to 20°, and the early noise is .2.

In the 30° and 90° flanker difference conditions, on a single trial the population responded to both the target θ_t and the flanker orientation θ_f. The population response to the flanker orientation was centred either near 30° or 90°, depending on the flanker condition. Figure 2.8 shows an example trial for the two flanker conditions (middle and bottom panels), with the same bandwidth and early noise values as the example uncrowded trial. This stage of the model could be considered as representative of the feature-detection stage outlined in many models of crowding.

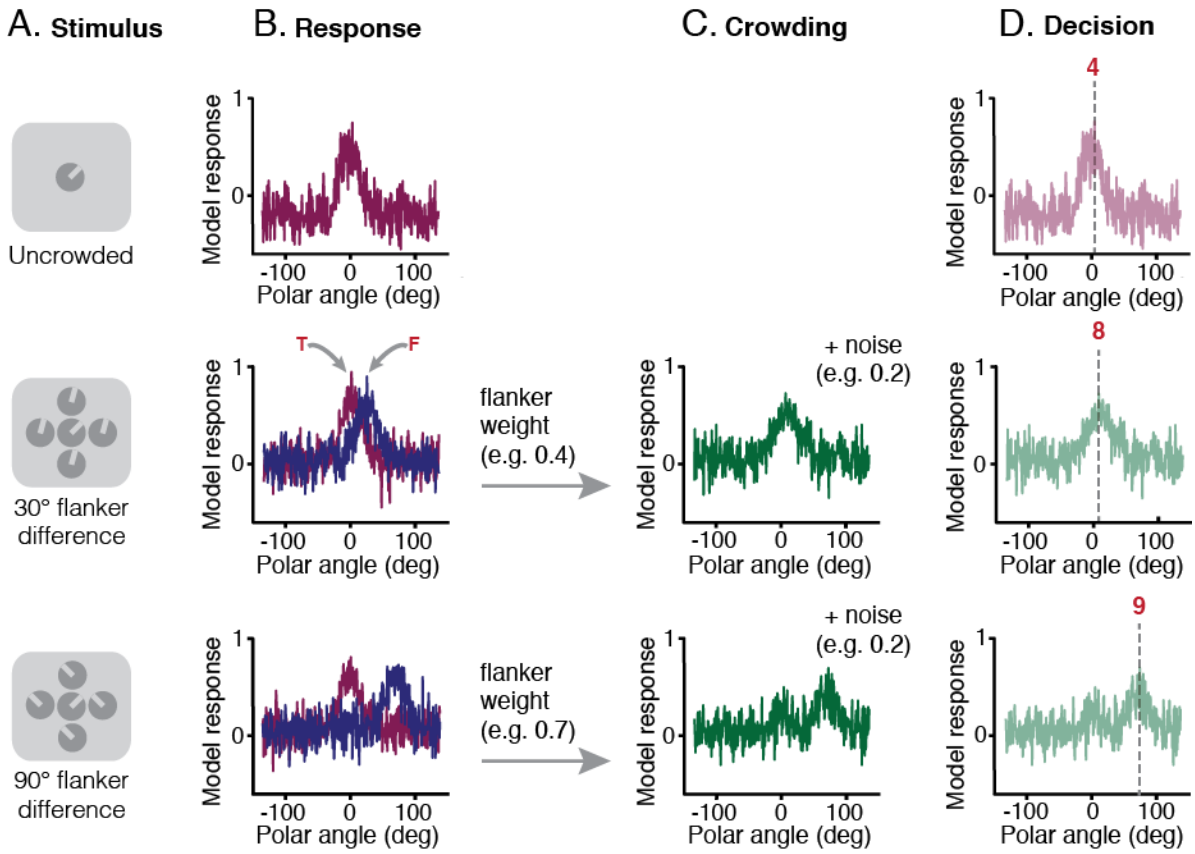

Figure 2-8 Illustration of the stages of the weighted population response pooling model

A. Example stimulus input for each of the 3 flanker conditions (uncrowded, flankers with a 30° difference from the target, flankers with a 90° difference).
B. Early response to the target (upper panel), and the target and flankers (middle and bottom panels). Arrows indicate the response to the target orientation ('T') and the response to the orientation of the flankers ('F').
C. Depiction of the crowding stage, modelled as the weighted combination (pooling) of the population responses to the target and flankers. Note that at this stage crowding noise is added.
D. Depiction of the decision stage, where the perceptual outcome is read as the peak of the combined population response to the target and flankers. The grey line indicates this peak, and the decision value on the x-axis is indicated in red.

At the second stage of the model, I simulated the effects of crowding on perceived target orientation. To achieve this, I followed other recent models that depict crowding as a pooling process resulting from the combination of population responses to the target and flanker elements (van den Berg et al., 2010; Harrison & Bex, 2015). I took a weighted average of the summed population response to the target and flankers that permitted me to modulate the precise combination of these population responses. The weighted combination of responses to the target and flankers was:

$$y_c = (y_t w_t + y_f w_f) + \beta\sigma_c$$

Where w_t and w_f were the weights for the population responses to the target and flankers, respectively. The flanker weight ranged from 0-1, with the weight of the

target being equal to one minus the flanker weight value. The flanker weight parameter was of particular interest, as it determined the magnitude of the contribution of the flanker response to the pooled response, and thus whether or not crowding occurred. As the results discussed above in section 2.2.2 indicated that the pattern of response error in the 30° flanker difference condition was not identical to the pattern of response error in the 90° flanker difference condition, the flanker weight was independent in these conditions. This resulted in two additional free parameters, the flanker weight for the 30° flanker difference condition ($w_{f_{30°}}$) and the flanker weight for the 90° flanker difference condition ($w_{f_{90°}}$).

Additional crowding noise σ_c of a magnitude β was also added at this combinational stage, and β was the fifth and final free parameter of the model. The inspiration for the inclusion of this parameter was the noisy response error pattern of children that showed increased variability (section 2.3.2.2.2). However, in Appendix A section 6.1.3.2.2 I show that this parameter was required not only to simulate the pattern of response error of children with typical vision and amblyopia, but also of adults in the visual periphery.

Figure 2.8 section C shows the second stage for each flanker condition on an example single trial. With a 30° target-flanker difference, the flanker weight is relatively low (e.g. 0.4), so the combined population response distribution remains centred closer to 0° than 30°. In contrast, with a 90° target-flanker difference, the flanker weight is higher (e.g. 0.7), and the combined population response distribution is centred closer to 90°. Note that the combined population response in both flanker conditions is flatter than the individual population responses to either the target or the flanker. This is not only due to the combination of the target and flanker population responses, but also due to the addition of crowding noise σ_c (in Figure 2.8 crowding noise is of a magnitude of 0.2).

In the third and final stage of the model, a “decision” on the perceived target orientation was made by extracting the peak population response on each trial as the maximum point in the response distribution. As can be seen in Figure 2.8 (section D, top panel), for uncrowded targets, the population response to the target is carried through the final stage, and the peak of the response is near the target (e.g. 4°). For the 30° and 90° flanker offset conditions, the peak of the combined target and flanker population responses was taken (Figure 2.8, section D middle and bottom panel). The relatively low flanker weight for the 30° flanker difference condition in this example trial results in a peak near the target (e.g. 8°), whereas the

higher flanker weight in the 90° flanker offset condition results in a peak closer to the flankers (e.g. 96°).

One model simulation included 1000 trials per flanker condition. I determined the best fitting parameters for the group and individual response distributions of children with unaffected vision and amblyopia, and adults using a two-stage coarse-to-fine fitting procedure. The coarse fit (first stage) involved a grid search that provided the parameters that best fit the data in the grid. In the fine fit (second stage), the best parameters from the coarse fit were inputted, and the best-fitting parameters were determined by minimising the least squared error (LSE) between the response distributions and the simulated response distributions over 1000 trials. This procedure was used as it minimises processing time and increases the likelihood of finding a global minimum. We then ran 1000 iterations of the model with the best-fitting parameters for each dataset.

2.3.3.1 Model Simulations of Group Data

2.3.3.1.1 Adults

Figure 2.9 shows the results after 1000 iterations of the model with the best fitting parameters for the adult group data at 2.5°, 5°, 10° and 15° eccentricity. For uncrowded targets, the response error distributions were almost identical across eccentricities, and the model almost perfectly captured the uncrowded data. For uncrowded targets, two free parameters were used in the model: the bandwidth and the early noise. The bandwidth of the underlying detectors was 30.03° at 2.5°, 33.3° at 5°, 28.16° at 10°, and 32.37° at 15° eccentricity, and the values for the early noise were .38 at 2.5° and 5° eccentricity, and .40 at 10° and 15° eccentricity.

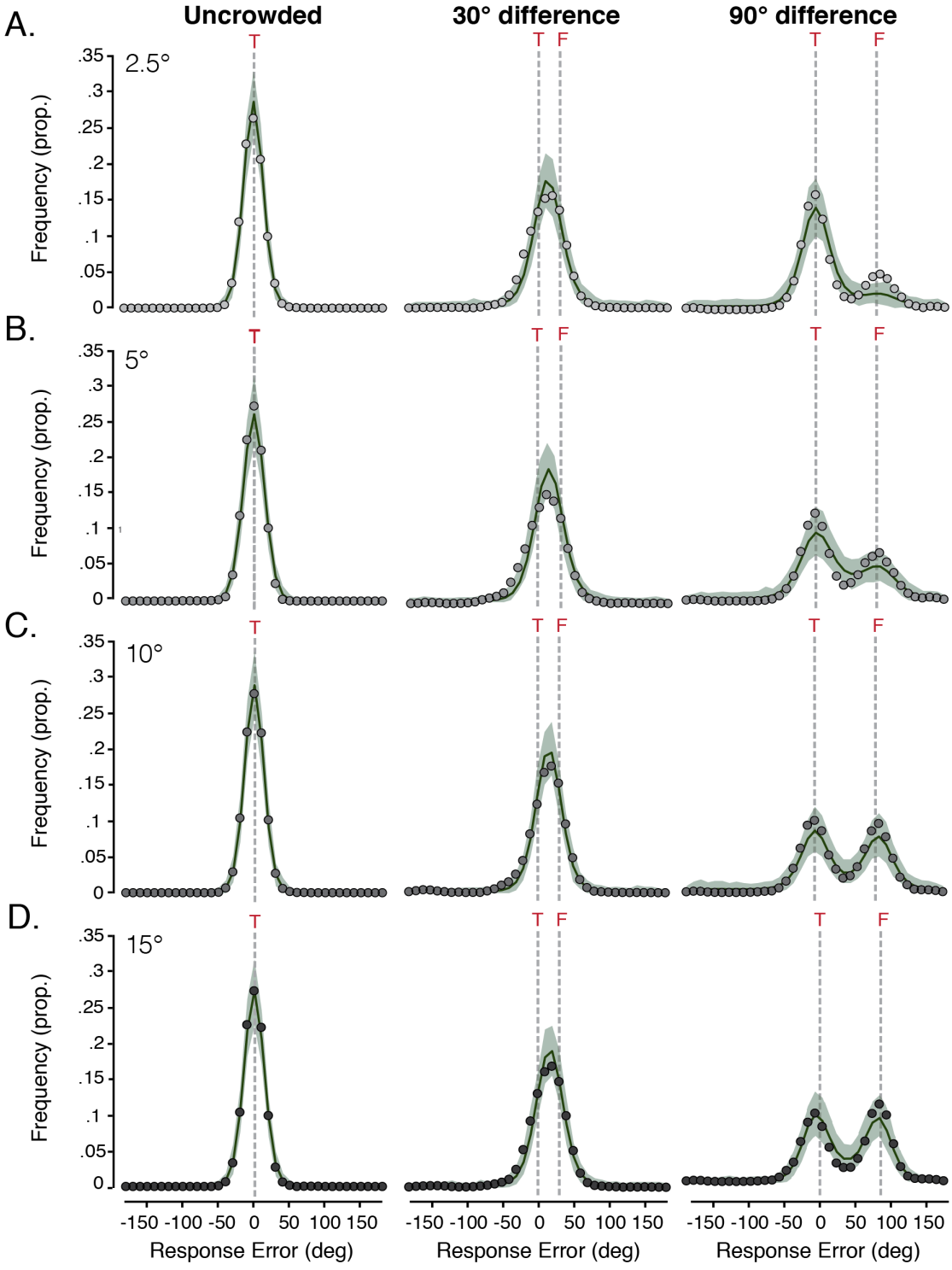

Figure 2-9 Model simulations on the group distributions from the adult periphery

A. Model simulations for the response error distributions at 2.5° eccentricity. The dark green line indicates the mean distribution of the population response pooling model. The light green shaded areas represent the range of simulated distributions for 1000 model iterations. The group response error distribution from the orientation-matching task is presented by the dots. The grey line indicates the target location ('T'), and for the two flanker conditions in which flankers were present, the flanker location ('F').
B-D. Model simulations for the response error distributions at 5°-15°, plotted in the same conventions as in A.

When the target was crowded, due to the combination of the population responses to the target and flankers and the addition of crowding noise, the bandwidth of the simulated response distributions increased and the peaks were lower. The distributions were wider and flatter in the conditions with flankers for all eccentricities, with no differences in the periphery. This was captured by the crowding noise that was applied to the combined population response distribution to the target and flankers and was similar across eccentricities: .50 at 2.5°, .43 at 5°, .41 at 10°, and .42 at 15° eccentricity. As such, there was no increase in crowding noise with eccentricity. The differences between eccentricities were thus captured by the flanker weight parameters. With flankers of a 30° difference from the target, at 2.5° eccentricity the flanker weight was .47, mirroring the response error distribution in Figure 2.9A that clearly shows that the majority of the errors were between the target and flanker orientations. There was a tendency for this shift towards the flankers to increase with eccentricity, and this was captured by the model as an increase in the flanker weights at 5°, 10° and 15° eccentricity, that were .42, .53 and .55, respectively.

With flankers of a 90° difference from the target, there was a marked effect of eccentricity, with the proportion of flanker responses increasing with eccentricity, as discussed in section 2.3.2.1.1. This was captured well by the model: the model followed the pattern of increased flanker responses with eccentricity seen in the data, and the flanker weights increased from .28 at 2.5°, to .43 at 5°, .49 at 10°, and finally .51 at 15° eccentricity. It is worth mentioning that the model undershoots the second peak at 2.5° eccentricity. It appears that this undershooting is specific to cases in which the second peak is low, as it is not evident at any other eccentricities. As the crowding noise parameter flattens the population response distributions to crowded targets, this “flattening” is manifested more clearly when the flanker weight is lower in this flanker condition. The overall success of the model in capturing the group response error distributions from the adult periphery is demonstrated with the low LSE values at all eccentricities, that were .0026 for 2.5°, .0027 for 5°, .0014 for 10°, and .0024 for 15° eccentricity.

The results from the model simulations on the group response error distributions of the adults showed that the bandwidth was similarly narrow across eccentricities, and there was little difference in the early noise parameter. These two parameters captured well the narrow distributions of target responses to uncrowded targets. When flankers differed by 30° from the target, at 2.5° eccentricity the

population response to the target and the flanker orientations had almost an equal weight when pooled, matching the response error distributions centred at intermediate orientations between the target and flankers. The population response to the flankers gained more influence at the pooling stage with eccentricity, modelling the small shift of the response error distributions towards the flankers with eccentricity. With flankers that differed 90° from the target, at 2.5° eccentricity the population response to the flanker orientation contributed less than that to the target orientation at the pooling stage, simulating the higher proportion of target than flanker responses at this eccentricity. Similarly to when the flankers differed by30° from the target, the contribution of the population response to the flanker orientation at the pooling stage increased with increasing eccentricity. Overall, the model results show that with increasing eccentricity, the contribution of the population response to the flanker orientation at the pooling stage increases.

2.3.3.1.2 Children

Figure 2.10 shows the simulated response distributions, computed as the mean of the 1000 model iterations, for the group data of children with typical vision (Figure 2.10A) and children with amblyopia (Figure 2.10B). For the two parameters that applied in all three flanker conditions, the bandwidth and the early noise, there were minimal differences between the two groups. The bandwidth of the underlying detectors was 38.75° for the group with typical vision and 39.49° for the amblyopic group. These bandwidth values are approximately 6° larger than the bandwidth applied to the group data of adults at 5° and 15° eccentricity, that had the greatest bandwidth values in the adult periphery. The magnitude of the early noise was .59 for the control group and .52 for the amblyopic group, and were thus very similar to those applied by the model at the group data of adults. Based on these two parameters only, the model clearly followed the group data with uncrowded targets for both the control group and the amblyopic group– the simulated response distribution was centred on zero with a relatively narrow bandwidth.

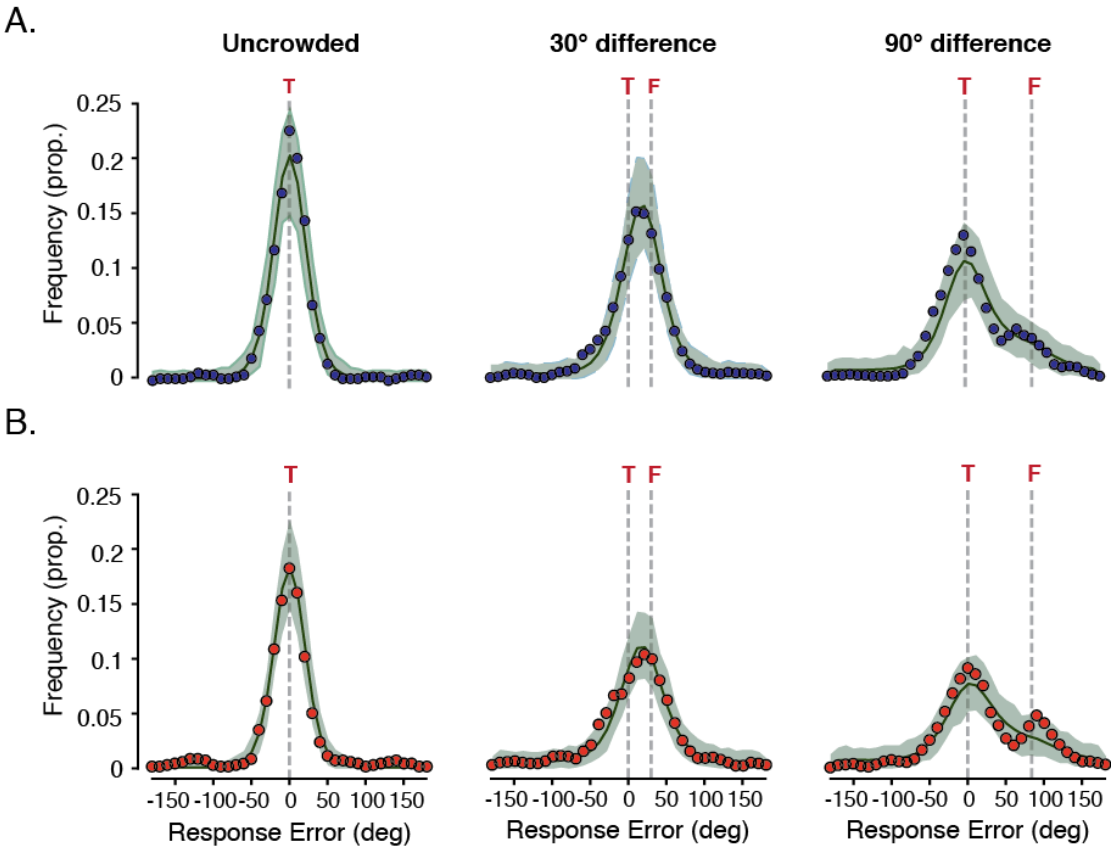

Figure 2-10 Model simulations on the group distributions from the typically developing and amblyopic fovea

A. Model simulations for the response error distributions of the group of children with typical vision. The dark green line indicates the mean distribution of the population response pooling model. The light green shaded areas represent the range of simulated distributions for 1000 iterations of the model. The group response error distribution from the orientation-matching task is presented by the dots. The grey line indicates the target location ('T'), and for the two flanker conditions in which flankers were present, the flanker locaion ('F').
B. Model simulations for the response error distributions of the group with amblyopia, plotted in the same conventions as in A.

When the target was crowded, due to the combination of the population responses to the target and flankers, as well as the addition of crowding noise, the bandwidth of the simulated response distributions increased and the peaks were lower. Although this was the case for both groups, the amblyopic group showed relatively flatter and wider response distributions than the control group. This was captured by the model in the crowding noise parameter, as crowding noise was .56 for the control group, and .77 for the group with amblyopia.

When considering each crowding condition individually, with flankers of a 30° difference from the target, the model almost perfectly captures the data from both groups – the simulated response distributions were centred between 0° and 30°. The flanker weights were .56 for the control group and .54 for the amblyopic group,

reflecting the balanced influence of the population response to the target and flankers when the responses were combined.

With flankers of a 90° difference, the model captured well the first peak of the bimodal response distribution of both groups near 0°, indicative of target responses. For the group with typical vision, the model also follows the second peak of response errors near 90°, that correspond to reports of the flanker orientation. For the group with amblyopia, this second peak near 90° was muted. This muting was due to a trade-off between the target and flanker weights. A low flanker weight captured well the first peak of the bimodal response error distribution indicative of target responses, as the population response to the target dominated at the crowding stage. This also means that a low flanker weight underperformed on the second peak of the response error distribution that contains flanker responses, as the contribution of the population response to the flankers at the crowding stage was low. Increasing the flanker weight with the aim of better capturing the second peak near 90° would reduce the contribution of the population response to the target at the crowding stage, and lead to an undershooting of the model on the first peak. As the first peak contained the majority of the response errors in this condition, undershooting on the first peak would lead to a larger difference between the response error distribution and the combined population response of the model, and result in a greater LSE. As such, the flanker weight for this condition was .31 for both control and amblyopic group. Overall, the model captured the data from both groups of children well, as indicated by the LSE that was .0035 for the control group and .0045 for the group with amblyopia.

Overall, the model results showed that the two groups of children did not differ substantially in the model parameters required to simulate the group data. The group with typical vision and amblyopia were similar in the two parameters that determined the population response to uncrowded targets, the bandwidth and the early noise. However, the values on these parameters were higher than those used to simulate the group response errors of adults, indicating that children's response error distributions were flatter and narrower than those of adults in the periphery. The group with amblyopia also required a higher crowding noise than both the group with children with typical vision and the adults, indicating that amblyopic crowding led to a greater spread of the response error distributions than developing and peripheral crowding. The flanker weights were very similar in the two groups in both crowding conditions. When the flankers differed by 30° from the target, the

population response to the target and the flanker orientation had almost an equal weight when pooled, matching the response error distributions centred at intermediate orientations between the target and flankers. When the flankers differed by 90° from the target, the population response to the target contributed more to the pooled response, matching the response error distributions that showed a majority of target responses. These flanker weights were most similar to those used in the near adult periphery at 2.5° eccentricity.

2.3.3.2 Model Simulations of Individual Data

As I discussed in previous sections (2.3.2.1.2 and 2.3.2.2.2), adults in the visual periphery and children in both the group with typical vision and the group with amblyopia show some variability in the pattern of response errors they make, that deviates from the group response errors. In the following section, I present the values of the five free model parameters for the individual observers in each group. For the response error distributions of each individual in each group (both children and adults), the best fitting model parameters were determined by the coarse-to-fine fitting procedure. Figure 2.11 shows the values for each of the five free parameters of the model.

In Figure 2.11A, the best-fitting values for the free parameter of the bandwidth of the detectors are presented. Children with typical vision required larger bandwidth values than children with amblyopia. No substantial differences were observed in the range of bandwidth values required for the response error distributions of adults across eccentricities in the visual periphery. For both groups of children, the range of bandwidth values required to simulate their individual response error distributions was larger than that for adults. Eleven children with typical vision and nine children with amblyopia required bandwidth values larger than the maximum bandwidth (SD= 42.29° at 2.5 eccentricity°) used to simulate the response error distributions of adults in the periphery. The remaining children had a similar bandwidth to adults. No children required bandwidth values smaller than those of adults.

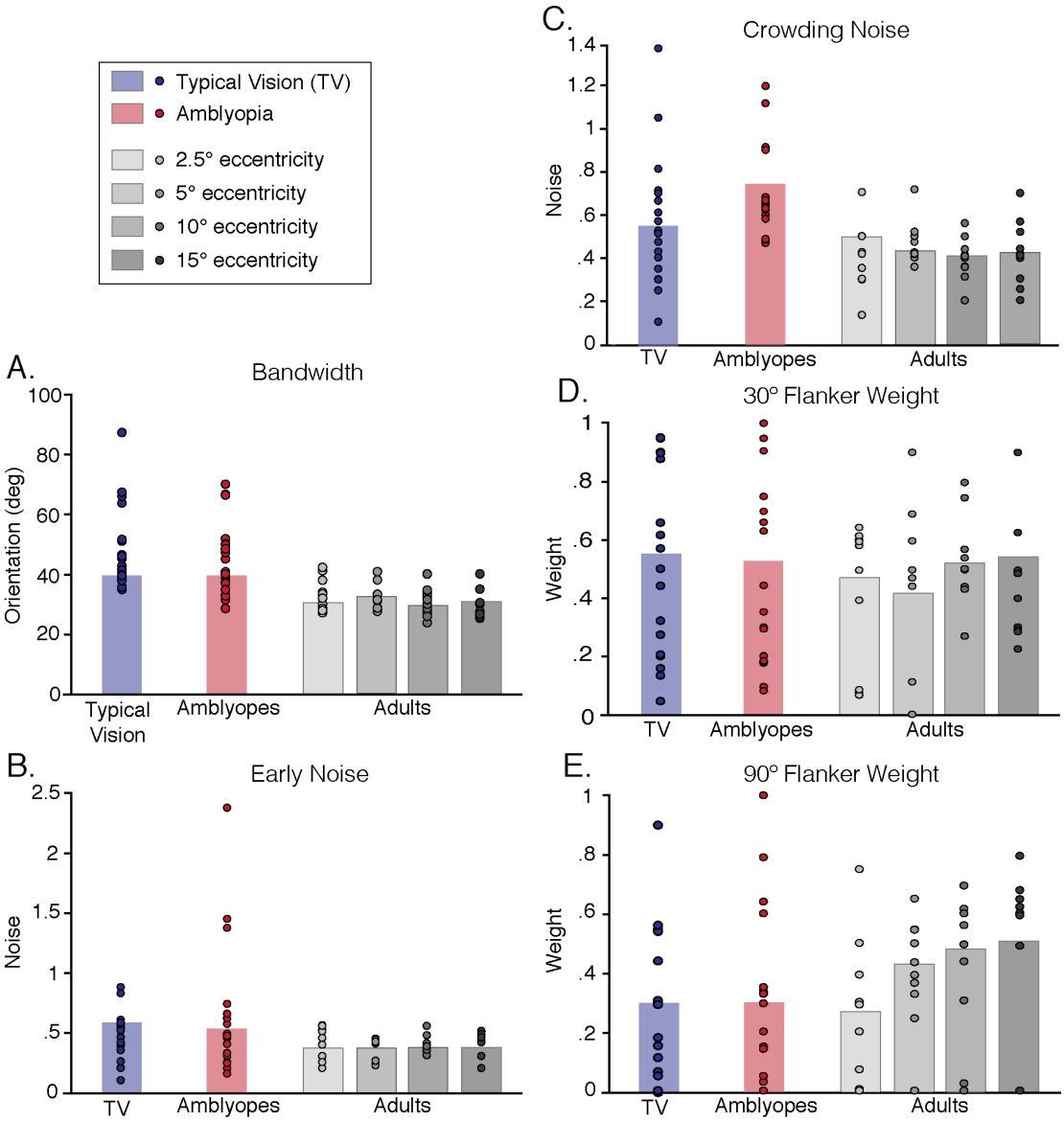

Figure 2-11 Individual best-fitting values of the model parameters in the adult periphery and the typically developing and amblyopic fovea

A. Best-fitting values from the coarse-to-fine fitting procedure for the bandwidth free parameter. Dots indicate individual observers, and bars the mean of the individual observers. Values are in degrees of visual angle.
B. Best-fitting values for the early noise free parameter. Plotted in conventions as in A.
C. Best-fitting values for the crowding noise free parameter. Plotted with conventions as in A.
D & E. Best fitting values for the flanker weights when the flankers differed by 30° from the target (D), and when the flankers differed by 90° from the target. Values range from 0 to 1. Note that the target weight was 1-the flanker weight. Plotted with conventions as in A.

In terms of the values required for the early noise parameter (Figure 2.11B), both children with typical vision and amblyopia required similar values to simulate their response error distributions. There was however an exception of three children with amblyopia who required early noise values that were substantially higher than those of children with typical vision, and from the rest of the amblyopic group. It is

worthwhile mentioning that the highest crowding noise for the children with typical vision, although lower than that of these three children with amblyopia, corresponded to observer C17, whose response error distributions were discussed in the section 2.3.2.2.2 as an example of highly noisy and variable distributions in both uncrowded and crowded cases. The early noise parameter was applied to the population response to the target (and the flankers when present) in all flanker conditions including for uncrowded targets. As such, the increased early noise value for these children in the amblyopic group cannot be considered a consequence of crowding. For the adults tested in the visual periphery, the early noise values were very similar across eccentricities, with a range of values between .20 and .56. The range of parameter values of children with typical vision and amblyopia was greater than that of adults, with the greatest early noise value for the children with typical vision being .87, and for children with amblyopia substantially larger, 2.35. However, the majority of children in both groups required values within the range of those of adults, with one child in each group requiring even a lower value than adults. In fact, the minimum noise value in the group of children with typical vision was .10 and in the group of children with amblyopia .15.

Crowding noise (Figure 2.11C) was applied at the crowding stage of the model, when the population response to the target and flankers was combined. Children in the amblyopic group required greater values of crowding noise than children with typical vision. In fact, one of the children with amblyopia with the highest crowding noise values was observer A04, whose response distributions were discussed in section 2.3.2.2.2 as an example of crowding significantly increasing response error variability. The range of the values for the amblyopic group were between .47 and 1.20, whereas for the children with typical vision the minimum value was much lower, at .11, and nine children with typical vision were below the minimum crowding noise value of children with amblyopia. However, the maximum crowding noise value for children with typical vision was 1.38, as one child had a value outside the range of those of children with amblyopia. Similarly to the early noise, the child with typical vision with the highest crowding noise was observer C17, whose individual response distributions were discussed in section 2.3.2.2.2 as an example of highly variable distributions in both uncrowded and crowded conditions. In the adult periphery, adults required slightly larger crowding noise values at 5° eccentricity (.36-.72) and somewhat lower at 10° (.20-.56), but showed no differences between 2.5° (.14-.70), and 15° (.20-.76). Children in both

groups had a greater range of crowding noise values than adults. However, most children with typical vision were similar to adults in this parameter, as all of the children with typical vision with the exception of three, required values within the range of those of adults in the typical periphery. Although most children with amblyopia also had crowding noise values within the range of adults, their values were primarily near the maximum values for adults, with 11 children having crowding noise values between .60 and .76.

The flanker weights determined the magnitude of the effect flankers had on the combined population response. In the condition in which the flankers differed by 30° from the target (Figure 2.11D), in both children with typical vision and amblyopia the simulated individual response error distributions required flanker weights that covered almost the entire range of possible values. The range of values for children with typical vision was .12 to .95, and for children with amblyopia 0-1. This is line with the heterogeneity of the underlying individual distributions discussed in 2.3.2.2.2. In the adult periphery, there was an effect of eccentricity with the range of flanker weights being lower on average at 2.5° and 5° eccentricity compared to 10° and 15. At 2.5° all adult observers required a flanker weight value below .65, with three adult observers requiring weights close to zero. At 5° eccentricity, there were four observers with a weight of 0, indicative of no effect of the flankers at the crowding stage of the model. Although the remaining observers all had a flanker weight below .65, there was an observer with a flanker weight of .9, deviating for the adult group in that eccentricity. This was observer P10 whose individual response distributions were discussed in section 2.3.2.1.2 as an example of increased flanker responses across all eccentricities. At 10° and 15° eccentricity, there were no observers that had a flanker weight of 0, as all observers required weights above .2. The highest flanker weights at 10° and 15° eccentricity were .8 and .9, respectively, which again corresponded to the response error distributions of observer P10.

In the condition in which the flankers differed by 90° from the target (Figure 2.11E), the range of flanker weight values was similar to the 30° flanker difference condition for both children with typical vision (0-.9) and amblyopia (0-1). However, the number of children with flanker weights near zero increased, indicative of a minimal to no contribution of the population response to the flanker orientation at the crowding stage. In the group with typical vision, nine children had flanker weights below .10 in this condition, whereas none had flanker values that low in the 30° flanker difference condition. In the group of children with amblyopia, six children

had values below .10, compared to two in the 30° flanker condition. In fact, fifteen children with typical vision and sixteen children with amblyopia had flanker weight values below .40, indicating a greater contribution of the population response to the target orientation in the pooling stage. In the adult periphery there was an effect of eccentricity, with the flanker weight values gradually gaining an increase with eccentricity for most observers. Flanker weight values below .40 were eight at 2.5° eccentricity, six at 5° eccentricity, four at 10° eccentricity and only three at 15° eccentricity. As such, although children in both groups had a greater range of flanker weight values, the trend for low flanker weights in this condition was most similar to adults at 2.5° eccentricity.

From this section discussing the parameters of the fits to the individual response error distributions a number of conclusions can be drawn. First, children required larger bandwidth and early noise values than adults in order for the model to simulate their response error distributions. Although children from both groups also required greater crowding noise parameter values than adults, this was especially the case for children with amblyopia. This suggests that in amblyopia, crowding increases response error variability more than in typically developing and peripheral vision. When the target was crowded with flankers of a 30° orientation difference from the target, flanker weights were higher in children from both groups compared to when the flankers differed by 90° from the target. This indicates that the contribution of the population response to the flankers to the pooling stage was greater with flankers of a 30° difference from the target than flankers of a 90° difference. This was also the case for adults in the near periphery. However, at greater eccentricities, the contribution of flankers to the pooled population response with 90° flankers clearly increased.

2.4 **Discussion**

In this chapter, my aim was to investigate the perceptual effects of crowding in amblyopic and developing vision, and examine whether they match those found in the adult typical periphery. To achieve this, I used a method of adjustment to quantify the perceptual error in the orientation of a crowded Landolt-C stimulus in the fovea of typically developing children and children with amblyopia, as well as adults in the visual periphery. Based on the perceptual errors of children, I determined that the effects of crowding on the appearance of objects were systematic in the typically developing and amblyopic fovea. When the target object

was crowded, children reported either the orientation of the target, intermediate orientations between the target and flankers (assimilation errors), or substituted the orientation of the target with that of the flankers (substitution errors). These classes of errors matched those made by adults in the typical periphery, both in this chapter and in previous reports (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010; van den Berg et al., 2010; Ester et al., 2015; Harrison & Bex, 2015). I showed that a weighted population response pooling model, consistent with prior models based on population coding (van den Berg et al., 2010; Harrison & Bex, 2015), can predict this complex pattern of observers' reports in all three instances of crowding: the typically developing and amblyopic fovea, and the adult periphery, suggesting a common underlying mechanism.

To my knowledge, this is the first attempt to determine the effect that amblyopic crowding and crowding in the typically developing visual system have on the appearance of a target object in clutter. My findings show that in both children with typical vision and amblyopia, crowding has systematic effects on the perceived orientation of the target. With 30° flanker differences, children primarily reported orientations between the target and flankers, making assimilation errors. On an individual level there were differences in this flanker condition, with some children reporting intermediate orientations closer to the target whereas others reported orientations that were more similar to the flankers. When the flankers differed by 90° from the target, children primarily either reported the orientation of the target, or reported the flanker orientation instead of the target, making substitution errors. Substitution errors were less frequent than target reports in this condition. This pattern of responses matched that in the adult periphery: adults made primarily assimilation errors with flankers of a 30° difference, and either reported the target or made substitution errors with flankers of a 90° difference, with an increase in the frequency of these latter errors that increased with eccentricity. Altogether, the findings show that both children with amblyopia and typical vision make the same perceptual errors as adults in the visual periphery. As such, despite differences in the extent of crowding between the amblyopic and typically developing fovea, and between children and adults, when stimuli were within the crowding zone, the effect of crowding on target appearance was the same.

In light of these systematic perceptual effects in peripheral, developing, and amblyopic vision I argued that there is no averaging or substitution crowding mechanism per se, but assimilation and substitutions errors. Instead, I followed

recent models of crowding (van den Berg, Roerdink, & Cornelissen, 2007; Harrison & Bex, 2015) and proposed that the perceived orientation of a crowded object is drawn from the response of a population of detectors selective to orientation, similarly to neurons in V1 (Schiller et al., 1976; Mazer et al., 2002). I simulated crowding as the weighted pooling of the population responses to the target and flanker orientations. The weights determined the magnitude of the contribution of the population response to the flankers to the pooled response, and thus to the stimulated perceptual effects of crowding. Therefore, by adjusting the weights it was possible to simulate the differential effect of the flankers depending on their similarity to the target and the target eccentricity. In line with previous models of peripheral crowding (Greenwood et al., 2009; Ester et al., 2014; Ester et al., 2015), noise was added to the pooled population response to the target and flankers. Overall, this model is consistent with two-stage theories of crowding (Levi, Hariharan, et al., 2002b; Pelli et al., 2004; Levi, 2008) – the first feature-detection stage could arise when the population in my model responds separately to the target and flanker orientations, with the second feature-integration stage arising when the population responses are pooled.

An important finding is that both children and adults in the near periphery showed a greater proportion of target responses when the flankers had a larger 90° orientation difference from the target compared to a 30° orientation difference. The greater proportion of target responses when the flankers were more dissimilar to the target is consistent with findings showing that peripheral crowding is selective for target flanker similarity in orientation (Andriessen & Bouma, 1976; Leat, Li, & Epp, 1999; Levi, Hariharan, et al., 2002b; Hariharan et al., 2005): the more dissimilar in orientation the flankers are from the target, the easier it is to identify the target orientation, and the weaker the effect of crowding. Importantly, the findings in this chapter suggest that the similarity between the target and the flankers also matters in developing and amblyopic crowding. This is a significant finding, as previous studies have shown that amblyopic crowding disrupts target recognition regardless of the similarity between the target and the flankers (Hess, Dakin, et al., 2001; Hariharan et al., 2005). This selectivity for target-flanker similarity, captured in the model by differences in the flanker weights, could arise due to differences in the cortical proximity of the target and flankers. Mareschal, Morgan, and Solomon (2010) found that in the typical adult periphery, increasing eccentricity and reducing the separation between target and flankers increased assimilation errors. Increasing

eccentricity and reducing target-flanker separation both reduce the cortical separation between the target and flankers, since less cortex is dedicated to the periphery than the fovea (Dow et al., 1981; Van Essen et al., 1984; Sereno et al., 1995; Engel, Glover, & Wandell, 1997). Orientation differences may modulate the flanker weights in a similar way by differences in the cortical separation of the gaps: an orientation difference of 30° between the target and flankers places their gaps in closer cortical proximity than an orientation difference of 90°. Therefore, this difference in cortical gap proximity could account for the increased assimilation errors with 30° flanker differences.

Interestingly, at higher eccentricities in the adult periphery substitution errors increased with both 30° and 90° flanker differences and selectivity for target-flanker similarity in orientation was less evident. However, it should be noted that there was substantial individual variation in the rate of increase in substitution errors, with some adult observers shifting to primarily flanker reports at 5° eccentricity whereas others later at 15°. The model captured this response pattern by increasing flanker weights with eccentricity, thus increasing the contribution of the population response to the flanker orientation to the pooled response. As eccentricity increases, the cortical magnification factor (the amount of cortex dedicated to 1° of visual angle) (Daniel & Whitteridge, 1961) is reduced. As such, the cortical distance between the target and flankers may become so small with increasing eccentricity than any differences in orientation are not substantial enough to increase their separation. This in turn may result in similarly high weights for 30° and 90° target-flanker differences. Individual differences in cortical magnification across the visual field may explain why this increase in flanker weights for both target-flanker orientation differences occurs at earlier eccentricities for some observers than others. Similarly, individual differences in functional architecture could account for this pattern of response error in the subset of children with amblyopia and typical vision. For example, it has been shown that individual differences in perceived object size across the visual field correlate with pRF size in V1 (Moutsiana et al., 2016). Although speculative, it may be that individual differences in the mapping of the foveal representation, such as localised distortions of the retinotopic map, results in high flanker weights regardless of the flanker orientation in children.

An alternative explanation for the increased substitution errors and thus flanker weights with eccentricity is positional uncertainty. Positional uncertainty could create source confusion, where flanker letters intrude into the target's percept

(Wolford, 1975) resulting in the observer being unable to allocate which feature belongs to the target and which to the flanker. Such source confusion could either involve mislocalised features or mislocalised whole letters (Strasburger, Rentschler, & Jüttner, 2011). Consistent with this interpretation of the findings is evidence showing that both adult peripheral vision and amblyopic foveal vision are characterised by increased positional uncertainty (Bedell & Flom, 1981; Levi & Klein, 1983; Rentschler & Treutwein, 1985; Levi et al., 1987; Hess & Hayes, 1994; Hess, McIlhagga, & Field, 1997). Additionally, during development vernier acuity does not reach adult levels until the early teens (Carkeet, Levi, & Manny, 1997; Skoczenski & Norcia, 2002), indicating developmental difficulties with positional uncertainty. As such, increased positional uncertainty could explain the substitution errors found with orthogonal flankers in in the adult periphery, as well as in the subset of children with amblyopia and typical vision. Note however that such an explanation does not necessarily require an additional mechanism beyond pooling, as the integration of features over increasingly large crowding zones would likely also increase positional uncertainty.

Although the perceptual errors made by children with typical vision matched those of adults, they showed increased variability in their responses compared to the adults. This suggests a reduction in the precision of reports of orientation. Importantly, this variability was present in the children's responses both to isolated targets and targets surrounded by flankers, and was captured in the model by an increase in early noise compared to the adults. This suggests that rather than being a consequence of crowding, it reflects a general developmental difficulty with orientation matching. As the selectivity of neurons to orientation emerges early in postnatal life (Braddick & Atkinson, 2011), this difficulty could be unrelated to visual function but associated with difficulties with engaging with psychophysical tasks. Although I took I number of precautions to ensure that children were engaged during the experiment, such as making the stimuli into cartoon characters, devising a story and including animations between trials, it is still likely that this subset of children were not fully engaged in the task. This would be consistent with previous studies showing that children are more likely to make errors in psychophysical tasks, even on easy trials, due to attentional lapses and poorer short-term memory skills (Witton et al., 2017; Manning et al., 2018).

Children with amblyopia also showed this increased variability for isolated targets compared to adults. However, they also showed an additional increase in

variability in their responses to crowded targets that was not evident in children with typical vision. Namely, with crowded targets, a subset of children with amblyopia made a disproportionate amount of reports to orientations that did not correspond to either the target or the flankers, and were thus classified as random. For these cases, a model that adds noise to the target population response but does *not* include any pooling with the flanker response performed best (Appendix A, section 6.1.4). Similarly to the children with typical vision, one explanation for this response variability could be attentional lapses. However, this is unlikely, as attentional lapses would have to be concentrated on crowded trials, since random responses were less frequent in uncrowded trials. Additionally, the disproportionate amount of random responses compared to children with typical vision would suggest an attentional deficit in amblyopia. This is unlikely as there is little evidence for attentional difficulties in amblyopia (Sharma et al., 2000), and it is unclear why they would only affect a subset of children while sparing others. Rather, as amblyopia is widely considered a disorder of spatial vision (Levi, 2013), the increased random perceptual errors children with amblyopia could be related to misperceptions of spatial structure or perceptual distortions that characterise amblyopic vision (Fronius & Sireteanu, 1989; Lagreze & Sireteanu, 1991; Sireteanu, Lagreze, et al., 1993; Sireteanu, Baumer, et al., 2008). Perceptual distortions can manifest as "misperceived orientations" (Barrett et al., 2003), and as such, they could have affected the perceived orientation of the crowded target by displacing the shape of the filled-in Landolt-C. Although distortions cannot account for the systematic effects of crowding on target appearance, they could have acted as a source of random errors in this subset of children with amblyopia.

However, I did not obtain measures of perceptual distortions for the children with amblyopia, and thus it is not possible to determine whether they were present in my sample and how they influenced the perceptual effects of crowding. One method that has been used to investigate perceptual distortions in adult observers is by asking them to sketch the appearance of stimuli, such as sinusoidal gratings of varying orientations and spatial frequencies (Hess, Campbell, & Greenhalgh, 1978; Barrett et al., 2003). Observers are first asked to compare the same grating presented monocularly in their fellow fixating eye and their amblyopic eye, and if they notice any differences in its appearance when viewing with their amblyopic eye, they are then asked to sketch the grating. For the purposes of testing children, sinusoidal gratings could be characterised as cartoon zebras, similar to prior tests

(Maione, Berardi, & Cerimele, 1983). Children can then be asked to sketch the "stripes" of the zebra when viewed with either their amblyopic or their fellow fixating eye. The sketches from the two eyes can then be compared to determine whether children experience similar distortions to adults, such as straight lines appearing wavy, jagged, or fragmented (Barrett et al., 2003). As not all observers with amblyopia experience such misperceptions of spatial structure (Piano et al., 2015), the use of such a task could establish how many children in the amblyopic group did. Importantly, it would allow to investigate whether these is a link between perceptual distortions and increased random responses in the orientation-matching task, and whether the children that experience such distortions are those whose responses are better accounted for by a noisy population response model.

Although the model presented in this chapter is agnostic with regards to the neural site of pooling, its suggestion of a common underlying crowding mechanism raises the question of common neural underpinnings (which is also explored in the next chapter of this thesis). In the periphery, the increase in the extent of crowding with eccentricity (Toet & Levi, 1992) has been attributed to the insufficient sampling of the peripheral visual field, with neurons with large receptive fields required for sufficient coverage (Parkes et al., 2001). During the development of the visual cortex, the post-natal maturation of connections in the primary visual cortex (Huttenlocher, de Courten, Garey, & Van der Loos, 1982; Huttenlocher & Dabholkar, 1997), and the later maturation of receptive fields in extrastriate cortex (Zhang et al., 2005), could result in enlarged receptive sizes until late childhood, thus resulting in increased crowding in the developing fovea. Similarly, strabismus during development results in fewer neurons responding to the amblyopic eye in V1 and V2 (Crawford & von Noorden, 1979; H. Bi et al., 2011; Shooner et al., 2015). As such, crowding in amblyopia could result from a diversion of neural resources from the amblyopic eye to the fellow fixating eye, creating the need for increased receptive field sizes and thus crowding in amblyopia. Advances in neuroimaging techniques have allowed for the estimation of population receptive field (pRF) size in the visual cortex of humans (Dumoulin & Wandell, 2008). Although studies employing this technique have failed to find an enlargement of pRF size in typically developing children (Dekker, Schwarzkopf, de Haas, Nardini, & Sereno, 2017), there is evidence demonstrating that pRFs are enlarged for the amblyopic eye in both striate and extrastriate areas of strabismic amblyopes (Clavagnier et al., 2015). In the following

chapter, I investigate whether there is a relationship between pRF size and crowding in amblyopia and typical vision.

3 **Chapter 3:** Population Receptive Field Size in Strabismic Amblyopia and Typical Vision

3.1 **Introduction**

Crowding in the visual periphery has been described as a mechanism that systematically promotes perceptual similarity among adjacent regions in the visual field, creating a homogeneous representation of the visual scene (Greenwood et al., 2010). In Chapter 2, the perceptual effects of crowding in amblyopic vision were investigated, to determine whether they matched those observed in the typical visual periphery. It was shown that when tasked with reporting the identity of a crowded target, children with amblyopia made the same systematic errors as adults in the typical periphery. Both children with amblyopia and adults reported average or intermediate identities between the target and flankers, or substituted the flanker for the target identity. As such, the findings from Chapter 2 showed that amblyopic and peripheral crowding share common perceptual effects. Most importantly, in both instances of crowding both assimilation and substitution errors were successfully accounted for by a population response pooling model. This suggests that the typical periphery and the amblyopic fovea share the same underlying pooling mechanism, and raises the question of whether peripheral and amblyopic crowding also have common neural underpinnings.

Pooling accounts have proposed that crowding arises in the typical visual periphery due to the undersampling of the peripheral visual field (Parkes et al., 2001). In the typical visual system, the first cortical areas in the visual pathway (V1-V4) are retinotopically organized (DeYoe, Bandettini, Neitz, Miller, & Winans, 1994; Wandell, Brewer, & Dougherty, 2005): nearby regions of the visual field on the retina project to nearby regions in the cortex, forming cortical maps of the visual field. The greater emphasis in the retina towards the fovea is also captured in these cortical maps. In the retina, there is an oversampling of the fovea, with the cone mosaic being almost forty times more dense in the fovea compared to the periphery (Curcio, Sloan, Kalina, & Hendrickson, 1990; Chui, Song, & Burns, 2008). This foveal oversampling is carried through the visual pathway, with ganglion cells oversampling cones in the fovea by a factor four relative to the periphery (Curcio & Allen, 1990), and the cells in the lateral geniculate nucleus (LGN) of the thalamus being four times as many for each ganglion cell afferent in the fovea than the periphery (Connolly & Van Essen, 1984). In V1 of non-human primates, the amount of cortex (in mm) dedicated to one degree of visual angle, called the cortical

magnification factor (CMF) (Daniel & Whitteridge, 1961), and thus the number of neurons, is greater for the fovea and decreases with eccentricity (Dow et al., 1981; Van Essen et al., 1984). This cortical overrepresentation of the fovea has been confirmed by fMRI studies in humans, showing that CMF is greater for the fovea and decreases with eccentricity in V1-V4, with human visual areas having a greater foveal emphasis than the monkey counterparts (Sereno et al., 1995; Engel et al., 1997). The cortical overrepresentation of the fovea can be associated with the better visual function in the fovea relative to the periphery. In support of this, Duncan and Boynton (2003) have shown that in human observers, larger CMFs correlate with higher visual acuity.

This oversampling of the fovea and results in an undersampling of the visual periphery, and thus a difference in the allocation of neural resources between the fovea and the periphery. Due to this undersampling, the fewer number of neurons that respond to peripheral stimulation must respond to larger regions of the visual field in order to ensure adequate coverage, and thus have larger receptive fields. In monkeys, neural receptive field (RF) size is smaller for the fovea and increases with eccentricity (Dow et al., 1981; Van Essen et al., 1984), and increases along early to late visual processing areas (Rosa & Tweedale, 2005). Advances in fMRI have allowed for the measurement of the aggregate RF of populations of neurons (pRF) within an fMRI voxel (Dumoulin & Wandell, 2008; Amano et al., 2009; Harvey & Dumoulin, 2011) in humans. These studies have found that pRF size increases with eccentricity in V1-V4, in line the neurophysiological studies in non-human primates. The increase in receptive field size with eccentricity in non-human primates and humans, parallels the increase in the extent of crowding with eccentricity (Bouma, 1970; Toet & Levi, 1992). As such, pooling accounts of crowding have proposed that the increased extent of crowding in the visual periphery may arise due to the increased pooling of the target and flanker features within large peripheral receptive fields (Parkes et al., 2001).

The exact cortical site where this pooling occurs is unclear. Psychophysical measurements have shown that the strength of crowding correlates with the cortical distance between the target and flankers in V1 (Pelli, 2008). A number of studies using fMRI have also pointed to the involvement of V1 by demonstrating that crowding has an influence on BOLD responses in V1, with crowded targets reducing BOLD activation relative to isolated targets (Chen et al., 2014; Millin et al., 2014). Although this suggests an early cortical locus for crowding, Motter (2002) has shown

that receptive fields capable of pooling information across Bouma's (1970) critical spacing of half the target eccentricity are first encountered in V4, thus showing that the scaling of V1 neurons is too small to account for the large extents of crowding (Motter, 2002, 2009). Additionally, there is evidence from neuroimaging studies for the involvement of areas beyond V1 in crowding, as fMRI studies have found that crowding-related changes in BOLD can be seen in V2, V3, and V4 (T. Bi et al., 2009; J. Freeman et al., 2011b; Millin et al., 2014). For example, Anderson et al. (2012) showed that crowding-induced changes in the appearance of targets modulated BOLD signal throughout visual areas V1-V4, with the strength of this effect increasing from early to late visual areas. As RFs and pRFs increase from V1 to later visual areas (A. T. Smith et al., 2001; Rosa & Tweedale, 2005; Dumoulin & Wandell, 2008), evidence of the involvement of multiple visual areas in crowding could be indicative of an increase in pooling across receptive fields that increase from early to late visual processing areas.

Crowding in the amblyopic eye of individuals with strabismic amblyopia could similarly arise due to a shift in the ocular dominance of neurons towards the fellow fixating eye, and thus a reduction in the number of neurons driven by the amblyopic eye. This in turn could result in increased receptive fields for the neurons driven by the amblyopic eye as a compensation mechanism. Evidence on such a reduction in the number of neurons driven by the amblyopic eye relative to the fellow fixating eye is unclear in V1 . Electrophysiological recordings on cats with artificial strabismus have shown a reduction in the number of neurons driven by the amblyopic eye relative to the fellow fixating eye (Berman & Murphy, 1981; R. D. Freeman & Tsumoto, 1983; Sireteanu & Best, 1992). Studies on non-human primates have yielded mixed results, with some showing balanced ocular dominance in V1 (E. L. Smith et al., 1997; Kiorpes et al., 1998), whereas others have shown an overall shift of eye dominance towards the fellow fixating eye (Crawford & von Noorden, 1979). However, in the neurons driven by the amblyopic eye of primates, the optimal spatial frequency and peak contrast sensitivity is better than what is predicted by the severity of the behavioural losses in the same animals (Kiorpes et al., 1998). Therefore, the growing consensus on the neural basis of strabismic amblyopia is that it involves cortical alterations beyond V1 (Kiorpes & McKee, 2006; Levi, 2006; Kiorpes & Daw, 2018)

Few studies have investigated cortical areas beyond V1 in strabismic amblyopia. A recent electrophysiological study on non-human primates showed that

the ocular dominance imbalance in favour of the fellow fixating eye was negligible in V1, whereas the shift in preference was large in V2, with more binocular neurons responding to the fellow eye (H. Bi et al., 2011). Shooner et al. (2015) recorded from populations of neurons in V1 and V2, and showed an ocular dominance bias towards the fellow fixating eye, that correlated with the severity of the behavioural visual deficits. Recently, Clavagnier et al. (2015) used fMRI to obtain estimates of pRF size responding to stimuli between 1-6° eccentricity for human observers with typical vision and amblyopia. They showed that the size of pRFs centred between 1-6° eccentricity were enlarged in visual areas V1, V2, and V3 relative to the fellow fixating eye and the eyes of observers with typical vision. The differences in pRF size in V1, V2, and V3 between the amblyopic and the fellow fixating eye could be due to fewer neurons with enlarged RFs for the amblyopic eye, indicative of a reduction in the allocation of neural resources. As such, these differences could be underlying the reduced visual function, and particularly the reduced acuity and increased extent of crowding. However, the relationship between fMRI estimates of pRF size and measures of acuity and crowding across the visual field has not been investigated in neither strabismic amblyopia nor typical vision.

Although it is clear that in the amblyopic fovea acuity is reduced and crowding is elevated (Levi & Klein, 1985; Greenwood et al., 2012), little is known about how the pattern of these deficits is manifested across the visual field, especially for crowding. Foveal measurements of pRF size can be unreliable and difficult to measure (Wandell, Dumoulin, & Brewer, 2007), it is necessary to characterise the pattern of the acuity and crowding deficits across the periphery in order to explore their relationship to variations in pRF size. In an early study, Hess and Jacobs (1979) investigated the variations in acuity and the extent of crowding across the visual field of strabismic amblyopes. For the four observers tested, they pattern of deficits differed. For two observers, the acuity and crowding deficits were primarily restricted to the fovea and parafovea, with acuity and crowding being similar to controls in farther periphery. In the periphery, these observers also showed a marked naso-temporal asymmetry in acuity and crowding, with the *trailing* visual field (the visual field in opposite the direction of the ocular deviation) showing greater impairment than the *leading* visual field (the visual field in the direction of the deviation). This is consistent with Sireteanu and Fronius (1981), who found that within 20° eccentricity, acuity for the amblyopic eye in the temporal visual field of esotropic amblyopes (i.e. observers with inward ocular deviation) showed a greater

impairment than in the nasal visual field. Note that such naso-temporal asymmetries are not evident in all observers with amblyopia, and thus are idiosyncratic.

In this chapter, I investigated whether the neural basis of amblyopic crowding can be considered in a similar way to the neural basis of crowding in typical vision. First, similarly to Clavagnier et al. (2015), I investigated whether the neural basis of amblyopia can be captured by estimates of pRF size in V1, V2, and V3. Second, I aimed to examine whether pRF size is associated with measures of acuity and crowding. In order to achieve this, I first measured acuity and crowding across the visual field in adult observers with strabismic amblyopia and controls with typical vision. Measures were obtained for 2°,4°,8° and 12° in both the nasal and temporal visual fields, in order to determine potential naso-temporal asymmetries. Secondly, using fMRI and pRF mapping, pRF size was measured for 1-19° eccentricity in the V1, V2, and V3 for the same observers. First, I hypothesise that if there is a reduction in the allocation of neural resources to the amblyopic eye, estimates of pRF size should be greater for the amblyopic eye of observers with amblyopia, relative to the eyes of observers with typical vision and to the fellow fixating eye. Additionally, if the amblyopic deficits in acuity and crowding arise due to this decrease in the allocation of neural resources, then there should be a common pattern across the visual field between acuity and crowding and pRF size. In strabismic amblyopia, regions of the visual field that show the greatest reduction in acuity and elevation in the extent of crowding relative to unaffected vision should also show the greatest elevation in pRF size. For example, if there are significant naso-temporal asymmetries in measures of acuity and crowding, these should also be manifested in measures of pRF size. This relationship should also be found in typical vision: with increasing eccentricity there should be an increase of pRF size, a reduction in acuity, and an increase in the extent of crowding.

3.2 **Methods**

3.2.1 Behavioural Testing

3.2.1.1 Observers

10 adult observers with strabismic amblyopia and 10 observers with unaffected vision were recruited. As this study had a neuroimaging component, adults with amblyopia were recruited because children are more likely to move when placed within the MRI scanner, and the ability to stay still is essential for obtaining

usable data. Due to technical difficulties in the MRI scanner that resulted in artefacts in the neuroimaging data, I had to exclude one observer with amblyopia. This resulted in a sample of nine observers in the group with amblyopia (2 males, mean age 31.8 years, range 19-44) and ten observers in the control group (4 males, mean age 32.6 years, range 19-44). Note that we age-matched each observer with unaffected vision with one observer with amblyopia (±1 year).

The observers with amblyopia were selected based on at least a two line difference in logMAR (logarithm of minimum angle of resolution) acuity between the eyes, normal or corrected-to-normal logMAR visual acuity in their fellow fixating eye, and a childhood history of amblyopia with strabismus (misalignment of the visual axes). Clinical details for the observers with amblyopia can be seen in Table 3.1 below. As indicated by their own report, the observers in the control group had normal or corrected-to-normal visual acuity and no history of binocular dysfunction. The experimental procedures were performed with the informed consent of the observers and were approved by the East of England – Cambridge South Research Ethics Committee of the National Health System (NHS) Health Research Authority.

	Age (years)	Sex	Ocular Alignment	Refractive Error		logMAR acuity	
				LE	RE	LE	RE
A1	22	F	n: L SOT 20$^{\Delta}$ d: L SOT 12$^{\Delta}$	Plano	Plano	1.00	-0.2
A2	28	F	n: L XOT, 50$^{\Delta}$ d: L XOT, 45$^{\Delta}$	Plano	Plano	0.48	0.18
A3	33	F	n: R XOT, 45$^{\Delta}$ d: R XOT 40$^{\Delta}$	-0.25	-0.5	-0.1	0.2
A4	43	M	n: R XOT 12$^{\Delta}$ d: R XOT 1$^{\Delta}$	-5.25/ 1.25x85°	-4/ 1.25x85°	-0.2	0.2
A5	18	F	R XOT 16$^{\Delta}$	-1.25/ -2.75x167°	-1.50/ -3.00x180°	-0.2	0.2
A6	43	M	n: R XOT 25$^{\Delta}$ d: R XOT 30$^{\Delta}$	Plano	Plano	-0.1	0.8
A7	35	F	n: L SOT 40$^{\Delta}$ d: L SOT 38 $^{\Delta}$	+7.5/- 0.75x120°	+7.25	0.3	-0.2
A8	41	F	n: R XOT 18$^{\Delta}$ L/R 12$^{\Delta}$ d: R XOT 16$^{\Delta}$ L/R 6$^{\Delta}$	+0.75/- 0.25x120°	+5.00/- 3.00x180°	0.0	0.6
A9	28	F	R XOT, 36$^{\Delta}$	Plano	Plano	0.0	0.48

Table 3-1 Clinical details of observers with amblyopia (N=9)

The "ocular alignment" column reports the outcome of both near (n) and distance (d) prism cover tests. When one value is denoted only, this indicates near cover test. SOT = esotropia (inward ocular deviation), XOT = exotropia (outward ocular deviation), L/R = left eye over right eye. The angle of deviation is shown in prism dioptres (Δ) and the amblyopic eye is denoted. Note that 2Δ = 1°. Acuity is denoted as measured by the logarithm of the minimum angle of resolution (logMAR) chart.

3.2.1.2 Apparatus

Experiments were programmed using Matlab (The Mathworks, Ltd., Cambridge, UK) on a Dell PC running PsychToolBox (Brainard, 1997; Pelli, 1997). Stimuli were presented on an ASUS VG278HE LCD monitor, with 1920x1080 resolution and 120Hz refresh rate. The monitor was calibrated using a Minolta photometer and linearised in software, to give a maximum luminance of 150 cd/m^2. A USB portable numeric keypad was used to register participant responses. Eye-tracking was performed using an Eyelink 1000 (SR Research, Mississauga, ON, Canada) with a level desktop camera. Figure 3.1A shows the experimental set-up.

3.2.1.3 Stimuli and Procedures

For all observers, one eye was patched using a plastic eye patch for monocular stimulus presentation. For the group with amblyopia, observers were tested in their amblyopic eye, and for the control group, observers were tested in

their dominant eye. Eye dominance was determined using the Miles test (Miles, 1928).

A Landolt-C target was presented either in isolation (acuity) or surrounded by four Landolt-C targets (crowding). Nine visual field locations were tested: the target was either in the fovea, or at 2°, 4°, 8°, and 12° eccentricity in the nasal and temporal visual fields. For foveal presentation, the target was presented in the centre of the screen (see Figure 3.1B). For peripheral presentation, a Gaussian fixation dot was presented near the bottom of the screen, and the target was presented diagonally at the relevant eccentricity either in a 135° angle on the left of the fixation dot or a 45° on the right of the fixation dot (thus falling either in the nasal or temporal visual field of the observer).

Target and flankers were Landolt-C stimuli. The gap of the Landolt-C was equal to one-fifth of the stimulus diameter, similarly to Sloan letters (Sloan, 1959). For both the acuity and crowding tasks, the observer had to respond with the orientation of the target in 4-alternative forced choice (4AFC). In the crowding task, the four flankers were located diagonally at the top left, top right, bottom left, and bottom right of the target. This position was chosen for the flankers because when the target is placed in diagonal locations, crowding has been shown to be significantly stronger (i.e. the extent of crowding is larger) when the target and flankers are horizontally rather than vertically arranged (Feng, Jiang, & He, 2007). The diagonal flanker positioning ensured that the four flankers had a similar effect on disrupting target recognition. Due to the position of the flankers, the target was presented in one of four oblique orientations (45°, 135°, 225° or 315°) in order for the gap of the Landolt-C to be always abutting with a flanker and thus be maximally crowded. Each of the flankers was randomly oriented either 45°, 135°, 225° or 315° (see Figure 3.1C). The relative separation between the target and flankers was fixed at 1.1 × the target size, as this separation has been shown to produce maximal crowding effects (Levi, Song, et al., 2007a; Song et al., 2014). All stimuli were rendered in black at 90% Weber contrast and presented against a mid-grey (45cd/m^2) background.

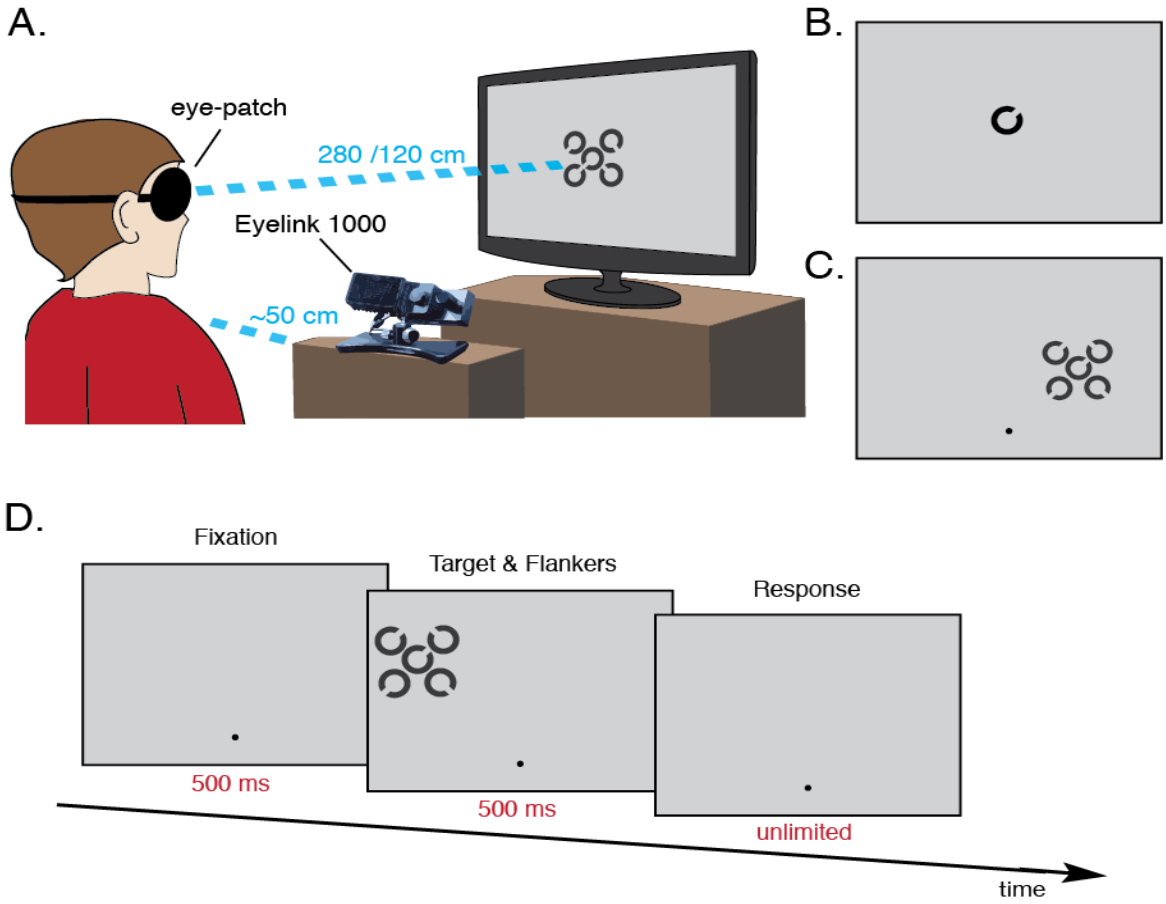

Figure 3-1 Stimuli and methods for the acuity and crowding tasks

A. Experimental set up for acuity and crowding tasks. Observers in the control group sat at a distance of 280cm from the screen monitor. Observers with amblyopia sat at a distance of 280cm or 214cm for acuity and crowding measurements in the fovea and at 2° and 4° eccentricity; for measurements at 8° and 12° eccentricity they sat 120cm from the screen. The Eyelink 1000 was positioned in front of the observers at a distance of approximately 50cm. Stimuli were presented to the amblyopic eye of observers with amblyopia and the dominant eye of observers with typical vision.
B. Example of Landolt-C stimulus for an acuity trial. The stimulus depicted here represents a trial for foveal measurement of acuity. For measurements in the visual periphery see C.
C. Example Landolt-C stimuli for a crowding trial. The central Landolt-C is the target, and the four surrounding Cs are the flankers. The stimulus array depicted here represents a trial for peripheral measurements of crowding, as observers had to maintain fixation on the dot in the bottom of the screen.
D. Stimulus presentation sequence: First the fixation dot was presented for 500ms. Then, the stimuli were presented for 500ms, during which time eye movements greater than the criterion distance invalidated the trial. Finally, a grey response screen was presented until the observer registered their response.

Figure 3.1D shows the stimulus presentation sequence for a given trial. Each trial started with the presentation of the Gaussian fixation dot for 500ms. Then, the target in isolation (acuity task, as in Figure 3.1B) or surrounded by the four flankers (crowding task, as in Figure 3.1 C) was presented for 500ms. During target presentation, the fixation dot was removed for foveal target presentation but remained on screen for peripheral target conditions to aid fixation. Distant line elements (one horizontal and one vertical line) were also presented during target

presentation to aid fixation. A grey screen then appeared until the observer made a response using the keypad to indicate the orientation of the target.

Gap size thresholds were measured by varying the overall size of the target (and thus the visibility of the gap of the Landolt-C that indicated orientation) using a QUEST algorithm procedure (Watson & Pelli, 1983) converging at 62.5% correct performance. On any given trial, a correct response resulted in a decrease in the QUEST gap size estimate for the next trial, and an incorrect response resulted in a decrease in the next gap size estimate. Note that for the crowding task, the change in the size of the target on each QUEST trial changed the absolute centre-to-centre separation between the target and flankers. QUEST terminated after 65 trials, and gave the size of the Landolt-C gap in degrees of visual angle as the output. Observers completed 2 runs of the QUEST per visual field target location (fovea + [4 eccentricities × 2 visual fields]), resulting in 18 blocks of data for each task.

QUEST is typically very quick to reach convergence and proceeds to present stimulus intensities (i.e. gap size values) at threshold for the remaining trials. Our aim was to use the intensities presented during QUEST and their corresponding proportion of correct responses to fit a psychometric function post-hoc, as this method has been shown to provide more robust threshold estimates to attentional lapses than QUEST which can be common in clinical populations (Witton et al., 2017; Manning et al., 2018). The narrow range of values at threshold typically provided by QUEST can make the later fitting of a psychometric function, which requires a range of values above and below threshold (Wichmann & Hill, 2001), problematic. To overcome this issue for the purposes of our later analysis by increasing the range of intensities presented, jitter was added to the QUEST threshold estimate for the next trial by adding ¼ of the estimate value multiplied by a random number from -1 to 1.

Note that distance from the screen was not fixed for all visual field target locations. In order for the extent of crowding to be measured at the higher eccentricities tested, the observers had to be brought closer to the screen to avoid the crowded stimulus array being limited by the screen boundaries. For 0°, 2°, and 4° eccentricity observers were tested at 280 cm from the screen, excluding two observers with amblyopia who were tested at 214 cm in order to measure crowding at 4° eccentricity. For 8° and 12° eccentricity, observers from both groups were tested at a distance of 120 cm from the screen.

3.2.1.3.1 Behavioural Data Treatment

For each observer, the data from the acuity and crowding tasks were pooled across blocks for each visual field location. For each visual field location, the Landolt-C gap size values presented were identified and the proportion of correct responses to each value was calculated. A psychometric function (weighted cumulative Gaussian) was fitted to the Landolt-C gap size values and their corresponding proportions of correct responses. The fit of the psychometric function, determined by the least squared error (LSE) between the data points and the fit, was weighted by the number of trials tested per gap size value, and had three free parameters: a mean, a standard deviation, and a lapse rate (key press error). From the psychometric function, threshold was derived as the gap size of the Landolt-C that yielded 62.5% correct performance. For acuity, this was the measure of interest, and analyses on gap size thresholds in minutes of arc (arcmins) are presented in the Results 3.3.1.1. For crowding, the measure of interest was the spatial extent over which the flankers interfered with performance. As we scaled the size of the target and thus the flankers using QUEST, we were able to also vary the separation between them. Therefore, as the crowding extent we took the centre-to-centre separation between the target and a flanker element. The spatial extent of crowding C in degrees of visual angle for each observer i was computed with the following formula:

$$C_i = 5t \times s$$

Where t is the gap size threshold (5 times the gap size is equal to the diameter of the Landolt-C), and s is the relative separation between the target and a flanker, which was set at 1.1 times the target size (as described in 3.2.1.5).

For three observers with amblyopia, there was very little variation in the gap size values tested, as repeated incorrect responses resulted in the presentation of the maximum gap size values set for QUEST. Note that the maximum gap size for the QUEST was set so that stimuli did not overlap with fixation or subtend over to the lower visual field. The concentration of gap size values at the maximum size meant the psychometric function computed gap size thresholds that resulted in estimates of crowding extent greater than the eccentricity tested (i.e. a crowding extent of 2.5° for a crowded target presented at 2° eccentricity). For these cases, the estimate from the psychometric function was replaced by the eccentricity, which represents the maximum stimulus size that would not interfere with the fixation dot. This adjustment was applied to three observers with amblyopia at specific

eccentricities. Two observers required one adjusted crowding extent measurement each, one at 2° in the temporal visual field, and the other at 4° in the nasal, whereas the third observer required three (2° in the temporal visual field, and 8° in both visual fields).

3.2.1.4 Eye-Tracking

Monocular eye-movements were recorded during the behavioural testing. Eye-calibration was performed in the start of the behavioural testing session. Subsequently calibration occurred every 4 QUEST blocks, unless there was significant head/body movement or eye drift (in which case calibration was performed immediately). A custom-coded calibration sequence was performed, consisting of the presentation of white circles on a grey background. The targets were randomly presented at five possible locations on a cross (centre, top, bottom, left, and right). Target size was increased if required for observers with amblyopia in order to ensure visibility.

Eye-tracking was used to ensure that observers fixated on the Gaussian fixation dot for peripheral stimulus presentation, and thus that acuity and crowding were measured for the desired eccentricities. A criterion distance (in degrees of visual angle) was used to define the area around the fixation dot over which the observer's gaze was could vary. In the start of a given trial, the target was presented only if the observers' gaze was within the window defined by the criterion distance. During stimulus presentation, if the observer's gaze deviated from the fixation dot for a distance greater than the criterion, then the trial was considered invalid. The observer was notified of the trial cancellation by a beep sound, and the trial was shuffled to the end of the block of trials in order to be repeated. The target did not appear for the repeated trial unless the observer's gaze had returned within the criterion-defined window. Invalid trials were repeated until 65 valid trials were completed per QUEST run.

The criterion distance for fixation was defined individually for each observer. For the control group, the radius of the criterion distance ranged from 2°-4° of visual angle (M=2.8°) across eccentricities. For the group with amblyopia, the criterion distance ranged from 2°-5° (M= 3.2°). The group with amblyopia included observers that required greater criterion distances due to increased fixation instability, a characteristic of amblyopia (Gonzalez, Wong, Niechwiej-Szwedo, Tarita-Nistor, & Steinbach, 2012; Subramanian, Jost, & Birch, 2013). However, it is clear that even

observers with typical vision required considerably large criterion distances. This was due to the poor accuracy of the calibration due to the large distance of the observer from the screen. The greater the distance from the screen on which the calibration targets are presented, the smaller the separation from one calibration target to the next in degrees of visual angle. When separation between targets is small, it becomes more likely for an eye position that does not correspond to the target to be picked up as a fixation sample. Calibrations were more prone to error for observers with amblyopia, as when fixating a target monocularly they show microsaccades of greater frequency and amplitude (Ciuffreda, Kenyon, & Stark, 1979), drift of increased of increased frequency and amplitude (Ciuffreda, Kenyon, & Stark, 1980), and greater variability of eye positions during fixation (Schor & Hallmark, 1978). Due to imprecise calibration, small criterion distances thus resulted in repeated trial cancellations while the observer was maintaining steady fixation as determined by the experimenter through the inspection of the real-time output of eye coordinates provided by Eyelink. To avoid increased experimental duration and observer fatigue, if the accuracy of the calibration could not be improved, the criterion distance was increased while eye coordinates were monitored throughout the experiment to ensure observers maintained fixation.

3.2.1.4.1 Eye-tracking Data Treatment

Post eye-tracking data collection, the X and Y positions (Cartesian coordinates) of the eye on each valid trial were converted to distance from the fixation dot in degrees of visual angle. This conversion was computed as the first step in determining fixation variability within the criterion distance for each observer. Next, eye blinks (pupil size equal to 0) and saccades (when eye velocity exceeded 3 standard deviations from a running mean) were detected for each valid trial, and the X and Y positions corresponding to a blink or saccade were removed. Eye blinks and saccades were removed because they are not indicative of fixational eye movements. Recordings of duration of ± ~30-50ms of eye blinks and saccades were also removed, as they could involve the beginning or end of an eye blink or saccade. To ensure the removal of voluntary saccades but not microsaccades, which are part of involuntary eye-movements during fixation of a target, the minimum duration threshold for a saccade was set to 15ms (Rolfs, 2009). The resulting X and Y eye distances from fixation were pooled across QUEST runs for each visual field location. Finally, the average standard deviation for that visual field

location across all valid trials was computed as a measure of fixation variability. For each observer, this resulted in one estimate of fixation variability for the X axis and one estimate for the Y axis per visual field location; these estimates are referred to in the following sections as horizontal and vertical fixation variability, respectively. The reason for computing fixation variability along these two axes was twofold. First, horizontal and vertical variability were chosen for ease of data manipulation, as the Eyelink output provides eye position along the X and Y axes. Secondly and most importantly, ignoring X and Y dimensions and computing eye position as absolute distance from the centre of the fixation dot instead would ignore the angle of deviation. For example, an eye movement from 2° eccentricity on the left of the centre to 2° on the right would result in 0 variability. Analyses on horizontal and vertical fixation variability, including a comparison between the amblyopia and the control group, are presented in the Results section 3.3.1.3 below.

3.2.2 Neuroimaging

3.2.2.1 Observers

The same observers took part in the neuroimaging part as in the behavioural part. See section 3.2.1.1 for details.

3.2.2.2 Stimuli

The stimuli were generated using MATLAB (Mathworks) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and displayed on a back-projection screen in the bore of the magnet via an LCD projector. The screen size was 42 × 24 cm with a resolution of 1920 × 1080. Observers viewed the back projection by means of a mirror mounted on the head coil at a distance of 63 cm from the screen.

Stimuli consisted of a drifting bar aperture with a dynamic, high-contrast pattern carrier (Figures 3.2A & 3.2B). The carrier was comprised of square tessellated blocks, with each block containing a drifting ripple-like pattern of concentric shapes that varied across time in spatial frequency and phase. The motion in neighboring blocks varied in a checkerboard-like fashion between expansion and contraction. While this pattern carrier differs from the conventional checkerboard design used in the literature (Wandell et al., 2007), it was chosen due to the motion energy and varying spatial frequencies that ensured maximal

stimulation of visually responsive neurons. As described in Schwarzkopf, Anderson, de Haas, White, and Rees (2014) and Alvarez, De Haas, Clark, Rees, and Schwarzkopf (2015) the pattern is defined by the following function:

$$I(x,y) = \sqrt{x^2 + y^2} \cos\left\{\frac{2\pi\left(\sin\frac{\delta\pi x}{180} + \cos\frac{\delta\pi y}{180}\right)}{4} + \theta\right\}$$

Where I is the pixel intensity at screen location (x, y) relative to the screen centre (in Cartesian coordinates), and θ and δ are the phase and spatial frequency, respectively. The θ parameter varied across time from 0 to 4π in 72 equal steps of 32ms duration, completing one cycle every 1.15s. The δ parameter was a function of θ:

$$\delta = \frac{\sin\theta}{4} + \frac{1}{2}$$

Pixel intensities were adjusted so that all positive values were set to white and all zero and negative values were set to black, presented against a grey background.

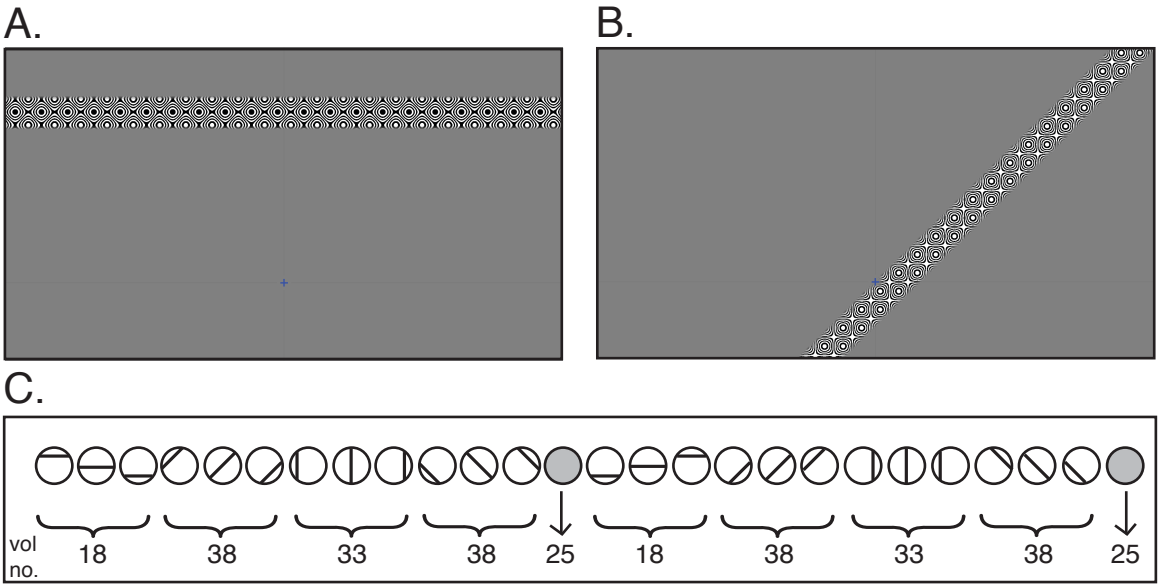

Figure 3-2 Illustrations of the stimuli used for pRF mapping

A. Image illustrating one volume (1 TR) for which the bar moved horizontally. The fixation cross located at the bottom half of the screen is also visible.
B. Image illustrating one volume (1TR) for which the bar drifted diagonally.
C. Schematic of stimulus change within a single fMRI scan. The bar aperture is represented by a line. For ease of illustration, the screen is represented as a circle. Three circles correspond to each orientation, to illustrate the direction of the bar. The grey circles indicate blank periods. Under each orientation or blank period, the number of volumes are denoted (vol no).

During the functional scans, the bar (2.28° in width) containing the ripple pattern moved along the visual field in steps, one step per fMRI image (volume) acquired. One functional scan consisted of 304 volumes. On each scan, first the bar

drifted along at four different orientations (0°, 45°, 90°, 135°). The duration of each bar step was 1s. Due to the rectangular shape of the screen, a different number of volumes was obtained per orientation: 18 volumes for 0°, 38 volumes for 45°, 33 volumes for 90°, and 38 volumes for 135°. This was followed by a blank period of 25 volumes where the gray screen with the fixation cross was presented. Subsequently, the direction of motion of the bar was reversed, and the bar drifted along at the four orthogonal orientations to the ones presented previously (180°, 225°, 270°, 315°) for the same number of volumes. Note that during presentation of the bar in the cardinal orientations (0° and 90°, 180° and 270°) the ripple pattern was not rotated, when in oblique orientations (45° and 135°, 225° and 315°) the pattern was rotated 45°. The functional scan ended with a second blank period of 25 volumes. Figure 3.2C shows a schematic of the bar steps through each of the four orientations in the two opposing motion directions during one functional scan. Each scan started with 10 blank volumes that were excluded from analyses to allow for the signal to reach equilibrium.

A fixation cross was presented in the centre of the screen horizontally, 5° above the bottom edge of the screen. The width of the fixation cross was 0.1°, whereas the length was 0.5°. For one observer with amblyopia the length of the fixation cross was increased to ensure visibility. This allowed for stimulus presentation that covered 16° eccentricity in the upper visual field along the vertical meridian and 19° eccentricity along the horizontal meridian. Note that in volumes in which the aperture overlapped with the fixation cross, a circular region around the fixation cross with a diameter of 0.5° did not contain the aperture pattern, but remained gray matching the background. This was done to prevent any disruptions from observers fixating the cross, and thus resulted in the central 0.5° eccentricity not being stimulated by the aperture.

3.2.2.3 Fixation Task

Observers were instructed to continually focus on the fixation cross and complete a simple task to ensure that they maintained fixation and remained alert. The fixation cross interchanged between blue and purple, and the observers had to watch out for the purple colour changes and press a key when this occurred. The probability of the fixation cross changing colour was 0.01 on every frame. The colour change periods lasted 200ms each. The observers' key presses were

recorded and checked at the end of each functional scan to ensure they kept fixating on the fixation cross.

3.2.2.4 Data Acquisition

Functional and anatomical scans were acquired using a Siemens Avanto 1.5 T MRI scanner (Siemens, Erlangen, Germany). The observers were lying on their back with a 32-channel surface coil, with one eye occluded using a patch. Head position was fixed using a foam head rest. For the anatomical scan, A T1-weighted anatomical magnetization-prepared rapid acquisition with gradient echo (MPRAGE) image was acquired (TR = 2730 s, TE = 3.57 ms) with a resolution of 1mm isotropic voxels. For the functional images, only the posterior section of the head coil was used, leaving 20 effective channels. The anterior section had to be removed because it did not allow for a full view of the stimulus across the visual field, and did not permit accurate eye-tracking. Functional T2*- weighted multiband 2D echo-planar images were taken with a multi-band sequence (Breuer et al., 2005) (voxel size= 2.3 mm isotropic voxels, repetition time (TR) = 1s, echo time (TE) = 55ms, 36 slices, flip angle = 75°, acceleration factor = 4). Each fMRI time series was made of 304 measurements (i.e. volumes). Five fMRI scans per eye were collected in a 2-hour long session. To minimize the movements observers made in the scanner, five fMRI scans for one eye were completed first, before the patch was switched to the other eye for the remaining five scans to be completed. The order between monocular viewing conditions alternated, with half the observers in each group starting with the five scans for the fellow fixating eye/ dominant eye, and the other half with the amblyopic/ non-dominant eye.

3.2.2.5 Eye-tracking

During the functional scans, monocular eye movements were recorded using an Eyelink 1000 MRI compatible eye-tracker (SR-Research, Mississauga, ON, Canada) sampling at 500 Hz. The eye-tracker was placed at the bottom part of the posterior aperture of the magnet bore. The same custom-coded five-point calibration used in the behavioural part of the study was performed before the first fMRI scan for each eye. Experimenters monitored the observers' fixation during the scans through the real-time Eyelink output. If the Eyelink output was noisy or there was significant eye-drift, calibration was performed in between scans to correct for this.

The eye-tracking data from the functional scans were treated identically to the eye-tracking from the behavioural part of the study, described in 3.2.1.4.1. X and Y eye distances from fixation (in degrees) were pooled across the five fMRI scans that were conducted per eye, and the average standard deviation for each eye across the five fMRI scans was computed. Analyses comparing horizontal and vertical fixation variability between controls and observers with amblyopia are reported in 3.3.2.2.

3.2.2.6 fMRI Data Pre-Processing

Pre-processing of the fMRI data was conducted using SPM12 (Ashburner et al., 2012). Functional images were mean-bias corrected, then re-aligned, un-warped, and finally co-registered to the anatomical scan. All the above was carried out using the default parameters of the SPM software. Freesurfer (Fischl, 2012) was used to generate a 3D reconstruction of the grey-white matter surface (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999). All further analyses were conducted using the SamSrf MATLAB toolbox for pRF analysis (Schwarzkopf, de Haas, & Alvarez, 2018).

3.2.2.7 pRF Analysis

Following Dumoulin and Wandell (2008), a similar forward-modelling approach was adopted to estimate pRF parameters using the fMRI data and the position of the bar stimulus in the visual field. The pipeline for the pRF modelling can be seen in Figure 3.3.

1. *Model creation:*

The model rested on the assumption of a simple Gaussian receptive field. The BOLD response of each voxel was modelled as a two-dimensional Gaussian g(x, y) defined by three parameters, x_0, y_0 and σ:

$$g(x,y) = e^{\wedge}\left(\frac{(x-x_0)^2 + (y-y_0)^2}{2\sigma^2}\right)$$

Where (x_0, y_0) is the centre of the pRF and σ is the standard deviation of the pRF, often called the spatial spread (in degrees of visual angle).

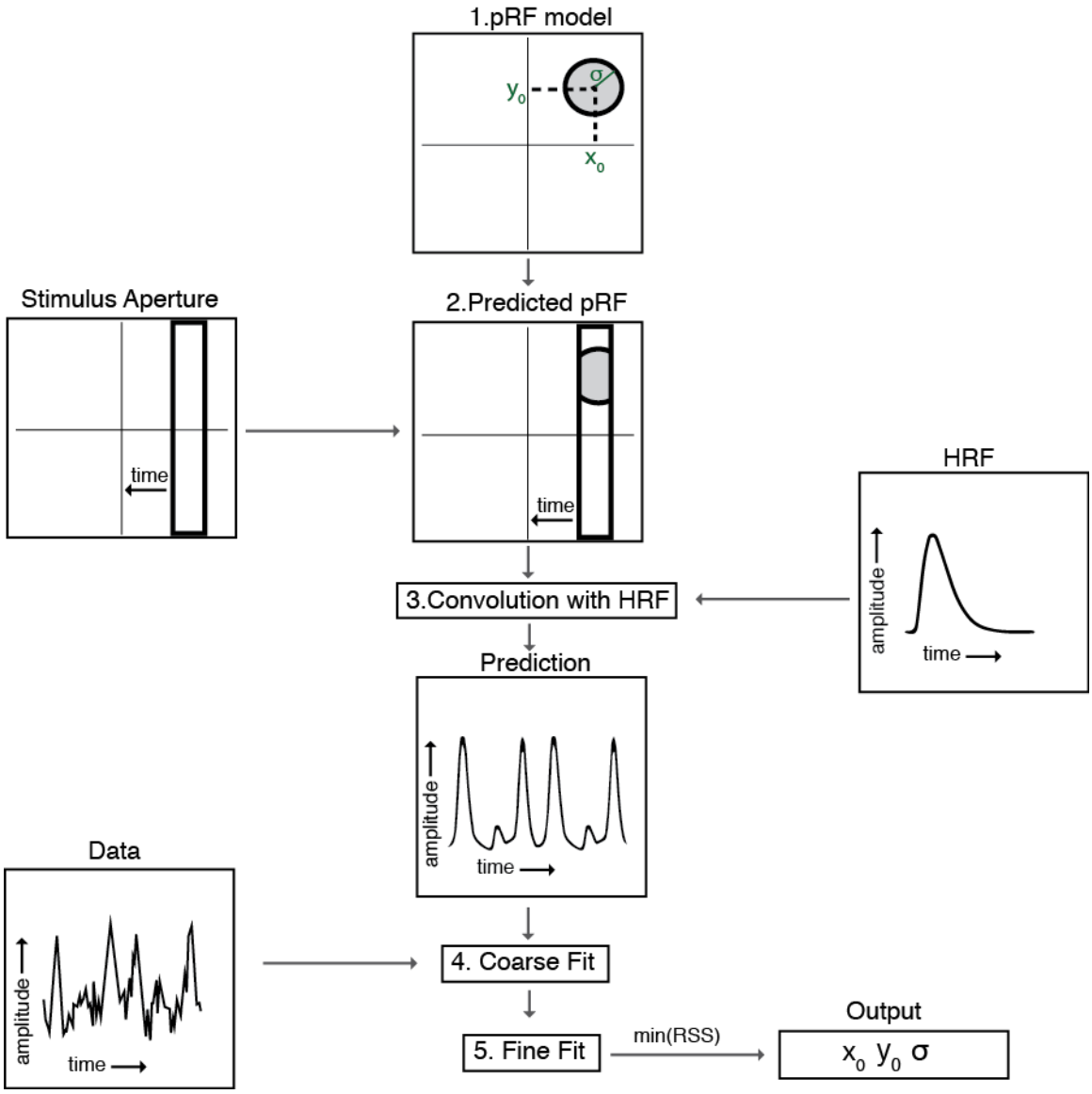

Figure 3-3 Pipeline of population receptive field (pRF) modelling.

1. Model creation: We assumed each pRF was a simple Gaussian with a peak located at (x_0, y_0) and a standard deviation σ.
2. Predicted pRF response: The stimulus aperture was convoluted with the pRF model to predict the response of the underlying neuronal population within each vertex.
3. Convolution with HRF: The model responses were convoluted with the canonical Haemodynamic Response Function (HRF).
4. Coarse Fit: An exhaustive grid search was conducted for the parameters that provided the highest correlation between the observed BOLD time course for each vertex ('Data') and the predicted time series ('Prediction').
5. Fine Fit: The best-fitting parameters from the coarse fit were used in an optimisation procedure. The optimal values for the parameters (x_0, y_0, σ)were determined by reducing the residual sum of squared errors (RSS)

2. *Predicted pRF Response:*

The linear overlap between the pRF model and the stimulus across time was used to predict the response of the underlying neuronal population at each vertex on the

cortical surface map (see Figure 3.3 step 2). The stimulus was defined in terms of the dimensions of the bar aperture and the moment in time (i.e. each 200 × 200 pixel frame of the aperture image corresponded to one of 304 TRs in one fMRI scan). The model predictions were based on the stimulus time course (by coding each vertex as stimulated or not, for a given point in time) and spatial sensitivity according to the assumed pRF parameters.

3. *Convolution with HRF:*

To compare model predictions with observed BOLD time course, the model predictions were convolved with the Haemodynamic Response Function (HRF). A canonical HRF was used to estimate parameters. The canonical HRF used was an average of individual HRFs (n=26) from de Haas, Schwarzkopf, Anderson, and Rees (2014). The use of a canonical HRF should not have affected our analyses, as previous studies (Dumoulin & Wandell, 2008; van Dijk, de Haas, Moutsiana, & Schwarzkopf, 2016) have shown that canonical and individual HRFs yield very similar results for pRF analyses.

4. Coarse fit:

An exhaustive grid search was then conducted for the set of the parameters providing the highest correlation between the observed BOLD time course for each vertex and the predicted time series. For this step, spatial smoothing was applied to an inflated spherical model of the cortical surface using a Gaussian kernel with a full-width at half maximum (FWHM= 5mm). The spatial smoothing served the purpose of increasing the signal-to-noise ratios by removing noise of high spatial frequency, and applying a spatial correlation between voxels. Using a 3D search space comprising of 15 × 15 × 34 combinations of location (x, y) and size (σ), the best fitting prediction was determined for each vertex by computing the Pearson correlation between the time series at that vertex and the search grid.

5. *Fine fit:*

The best fitting parameters from the coarse fit were used as initial values for an optimisation procedure. For this step unsmoothed data were used. The optimal parameter values were determined by reducing the residual sum of squared errors (RSS) between predicted and observed time series using a simplex-based method (Nelder & Mead, 1965; Lagarias, Reeds, Wright, & Wright, 1998). This fine fitting

stage also included a scaling parameter to estimate the overall response strength in addition to the three pRF parameters.

3.2.2.8 Definition of Visual Areas

pRF centre parameters values (x_0, y_0) were transformed into polar angle and eccentricity, colour-coded and projected on the inflated cortical surface of individual hemispheres using the SamSrf MATLAB toolbox (Schwarzkopf et al., 2018). The boundaries of retinotopic regions were delineated manually based on reversals in the polar angle map according to standard criteria (DeYoe et al., 1994; Sereno et al., 1995; Engel et al., 1997), assisted by eccentricity maps for identification of the foveal representation. Visual areas V1, V2, and V3 in the ventral stream were selected because our stimuli covered primarily the upper visual field, and there is a strong retinotopic bias with predominantly upper visual field representations in the ventral early visual areas (Kravitz, Saleem, Baker, Ungerleider, & Mishkin, 2013). Maps resulting from the monocular stimulation of the fellow fixating eye were used to delineate regions of interest for the observers in the amblyopic group, and maps resulting from stimulation of the dominant eye were used for the control group. After ensuring the polar angle maps between the two eyes were broadly similar, the boundaries from these delineations were then applied to the maps resulting from stimulation of the amblyopic eye and the non-dominant eye. Only vertices with a model fit of $R^2 > 0.1$ were included in any further analyses.

3.3 **Results**

3.3.1 Behavioural Testing

3.3.1.1 Acuity

Acuity gap size thresholds for the control group and the group with amblyopia for the nine visual field locations (fovea, 2°, 4°, 8°, and 12° in the nasal and temporal visual fields) are presented in Figure 3.4A. It is clear that thresholds increased with eccentricity in the nasal and temporal visual fields for both groups. The group with amblyopia had higher thresholds on average for all visual field locations, and showed more variability than the control group. To statistically compare acuity between the two groups, I conducted separate analyses on the foveal data. This was done in order to factor in the effect of visual field (nasal vs

temporal) when comparing the two groups for acuity measured in the visual periphery.

Figure 3.4A shows that acuity thresholds in the fovea were greater for observers with amblyopia than controls, with all observers with amblyopia having a value above the range of controls. An independent samples t-test (one-tailed, Bonferroni-corrected) confirmed the difference between the two groups in foveal acuity [t(8.07)= -2.42, *P*=.002; equal variances not assumed]. This demonstrates that the sample of observers with strabismic amblyopia had a significant reduction in acuity (i.e. elevated acuity thresholds) compared to control observers with unaffected vision, the defining characteristic of amblyopia. In fact, Figure 3.4B shows acuity thresholds for individual observers with amblyopia as a multiple of the mean acuity threshold of the control group. It is clear that the difference in acuity in the fovea between the two groups was large, being on average more than 4 times that of controls. Although this effect was magnified by one observer whose acuity was 16 times larger than the control mean, all observers with amblyopia had acuity thresholds substantially larger than the mean of observers with typical vision. Specifically, one observer had a foveal acuity threshold 7× control mean, two observers 5×, and the remaining five observers between 1.8× and 2×.

The difference in acuity between the two groups persisted in the periphery, with observers with amblyopia having overall higher thresholds than controls in all peripheral visual field locations. In order to compare acuity in the visual periphery between the two groups, a 4×2×2 mixed effects ANOVA was conducted, with eccentricity (2°,4°,8°,12°) and visual field (nasal and temporal) as the within subjects factors, and group (control and amblyopic) as the between subjects factor. Prior to conducting this analysis, the required assumptions were checked. Although as can be seen in Figure 3.4A, two or three observers with amblyopia had substantially larger thresholds at each eccentricity compared to the rest of the observers in the amblyopic group, these cases were not removed from the analyses as outliers, as they represent the richness typically found in clinical data. Additionally, there was greater variability in acuity thresholds at peripheral visual field locations in the amblyopia than in the control group (see SD for the two groups in Table 3.2 below). This violated the assumption of sphericity for eccentricity, and thus Greenhouse-Geisser corrections were used. The remaining assumptions for mixed effects ANOVA were met.

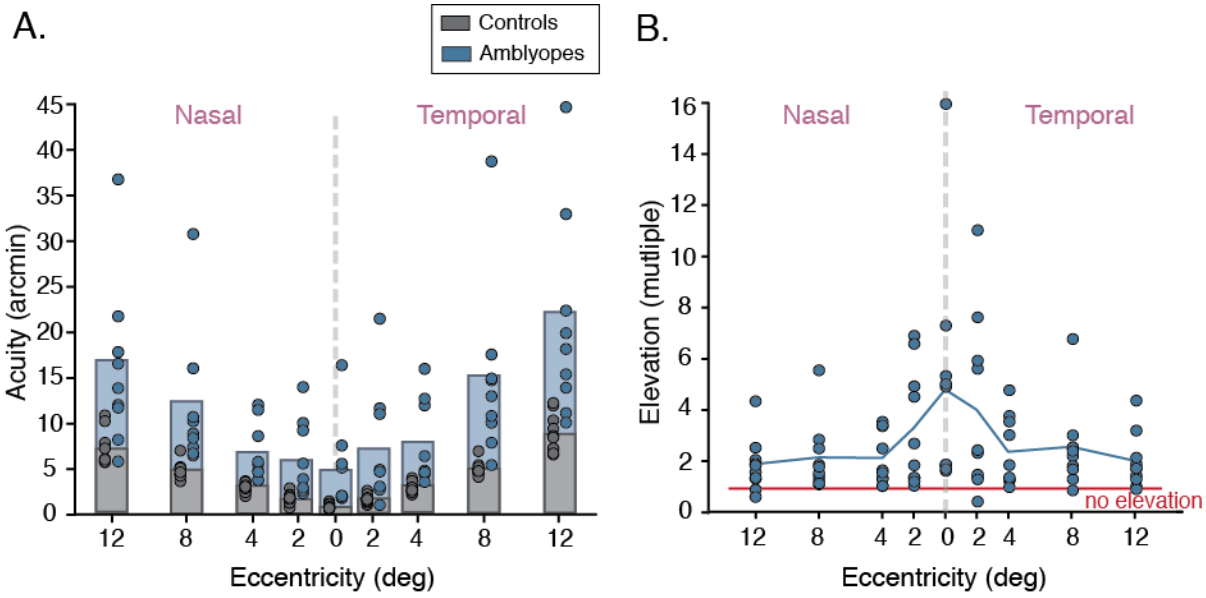

Figure 3-4 Acuity thresholds across the visual field in observers with amblyopia and controls

A. Acuity thresholds (in arcmins) at the nine visual field locations tested (fovea and 2°,4°,8°,12° in the nasal and temporal visual fields). The grey bar indicates the mean acuity for the control group (N=10), and the blue bar indicates the mean acuity for the amblyopia group (N=9). Each dot indicates the acuity threshold for an individual observer at that visual field location. The grey dashed line separates the nasal (left) and temporal (right) visual fields.
B. Elevations in acuity across the visual field for observers with amblyopia. Elevations are computed by dividing the individual thresholds of observers with amblyopia by the average acuity of controls at each visual field location, giving a measure of multiples of the average acuity in unaffected vision. Dots indicate individual multiples and the line indicates the mean for each visual filed location. The red line represents the average control acuity in controls, and any data point above it is indicative of an elevation. The grey dashed line separates the nasal (left) and the temporal (right) visual fields.

The analysis revealed a significant main effect of eccentricity [$F(1.4, 23.75)= 53.99$, $P< .00001$; Greenhouse-Geisser corrected], consistent with the well-documented increase acuity thresholds with eccentricity (Sloan, 1968; Rovamo & Raninen, 1990). There was also a main effect of visual field [$F(1,17) = 7.55$, $P= .014$], with acuity thresholds being higher in the temporal visual field (M= 7.90', SD= 1.06') compared to the nasal visual field (M=6.58', SD= .81'). There was also a significant interaction between eccentricity and visual field [$F(1.5, 25.62)= 4.65$, $P= .027$, Greenhouse-Geisser corrected], suggesting that the effect of visual field was dependent on eccentricity. In fact, Figure 3.4A indicates that the difference in acuity thresholds between the nasal and temporal visual field was larger for 8° and 12° eccentricity, where thresholds were greater in the temporal visual field even for controls (see also Table 3.2).

Of integral importance to the aims of this investigation was the difference in acuity in the peripheral visual field between the two groups. The analysis revealed a significant main effect of group [$F(1,17) = 9.54$, $P= .007$], demonstrating that the

amblyopic eye of observers with amblyopia showed a significant elevation in acuity thresholds compared to controls across the peripheral visual field. There was also a significant interaction between eccentricity and group [$F(1.4, 23.75) = 4.73$, $P = .029$], indicating that the differences in acuity between groups were not constant across eccentricities. Figure 3.4A suggests that this interaction could be driven by the difference in acuity between the two groups being reduced further into the periphery, as at higher eccentricities more observers with amblyopia fall within the range of acuity values for controls. Indeed, this reduction in the difference between the two groups is clear when acuity for observers with amblyopia is expressed as a multiple of the mean of controls (Figure 3.4B): at 2° eccentricity amblyopic acuity was more than 3 times the average of the control group, but dropped to 2 times the control average at 4° and 8° eccentricity, and less than 2 times at 12° eccentricity. The analysis did not reveal a significant interaction between visual field and group [$F(1,17) = 3.60$, $P = .075$], or between eccentricity, visual field, and group [$F(3, 25.62) = .83$, $P = .417$].

		12n	8n	4n	2n	0	2t	4t	8t	12t
Controls	M	7.35	4.83	2.93	1.77	.90	1.71	2.91	5.01	8.89
	SD	1.88	.90	.51	.57	.25	.45	.57	.77	2.08
Amblyopes	M	14.22	10.38	5.92	5.24	4.34	6.22	6.78	13.09	18.57
	SD	8.13	6.85	2.85	3.70	1.42	5.82	4.08	8.65	10.00

Table 3-2 Mean (M) and standard deviation (SD) of acuity thresholds.

Mean and standard deviation values presented in minutes of arc for the control group (dominant eye) and the group with amblyopia (amblyopic eye) for the nine visual field locations tested.

To further explore the differences in acuity between controls and observers with amblyopia, planned post-hoc independent samples t-tests (one tailed, Bonferroni-corrected including foveal acuity) were conducted comparing acuity thresholds at each peripheral visual field location. Due to the differences in variability between the two groups, Levene's test for the equality of variances was significant for all peripheral visual field locations, and values assuming unequal variances are reported. None of the t-tests yielded significant results. The differences in acuity were not significant in the nasal visual field at 2° [$t(8.35) = -2.78$, $P = .112$], 4° [$t(8.46) = -3.10$], 8° [$t(8.25) = -2.41$, $P = .164$], and 12° eccentricity [$t(8.77) = -2.47$, $P = .144$]. Similarly, the differences between the two groups were not significant in the temporal visual field at 2° [$t(.809) = -2.32$, $P = .196$], 4° [$t(8.28) = -$

2.82, *P*= .088], 8° [t(8.12)= -2.80, *P*= .092], and 12° eccentricity [t(8.62)= -2.85, *P*= .08]. As it is clear from Figure 4A that the group with amblyopia had greater acuity thresholds for all peripheral visual field locations, the lack of significant results in these comparisons was likely due a combination of the increased variability in the amblyopic group and the strict Bonferroni corrections. Note that the comparison of foveal acuity between the control and amblyopic group reported above was also included in these corrections, but was found to be significant as the difference between the two groups was substantially larger.

Overall, these results indicate that for both controls and observers with amblyopia, acuity thresholds increase with eccentricity, with this increase being greater in the temporal compared to the nasal visual field. Observers with amblyopia showed elevated acuity thresholds relative to controls for all visual field locations. However, this elevation was not uniform across the visual field. Particularly, the amblyopic fovea showed the largest elevation in acuity thresholds (and thus the greatest acuity deficit) relative to unaffected vision, and this elevation was gradually reduced with increasing eccentricity.

3.3.1.2 Extent of Crowding

Figure 3.5A shows the extent of crowding for the control group and the group with amblyopia at the nine visual field locations tested. It is clear that the extent of crowding increased with eccentricity for both groups. However, the amblyopic group had larger extents of crowding on average. Similarly to the acuity results reported above, in order to compare the extent of crowding between the two groups, analyses were conducted independently for the fovea and the periphery.

The extent of crowding in the fovea of controls was minimal (M =0.11° or 6.6 arcmins) or and thus barely noticeable in Figure 3.5A, whereas for observers with amblyopia it was substantially larger, exceeding 1° of visual angle on average (see Table 3.3 for descriptive statistics). To statistically compare the crowding extent in the fovea between the two groups, an independent samples t-test was conducted (one-tailed, Bonferroni-corrected), that revealed that the difference between the two groups was significant [t(8.01) = -2.73, *P*= .013]. It is clear from Figure 3.5A that the fovea was the visual field location for which the group with amblyopia showed the greatest elevation in the extent of crowding relative to the control group. Indeed, when expressed as a multiple of the average extent of crowding of the control group (Figure 3.5B), the extent of crowding in the fovea of observers with amblyopia was

11 times the average of controls. The magnitude of this difference in the fovea was in part due one observer having an extent of crowding equal to 31 times the mean extent of the control group. However, the difference from the control group was still large for the rest of the group with amblyopia, for two observers being approximately 21× control mean, for one observer 7×, for two observers 4×, and for the remaining three between 2 and 2.6. In fact, the fovea was the visual field with the greatest difference in the extent of crowding from the mean of controls for all observers with amblyopia. Note that this was consistent with the acuity results reported above, that also indicated that the fovea was the visual field location with the greatest acuity deficit.

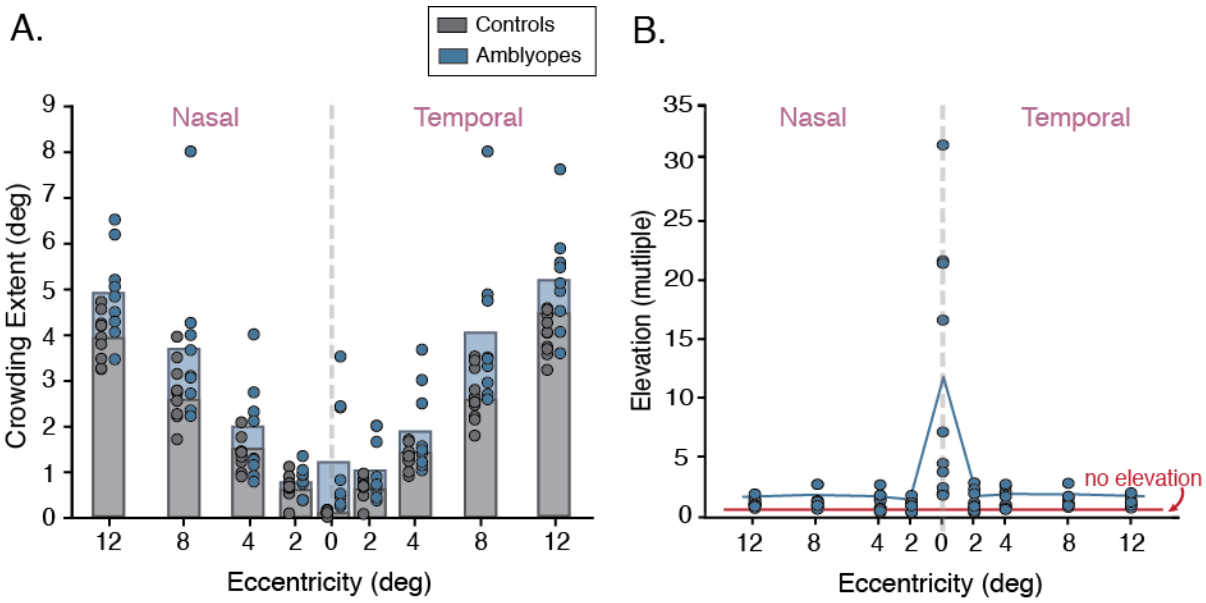

Figure 3-5 Extent of crowding across the visual field in observers with amblyopia and controls

A. Extent of crowding (in degrees) at the nine visual field locations tested (fovea and 2°,4°,8°,12° in the nasal and temporal visual fields). The grey bar indicates the mean crowding extent for the control group (N=10), and the blue bar indicates the mean crowding extent for the amblyopia group (N=9). Each dot indicates the extent of crowding for an individual observer at that visual field location. The grey dashed line separates the nasal (left) and temporal (right) visual fields.
B. Elevations in the extent of crowding across the visual field for observers with amblyopia. Elevations are computed by dividing the individual values for crowding extent of observers with amblyopia by the average crowding extent of controls at each visual field location. This gives a measure of multiples of the average crowding extent in unaffected vision. Dots indicate individual multiples and the line indicates the mean for each visual filed location. The red line represents the average extent of crowding in controls, and any data point above it is indicative of an elevation. The grey dashed line separates the nasal (left) and the temporal (right) visual fields.

The difference in the extent of crowding between the two groups persisted in the periphery, with the group with amblyopia having greater extents of crowding on average than the control group. However, it is worth noting that despite this difference, the majority of observers with amblyopia had an extent of crowding

within the range of controls. In order to compare the differences between controls and observers with amblyopia, a 4 (eccentricity) ×2 (visual field) ×2 (group) mixed effects ANOVA was conducted. The analysis revealed a significant main effect of eccentricity [F(2.05, 34.85) = 84.96, *P*<0.0001, Greenhouse-Geisser corrected], indicating that for both groups the extent of crowding increased with eccentricity. There was also a significant albeit small main effect of visual field [F(1,17) = 4.67, *P*= .045], suggesting a naso-temporal asymmetry in the extent of crowding. This asymmetry was in the same direction as the one reported for acuity above, with the temporal visual field having larger extents of crowding on average (M= 2.58°, SD= 1.75°) than the nasal visual field (M= 2.49°, SD= 1.69°). The interaction between eccentricity and visual field was not significant [F(2.43, 41.22)= .654, *P*= .553; Greenhouse-Geisser corrected].

		12n	8n	4n	2n	0	2t	4t	8t	12t
Controls	M	3.97	2.77	1.40	.68	.11	.66	1.33	2.68	4.02
	SD	.52	.65	0.34	.26	.005	.24	.25	.58	.45
Amblyopes	M	4.89	3.70	1.84	.76	1.21	1.04	1.88	4.01	5.19
	SD	.98	1.76	1.05	.34	1.24	.66	.94	1.70	1.17

Table 3-3 Mean (M) and standard deviation (SD) of the extent of crowding

Crowding extent values for the control group (dominant eye) and the group with amblyopia (amblyopic eye). The extent of crowding was measured as the centre-to-centre separation between the target and a flanker. Values are presented in degrees of visual angle.

The analysis also revealed a main effect of group [F(1,17) = 74.69, *P*< .00001], indicating that observers with amblyopia had larger extents of crowding in their amblyopic eye compared to the dominant eye of controls. However, the analysis also yielded a significant interaction between eccentricity and group [F(2.05, 34.85) = 49.85, *P*< .0001; Greenhouse-Geisser corrected], indicating that the difference between the groups was not constant across eccentricities. Indeed, Figure 3.5A suggests that there was some variability in the magnitude of the differences between the two groups depending on the eccentricity. For example, in the temporal visual field the difference in the extent of crowding between the two groups was noticeably smaller at 2° compared to 8° eccentricity. These small differences can also be seen in Figure 3.5B: at 2° eccentricity in the temporal visual field the extent of crowding was 1.2 times the average of controls, whereas at 8° eccentricity it was 1.5 times. The analysis did not show a significant interaction

between visual field and group [F(1, 17)= 2.45, *P*= .136], or between eccentricity, visual field, and group [F(3, 41.22)= .305, *P*= .77; Greenhouse-Geisser corrected].

In order to compare the extent of crowding between the dominant eye of controls and the amblyopic eye of observers with amblyopia at each peripheral visual field location, planned post-hoc independent samples t-tests were conducted (one-tailed, Bonferroni-corrected with the inclusion of the foveal comparison). Due to the differences in variance between the two eyes, Levene's test for the equality of variances was significant for 4° in the nasal visual field, and 2°, 4°, 8°, and 12° in the temporal visual field. Adjusted values assuming unequal variances are reported for those visual field locations. In addition to the significant difference in the extent of crowding between the two groups at the fovea, the t-tests revealed a significant difference between the control group and the group with amblyopia at 12° eccentricity in the temporal visual [t(17)= -2.96, *P*= .041]. For the remaining peripheral visual field locations, the difference in the extent of crowding between the control group and the group with amblyopia in the nasal visual field was not significant at 2° [t(17)= -.57, *P*>.99], 4° [t(9.53)= -1.18, *P*>.99], 8° [t(17)= -1.57, *P*= .603], or 12° eccentricity [t(17)= -2.61, *P*= .081], or in the temporal visual field at 2° [t(9.88)= -1.60, *P*= .0639], 4° [t(9.04)= -1.72, *P*= .054], and 8° eccentricity [t(9.66)= -2.32, *P*= .229].

Overall, the results demonstrate that in both controls and observers with amblyopia, the extent of crowding increased with eccentricity. Additionally, there was a small naso-temporal asymmetry for both groups, with the temporal visual field showing greater extents of crowding compared to the nasal visual field. Observers with amblyopia showed greater extents of crowding on average compared to controls at all the visual field locations tested, but similarly to acuity, the greatest elevation was found at the fovea.

3.3.1.3 Eye Movements

In order to ensure that observers fixated at the fixation dot during the acuity and crowding measurements, eye movements during stimulus presentation that were greater than the criterion distance led to the trial being deemed invalid (see Methods 3.2.1.4 for details). Although invalid trials did not contribute to QUEST estimates and were required to be repeated until valid, fixation variability could still be high for an observer whilst being within the criterion distance. As amblyopia is characterised by poor fixation stability (Gonzalez et al., 2012; Subramanian et al.,

2013), this raises the possibility that increased fixation variability during the *valid* trials could have contributed to the elevated acuity thresholds and increased crowding extents found for the amblyopic group in the previous sections 3.3.1.1 and 3.3.1.2. Therefore, it was important to rule out the influence of such a confounding effect in acuity and crowding measurements by comparing fixation variability during valid trials between observers with amblyopia and controls.

3.3.1.3.1 Eye Movements During Acuity Measurements

Based on my recordings of eye movements during acuity measurements at the nine visual field locations (fovea, 2°, 4°, 8°, and 12° in the nasal and temporal visual fields), I computed average fixation variability across the X and Y axes (see Methods 3.2.1.4.1). These measures are referred to as horizontal and vertical fixation variability respectively, and are plotted in Figure 3.6A and 3.6B. The group with amblyopia had greater horizontal and vertical fixation variability on average for most visual field locations compared to the control group, but this elevation was primarily driven by one or two observers with amblyopia. In order to statistically examine these differences, the same statistical tests were conducted as those described above for the behavioural measures of acuity and crowding.

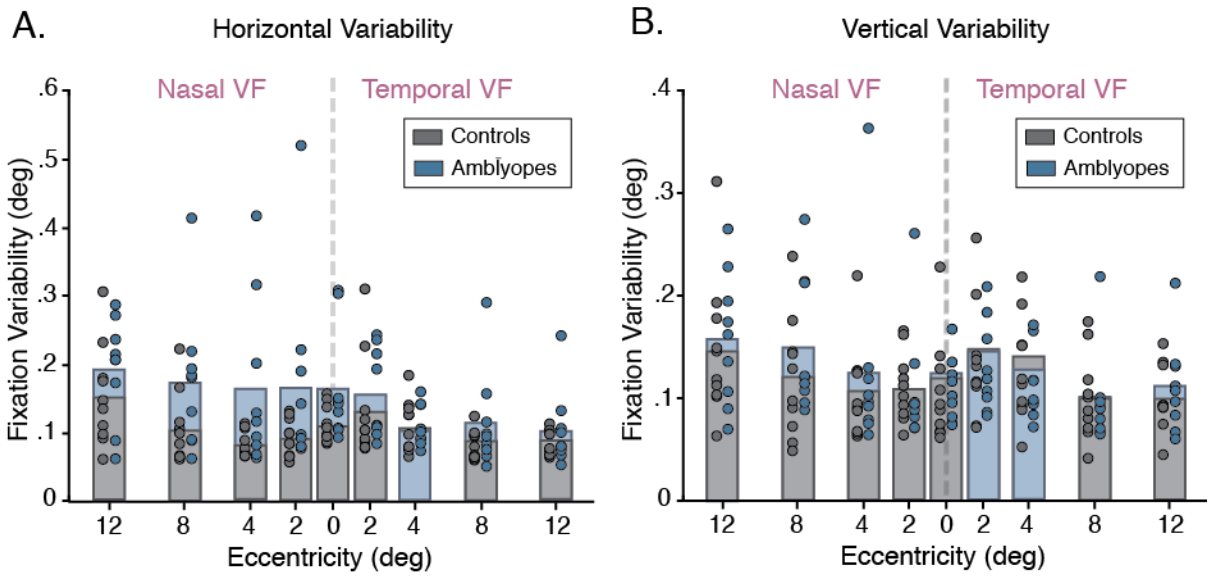

Figure 3-6 Fixation variability in the acuity task for observers with amblyopia and controls

A. Horizontal fixation variability (in deg) for the acuity task at the nine visual field locations tested (fovea and 2°,4°,8°,12° in the nasal and temporal visual fields) for the control group (grey; N=10) and the group with amblyopia (blue; N= 9). Bars indicate the average horizontal fixation variability for each group and dots indicate the fixation variability for each observer at that visual field location. Note that the bar that is presented at the front corresponds to the group with the lowest mean (e.g. at 8° in the temporal visual field the group with amblyopia has the lowest mean horizontal fixation variability). The dashed line separates the nasal (left) and the temporal (right) visual fields.
B. Vertical fixation variability (in deg) for the acuity task at the nine visual field locations tested for the control group (grey) and the group with amblyopia (blue). Plotted with the same conventions as in A.

I first consider horizontal fixation variability during the acuity task. For targets presented at the fovea, the group with amblyopia showed greater horizontal fixation variability on average than the control group (see Figure 3.6A). However, the majority of the observers with amblyopia showed horizontal fixation variability similar to that of observers in the control group for foveally presented targets, with the exception of two observers that showed greater fixation variability. An independent samples t-test comparing horizontal fixation variability in the fovea between the two groups revealed that the difference was not significant [t(9.59)= -1.96, *P*= 0.08; equal variances not assumed].

To compare horizontal fixation variability between the control group and the group with amblyopia for targets presented at peripheral visual field locations, a 4 (eccentricity) ×2 (visual field) ×2 (group) mixed effects ANOVA was conducted. As it is clear from Figure 3.6A, eccentricity did not have an effect on horizontal fixation variability, which appeared relatively constant across the peripheral visual field. This was confirmed by the ANOVA that did not show a main effect of eccentricity [F(1.62, 27.55)= 3.58, *P*= .05; Greenhouse-Geisser corrected]. However, Figure 3.6A points to differences between the nasal and temporal visual fields, with horizontal fixation

variability being greater for targets presented in the nasal visual field. This difference was confirmed by the significant main effect of visual field, [F(1,17)= 14.15, *P*= .002]. The greater fixation variability for the nasal visual field was particularly clear for the control group at 12° eccentricity, and for the group with amblyopia at all eccentricities. The analysis did not yield a significant interaction between eccentricity and visual field [F(2.05, 34.86) = .086, *P*= .922; Greenhouse-Geisser corrected].

Importantly, the differences in horizontal fixation variability between the control group and the group with amblyopia were not significant, as the main effect of group did not reach significance [F(1,17)=4.17, *P*= .057]. However, the group with amblyopia showed increased fixation variability and thus poorer fixation stability in the nasal visual field (M= .18°, SD= .021°) compared to the temporal visual field (M= .12°, SD= .11°). This difference between the visual fields was not as evident for the control group, which showed more balanced variability on average in the nasal (M= .11°, SD= .02°) and the temporal visual field (M= .10°, SD= .01°). Indeed, the analysis revealed a significant interaction between visual field and group [F(1, 17)= 9.02, *P*= .008]. There interaction between eccentricity and group was not significant [F(1.62, 27.55)= .262, *P*= .725; Greenhouse-Geisser corrected], neither was the interaction between eccentricity, visual field, and group [F(2.05, 34.86)= .815, *P*= .454; Greenhouse-Geisser corrected].

Identical analyses were conducted to compare vertical fixation variability in the acuity task between the two groups. For targets presented foveally, it appears from Figure 3.6B that observers with amblyopia showed very similar vertical fixation variability to controls. An independent samples t-test confirmed this, revealing no significant difference between the control group (M= .10°, SD= .05°) and the group with amblyopia (M= .11°, SD= .03°) in vertical fixation variability for foveally presented targets [t(17)= -.21, *P*= .835]. As the fovea was the visual field location for which observers with amblyopia showed the greatest elevation in acuity thresholds relative to controls, the lack of differences in horizontal and vertical fixation variability suggests that eye movements cannot account for this acuity elevation.

To compare vertical fixation variability for targets presented at peripheral visual field locations during the acuity task, a 4 (eccentricity) ×2 (visual field) ×2 (group) mixed effects ANOVA was conducted. As can be seen in Figure 3.6B, there was no effect of eccentricity on vertical fixation variability, as there was little variation across the visual field for both controls and observers with amblyopia. This

was supported by the ANOVA that revealed that the main effect of eccentricity was not significant [F(1.43, 24.34)= 3.54, P= .06; Greenhouse-Geisser corrected]. However, there was a trend for greater vertical fixation variability in the nasal visual field, an effect that was also observed for horizontal fixation variability (see Figure 3.6A). Indeed, the analysis revealed a main effect of visual field [F(1,17)= 5.72, *P*= .029], demonstrating that fixation stability was poorer on average in the nasal visual field (M= .13°, SD= .04°) compared to temporal visual field (M= .11°, SD= .03°). The interaction between eccentricity and visual field was not significant [F(3, 51)= .402, *P*= .752].

When considering differences between the control group and the group with amblyopia in vertical fixation variability, Figure 3.6B suggests that the two groups showed similar fixation variability across eccentricities and the analysis did not yield a main effect of group [F(1,17)= 0.24, *P*= .633]. Similarly, the analysis did not show a significant interaction between eccentricity and group [F(1.43, 24.34)= .016, *P*= .957, Greenhouse-Geisser corrected], between visual field and group [F(1, 17)= 2.21, *P*= .156], or between eccentricity, visual field, and group [F(3, 51)= 1.32, *P*= .277].

Overall, these results indicate that horizontal and vertical fixation variability did not differ between the control group and the group with amblyopia in the fovea, where the greatest difference in acuity between the two groups was found. For targets presented in the visual periphery, eccentricity did not have an effect on either horizontal or vertical fixation variability. However, when considering the effect of visual field, I found the *opposite* effect to acuity thresholds, with both horizontal and vertical fixation variability being greater in the nasal compared to the temporal visual field. Crucially, the control group and the group with amblyopia did not show overall differences in neither horizontal nor vertical fixation variability. The amblyopic group did however show greater horizontal fixation variability in the nasal visual field. I conclude that fixation variability cannot account for the differences in acuity found between controls and observers with amblyopia.

3.3.1.3.2 Eye Movements During Crowding Measurements

Horizontal and vertical fixation variability during measurements of the extent of crowding at the nine visual field locations tested can be seen in Figure 3.7A and 3.7B, respectively. It is clear from Figure 3.7A that for most visual field locations, horizontal fixation variability was on average greater for the group with amblyopia

compared to the control group. To further explore these differences, I conducted the same analyses reported above for eye movements during the acuity task.

For crowded targets presented foveally, Figure 3.7A shows that the group with amblyopia showed greater horizontal fixation variability on average compared to the control group. However, an independent samples t-test showed that this difference in horizontal fixation variability between the control group (M= .11°, SD= .02°) and the group with amblyopia (M= .17°, SD= .10°) was not significant (equal variances not assumed).

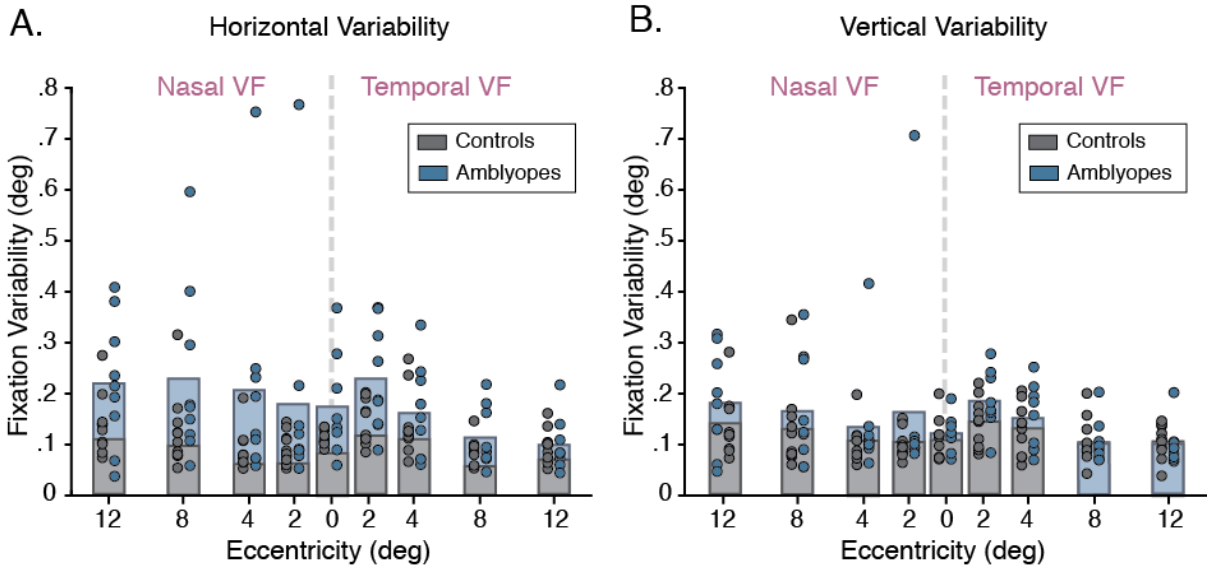

Figure 3-7 Fixation variability in the crowding task for observers with amblyopia and controls

A. Horizontal fixation variability (in deg) for the crowding extent task at the nine visual field locations tested (fovea and 2°,4°,8°,12° in the nasal and temporal visual fields) for the control group (grey; N=10) and the group with amblyopia (blue; N= 9). Plotted in conventions as in Figure 3.6
B. Vertical fixation variability (in deg) for the crowding task at the nine visual field locations tested for the control group (grey) and the group with amblyopia (blue). Plotted with the same conventions as in A.

For crowded targets presented at peripheral visual field locations, a 4 (eccentricity) × 2 (visual field) × 2 (group) mixed effects ANOVA was conducted. As can be seen in Figure 3.7A, the eccentricity of the crowded target does not appear to affect horizontal fixation variability for either observer group. This was confirmed by the analysis that did not show a significant main effect of eccentricity [$F(1.46, 24.86) = 3.6$, $P = .055$; Greenhouse-Geisser corrected]. Although it appears from Figure 3.7A that horizontal fixation variability was greater for crowded targets presented in the nasal compared to the temporal visual field, the analysis did not yield a main effect of visual field [$F(1,17) = 3.67$, $P = .072$]. This is likely due to the

difference in horizontal fixation variability between the two visual fields being large for the group with amblyopia, but minimal for the control group.

It is clear from Figure 3.7A that the group with amblyopia showed greater fixation variability, and thus poorer fixation stability, than the control group. The analysis yielded a main effect of group [$F(1,17)= 4.82$, $P= .042$], suggesting that although this difference in horizontal fixation variability between the two groups was significant. Additionally, this difference appeared to depend on the visual field to which the crowded target was presented, as it is clear from Figure 3.7A that fixation variability was more similar between the two groups in the temporal visual field than the nasal. Indeed, the analysis revealed a small albeit significant interaction between group and visual field [$F(1, 17)= 4.91$, $P= .041$], demonstrating that horizontal fixation variability was significantly larger for the group with amblyopia compared to controls when the crowded targets were presented in the nasal visual field. The analysis did not yield a significant interaction between eccentricity and group [$F(1.46, 24.86)= .22$, $P= .736$; Greenhouse-Geisser corrected], or between eccentricity, visual field, and group [$F(1.51, 25.72)= .78$, $P= .434$; Greenhouse-Geisser corrected].

When considering vertical fixation variability for crowded targets presented at the fovea (Figure 3.7B), it was clear that the two groups showed very similar vertical fixation variability. Therefore, the independent samples t-test between the group with amblyopia (M= .12°, SD= .04°) and the control group (M= .11°, SD= .04°) was not significant [$t(17)= -0.737$, $P= 0.471$]. Note that the greatest difference in the extent of crowding between the two groups was found in the fovea. The lack of differences in both horizontal and vertical fixation variability indicates that the greater extent of crowding found for observers with amblyopia foveally cannot be explained by increased fixation variability relative to controls.

For crowded targets presented in the visual periphery, a 4 (eccentricity) × 2 (visual field) × 2 (group) mixed effects ANOVA was conducted. Figure 3.7B indicates that there was little variation in vertical fixation variability across eccentricities, suggesting that eccentricity did not affect vertical fixation variability. As such, the analysis revealed that the main effect of eccentricity was not significant [$F(1.42, 24.14)= 0.21$, $P= .080$;Greenhouse-Geisser corrected]. Similarly, although some observers from both groups showed substantially greater vertical fixation variability for crowded targets presented in the nasal visual field, overall there did not appear to be substantial differences between the nasal and temporal visual field.

This was supported by the results of the analysis, that did not yield a main effect of visual field [$F(1,17)= 1.52$, $P= .235$].

Although the group with amblyopia showed significantly greater horizontal fixation variability for crowded targets for most eccentricities relative to the control group, this difference was small. In fact, at 8° and 12° eccentricity in the temporal visual field, the control group showed greater vertical fixation variability on average than the group with amblyopia. The analysis thus showed that the main effect of group on vertical fixation variability was not significant [$F(1,17)= 1.85$, $P= 0.19$].

Similarly to the results on fixation variability during the acuity measurements, these results show that the eccentricity of the crowded target did not affect horizontal and vertical fixation variability. There was however a difference in horizontal but not vertical fixation variability between the control group and the group with amblyopia. Although this difference was small, the group with amblyopia had greater horizontal fixation variability relative to controls, with this difference being larger in the nasal visual field. The effect of visual field here was *opposite* to the results showing greater crowding extents for the temporal visual field, and thus cannot account for this effect. Importantly, fixation variability did not differ between the two groups in the fovea, where we observed the greatest difference in the extent of crowding.

3.3.1.4 Summary of Behavioural Results

The findings from behavioural measurements point to important variations in acuity and the extent of crowding across the visual field, as well as differences in these measures between control observers with unaffected vision and observers with strabismic amblyopia. For both control observers and those with amblyopia, acuity thresholds increased, and thus acuity was reduced, with eccentricity. There was also a small naso-temporal asymmetry in acuity, with the temporal visual field showing reduced acuity for both the control and the amblyopic group, especially for the higher eccentricities tested. When considering differences in acuity between typical and amblyopic vision, I found that acuity was reduced across the amblyopic visual field. However, this reduction was not uniform, with the amblyopic fovea showing the greatest reduction in acuity thresholds relative to controls. The difference between the two groups in acuity was reduced in the parafovea (i.e. 2° eccentricity), and even further reduced at the higher eccentricities tested.

The results on the extent of crowding followed a very similar pattern to acuity. For the control group, the extent of crowding increased with eccentricity. For observers with amblyopia, crowding also increased with eccentricity, but this increase was seen beyond 2° eccentricity, as the extent of crowding was very similar at the fovea and parafovea. Similarly to acuity, the extent of crowding showed a small naso-temporal asymmetry for both groups, with the temporal visual field having greater crowding extents. When I examined differences between the amblyopic and the control group in the extent of crowding, I found that the observers with amblyopia had increased crowding extents across the visual field. However, as was the case with acuity, this increase in the extent of crowding was non-uniform, but was disproportionately large in the fovea and substantially reduced in the periphery. The foveal elevation in crowding was greater that the foveal reduction in acuity, being 11× the average of observers with typical vison compared to 4×.

I then considered the effect of eye movements on these results. During acuity measurements, there was a tendency of observers to have greater fixation variability, and thus poorer fixation stability, for targets presented in the nasal visual field – this was especially true for observers with amblyopia. During the crowding task, observers with amblyopia had poorer fixation stability than controls. Similarly to acuity measurements, this was particularly evident for crowded targets presented in the nasal visual field. Note that in neither tasks did I find a difference between the two groups in fixation stability for foveally presented targets, where the greatest differences in acuity and crowding were found. It is clear that this pattern of fixation variability does not follow the pattern of acuity and crowding measurements across the visual field. Therefore, poor fixation stability cannot account for the differences in acuity and crowding between observers with typical vision and those with amblyopia.

3.3.2 Neuroimaging

In this section I consider the results from pRF mapping. Similarly to the acuity and crowding sections above, my aim here was to compare pRF size (i.e. the σ parameter) between controls and observers with amblyopia. Note that in contrast to the behavioural results where I had measures only for the dominant eye of controls and the amblyopic eye of observers with amblyopia, I obtained estimates of pRF size for both eyes. For observers in the control group, the dominant and non-

dominant eyes are referred to as DE and nDE, respectively. For observers with amblyopia, the amblyopic eye and fellow fixating eye are referred to as AME and FFE, respectively.

3.3.2.1 pRF Mapping

Figure 3.8 shows maps for polar angle (Figure 3.8A), eccentricity (Figure 3.8B), and pRF size (Figure 3.8C) for the DE and nDE of one example observer from the control group (left) and the FFE and AME of one example observer from the group with amblyopia (right). These maps were spatially aligned to a spherical template of the cortical surface of the left hemisphere. Reliable retinotopic organisation was evident in both groups of observers. For example, when stimuli were shown to the DE of the example control observer (Figure 3.8A), the upper section of V1 responded more strongly to the lower vertical meridian (green), the mid-section the horizontal meridian (blue), and the lower section represented the upper vertical meridian (red) which formed the border to area V2. In V2, the organisation then reversed back to the horizontal meridian that corresponded to the border between V2 and V3, and the anterior border of V3 was formed by a representation of the upper vertical meridian. Note that since the left hemisphere is presented in Figure 3.8A, the right half of visual space is encoded. V1, V2 and V3 shared a foveal confluence, and the central fovea was presented over a larger fraction of cortical surface than the peripheral visual field (Figure 3.8B). The same organisation was evident in the nDE of the same observer, as well as the both the FFE and the AME of the example observer with amblyopia. Overall, I did not observe any evident qualitative differences in terms of the macroscopic architecture of the visual regions between the two groups.

The primary aim in this section was to establish whether there were differences in pRF size in the retinotopic regions of interest between the DE of controls and the AME of observers with strabismic amblyopia. The behavioural deficits found in acuity and crowding for the AME relative to the DE of controls predict a difference in pRF size between the AME and DE, with larger pRFs for the AME. In order to better characterise the differences in pRF size, average pRF size for the DE of controls, and the FFE and AME of observers with amblyopia are plotted in Figure 3.9.

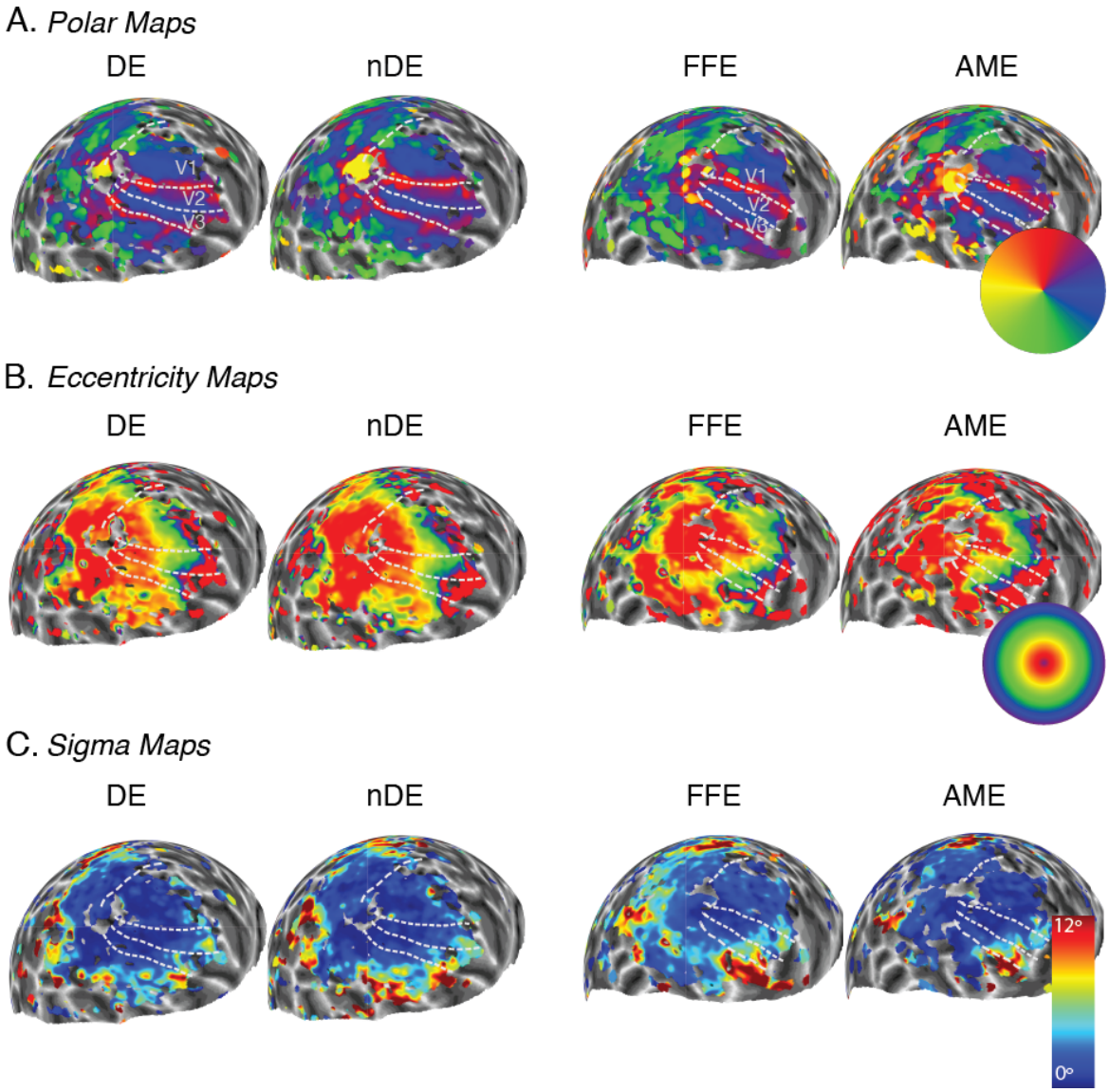

Figure 3-8 Activation patterns from one example control observer and one example observer with amblyopia for polar angle, eccentricity, and pRF size (σ)

A. Maps for polar angle from the dominant eye (DE) and non-dominant eye (nDE) of a control observer with unaffected vision (left), and the fellow fixating eye (FFE) and the amblyopic eye (AME) of an observer with amblyopia (right). Maps are superimposed onto a spherical surface of the left hemisphere. Reference icon on the right indicates the colour-correspondence for the polar co-ordinates.
B. Maps for eccentricity, plotted in the same conventions as in A. Reference icon on the right indicates the colour-correspondence for the eccentricities tested. Note that the eccentricity map is cyclical and wraps around 19° eccentricity.
C. Maps for pRF size (σ), plotted in the same conventions as in A. Reference icon indicates the colour-correspondence for pRF size, ranging from 0° to 12°.

Figure 3.9 clearly shows that pRF size increased with eccentricity for V1 (Figure 3.9A), V2 (Figure 3.9B), and V3 (Figure 3.9C) for the DE of controls and the eyes of observers with amblyopia. The AME had larger pRF sizes compared to the DE, but the magnitude of this difference was variable across eccentricities, visual fields, and regions of interest. For example, in V1, the difference between the DE

and the AME was greater in the temporal than the nasal visual field. In V2, the difference between the DE and the AME was greater in the near compared to the far periphery. Additionally, overall differences between the DE and the AME were larger in V2 than V1. Interestingly, the FFE also showed larger pRFs than the DE for the regions of interest, and this difference was also non-uniform. This indicates that there were multiple sources of variation that determined the differences in pRF size between controls and observers with amblyopia.

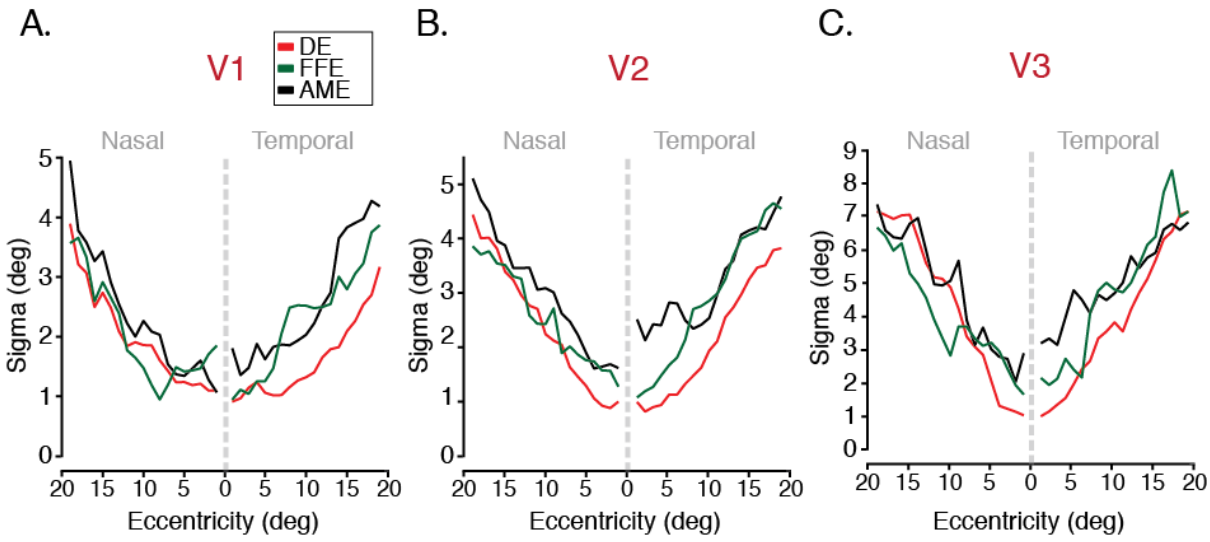

Figure 3-9 pRF size across eccentricity in V1-V3 for the DE, the AME, and the FFE

A. Average pRF size (sigma) in V1 across eccentricities (1°-19°) in the nasal and temporal visual fields. Solid lines indicate the mean pRF size in degrees of visual angle for the dominant eye (DE) of control observers (N=10), and the amblyopic eye (AME) and fellow-fixating eye (FFE) of observers with amblyopia (N=9). The dashed grey line separates the nasal and temporal visual fields.
B. Average pRF size (sigma) in V2. Plotted with the same conventions as in A.
C. Average pRF size (sigma) in V3. Plotted with the same conventions as in A.

In order to better understand the sources of variation in pRF size between the DE and the AME, and the DE and the FFE, stepwise linear regression analyses were conducted. This analysis deviates from the analysis of variance conducted in sections 3.3.1.1 and 3.3.1.2 for measures of acuity and crowding, because for pRF size the range of eccentricity was continuous. pRF size in V1, V2, and V3 was included as the dependent variable. In order to account for the effects of visual field, pRF size estimates for the eccentricities tested were separated by nasal and temporal visual field. This resulted in the inclusion of eccentricity (1°-19°), visual field (coded 0 for nasal and 1 for temporal), and eye (DE and FFE, DE and AME) as predictors in the analyses. Note that I only use estimates of pRF size for the DE of controls as this was the eye for which I had acuity and crowding measurements for

this group. In order to match these analyses to the those conducted on the behavioural estimates of acuity and crowding, the interactions between the predictor variables were also included in the models. A forward-selection procedure was applied, where one predictor was added at a time to the model with the goal of deriving the minimal set of predictors that account for the maximum source of variation in pRF size. Due to an insufficient number of voxels, in some observers pRF size could not be estimated for a particular eccentricity in a particular visual ares. For these cases, missing values were replaced with the mean.

The relevant assumptions for these analyses were met. Examination of the correlations between the predictors included in each model revealed that none were highly correlated to cause concern for multicollinearity. No extreme univariate outliers were identified in initial data screening and Mahalanobis distance scores (D^2) did not indicate multivariate outliers. Residual and scatter plots were checked and the assumptions of normality, linearity, and homoscedasticity were all satisfied.

3.3.2.1.1 pRF size: Dominant Eye of Controls and Amblyopic Eye

For the aims of this investigation, it was integral to establish whether the amblyopic eye showed enlarged pRF size compared to unaffected vision, and if it did, whether this elevation was constant across the visual field or whether some regions were more affected than others. Results from stepwise regression analyses comparing pRF size for the DE and the AME (coded as 0 and 1) in V1, V2, and V3 are reported below. The full progression of the stepwise regression models can be seen in Appendix B section 6.2.1.1.

3.3.2.1.1.1 pRF size in V1

The final model contained three out of the six predictors: eccentricity, eye, and the interaction between eccentricity and eye. This model was reached in three steps, with no variables removed. The model was statistically significant in predicting pRF size in V1, $F(3,721) = 147.74$, $P < .001$, and accounted for approximately 38% of the variance ($R^2 = .382$, adjusted $R^2 = .379$). The regression coefficients of the predictors together with their correlations with pRF size, their squared semi-partial correlations, and their structure coefficients are shown in Table 3.4.

Model	B	SE B	β	Pearson R	sr^2
Intercept	.599	.108			
Eccentricity***	.117	.009	.501	.558	.133
Eye*	.377	.157	.147	.257	.005
Eccen. × Eye*	.028	.014	.138	.463	.004

Table 3-4 Stepwise regression results for variables significantly predicting pRF size in V1 (DE & AME)

The dependent variable was pRF size in V1 in the DE and the AME. R= .382, adjusted R= .379. B= unstandardized coefficient; SE B= standard error of unstandardized coefficient; β= standardised coefficient; Pearson R= correlation between the predictor and the DV; sr = squared semi-partial correlation (*P<.05, ***P< .001).

As expected from the well-documented increase of receptive field size with eccentricity, pRF size in V1 was primarily predicted by eccentricity, with pRF size increasing by .117° for every 1° eccentricity. The AME had larger pRF sizes by .377° on average compared to the DE of controls. The interaction between group and eccentricity was also a significant predictor, indicating that for the AME pRF size in V1 increased by a factor of .028° for every 1° increase in eccentricity. This is clear in Figure 3.9A where the AME shows greater pRF size compared to the DE for all the eccentricities tested. The unique variance indexed by the squared semi-partial correlations was highest for eccentricity, which accounted for approximately 13.3% of the variance in pRF size, whereas eye and the interaction between eccentricity and eye accounted for less than 1%.

3.3.2.1.1.2 pRF size in V2

Identical analyses were carried out with pRF size in V2 as the dependent variable. The final model in this case consisted of the following significant predictors: eccentricity, eye, and the interactions between eccentricity and eye, eccentricity and visual field, and visual field and eye. This model was reached in five steps, with one predictor variable added at each step and no variables removed. pRF size in V2 was significantly predicted by the model, $F(5, 721)= 156.26$, $P< .001$, with the model accounting for approximately 52% of the variance in V2 estimates of pRF size, ($R^2= .522$, adjusted $R^2= .518$). Table 3.5 below shows the relevant information for each predictor variable.

Model	B	SE B	β	Pearson R	sr^2
Intercept**	.373	.109			
Eccentricity***	.202	.010	.754	.646	.251
Eye***	.968	.172	.330	.303	.021
Eccen × Eye*	.030	.014	.126	.486	.003
Eccen × VF ***	-.030	.008	-.128	.231	.009
VF × Eye**	.435	.138	.126	.204	.007

Table 3-5 Stepwise regression results for variables significantly predicting pRF size in V2 (DE &AME)

The dependent variable was pRF size in V2 in the DE and AME. R^2= .522, adjusted R^2= .518. B= unstandardized coefficient; SE B= standard error of unstandardized coefficient; β= standardised coefficient; Pearson R= correlation between the predictor and the DV; sr^2 = squared semi-partial correlation (* *P*<.05, ** *P*< .01, ***P< .001).

Similarly to the previous analysis, eccentricity was a highly significant predictor of pRF size. pRF size increased by .202° for every 1° eccentricity. The eye tested was also a significant predictor, with the amblyopic eye showing a .968° in pRF size increase on average compared to the fellow eye, demonstrating a substantial effect of amblyopia on pRF size in area V2. The significant interaction between eccentricity and eye indicates that pRF size for the AME was larger by a factor of .03° for every 1° increase in eccentricity compared to the DE. Interestingly, the effect of eccentricity was also dependent on the visual field, with the temporal visual field showing smaller pRF sizes by a factor of .03° for every 1° eccentricity, compared to the nasal visual field. The effect of eye was dependent on the visual field, with the AME showing larger pRF sizes by .435° in the temporal visual field compared to the DE of controls. The semi-partial correlations presented on Table 3.5 indicate that eccentricity had the greatest amount of unique variance, accounting for 25.1% of the variance in V2 pRF size, whereas eye explained 2.1% of the variance, and the remaining predictors accounted for less than 2% in total.

3.3.2.1.1.3 pRF size in V3

The final model from a stepwise multiple regression analysis with pRF size in area V3 contained the same predictor variables as the model for V2: eccentricity, eye, and the interactions between eccentricity and eye, eccentricity and visual field, and visual field and eye. This model was reached in five steps without the removal of a predictor. Examination of the correlation matrix did not reveal any significant correlations between the predictors. The model overall explained a significant

proportion of the variance in V3 pRF size, F(5, 721)= 148.29, *P*< .0001, accounting for approximately 51% of the variance (R^2= .509, adjusted R^2= .505). Table 3.6 presents information on the model coefficients.

Model	B	SE B	β	Pearson R	sr^2
Intercept ***	.711	.169			
Eccentricity***	.354	.016	.863	.662	.330
Eye***	1.724	.268	-.383	.193	.029
Eccen × VF ***	-.049	.013	-.139	.233	.010
Eccen × Eye***	-.121	.022	-.338	.364	.002
VF × Eye**	.713	.214	.135	.140	.008

Table 3-6 Stepwise regression results for variables significantly predicting pRF size in V3 (DE &AME)

The dependent variable was pRF size in V3 in the DE and the AME. R^2=.509, adjusted R^2=.505. B= unstandardized coefficient; SE B= standard error of unstandardized coefficient; β= standardised coefficient; Pearson R= correlation between the predictor and the DV; sr^2 = squared semi-partial correlation (** *P*< .01, ****P*< .001).

Eccentricity was a significant predictor of pRF size in V3, with a .354° increase in pRF size for every 1° increase in eccentricity. Eye was a significant predictor, with the AME having larger pRF sizes by 1.724° on average compared to the DE. The interaction between eccentricity and visual field indicated that the temporal visual field showed a smaller increase with 1° eccentricity compared to the nasal visual field by .049°. The significant interaction between eccentricity and eye indicated that the AME showed a smaller increase in pRF size by .121° for every 1° eccentricity compared to the DE. The interaction between visual field and eye suggested that in the temporal visual field the AME had greater pRF sizes by .713° compared to the DE. Consistently with the previous models, eccentricity accounted for the greatest percentage of unique variance, accounting for 33% of the variance in pRF size, whereas eye accounted only for approximately 3% and the interactions for 2% in total.

3.3.2.1.2 pRF size: Dominant Eye of Controls and Fellow Fixating Eye

While visual acuity is considered clinically normal in the FFE of observers with amblyopia (see Clinical details table 3.1), it was important to consider whether pRF size for the FFE in the regions of interest was also “normal”, and thus similar to the DE of controls. In fact, Figure 9 shows that this wasn’t the case: for all regions of interest there was a difference in pRF size between the DE and the FFE, but this

difference was not uniform across the visual field. Particularly, the FFE showed similar pRF size to the DE in the nasal visual field in all regions of interest, but substantially larger pRF size in the temporal visual field. In order to investigate these differences and establish whether the FFE significantly deviated from estimates of pRF size for the DE of controls, stepwise regression analyses were conducted comparing pRF size in V1, V2, and V3 between the DE of controls and the FFE (coded as 0 and 1, respectively). The full progression of the stepwise regression models can be seen in Appendix B section 6.2.1.2.

3.3.2.1.2.1 pRF size in V1

The best model was reached in three steps with no variables removed, and included eccentricity, visual field, and the interaction between visual field and eye as predictors. The model was significant in predicting pRF size in V1, $F(3, 721)= 121.65$, $P< .001$, and accounted for approximately 34% of the total variance ($R^2= .337$, adjusted $R^2= .334$). Table 3.7 shows the relevant information on the regression coefficients of the predictors.

Model	B	SE B	β	Pearson R	sr^2
Intercept***	.774	.082			
Eccentricity***	.116	.006	.544	.544	.264
VF***	-.325	.085	-.139	-.003	.014
VF × Eye***	.673	.100	.246	.168	.042

Table 3-7 Stepwise regression results for variables significantly predicting pRF size in V1 (DE & FFE)

The dependent variable was pRF size in V3 in the DE and FFE. R^2=.337, adjusted R^2=.334. B= unstandardized coefficient; SE B= standard error of unstandardized coefficient; β= standardised coefficient; Pearson R= correlation between the predictor and the DV; sr^2 = squared semi-partial correlation (***$P< .001$).

The analyses revealed that eccentricity was a significant predictor of pRF size in V1. As indicated by the unstandardized B coefficient in Table 3.7, for every 1° increase in eccentricity there was .116° increase in V1 pRF size. Visual field was also a significant predictor, with the B coefficient indicating that the nasal visual field had larger pRFs by .325° compared to the temporal. Figure 3.9 shows that this difference was driven by the DE of controls that had larger pRFs for all eccentricities in the nasal visual field. Indeed, the significant interaction between eye and visual field indicated that the effect of visual field was different for the DE and the FFE. Particularly, the FFE had larger pRFs by .673° in the temporal visual field compared

to the DE. This is clear in Figure 3.9A, as the difference in pRF size between the DE and the FFE was much larger in the temporal visual field, especially beyond 5° eccentricity. The proportion of unique variance explained, indicated by the squared semi-partial correlation, was greatest for eccentricity which accounted for 26.4% of the variance whereas visual field accounted for 1.4%. The interaction between visual field and eye also explained a substantial amount of the variance, accounting for 4.2%. Overall, these analyses suggest that although there were no significant differences in pRF size between the DE and the FFE on average, the FFE showed larger pRFs in the temporal visual field compared to the DE.

3.3.2.1.2.2 pRF size in V2

Identical analyses were conducted for pRF size in V2. The best model was reached in five steps and similarly to V1 included eccentricity, visual field, and the interaction between visual field and eye as predictors. Eye was included as a predictor in the second step but removed in fourth when visual field was added (see Appendix B, section 6.2.1.2.2). The final model was significant in predicting pRF size in V2, $F(3, 721)= 226.79$, $P< .001$, and explained approximately 49% of the total variance ($R^2= .487$, adjusted $R^2= .484$). Table 3.8 shows the information on the regression coefficients of the significant predictors.

Model	B	SE B	β	Pearson R	sr^2
Intercept***	.652	.091			
Eccentricity***	.180	.007	.670	.670	.449
VF***	-.362	.095	-.123	.006	.010
VF × Eye***	.804	.111	.233	.164	.027

Table 3-8 Stepwise regression results for variables significantly predicting pRF size in V2 (DE & FFE)

The dependent variable was pRF size in V2 in the DE and FFE. R^2=. 487, adjusted R^2=.484. B= unstandardized coefficient; SE B= standard error of unstandardized coefficient; β= standardised coefficient; Pearson R= zero-order correlation between the predictor and the DV; sr^2 = squared semi-partial correlation (****P*< .001).

As indicated by the B value for eccentricity in Table 3.8, for every 1° increase in eccentricity there was an increase of .180° on average in pRF size in V2. Visual field was also a significant predictor, with the nasal visual field having larger pRFs by .362° compared to the temporal visual field. This effect was primarily driven by the DE of controls that had larger pRFs across the nasal visual field, and to a lesser degree by the FFE that showed larger pRFs for parafoveal eccentricities in the nasal

visual field. In fact, similarly to V1, the inclusion of the interaction between eye and visual field in the model demonstrated that the effect of visual field was different for the DE and the FFE. Particularly, the FFE had larger pRF sizes in the temporal visual field by .804° compared to the DE. This is clear in Figure 3.9B, where the FFE shows larger pRF size in the temporal visual field, resulting in a marked elevation in pRF size relative to the DE. Eccentricity accounted for the greatest unique variance in the model by 44.9%, visual field accounted for 1%, and the interaction between visual field and eye for 2.7%. Similarly to the results for V1, although the DE and the FFE did not differ on average pRF size, there were significant naso-temporal asymmetries for the FFE, with the temporal visual field having larger pRFs.

3.3.2.1.2.3 pRF size in V3

For pRF size in V3, the best model was reached in four steps and included eccentricity, visual field, and the interactions between eccentricity and eye and visual field and eye, with no variables removed. The model was significant in predicting pRF size in V3, $F(4, 721)= 213.22$, $P<.001$, and explained approximately 54% of the variance ($R^2= .543$, adjusted $R^2= .541$).

Model	B	SE B	β	Pearson R	sr^2
Intercept***	1.211	.130			
Eccentricity***	.311	.012	.762	.718	.434
VF***	-.618	.149	-.138	.003	.011
Eccen. × Eye**	-.038	.013	-.106	.328	.006
VF × Eye***	1.331	.1204	.253	.126	.027

Table 3-9 Stepwise regression results for variables significantly predicting pRF size in V3 (DE & FFE)

The dependent variable was pRF size in V3 in the DE and FFE. $R= .543$, adjusted $R= .541$. B= unstandardized coefficient; SE B= standard error of unstandardized coefficient; β= standardised coefficient; Pearson R= correlation between the predictor and the DV; sr^2 = squared semi-partial correlation (**$P< .01$, ***$P< .001$).

pRF size in V3 significantly increased with eccentricity, and as indicated by the B value in Table 3.9, for every 1° increase in eccentricity there was an increase of .311° in pRF size in V3. Visual field was also a significant predictor, with the nasal visual field having larger pRFs overall than the temporal visual field by .618°. The effect of visual field can be seen in Figure 3.9 for the DE beyond 10° and the FFE in the parafovea. The inclusion of the interaction between eccentricity and eye in the model suggests that the effect of eccentricity was different for the DE of controls

and the FFE of observers with amblyopia. Interestingly, it was the DE that showed larger pRFs by .038° for every 1° increase in eccentricity, compared to the FFE. Based on Figure 3.9, this difference appears to be primarily driven by the DE having larger pRFs in the nasal visual field beyond ~8° eccentricity. In contrast, in the temporal visual field, as the significant interaction between visual field and eye suggests, that the FFE has larger pRFs by 1.331° compared to the DE. A very high percentage of unique variance in the model was accounted by eccentricity that accounted for 43.4%, wand visual field accounted for 1.1%. The interaction between eccentricity and eye accounted for less than 1%, and the interaction between visual field and eye accounted for 2.7%. These results suggest that there is a significant naso-temporal difference in pRF for all regions of interest, with pRFs for the temporal visual field being significantly larger in the FFE than in the DE.

3.3.2.1.3 Summary of pRF Mapping

In this section, I investigated differences in pRF size between controls and observers with amblyopia in V1, V2, and V3. I found that pRF size increased with eccentricity for both groups in all the regions of interest. When considering differences between the DE of controls and the AME of observers with amblyopia, I found that pRF size for the AME was overall larger relative to the DE in V1, V2 and V3. However, the differences in pRF size between the DE and the AME were not uniform across the visual field. There was a very clear difference between the nasal and temporal visual fields in the elevation in pRF size for V2 and V3. In those areas, there was a substantial elevation in pRF size for the AME relative to the DE in the temporal visual field.

I also investigated differences between the DE of controls and the FFE of observers with amblyopia. The results showed that there was no overall difference in pRF size between the DE and the FFE in V1, V2, or V3, and pRF size was larger in the nasal visual field for both the DE and the FFE in all regions of interest. However, there were substantial differences in the FFE between the two visual fields: the FFE showed larger pRF size for the temporal visual field compared to the DE. This was found in V1, V2, and V3, and thus was consistent across regions of interest. Note that I also conducted analyses comparing pRF the AME to the FFE. These are reported in section 2.2 of Appendix B.

3.3.2.2 Eye Movements during pRF Mapping

Unsteady fixation and increased eye-movements during the fMRI scans can bias estimates of pRF size. For the pRF model, I assumed that observers fixated on the cross, and thus that there was a constant correspondence between the screen on which the stimulus aperture was presented and the observer's visual field. If the position of the visual field was significantly shifted due to increased fixation instability, this would have resulted in a lack of correspondence between the stimulus aperture and its assumed location in the observer's visual field. Particularly, repeated eye movements would spread the activation across the cortical surface resulting in enlarged pRF estimates. As increased fixation instability is a characteristic of amblyopia (Gonzalez et al., 2012; Subramanian et al., 2013), if observers with amblyopia showed greater fixation variability and thus poorer fixation stability than controls, this could have led to artificially enlarged estimates of pRF size. Crucially, if this were the case increased fixation variability could account for the differences in pRF size reported above between the DE of controls and the AME, as well as the DE and the FFE. Therefore, here I compared fixation variability between the DE and the eyes of observers with amblyopia to establish whether there were any significant differences.

Individual and average estimates for each eye can be seen below in Figure 3.10A for horizontal variability and Figure 3.10B for vertical variability. Due to technical difficulties I was unable to record eye-tracking data for one control observer, and thus Figures 3.10A and 3.10B and the analyses reported below refer to eye-tracking data from 9 observers in the control group.

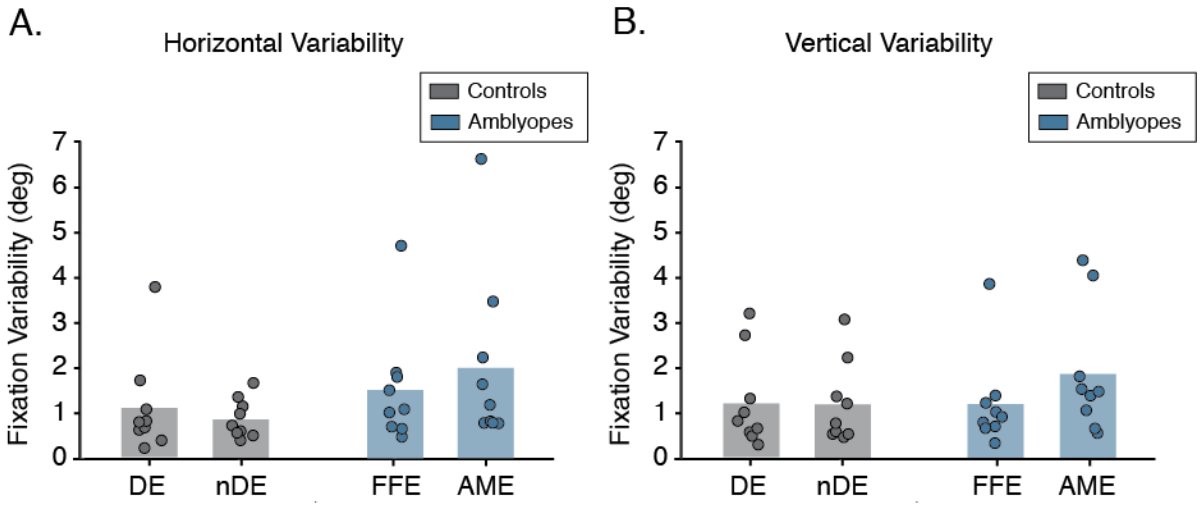

Figure 3-10 Fixation variability of observers with amblyopia (N=9) and controls (N=9) during the fMRI scans

A. Horizontal fixation variability during the five fMRI scans conducted for each eye. Bars indicate the group mean horizontal fixation variability for each eye, and dots indicate the individual mean fixation variability for each eye in degrees of visual angle. DE= dominant eye; nDE=non-dominant eye, FFE = fellow fixating eye, AME= amblyopic eye.
B. Average vertical fixation variability during the five fMRI scans conducted for each eye. Plotted with conventions as described in A.

In order to investigate differences in horizontal and vertical fixation variability, independent samples t-tests were conducted (Bonferroni-corrected). Fixation variability was only compared between the DE and the AME, and the DE and the FFE, to match the analyses above. The comparison that was of most interest was between the DE and the AME as in 3.3.2.1.1 regression analyses show significant differences between the two eyes in pRF size for all regions of interest. Although it is clear in Figure 3.10A that the AME shows almost twice the horizontal fixation variability than the DE (M= 2.02°, SD= 1.95°), this difference was not significant [t(16)= -1.21, *P*= .732]. The FFE also showed greater horizontal fixation variability than the DE on average (M= 1.50°, SD= 1.29°), but similarly the t-test did not reveal significant differences [t(16)= -.68, *P*<.99].

Although Figure 3.10B shows that the AME showed greater vertical fixation variability than the DE on average (M= 1.84°, SD= 1.37°), similarly to horizontal fixation variability, this difference was not significant [t(16)= -1.07, *P*= .897]. The FFE showed almost identical fixation variability to the DE on average (M= 1.19, SD= 1.03), and as such the t-test did not yield significant results, t(16)= .93, *P*<.99.

The results of these analyses suggest that although there were differences in fixation variability between the DE and the AME, these differences were not statistically significant. Note that the small elevation in average horizontal and vertical fixation variability for the AME relative to the DE was primarily driven by two

observers with amblyopia, who showed the greatest fixation variability. As the differences in pRF size between the DE and the AME were central to the aims of this study, it was important to rule out the influence of increased fixation variability in these observers as the source of the differences in pRF size. In Appendix B section 6.2.3 I show that stepwise regression analyses comparing pRF size in the regions of interest between the DE to the AME yield very similar results to 3.3.2.1.1 when these observers are excluded. I thus conclude that poor fixation stability for observers with amblyopia cannot account for the differences in pRF size between the control group and the group with amblyopia.

3.3.2.3 Acuity, Crowding, and pRF size

In order to investigate whether behavioural measures of acuity and crowding can predict pRF size, a linear mixed effects model was constructed using Matlab (Mathworks) and the *fitlme* function using maximum likelihood estimation. pRF size in each region of interest was the dependent variable. For comparison purposes, pRF size estimates were required to have the same data structure as the acuity and crowding measures. Therefore, for the analyses below, pRF estimates were taken for 2°, 4°, 8°, and 12° eccentricity in the nasal and temporal visual fields. Note that for one control observer and two observers with amblyopia, it was not possible to estimate pRF size in some eccentricities in V2 and V3 due to a lack of responsive voxels. In these cases, the missing value was replaced with the mean. Eccentricity, visual field, group (DE and AME) and the two-way interactions between them were taken as fixed effects factors. The three-way interaction was not included as they were not found to be significant in accounting for variability in pRF estimates of size. Visual field was coded as 0 for nasal and 1 for temporal, and group was coded as 0 for the DE of controls and 1 for the AME of observers with amblyopia. Acuity or crowding were also included as a fixed effect in the model. To account for between-observer variability, observer was included as a random effect. This was a random intercept and slope model, as observers had different intercepts and different slopes for the effect of acuity or crowding. For the estimation of model parameters, maximum likelihood (ML) was used. To test the significance of each fixed effect predictor, we ran 1000 simulations of the full model including all predictors and a reduced model that did not include the selected predictor. Random effects were kept constant in both models. We determined whether each fixed effect contributed in the change in residual deviance by comparing the Akaike Information Criterion

(AIC) (Akaike, 1973, 1974) for each model. Note that smaller AIC values indicated a better fit. Results are outlined below separately for the inclusion of acuity and crowding in predicting pRF size in each region of interest.

3.3.2.3.1 Acuity & pRF size

3.3.2.3.1.1 Acuity and pRF size in V1

For pRF size in V1, the AIC value for the full model (AIC_{full}) was 325.48 whereas for the alternative model that did not include acuity the AIC value (AIC_{alt}) it was 320.48. This demonstrates that acuity did not contribute in accounting for variance in pRF size in V1 (β= .019, SE= .016, *P*= .214). I thus did not proceed further with the analysis by evaluating the contribution of the other variables in the model.

3.3.2.3.1.2 Acuity and pRF size in V2

For pRF size in V2, the results of the model comparison indicated that acuity did not account for a significant amount of variance in pRF size in V2 (AIC_{full}= 371.21, AIC_{alt} = 369.79) and was not a significant predictor for pRF size in V2 (β= .026, SE= .020, *P*= .186).

3.3.2.3.1.3 Acuity and pRF size in V3

For V3, acuity significantly accounted for a proportion of the variance in pRF size, (AIC_{full}= 446.24, AIC_{alt} = 454.65), with pRF size increasing with acuity (β= .114, SE= .056, *P*= .044). As in the stepwise regression analyses above, eccentricity was also a significant fixed effect in predicting pRF size in V3 (AIC_{alt} = 483.39), with pRF size significantly increasing with eccentricity (β= 1.743, SE= .038, *P*< .001). The interaction between eccentricity and visual field was also a significant predictor, (AIC_{alt} = 457.29), with eccentricities in the nasal visual field having larger pRF size on average than in the temporal visual field (β= -.337, SE= .153, *P*= .021). The interaction between visual field and group led to a small increase in the AIC value when removed from the model (AIC_{alt} = 447.93) thus indicating that it explained a small amount of variance in pRF size, but did not reach significance as a fixed effects predictor (β= .653, SE= .336, *P*= .055). Visual field did not significantly contribute in explaining variance in pRF size in V3 (AIC_{alt} = 444.34; β= -.201, SE= .638, *P*= .753), neither did group (AIC_{alt} = 444.74; β= .558, SE= .775, *P*= .473).

Finally, the interaction between eccentricity and group did account for any variance in pRF size in V3 (AIC_{alt} = 445.42; β= -.219, SE= .196, *P*= .266). As such, although this is a small effect, acuity predicts pRF size in V3 when included in a model that includes eccentricity, visual field, eye, and the interactions between them.

3.3.2.3.2 Crowding & pRF size

3.3.2.3.2.1 Crowding and pRF size in V1

Similarly to acuity and pRF size in V1, crowding did not significantly account for variance over and above the other predictors (AIC_{full}= 326.26, AIC_{alt}= 324.46; β= -.039, SE= .084, *P*= .639). I thus did not proceed further with the analysis.

3.3.2.3.2.2 Crowding and pRF size in V2

The addition of crowding to the model did not significantly account for variance in pRF size in V2 (AIC_{full}= 379.38, AIC_{alt}= 378.28), and crowding was not a significant fixed effects predictor (β= -.099 , SE= .102, *P*= .337).

3.3.2.3.2.3 Crowding and pRF size in V3

Crowding was found to not significantly account for variance in the model (AIC_{full}= 456.02, AIC_{alt}= 454.65; β= .318, SE= .192, *P*= .102), and thus the contribution of the remaining variables was not assessed.

3.4 **Discussion**

In this chapter, I investigated whether there is a shared neural basis of crowding between the typical and amblyopic visual system. Pooling models of crowding (Parkes et al., 2001) have attributed the reduction of acuity and the increase in the extent of crowding with eccentricity in the typical visual field (Bouma, 1970; Toet & Levi, 1992) to the cortical undersampling of the visual periphery, as fewer neurons with large receptive fields are dedicated to the periphery compared to the fovea (Dow et al., 1981; Van Essen et al., 1984). Following pooling models, I hypothesised that if crowding in amblyopia similarly arose from an undersampling of the amblyopic eye, acuity should be reduced, the extent of crowding should be greater, and pRF size should be elevated relative to unaffected vision. In characterising the deficit in visual function, I showed that acuity was reduced and the extent of crowding was greater relative to unaffected vision across the

amblyopic visual field. This deficit in visual function was non-uniform, affecting the fovea disproportionately compared to the periphery, and the temporal visual field more than the nasal. The amblyopic deficit was reflected in pRF size, as V1-V3 pRFs for the amblyopic eye were greater relative to typical vision across the visual field. Matching the pattern of acuity and crowding, the elevation in pRF size was non-uniform in V2 and V3, with pRF size being greater at parafoveal eccentricities, and affecting disproportionately the temporal visual field. Together, these findings demonstrated that the general pattern of acuity and crowding across the visual field matched the pattern of pRF size in typical and amblyopic vision, consistent with predictions of pooling models of crowding.

In typical vision, the findings on acuity were in line with the well-documented dependency of acuity on eccentricity (Sloan, 1968; Rovamo & Raninen, 1990): acuity was best in the fovea but was gradually reduced with eccentricity. In amblyopia, as expected by the clinical definition of the disorder, there was a significant acuity reduction in the amblyopic fovea relative to the typical fovea. In the amblyopic periphery, acuity was reduced across the peripheral visual field. Additionally, there was a naso-temporal difference in the peripheral reduction in acuity, with the temporal visual field showing disproportionately reduced acuity relative to the nasal. Although this visual field difference was found both in observers with amblyopia and typical vision, the asymmetry was markedly more pronounced in the amblyopic visual field. Crucially, when comparing the acuity deficit across the visual field in amblyopia, a significant non-uniformity was evident: the overall magnitude of the acuity reduction was disproportionate for the fovea relative to the periphery. Specifically, whereas acuity in the amblyopic fovea was four times worse on average compared to unaffected vision, in the periphery it was approximately twice as poor. This suggests that the visual field location with the greatest acuity deficit was the fovea, consistent with early studies demonstrating that acuity and contrast sensitivity losses are greater in the central visual field of strabismic amblyopes compared to the periphery (Kirschen & Flom, 1978; Thomas, 1978; Avetisov, 1979; Hess & Jacobs, 1979). Overall, the results clearly demonstrated that there was a reduction in acuity across the amblyopic visual field, with the greatest deficit found at the fovea, and the temporal visual field being more affected than the nasal.

When it came to crowding, the findings in typical vision were consistent with previous studies (Bouma, 1970; Toet & Levi, 1992): the extent of crowding was minimal in the fovea and increased with eccentricity. In amblyopia, the pattern of the

crowding deficit was very similar to acuity. In line with the well-documented crowding deficit in the amblyopic fovea (Levi & Klein, 1985; Greenwood et al., 2012), foveal crowding extended over large distances. Similarly to acuity, the crowding deficit persisted into the periphery, and showed the same naso-temporal difference, with the extent of crowding being larger in the temporal visual field.
In addition to the naso-temporal difference, amblyopic crowding also showed the foveal-peripheral non-uniformity. Specifically, although crowding in the amblyopic fovea extended over distances eleven times larger on average than in the typical fovea, in the periphery this difference was reduced, with the extent of crowding being less than twice that of the typical periphery. Crucially, the foveal-peripheral non-uniformity was greater in crowding compared to acuity, in line with previous studies showing a disproportionate crowding deficit relative to acuity in amblyopia (Hess, Dakin, et al., 2001; Levi, Hariharan, et al., 2002a; Bonneh et al., 2004; Greenwood et al., 2012; Song et al., 2014). As such, the disproportionate foveal deficit clearly demonstrates that crowding in amblyopia is not merely a result of reduced acuity.

The findings from pRF mapping showed that pRF size increased with eccentricity in areas V1, V2, and V3, in line with previous studies (Dumoulin & Wandell, 2008; Amano et al., 2009; Harvey & Dumoulin, 2011; Clavagnier et al., 2015). This increase in pRF size with eccentricity could be associated with the cortical undersampling of the visual periphery (Dow et al., 1981; Van Essen et al., 1984). In non-human primates, electrophysiological studies have shown the number of V1 neurons responding to peripheral stimulation is dramatically smaller than the number neurons responding to foveal stimulation (Dow et al., 1981; Van Essen et al., 1984). fMRI studies have confirmed this cortical undersampling of the periphery in V1-V4, showing that CMF is greater for the fovea and is reduced with increasing eccentricity (Sereno et al., 1995; Engel et al., 1997). As there is a relationship between larger CMF and smaller pRFs (Harvey & Dumoulin, 2011), the undersampling of the visual periphery could result in neurons that have larger receptive fields in order to ensure adequate coverage of the visual field.

In amblyopia, pRF size in V1, V2, and V3 for the amblyopic eye was greater relative to typical vision, consistent with previous findings by Clavagnier et al. (2015). Similarly to peripheral vision, this elevation in pRF size could also arise due to a reduction in the cortical resources dedicated to the amblyopic eye, resulting in an undersampling of the amblyopic visual field. Such a reduction in neural resources

would be consistent with the previous studies in non-human primates showing a shift in ocular dominance towards the fellow fixating eye, with fewer neurons responding preferentially to the amblyopic eye in V1 (Crawford & von Noorden, 1979; Shooner et al., 2015). Recently it was found that the shift in ocular dominance is small in V1 of macaques but large in V2 (H. Bi et al., 2011), providing a potential explanation as to why some earlier studies had not found evidence of such a shift in V1 (E. L. Smith et al., 1997; Kiorpes et al., 1998). The increased pRF size in V1, V2, and V3 found in this chapter could is thus consistent with a reduction of neurons driven by the amblyopic eye, resulting in fewer neurons with large receptive fields.

An important question concerns whether the deficit in pRF size in V2 and V3 in amblyopia is a reflection of earlier V1 losses. Evidence from electrophysiological studies in primates show that the disproportionate shift in ocular dominance found in V2 relative to V1 (H. Bi et al., 2011) is indicative of additional V2 losses. In the context of the findings in this chapter, this suggests that the increase in pRF size found in later visual areas should be reflective of additional losses from V1, and not just be a mere amplification of the V1 losses. In fact, this is supported by previous neuroimaging findings using retinotopic mapping and investigating pRF size in amblyopia. In typical vision, when the rate of expansion in pRF size with eccentricity is expressed in V1 cortical surface area, it was found to be constant across eccentricity in V2 and V3 (Harvey & Dumoulin, 2011). This demonstrated that regardless of eccentricity, these extrastriate visual areas pool from a constant surface area in V1. In contrast, Clavagnier et al. (2015) showed that in amblyopia the elevation in pRF size in V2 and V3 for the amblyopic eye could not be explained by constant sampling from V1. The involvement of the extrastriate cortex in the neural basis of amblyopia is in line with the wide range of visual functions affected in amblyopia, including higher-level deficits like motion processing (Simmers et al., 2003; Simmers et al., 2005; Levi, Yu, Kuai, & Rislove, 2007; Rislove, Hall, Stavros, & Kiorpes, 2010), and is consistent with the amblyopic deficit being involving multiple stages of abnormal processing in the cortex (Kiorpes & Daw, 2018).

In addition to an overall increase in pRF size in V1-V3, central to the aims of this chapter was the characterisation of the amblyopic deficit in pRF size across the visual field. As such, findings by Clavagnier et al. (2015), who only mapped pRF size up to 6° eccentricity, were extended by mapping a larger portion of the visual field and independently examining pRF size in the nasal and temporal visual fields. Following this approach, the findings showed that the elevation in pRF size for the

amblyopic eye persists up to 19° eccentricity. Additionally, by considering naso-temporal differences, the lack of uniformity of the amblyopic deficit across the visual field was revealed. Particularly, in V2 and V3, there was a naso-temporal difference for the amblyopic eye, with larger pRFs responding to visual stimulation of the temporal visual field relative to the nasal visual field. Crucially, the naso-temporal difference in pRF size in V2 and V3 was especially pronounced for parafoveal eccentricities, for which the elevation in pRF size for the temporal visual field was greater compared to more eccentric locations. In fact, the difference in pRF size between the amblyopic eye and the eye of observers with typical vision was larger in parafoveal eccentricities (up to 5° eccentricity), regardless of visual field or visual area. Together, these findings demonstrated that the amblyopic deficit in pRF size was non-uniform across the visual field.

In addition to larger receptive fields of individual neurons in each voxel that comprise a pRF, it is important to consider other factors that could be underlying the enlargement in pRF size observed for the amblyopic eye relative to unaffected vision. Poor fixational stability, a characteristic of amblyopia (Gonzalez et al., 2012; Subramanian et al., 2013), could have artificially enlarged the estimates of pRF size. By recording eye movements in the scanner, I was able to compare fixation stability between the observers with amblyopia and unaffected vision and show that the two groups did not substantially differ. To further ensure that fixation variability did not contribute in the increased pRF size for the amblyopic eye in the visual areas of interest, I also demonstrated that excluding observers with amblyopia who showed the greatest fixation variability relative to controls maintained the difference between in pRF size between the amblyopic eye and unaffected vision (see Appendix B, section 6.2.3). I thus conclude that the observed increased pRF size for the amblyopic eye in the regions of interest cannot be accounted for by unsteady fixation.

Increased positional scatter could have also contributed to an increase in estimates of pRF size for the amblyopic eye. Using the correspondence in pRF centres for a given voxel between the two eyes of observers with amblyopia, Clavagnier et al. (2015) determined that there was more variability (i.e. scatter) in the centres of the pRF responding to the amblyopic eye relative to the fellow eye. In line with these findings, animal models of amblyopia have shown that V1 neurons have reduced spatial resolution even though their thalamic inputs are normal (Movshon et al., 1987), potentially reflecting a disordering of afferent connectivity. It has been

argued that such disordered connectivity could lead to a disruption of the receptive field structure and abnormal pooling of visual signals, with these changes cascading through the cortical pathway and amplifying the differences in neural resources between the amblyopic and the fellow eye from one cortical area to the next (Shooner et al., 2015). Therefore, it could be that this disordered connectivity results in the increased scatter in the centres of the pRFs for the amblyopic eye found by Clavagnier et al. (2015). Although I did not detect differences in the polar maps between the amblyopic eye and the dominant eye of observers with unaffected vision that could be indicative of a disorganisation of the cortical map and thus scatter in the pRFs, I did not quantify the correspondence between the visual field locations to which pRFs in the amblyopic and fellow eye respond to. Increased positional scatter could thus be a contributing factor in the increased pRF size found for observers with amblyopia.

The central aim of this chapter was to investigate whether there is a common neural basis in pRF size for crowding in typical vision and amblyopia. As such, having characterised visual function and V1-V3 pRF size in typical and amblyopic vision across the visual field, the relationship between acuity, crowding and pRF size was explored. The combined behavioural and neuroimaging findings clearly indicated that there were commonalities in the overall pattern of variation in acuity, crowding and pRF size. In typical vision the reduction in acuity and increase in the extent of crowding with eccentricity paralleled the increase in pRF size with eccentricity in V1-V3. In amblyopia, the overall increase in pRF size in V1-V3 relative to unaffected vision also matched the reduction in acuity and elevation in the extent of crowding. The non-uniformity of the amblyopic deficit across the visual field was found in both measures of visual function and pRF size. In V2 and V3, there was a naso-temporal difference in pRF size that matched behavioural deficits: in the temporal visual field, pRF size was larger, acuity was reduced, and the extent of crowding was greater for the amblyopic eye relative to typical vision. Additionally, the elevation in pRF size for the amblyopic eye was larger in parafoveal relative to more eccentric visual field locations. This is broadly consistent with the foveal-peripheral asymmetry found for acuity and crowding, an asymmetry that was disproportionately large for crowding compared to acuity. The difference in the magnitude of this asymmetry between acuity and crowding could reflect the involvement of different visual areas. The acuity deficit in amblyopia could be associated with an enlargement in V1 pRF size, consistent with previous studies

linking acuity to measures of V1 functional architecture (Duncan & Boynton, 2003{Srinivasan, 2015 #544)}, whereas crowding could be linked to pRF size in V2 and V3 where the parafoveal deficit in pRF size was most pronounced. Overall, the general pattern of variation in acuity and crowding across the typical and amblyopic visual field matched the pattern of variation in pRF size. However, when exploring the relationship between individual variations in pRF size and acuity and crowding across the visual field, only a weak relationship between pRF size in V3 and acuity was found. The lack of a statistically significant relationship does not necessarily signify a lack of relationship between pRF size and acuity and crowding, but could be due to insufficient power because of the small sample size.

As the focus of this thesis is on visual crowding, it is important to consider what the results from this chapter fit within previous findings on the cortical locus of crowding. The overall pattern in pRF size in V1, V2, and V3 and crowding was broadly similar across the visual field. Prior neuroimaging studies have also pointed to the involvement of multiple visual areas in crowding, from V1 to V4 (T. Bi et al., 2009; J. Freeman et al., 2011b; Anderson et al., 2012; Millin et al., 2014). The increase in pooling across receptive fields that increase from V1-V4 could be underlying the involvement of these areas in visual crowding. In fact, V4 receptive field sizes have been found to fit with Bouma's (1970) critical spacing stating that crowding occurs when the flankers are at a distance of half the eccentricity of the target (Motter, 2006, 2009), and show the characteristic naso-temporal asymmetry of crowding (Motter, 2018). If V4 is the site of maximal pooling between target and flankers, it could be the area for which the extent of crowding shows a relationship with pRF size in V4. However, it was not possible to assess this relationship as estimates of pRF size in V4 were not computed. Nonetheless, the general commonalities between V1-V3 pRF size and crowding are consistent with crowding being a process that involves multiple visual processing stages.

A limitation of the study reported in this chapter concerns the lack of an estimate of pRF size for the fovea. An estimate of foveal pRF size would have been especially important for assessing the relationship between measures of visual function and pRF size in amblyopia. Although parafoveal eccentricities were considered when examining this relationship, it was not possible to factor in pRF size for the location that showed the greatest behavioural deficits – the fovea. Therefore, a foveal pRF estimate could have contributed significantly in better characterising the relationship between pRF size and acuity and crowding, as the

behavioural measurements predict it would show the greatest elevation in pRF size. The lack of a foveal estimate of pRF size was due to the central 0.5° eccentricity of the display containing the fixation cross, resulting in foveal neurons not being stimulated by the bar stimulus. This methodological decision was made because the measurement of pRF size at the fovea is very difficult due to the foveal confluence: the foveal parts of the retinotopic maps in V1, V2, and V3 converge at a common centre, similarly to wedges of a pie meeting (Zeki, 1969). The precise layout of the fovea in these visual areas thus remains a matter of debate. As such, pRF estimates typically begin at 0.5° eccentricity as the closest visual field location to the fovea. To overcome this limitation for the purposes of this study, behavioural measures of acuity and the extent of crowding could be obtained at 0.5° eccentricity. This would allow for the examination of the relationship between acuity, crowding and pRF size at the nearest visual field location to the fovea for which pRF estimates can be reliably obtained.

Although the fellow eye in amblyopia was not the primary focus of this study, it is worth noting that differences between unaffected vision and amblyopia were not restricted to the amblyopic eye. I did not find an overall increase in pRF size for the fellow fixating eye, but there was an increase in V1-V3 pRF size for the temporal visual field in the fellow eye, relative to unaffected vision. Supplemental analyses reported in Appendix B (section 6.2.2) in fact suggest that pRF size for the fellow eye was more similar to the amblyopic eye than to unaffected vision. The similarity between the two eyes in amblyopia was primarily driven by both showing the disproportionate elevation in pRF size for the temporal visual field. The lack of a greater difference in pRF size between the two eyes in amblyopia is not inconsistent with a shift of neural resources to the fellow eye, or crowding arising due to increased receptive field size. However, within a pooling framework the moderate increase in pRF size related to typical vision would predict a crowding deficit for the fellow eye. Here I did not obtain psychophysical measures for the fellow eye, and thus I cannot evaluate whether the enlargement in pRF was accompanied by potential acuity and crowding deficits. However, previous studies have found that the extent of crowding in the fellow fixating eye of children with amblyopia is similar to that of children with typical vision (Greenwood et al., 2012). Still, as reported in Chapter 2 and previous studies (Atkinson & Braddick, 1983; Atkinson et al., 1987; Jeon et al., 2010; Greenwood et al., 2012), children with typical vision show an increase in foveal crowding relative to adults. Crowding in the developing fovea

could be indicative of the late maturation of the extrastriate visual cortex (Hensch, 2004), with receptive fields reaching adult levels in adolescence. In amblyopia the absence of binocularity may result in a form of “arrested development”, with receptive fields remaining fixed at a large state rather than reducing in size. Both eyes could thus be affected, but the amblyopic eye disproportionately so due to suppressed (Harrard, 1996). In adulthood, this would result in moderate deficits in pRF size and crowding for the fellow eye, and greater deficits for the amblyopic eye, consistent with the results reported in this chapter and in Appendix B. Future studies could focus on determining whether the crowding deficit affects both eyes in amblyopia, and its association to pRF size. Treatment for amblyopia involves the patching of the fellow eye with the aim of improving visual acuity in the amblyopic eye, and thus obtaining a better understanding of fellow eye deficits that are not currently targeted by treatment is of critical importance.

To conclude, in this chapter I found that the general pattern of increase in V1-V3 pRF size in the typical and amblyopic visual fields broadly matched the increase in the extent of crowding. These findings are thus consistent with a pooling account of crowding (Parkes et al., 2001), that would predict that the visual field locations with the largest extents of crowding should also an increase in pRF size. Together with Chapter 2, up to this point the findings in this thesis demonstrate that crowding in typical vision and amblyopia shares common perceptual effects described by the pooling of population responses that could arise due to enlarged population receptive fields of visual neurons in V1-V3. However, the involvement of multiple visual areas in the amblyopic deficit raises the possibility of the involvement of higher-level processes in amblyopic crowding. In the next experimental chapter, I addressed this issue by investigating whether amblyopic crowding is modulated by grouping, and thus whether grouping (Herzog & Manassi, 2015; Herzog et al., 2015) could provide a successful framework for crowding in amblyopia.

4 **Chapter 4:** Grouping effects on crowding in strabismic amblyopia

4.1 **Introduction**

In Chapters 2 and 3, amblyopic crowding has been considered within the framework of pooling accounts of crowding. According to population response pooling accounts (van den Berg et al., 2010; Harrison & Bex, 2015), like the one presented in Chapter 2, crowding arises due to the pooling of the responses of a population of detectors to the target and flanker features. Whether or not the target and flanker features will be pooled is assumed to be determined by their separation (Harrison & Bex, 2015), that has been found to scale up to half the target eccentricity (Bouma, 1970). Within this spatially restricted region, the population of detectors responds selectively to a stimulus dimension, such as orientation. As such, population response pooling models can simulate the selectivity of crowding to target-flanker similarity: on a black target, black flankers strongly disrupt target recognition, whereas white flankers have only a weak disruptive effect (Chakravarthi & Cavanagh, 2007; Chung & Mansfield, 2009). However, studies using more than one flanker on each side of the target have shown that flankers outside Bouma's (1970) critical spacing can modulate crowding in ways that do not always adhere to selectivity for target-flanker similarity (Livne & Sagi, 2007; Malania et al., 2007; Sayim et al., 2008; Saarela & Herzog, 2009; Saarela, Sayim, Westheimer, & Herzog, 2009; Manassi et al., 2012, 2013; Manassi et al., 2016), posing a challenge to pooling accounts of crowding.

Additional flankers placed next to the ones neighbouring the target and extending outside Bouma's (1970) critical spacing can have an effect on target recognition, and thus modulate crowding. For example, Strasburger et al. (1991) found that when the number of flankers increased from two (one on each side of the target) to four (two on each side of the target), target recognition deteriorated, even though the second set of flankers was located beyond Bouma's (1970) critical spacing. Manassi et al. (2012) presented two vertically positioned lines with a small offset (i.e. a vernier) and asked observers to discriminate the direction of the offset (left or right). The vernier target was flanked by 2,4,8 or 16 flankers of either equal length to the target, smaller, or greater length. They found that crowding was weak with dissimilar flankers (i.e. of shorter or longer length to the target), and became weaker the more dissimilar flankers were added. These findings suggest that flankers outside Bouma's critical spacing can both further disrupt target recognition

strengthening the effects of crowding, and lead to improvements in target recognition relieving crowding.

What determines crowding in these instances? It has been proposed that crowding is determined by Gestalt principles of grouping (Herzog & Manassi, 2015; Herzog et al., 2015). Gestalt principles (Wertheimer, 1923) state that given any number of stimuli, the visual system tends to group these stimuli and perceive greater wholes instead of independent, individual parts. Applied to crowding, grouping accounts argue that the disruptive effect of the flankers on target recognition depends on their global configuration ("their whole"), and whether or not the target and flankers group based on their appearance (Herzog & Manassi, 2015; Herzog et al., 2015). When additional flankers are added that match *both* the previous local flankers *and* the target, as in Strasburger et al. (1991), a coherent configuration is formed complimenting the percept. In this case, *target-flanker grouping* occurs, and crowding is strong. On the other hand, when additional flankers "fit" *only* with the previous flankers but not with the target, they are segregated from the target, forming a coherent configuration with each other but not with the target. An example of this effect is the shorter- and longer-length flankers to the target in Manassi et al. (2012). In this case, *flanker-flanker grouping* occurs, and crowding is weak.

Grouping accounts of crowding argue that crowding is determined by how the objects look and whether or not they group (Herzog & Manassi, 2015; Herzog et al., 2015; Francis et al., 2017). The way in which grouping may operate on crowding was put forward by a model by Francis et al. (2017), that represents grouping as an automatic individuation process occurring after pooling and determining the boundaries between objects. When the target and flankers form a uniform configuration and there is target-flanker grouping, the boundaries between the target and flankers are connected. In contrast, when the flankers form a uniform configuration with each other but not with the target, the boundaries between the target and the flankers are not connected. Top-down mechanisms determine whether crowding occurs based on the perceptual grouping of the target and flankers. When the boundaries are connected, the target and flankers cannot be segmented, resulting in errors in target recognition and strong crowding. In contrast, when the boundaries between the target and flanker are not connected, top-down processes allow the target to be segmented from the surrounding flankers, thus releasing crowding and resulting in more accurate recognition of the target identity.

Contrary to pooling accounts that represent crowding as a primarily bottom-up process (Parkes et al., 2001; Greenwood et al., 2009; van den Berg et al., 2010; Harrison & Bex, 2015), grouping accounts of crowding involve both top-down and bottom-up processes, placing crowding at a higher-level stage of visual processing.

Many grouping cues have been found to be involved in crowding, including similarity (Malania et al., 2007; Manassi et al., 2013), spacing regularity (Saarela, Westheimer, & Herzog, 2010), and good Gestalt (Sayim et al., 2010; Manassi et al., 2012). A characteristic example of target-flanker grouping is when target and flankers follow the Gestalt principle of regularity, that states that elements are likely to be grouped together if they are parts of an orderly, regular, and balanced configuration. Rosen and Pelli (2015) measured the spatial extent of crowding in the periphery with flankers of the same polarity to the target, opposite polarity to the target, and mixed polarity (alternating black and white), whilst also varying the number of flankers (illustrated in Figure 4.1). With one flanker on each side of the

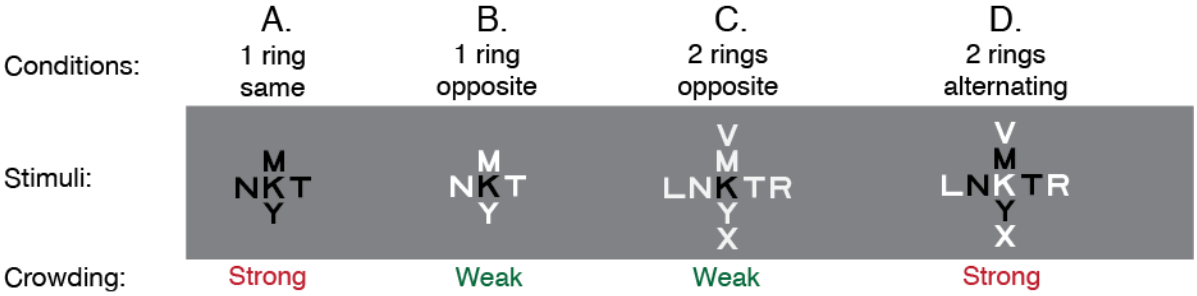

Figure 4-1 Illustration of stimuli and results by Rosen & Pelli (2015)

A: With one flanker of the same polarity on each side of the target, crowding is strong (i.e. the spatial extent of crowding is large). *B:* With one opposite-polarity flanker on each side of the target, crowding is weak (i.e. the spatial extent of crowding is small). *C:* With additional flankers of opposite polarity to the target, crowding is weak. Note that the crowding extent is smaller than in B. *D:* With additional flankers of alternating polarity, crowding is strong. Note that the extent of crowding is greater than in C and B here, but similar to A.

target, crowding was strong (the spatial extent of crowding was large) with flankers of the same polarity to the target, and weak (the spatial extent was reduced) with flankers of opposite polarity (Fig 4.1. conditions A-B). These results are in line with both crowding being selective for target-flanker similarity in polarity (Chakravarthi & Cavanagh, 2007; Chung & Mansfield, 2009), and target-flanker grouping. Whilst keeping the local flankers neighbouring the target of opposite polarity, they added two flankers of opposite polarity on each side (Fig 4.1. Condition C) or two flankers of alternating polarity (Fig. 4.1 condition D). Additional opposite polarity flankers amplified their facilitatory effect on target recognition, consistent with flanker-flanker

grouping. In contrast, when the additional flankers were of alternating polarity, crowding was strong, extending over distances similar to two flankers of the same polarity. This is consistent with target-flanker grouping by regularity, as the alternating polarity target and flankers form an orderly and balanced configuration. Such target-flanker grouping effects have also been shown with alternating colour stimuli in the visual periphery (Manassi et al., 2012), as well as alternating polarity and colour in the typical fovea (Sayim et al., 2008), suggesting that these grouping effects modulate crowding in both the typical periphery and the fovea.

In addition to objects forming global configurations based on features such as polarity, there is also evidence of different shapes forming such global configurations and modulating the effects of crowding. Manassi et al. (2013) asked observers to discriminate the offset direction of a target vernier presented in the periphery (illustrated in Figure 4.2). When the vernier was enclosed within a square flanker (Fig. 4.2.B), thresholds for identifying the vernier offset increased relative to acuity (Fig. 4.2 A), and crowding was strong. However, when additional squares were consecutively placed on each side of the square enclosing the target, thresholds were lower, and crowding was weak. This effect does not appear to depend on the collinearity of the vernier with the individual horizontal lines of the square flankers, as it has been replicated with a variety of different flanker shapes, including diamonds, hexagons and stars (Manassi et al., 2016). Rather, Manassi et al. (2013) argued that when more squares are added, the flankers from a configuration with each other, but ungroup from the target. These findings are thus taken to indicate flanker-flanker grouping.

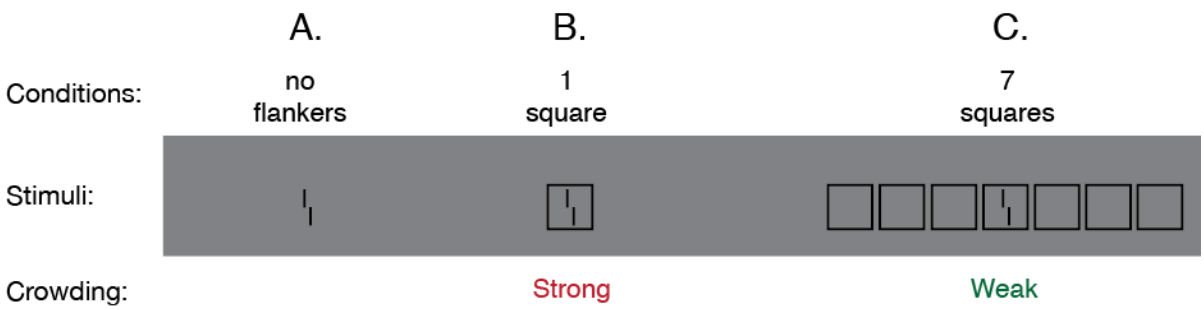

Figure 4-2 Illustration of findings by Manassi, Sayim, & Herzog (2013)

Condition A: Depiction of stimuli used to measure vernier acuity, where the offset of two lines is readily recognised and thresholds are low (no crowding). *Condition B:* With a local square flanker enclosing the target vernier, thresholds for identifying the vernier offset increase, and crowding is strong. *Condition C:* With three additional square flankers on each side of the local flanker enclosing the target, thresholds decrease relative to B, and crowding is weak.

Multiple studies have found evidence for grouping effects on crowding in both the typical fovea and periphery following Gestalt principles of closure and good continuation (Livne & Sagi, 2007; Levi & Carney, 2009) and good Gestalt (Sayim et al., 2010; Manassi et al., 2012). However, it is unclear whether global configuration can also determine crowding in strabismic amblyopia, and thus whether grouping can account for amblyopic crowding. In a rare study exploring the effects of the global configuration of the target and flankers on amblyopic crowding, Levi and Carney (2011) investigated the effect of varying the number of flankers on the recognition of a Gabor target. The flankers were segments of an annular surround, with eight flankers forming the full circle. When flankers were gradually increased from one to four individual segments of the annular surround, the identification of the orientation of a Gabor target was disrupted, and crowding became increasingly strong. However, when eight flankers surrounded the target that formed the full annular surround, target identification improved and crowding was weak. The reduction of the disruptive effect of clutter with eight flanker segments could be indicative of the flankers forming a uniform and regular configuration with each other and not the target, and thus flanker-flanker grouping.

However, findings from other studies cast doubt on the likelihood of amblyopic grouping being determined by the global configuration. Studies using only one local flanker on each side of the target have shown that flankers of the same and opposite polarity to the target were equally disruptive on target recognition for observers with strabismic amblyopia (Hess, Dakin, et al., 2001; Hariharan et al., 2005). Modulations of crowding based on differences in polarity between the target and flankers could be accounted for both by selectivity for target-flanker similarity in polarity, and thus pooling accounts, and by grouping based on similarity. The lack of such modulations in amblyopia could be due to the perceptual distortions that characterise amblyopic vision, and tend to disrupt perception in a non-systematic way (Pugh, 1958; Sireteanu, Wolf-Dietrich, et al., 1993; Sireteanu, Thiel, et al., 2008). These distortions could result in disordered boundaries between the target and flankers, creating boundaries between objects that do not belong in the same “whole”. This could lead to some form of improper grouping that would in turn result in improper segmentation, and thus performance would not improve when the flankers are dissimilar from the target.

In order to investigate whether grouping can determine crowding in strabismic amblyopia, in this chapter I examined the effects of the global

configuration of the target and flankers on amblyopic crowding. I used the characteristic flanker conditions from in Rosen and Pelli (2015) and Manassi et al. (2013) discussed above, and measured crowding in the amblyopic and the typical fovea. Based on these previous studies, the following predictions for crowding in the amblyopic and typical fovea can be made (illustrated in Figure 4.3 using my own stimuli). First, if crowding is modulated by similarity between the target and flankers, then crowding should strong (the extent of crowding should be greater) with flankers of the same polarity to the target and weak (the extent of crowding should be reduced) with flankers of opposite polarity (conditions B-C). This could be indicative either of crowding being selective for target-flanker similarity, or crowding being determined by target-flanker grouping. Second, if crowding is determined by grouping, then adding more flankers of opposite polarity to the target should reduce crowding more than two opposite-polarity flankers, showing flanker-flanker grouping (conditions C-D). Third, flankers of alternating polarity should strongly crowd the target, more than additional flankers of opposite polarity (conditions D-E), showing target-flanker grouping. Fourth, crowding should be weaker with seven flanker diamonds than one flanker diamond (conditions F-G), showing flanker-flanker grouping. If amblyopic crowding is determined by target-flanker similarity but not grouping, then differences should only be found between two same and opposite polarity flankers (B-C).

Prior to testing these flanker configurations in the amblyopic and typical fovea, I aimed to replicate previous findings in the periphery using my own stimuli and paradigm. The method used in this chapter to measure grouping effects differs from previous studies in two ways. First, as seen in Figure 4.3 Landolt-C stimuli were used, similar to the letters used in Rosen and Pelli (2015), but different from the vernier used in Manassi et al. (2013). This decision was made because amblyopic vision is characterised by increased positional uncertainty (Levi et al., 1987; Hess et al., 1997), and a vernier task would rely more heavily on fine-scale position judgements than a task involving the identification of the Landolt-C orientation. Second, the method used to measure crowding was different than what was used in Manassi et al. (2013). The majority of studies investigating grouping effects have measured crowding as threshold elevation from acuity (Sayim et al., 2008; Saarela & Herzog, 2009; Sayim et al., 2010; Manassi et al., 2012, 2013). This requires first the measurement of threshold for isolated targets, and then determining how many times this threshold is elevated with different flanker

configurations, whilst using a fixed spacing between the target and the neighbouring flankers. Here, I used the approach by Levi, Song, et al. (2007a) that I also used in the previous two chapters, where target size and flanker spacing are varied simultaneously using QUEST (Watson & Pelli, 1983). The advantage of this method is that it measures just one dimension: spacing as a multiple of target size. This reduces the time taken to separately determine the size of the target needed overcome acuity limitations, and then the appropriate spacing between the target and flankers for crowding to occur. This method is especially useful for testing observers with amblyopia, as contrary to the periphery, one size (and spacing) does not fit all; there is great variability in the required target size and spacing between observers due to variations in deficit. In Experiment 1, I aimed to replicate results of previous studies in the periphery (Manassi et al., 2013; Rosen & Pelli, 2015) using my stimuli and method. In Experiment 2, I tested the same conditions in the amblyopic and typical fovea.

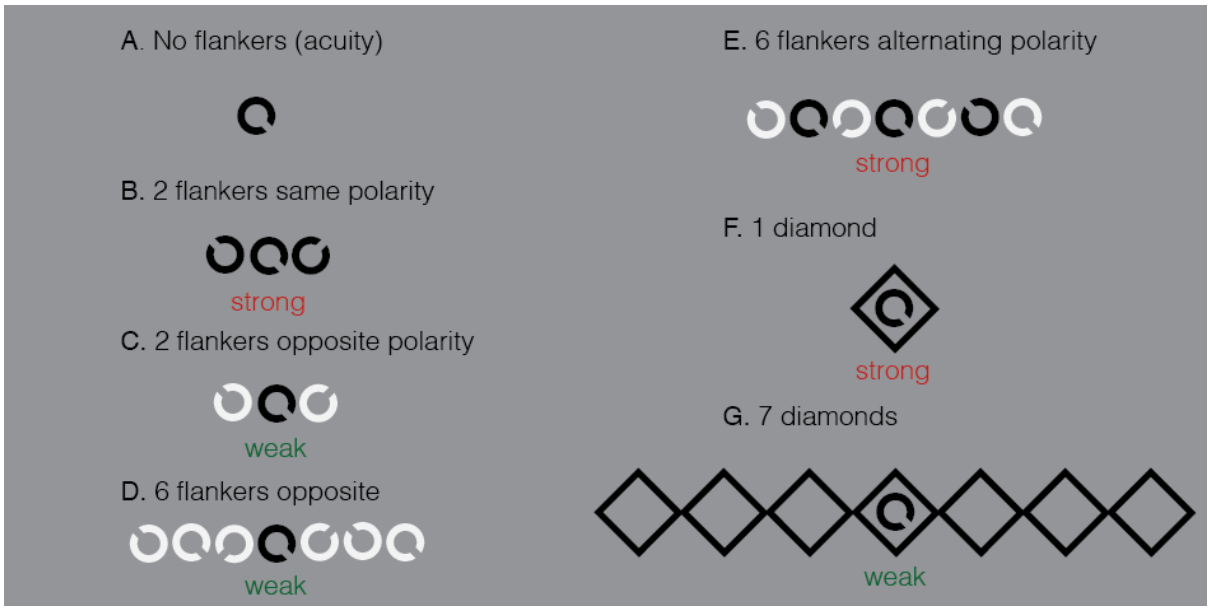

Figure 4-3 Illustration of stimuli in the seven flanker conditions in Experiments 1 and 2

Condition A: the Landolt-C target (black) was presented in isolation (no flankers) in order to measure uncrowded acuity. *Condition B*: the target was presented with two Landolt-C flankers, one on each side. Target and flankers were the same polarity (black). *Condition C*: the target was presented with two Landolt-C flankers, one on each side. The flankers were of opposite polarity (white) to the target. *Condition D*: the target was presented with six Landolt-C flankers, three on each side. The flankers were of opposite polarity to the target (white). *Condition E:* the target was presented with six flankers of alternating polarity, with three flankers on each side. This created an array of Landolt-Cs of alternating polarity. *Condition F:* the target was enclosed within a diamond flanker. Target and flanker polarity was the same (black). *Condition G*: the target was enclosed within a diamond flanker, with three flankers on each side of the diamond enclosing the target. Target and flanker polarity was the same (black). Note that conditions A-E are modified versions from Rosen and Pelli (2015), and conditions F and G are modified versions from Manassi, Sayim, and Herzog (2013). Under each condition “strong” or “weak” refer to the predicted effects of crowding based on these previous studies.

4.2 **Methods**

4.2.1 Observers

In Experiment 1 that measured flanker configuration effects in the typical visual periphery, ten adult observers with typical vision took part (4 males, mean age 28.2 years, range 23-37). In Experiment 2 that measured flanker configuration effects in the amblyopic and typical fovea, six observers with typical vision from Experiment 1 took part along with 4 observers that did not take part previously (7 males, mean age 32.3 years, range 24-49). For both experiments, observers with typical vision had normal or corrected-to-normal visual acuity and no history of binocular dysfunction.

In Experiment 2, ten observers with strabismic amblyopia (1 male, mean age 33 years, range 24-50) took part. The observers with amblyopia were selected based on at least a two line difference in logMAR (logarithm of minimum angle of resolution) acuity between the eyes, normal or corrected-to-normal visual acuity in their fellow unaffected eye, and a childhood history of amblyopia. Clinical details for the observers with amblyopia are presented in Table 4.1 below. The experimental procedures were performed with the informed consent of the observers and were approved by the East of England – Cambridge South Research Ethics Committee of the National Health System (NHS) Health Research Authority.

Subject	Age (years)	Sex	Ocular Alignment	Refractive Error		logMAR acuity	
				LE	RE	LE	RE
A1	39	F	n: R SOT, 16Δ d: R SOT, 16Δ	Plano	Plano	-0.1	0.2
A2	29	F	R XOT, 36Δ	Plano	Plano	0.0	0.48
A3	25	F	n: L XOT, 50Δ d: L XOT, 45Δ	Plano	Plano	0.3	-0.1
A4	37	F	n: L SOT,40Δ d: L SOT 25Δ	Plano	Plano	0.8	0.0
A5	31	F	n: L SOT 18Δ d: L SOT 25Δ	Plano	Plano	0.18	-0.18
A6	36	F	L SOT 8Δ	-3.00/ -0.25×140°	-3.75/ -0.75 ×166°	0.48	-0.1
A7	50	F	n: L SOT 10Δ d: L SOT 8Δ	Plano	Plano	0.6	-0.18
A8	26	M	n: R XOT 18ΔR/L 3Δ d: R XOT 8ΔR/L 4Δ	+2.50/ -1.25 ×82°	+6.25/ -1.75 ×133°	-0.1	0.3
A9	24	F	n: R XOT 45Δ d: R XOT 52Δ	Plano	Plano	0.1	1.0
A10	33	F	L SOR 25Δ Plano		Plano	0.42	0.0

Table 4-1 Clinical details of observers with amblyopia (N=10).

The "ocular alignment" column reports the outcome of both near (n) and distance (d) prism cover tests. When one value is denoted only, this indicates near cover test. SOT = esotropia (inward ocular deviation), XOT = exotropia (outward ocular deviation), R/L = right eye over left eye. The degree of deviation is shown in prism diopters (Δ) and the amblyopic eye is denoted. Note that 2Δ = 1°. Acuity is denoted as measured by the logarithm of the minimum angle of resolution (logMAR) chart.

4.2.2 Apparatus

Experiments were programmed using Matlab (The Mathworks, Ltd., Cambridge, UK) and the Psychophysics ToolBox (Brainard, 1997; Pelli, 1997). For Experiment 1, a Viglen Genie PC running was used. Stimuli were presented on a 24-inch Sony GDM-FW900 cathode ray tube (CRT) monitor (2304 × 1440 resolution, 80Hz refresh rate) to ensure uniform luminance across the screen. Liquid crystal display (LCD) monitors often suffer from backlight bleeding at the screen edges, and this could be especially problematic for the purposes of Experiment 1 involving the presentation of stimuli in the visual periphery near the screen edges. The monitor was calibrated using a Minolta photometer and linearised to give a maximum luminance of 122 cd/m^2. Observers rested their head on a chinrest that was positioned 57cm from the screen monitor at eyelevel, and used a USB portable numeric keypad to register their responses.

For Experiment 2, a Dell PC was used to run the experiment and stimuli were presented on an ASUS VG278HE LCD monitor (1920 x 1080 resolution, 120Hz refresh rate). The monitor was calibrated with a Minolta photometer and linearised in software to give a maximum luminance of 150 cd/m^2. Observers rested their head on a chinrest and used a USB portable numeric keypad to register their responses. The distance of the chinrest from the monitor at eye-level was 272cm for the observers in the control group. For some observers with amblyopia the required stimulus size exceeded screen borders at 272 cm. Therefore, for the amblyopia group I varied the distance from the screen to ensure we obtained reliable estimates of configuration effects: four observers were tested at 272cm, four observers were tested at 200cm, and two observers at 125cm. Eye-tracking was performed using an Eyelink 1000 (SR Research, Mississauga, ON, Canada) with a level desktop camera positioned in front of the chinrest at an approximate distance of 50cm from the eye that was being tracked.

In both experiments, one eye was patched using a plastic eye patch for monocular stimulus presentation. For observers with typical vision, the non-dominant eye was patched in order to test the dominant eye. Eye dominance was determined using the Miles test (Miles, 1928). For observers with amblyopia the fellow fixating eye was patched in order to test the amblyopic eye.

4.2.3 Stimuli

The target was a Landolt -C with a gap equal to one-fifth of the stimulus diameter, similarly to Sloan letters (Sloan, 1959). The target was always black (2 cd/m^2) and presented at 90% Weber contrast. Depending on the flanker condition, the flankers were either black or white (98 cd/m^2). Stimuli were presented against a mid-grey background (45cd/m^2). I tested seven flanker conditions (which can be seen in Figure 4.3): (A) no flankers (acuity); (B) two Landolt-C flankers, same polarity; (C) two Landolt-C flankers, opposite polarity; (D) six Landolt-C flankers, opposite polarity; (F) six Landolt-C flankers, mixed polarity; (G) one diamond flanker, same polarity; (F) seven diamond flankers, same polarity.

On each trial, the target was randomly oriented either 45°, 135°, 225°, or 315°. These orientations were chosen because during pilot testing it was determined that in conditions G and F, the diamond flankers were more disruptive on target recognition when the target had one of the oblique orientations, than when it was oriented in one of the cardinal orientations. This effect likely resulted because

when the target had one of the four oblique orientations, the separation between the Landolt-C gap and the flanker edge was smaller (see Figure 4.4. i). In contrast, when the Landolt-C target was oriented in one of the cardinal orientations, the gap coincided with the angle of the diamond flanker, and the separation between the Landolt-C gap and diamond flanker edge was greater (see Figure 4.4. ii), making the gap more easy to identify. When the condition included Landolt-C flankers (conditions B-E), an equal number of flankers was positioned on each side of the target horizontally (either one or three). With two Landolt-C flankers, two orientations were randomly chosen from 45°, 135°, 225° and 315° on each trial. With six Landolt-C flankers, 45°, 135°, 225° and 315° were each allocated to one flanker, and two orientations were repeated for the remaining two flankers. The target and flankers were evenly spaced by a relative centre-to-centre separation of 1.1 times the target size, the optimal spacing to use for this scaling method (Levi, Song, et al., 2007a). Thus, when three Landolt-C flankers were presented on each side, the second flanker was twice as far from the target as the innermost flanker, and the third flanker was three times as far from the target as the innermost flanker.

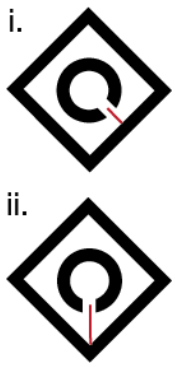

Figure 4-4

Differences in separation between Landolt-C gap and flanker edge (conditions G-F) depending on target orientation

For the conditions with diamond flankers (condition F and G), the orientation of the diamonds did not vary across trials. I used diamonds as flankers, that have also been used by Manassi et al. (2016) and shown to have the same effect as square flankers on a vernier. I selected diamonds because during pilot testing I found that a single diamond crowded the Landolt-C target more (i.e. had greater extents of crowding) than a single square. The width of the diamond outline was equal to the size of the gap of the Landolt-C. Similarly, the space between the diamond outline enclosing the target and the target was equal to the Landolt-C gap (excluding the diamond corners for which the space was larger). The relative centre-to-centre separation between the target and the innermost diamond flanker was 1.3 times the target size. The separation was greater than for the Landolt-C flankers, as the diamonds were larger in size in order for the diamond to enclose the target, and smaller separations resulted in the diamonds overlapping. This particular separation was chosen because it resulted in the gap between the Landolt-C target and the flanker edge being equal in size with the gap of the Landolt-C target.

4.2.4 Procedures

Gap size thresholds were measured by varying the overall size of the target (and thus the visibility of the gap of the Landolt-C that indicated orientation) using a QUEST algorithm procedure (Watson & Pelli, 1983) converging at 62.5% correct performance. The observers' task was a four-alternative-forced choice (4AFC) that required them to report the orientation of the Landolt-C target.

I used the intensities (i.e. case gap size values) presented during the QUEST runs and their corresponding proportion of correct responses to fit a psychometric function post-hoc. This was done as the post-hoc fitting of a psychometric function has been shown to provide more robust threshold estimates to attentional lapses than QUEST (Manning et al., 2018), and as such can also be useful in cases of high noise such as clinical groups (and thus observes with amblyopia). As QUEST is typically very quick to converge and proceeds to present stimulus intensities at threshold for the remaining trials, it can make the later fitting of a psychometric function, which requires a range of values above and below threshold, problematic (Wichmann & Hill, 2001). Therefore, in order to overcome this issue and increase the range of intensities presented, jitter was added to the QUEST threshold estimate for the next trial by adding noise of a Gaussian range of values between -0.25 and +0.25 of the current QUEST estimate (rounded to whole pixels). This increased the variability of estimates of intensities on a trial-to-trial basis, either presented values above threshold or below the QUEST threshold estimate for the next trial, depending on the sign of the random number added.

Figure 2 shows the stimulus presentation sequence for Experiments 1 and 2. Each trial progressed as follows: First, a Gaussian fixation dot appeared for 500ms. For peripheral stimulus presentation in Experiment 1 (Figure 4.5A), the Gaussian fixation dot was presented in the centre of the screen horizontally and was offset vertically by 300 pixels towards the bottom of the screen. In Experiment 2, the fixation dot was presented in the centre of the screen (Figure 4.5B). This was followed by the presentation of the target for 150ms, either in isolation (acuity) or with the flankers, depending on the flanker configuration condition.

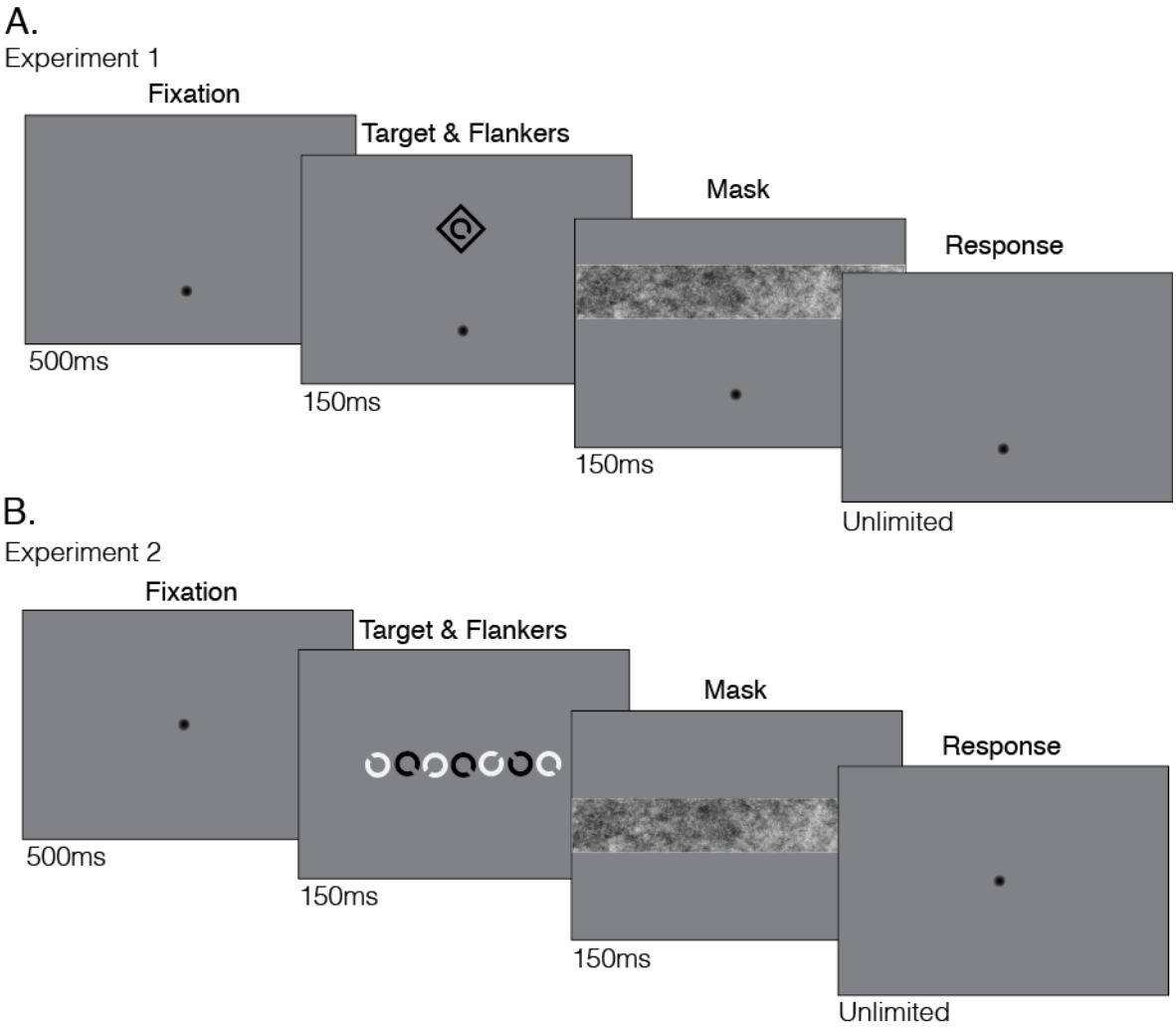

Figure 4-5 Stimulus presentation sequence during a single trial in Experiments 1 and 2

A. Stimulus presentation sequence on each trial in Experiment 1, where the adult periphery was tested. Each trial started with the presentation of the fixation dot near the bottom of the screen for 500ms. This was followed by the fixation dot and the target at 15° eccentricity, either in isolation (condition A) or with flankers (conditions B-G) for 150ms. Here condition F is depicted. A noise on f mask was finally presented for 150ms. While the observers made their response, the fixation dot was on screen. This presentation sequence was repeated until 65 valid trials were completed per QUEST run.
B. Stimulus presentation sequence on each trial in Experiment 2, where the typical and amblyopic fovea were tested. Each trial started with the presentation of the fixation dot in the centre of the screen for 500ms. This was followed by the target, either in isolation (condition A) or surrounded by flankers (conditions B-G) for 150ms . Here, condition E is depicted. A noise on f mask was finally presented for 150ms. While the observers made their response, the fixation dot was on screen. This presentation sequence was repeated for 65 valid trials.

In Experiment 1 the target was presented in the upper visual field, at 15° eccentricity from the fixation dot. In Experiment 2 the target was presented in the centre of the screen, and the fixation dot did not appear on screen during target presentation. To restrict stimulus processing time, a 1/f mask was presented for 150ms after the stimulus presentation. For all flanker conditions, the mask was the maximum size of the target in width, with a Gaussian fringe of 10 pixels on each side, and covered the entire screen in length (see Figures 4.5A and 4.5B). Finally,

the fixation dot was presented until the observers registered their response with a key press. The next trial was presented at an intertrial interval of 500ms. Two distant grey lines (one horizontal and one vertical) were on screen during the entire trial sequence as fixation guides. Each QUEST consisted of 65 trials. For each flanker condition, 3 QUEST blocks were completed by each observer, resulting in 195 trials for each of the seven conditions.

After data gathering was completed, trials from the 3 QUEST blocks for each flanker condition was amalgamated. For each condition, the gap size values presented were identified and the proportion of correct responses for each value was computed and a psychometric function was fitted (cumulative Gaussian). Note that gap sizes were binned as unique values, and thus involved the added jitter. The psychometric function was weighted by the number of trials tested per value, and had three free parameters: a mean, a standard deviation, and a lapse rate (key press error). From the psychometric function, threshold was derived as the gap size of the Landolt-C that yielded 62.5% correct performance. Analyses on gap size thresholds in minutes of arc (arcmins) are presented below in the Results section. Following the approach by Levi, Song, et al. (2007a), the extent of crowding (centre-to-centre separation between the target and flankers) corresponded to 5 times the gap size threshold (equal to one whole Landolt-C), multiplied by 1.1 (the relative constant separation between the target and a neighbouring flanker).

4.2.5 Eye-tracking

For Experiment 2, monocular eye-movements were recorded. Eye-calibration was performed before the first QUEST block of the experiment. Subsequently, calibration occurred every 4th QUEST block, unless there was significant head/body movement or eye drift (in which case the QUEST was paused and calibration was performed immediately). A custom-coded calibration sequence was used that involved the presentation of a white target circle on a gray background. The target was presented at five possible locations on a cross (centre, top, bottom, left, and right) in random order, and the observers were required to look at the target location. Target size was increased if required for observers with amblyopia in order to ensure visibility.

Eye-tracking was used to ensure that observers looked at the target during stimulus presentation. This was especially important for the conditions including flankers, as errors could arise due to fixating on a flanker element instead of the

target, resulting in large gap size thresholds not necessarily due to crowding but unstable fixation. A criterion distance set at 2° was used to restrict the area around the fixation dot (and thus the target during stimulus presentation, see Figure 2B) over which the observer's gaze was could vary. As amblyopia is characterized by increased fixation instability (Gonzalez et al., 2012; Subramanian et al., 2013), the criterion distance was increased if needed for observers with amblyopia. In the start of a given trial, the target was presented only if the observers' gaze was within the window defined by the criterion distance. During target presentation, if the observer's gaze deviated from the centre of the target for a distance greater than the criterion, then the trial was considered invalid. The observer was notified of the trial cancellation by a beep sound, and the trial was added to the end of the QUEST. Note that the target did not appear for the next trial unless the observer's gaze had returned within the criterion-defined window. Invalid trials were repeated until 65 valid trials were completed per QUEST run.

Post experiment completion, from the eye movement recordings on valid trials the X and Y positions of the eye (Cartesian coordinates) were obtained. For each valid trial, these positions were converted to distance from the fixation dot in degrees of visual angle. Next, eye blinks and saccades were detected and the X and Y positions corresponding to a blink or saccade were removed. To ensure only voluntary saccades were removed and not microsaccades, the minimum duration threshold for a saccade was set to 15ms (Rolfs, 2009). The resulting X and Y eye distances from fixation were pooled across the three QUEST blocks for each flanker condition. Finally, the average standard deviation of X and Y distances from fixation across all valid trials was computed, giving one measure of fixation variability for the X axis (horizontal fixation variability) and one estimate for the Y axis (vertical fixation variability). In the Results section, analyses on these measures of fixation variability, including a comparison between the control and amblyopic group, are presented. Note that eye-tracking failed for one observer with amblyopia, and thus data from only 9 out of 10 observers with amblyopia are presented in section 4.3.2.2.

4.3 **Results**

4.3.1 Experiment 1: Typical Visual Periphery

The group mean and individual gap size thresholds for the seven flanker conditions tested in Experiment 1 can be seen in Figure 4.6. Descriptive statistics (mean and standard deviation) for each flanker condition can be seen in Table 4.2,

expressed both as gap size thresholds (arcmin) and crowding extent values (centre-to-centre separation between the target and flankers in deg). It is clear that there are large differences in threshold estimates between the seven conditions. To further investigate these differences, I conducted a repeated measures ANOVA on gap size thresholds. The analysis showed a main effect of flanker condition, $F(2.10, 18.91)= 35.78$, $P<.001$ (Greenhouse-Geisser corrected). To see how this main effect of flanker condition relates to the individual hypotheses, I conducted planned repeated measures t-tests. Note that the conditions tested with t-tests corresponded to the comparison stated in the hypotheses, and were planned a-priori. I thus did not Bonferroni correct for multiple comparisons.

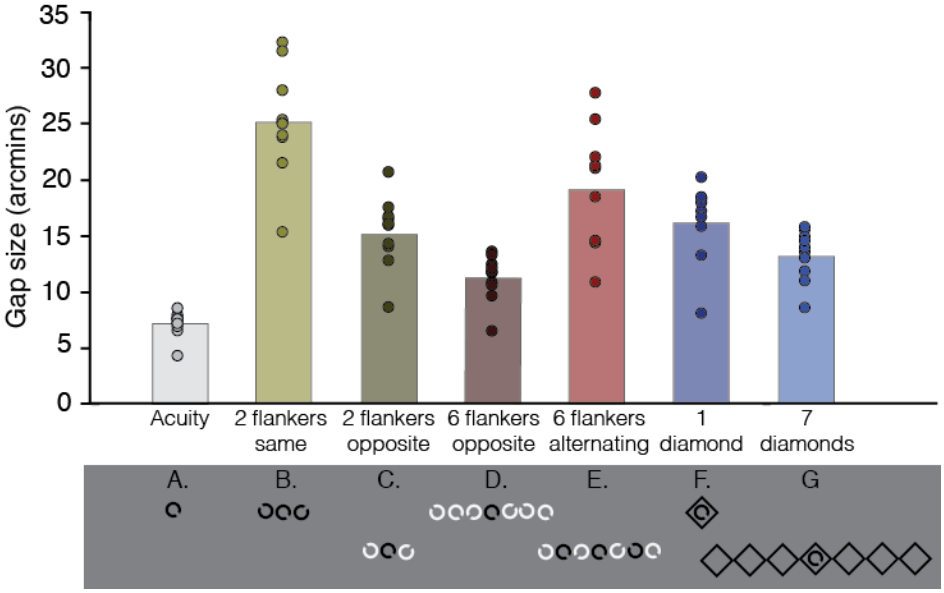

Figure 4-6 Gap size thresholds for the seven flanker conditions tested in the visual periphery

Gap size thresholds (in minutes of arc) for the control observers (N=10) tested at 15° in the visual periphery. Bars indicate the group average gap size for each flanker condition, whereas dots indicate the average gap size for each individual. Below the condition labels on the x-axis the stimulus configuration for each flanker condition is presented.

First, it was important to establish that the addition of flankers to an isolated target Landolt C led to an impairment in target recognition, which is a marker of crowding. Figure 4.6 shows a large increase in thresholds with two flankers of the same polarity compared to acuity, which is also clear in the descriptive statistics presented in Table 4.2. Indeed, two flankers of the same polarity appear to be the most disruptive of all flanker conditions. The repeated measures t-test showed that with these flankers, the impairment in performance compared to acuity was significant, as observers required larger gap size thresholds to identify targets

surrounded by flankers of the same polarity than isolated targets, t(9)= -14.102, *P* < .0001.

Observers' performance improved and gap size thresholds are smaller with two flankers of opposite polarity compared to two flankers of the same polarity. The repeated measures t-test showed this difference was significant, t(9)= 12.53, *P* <.0001, indicating that crowding was relieved with two flankers of opposite polarity, and demonstrating that crowding was modulated by the local similarity between target and flankers

		Acuity	2 flankers same	2 flankers opposite	6 flankers opposite	6 flankers mixed	1 diamond	7 diamonds
Gap size (arcmin)	M	7.16	25.17	15.34	11.27	19.04	16.44	13.29
	SD	1.08	4.60	3.03	1.95	5.13	3.29	2.14
Crowding extent (deg)	M	.66	2.31	1.41	1.03	1.75	1.51	1.22
	SD	.10	.42	.28	.18	.47	.30	.20

Table 4-2 Descriptive statistics on gap size thresholds in the periphery for acuity and the six flanker configurations.

Gap size thresholds are presented in arcmins as well as values for the extent of crowding in degrees of visual angle (5 × gap size threshold in deg × 1.1 the centre-to-centre separation between the target and a neighbouring flanker).

Figure 4.6 shows that there was also a clear difference in thresholds between two and six flankers of opposite polarity, with thresholds being lower with six flankers of opposite polarity to the target. In fact, gap size thresholds were significantly lower with six flankers of opposite polarity compared to two, t(9)= 7.87, *P*= <. 0001, demonstrating that peripheral crowding was further relieved by additional opposite polarity flankers, and suggesting that both polarity and flanker number matter. When the six flankers were of alternating polarity, gap size thresholds increased compared to six flankers of opposite polarity and this difference between the two conditions was significant, t(9)= -5.47, *P*= .0004. Note that the nearest flankers to the target were identical in these two conditions (D and E), as well as condition C. This suggests that the differences between conditions C and D, and D and E are not due to any differences in the local flankers neighbouring the target, but due to the global configuration of the target and flankers.

Figure 4.6 shows that a single diamond flanker disrupted target recognition, as thresholds were substantially higher than the acuity condition. Importantly, there

was a also difference between conditions F and G, with gap size thresholds being greater with one diamond flanker enclosing the target, compared to an array of seven diamond flankers. The t-test showed that this difference was significant, $t(9)= 3.92$, $P= .004$, indicating that the addition of extra diamonds relieved crowding.

Overall, in the typical visual periphery, I found that crowding was strong with two flankers of the same polarity to the flanker and weak with flankers of opposite polarity, and thus that crowding was modulated by the similarity between the flankers local to the target. The results were also indicative of flanker-flanker grouping, as crowding was further relieved when additional opposite polarity flankers were added to the two local opposite-polarity flankers. I also found that alternating polarity flankers strongly crowded the target compared to opposite polarity flankers, suggesting target-flanker grouping. Finally, the results showed that strong crowding with one flanker diamond was relieved when additional flanker diamonds were added, consistent with flankers ungrouping from the target and flanker-flanker grouping. All together, I replicated the findings by Rosen and Pelli (2015) and Manassi et al. (2013).

4.3.2 Experiment 2: Typical & Amblyopic Fovea

4.3.2.1 Gap Size Thresholds

Figure 4.7 shows the results from Experiment 2 for the control group and the group with amblyopia. Both groups were tested foveally. As the range of threshold values was substantially greater for the amblyopic group, gap size thresholds are plotted separately for each group to ensure the pattern of flanker configuration effects was visible for the control group. In order to compare gap size thresholds between the two groups, I conducted a 7 (flanker condition) × 2 (group) mixed effects ANOVA. This revealed a significant main effect of condition, $F(1.48, 26.58)= 4.5$, $P= .030$ (Greenhouse-Geisser corrected), and group $F(1,18) = 8.06$, $P= .011$, suggesting that performance differed both between flanker conditions, and between groups. Additionally, the effect of the flanker condition differed between the two groups, as indicated by the significant interaction between group and flanker condition $F(1.48, 26.58) = 4.02$, $P= .041$ (Greenhouse-Geisser corrected). To establish the exact pattern of flanker effects on crowding within each group, I conducted the same comparisons between flanker conditions reported in section 4.3.1 for the typical periphery in the typical and amblyopic fovea using repeated measures t-tests.

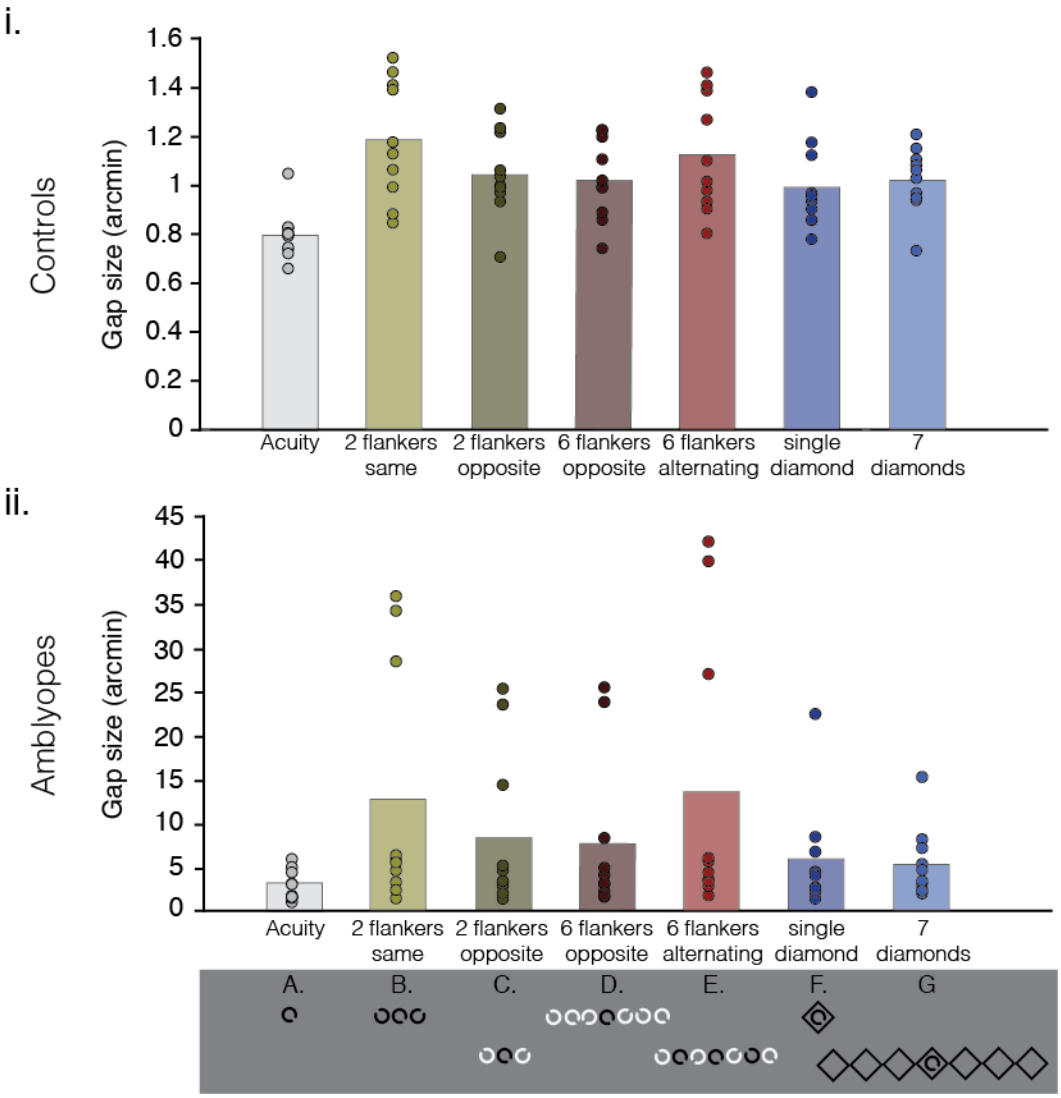

Figure 4-7 Gap size thresholds for the seven flanker conditions tested in the fovea for the control and amblyopic groups

i. Gap size thresholds (in minutes of arc) for the control observers (N=10) tested foveally. Bars indicate the group average gap size for each flanker condition, whereas dots indicate the average gap size for each individual.
ii. Gap size thresholds for the amblyopic observers (N=10), plotted with conventions as in A. Below the condition labels on the x-axis the stimulus configuration for each flanker condition is presented.

Figure 4.7.i shows variations in thresholds between the different flanker conditions, indicating differential effects on crowding in the fovea. In order to establish whether flankers impaired target recognition in the typical fovea, I first compared gap size thresholds between the acuity condition and two flankers of the same polarity to the target. Figure 4.7.i clearly shows that with two flankers of the same polarity, thresholds increased compared to the acuity condition where flankers were presented in isolation (see also Table 4.3 for descriptive statistics). This increase in thresholds was significant, $t(9)= -6.04$, $P= <.001$, demonstrating that flankers significantly impaired target recognition in the typical fovea – a marker of

crowding. Note however that although significant, the difference in thresholds between these conditions A and B was on average less than 0.5 a minute of arc. It is important to bear in mind the small magnitude of the differences when considering the remaining comparisons. Thresholds were lower with two flankers of opposite polarity to the target compared to two flankers of the same polarity, suggesting that reversing flanker polarity relieves crowding, similarly to the typical periphery. Indeed, the repeated measures t-test revealed that the difference between conditions B and C was significant, t(9)= 3.17, *P*= .011, suggesting that local target-flanker similarity modulates the strength of crowding in the typical fovea.

Gap size (arcmin)		Acuity	2 flankers same	2 flankers opposite	6 flankers opposite	6 flankers mixed	1 diamond	7 diamonds
CON	M	.80	1.19	1.05	1.02	1.13	0.99	1.02
	SD	.10	.23	.17	.16	.23	.17	.13
AMB	M	3.06	12.93	8.73	7.99	13.79	6.10	5.55
	SD	1.61	13.28	8.62	8.58	15.27	5.85	3.83

Table 4-3 Descriptive statistics on gap size thresholds in the seven flanker conditions for the control (CON) and amblyopic (AMB) groups

However, in contrast to the typical periphery, there were no differences between two flankers of opposite polarity to the target and six flankers of opposite polarity, t(9)= .73, *P*= .487, demonstrating that increasing the number of opposite polarity flankers did not have an effect on crowding in the typical fovea. Although as previously stated the differences between conditions were small, thresholds for two and six opposite polarity flankers were greater than the acuity threshold. This suggests that there was still some range for improvement in performance when adding more opposite polarity flankers. The lack of a difference between two and six opposite polarity flankers thus likely does not result from two opposite polarity flankers completely releasing crowding (i.e. performance reaching acuity levels). Figure 4.7.i shows that there was a difference in gap size thresholds between six flankers of opposite polarity and six flankers of alternating polarity. Particularly, gap size thresholds were greater with alternating polarity flankers compared to six flankers of opposite polarity, and this difference was significant, t(9)= -2.68, *P*= .025, matching the typical periphery.

Figure 4.7.i clearly shows that one diamond flanker had a disruptive effect on target recognition, as thresholds were elevated relative to acuity. However, gap size thresholds were very similar between the condition in which one diamond enclosed the target and with an array of seven diamond flankers. This was confirmed by the paired samples t-test that showed no significant differences between the diamond conditions, t(9)= -.52, *P*= .616, contrary to the visual periphery where the addition of diamonds alleviated crowding.

Figure 4.7.ii shows a similar pattern to the typical fovea of the different flanker conditions on the amblyopic fovea. The addition of two flankers of the same polarity impaired performance, and gap sizes thresholds for this condition were significantly larger than for acuity measurements where there were no flankers, t(9)= -2.44, *P*= .037. This demonstrates that flankers had a significant effect on target recognition in the amblyopic fovea, a marker of crowding.

Similarly to both typical periphery and fovea, gap size thresholds were smaller in the amblyopic fovea with two flankers of opposite polarity to the target than two of the same polarity. Indeed, this difference in gap size thresholds between the two conditions was significant, t(9)= 2.49, *P*= .034, demonstrating that local target-flanker similarity in polarity modulates crowding in the amblyopic fovea.

However, similarly to the typical fovea, thresholds were very similar between two and six flankers of opposite polarity, suggesting that the number of opposite polarity flankers was irrelevant to amblyopic crowding. This was confirmed by the repeated measures t-test that showed no significant difference between these conditions, t(9)= 1.15, *P*= .282, demonstrating that crowding was not relieved by increasing the number of opposite polarity flankers. Similarly to the typical fovea, the lack of an improvement in performance with additional opposite polarity flankers is unlikely to result from crowding already being fully relieved with two opposite polarity flankers. In fact, the average threshold with two opposite polarity flankers was more than twice that for acuity (see Figure 4.3). This clearly suggests that there was a large range for improvement with the addition of more opposite polarity flankers. However, when flankers were of mixed polarity, performance deteriorated, and thresholds were higher on average with six flankers of mixed polarity compared to six flankers of opposite polarity to the target. This suggests that the former flanker condition was more disruptive to target recognition. However, contrary to the typical fovea, the difference between these two conditions was not significant, t(9)=

-1.69, *P*= .126, even though the elevation in thresholds with mixed polarity flankers was high on average (Figure 4.3).

Finally, Figure 4.7.ii shows a very small difference between conditions F and G, with one diamond flanker being only slightly more disruptive on average than seven diamond flankers. However, this difference was minimal and the t-test did not yield a significant result, t(9)= .73, *P*= .486, suggesting that similarly to the typical fovea the number of diamond flankers did not affect crowding in the amblyopic fovea.

Although the planned comparisons between the different flanker conditions did not yield significant results for the amblyopic fovea when changes in the global flanker configuration were considered, it is important to consider the significant variability in thresholds for the amblyopic group. It is evident from Figure 4.7 B that although acuity thresholds did not show much variability between observers with amblyopia, when flankers were present there was a subset of observers that showed substantially higher thresholds compare to the rest of the amblyopic group. Such variability could have muted effects of the global flanker configuration. To further explore whether there was a trend indicative of observers with amblyopia performing worse in one condition over the over, in Figure 4.8.A I plot thresholds from one condition against the other for each of the pre-planned comparisons. For comparison purposes, I also present the comparisons in the typical periphery where differences between conditions are large, and the typical fovea.

In Figure 4.8.A gap size threshold values from the condition in which flankers were of the same polarity to the target are plotted against the condition with two flankers of opposite polarity. Note that the rectangle half on which each dot falls (i.e. whether it is above or below the red line of unity) indicates the condition in which that observer had a higher threshold. It is clear that in the typical periphery all observers had higher thresholds with two flankers of the same contrast polarity than with two flankers of opposite polarity. This was also the case for the typical fovea, where only one observer showed a greater gap size threshold with flankers of the opposite polarity. For the amblyopic fovea, all the observers showed larger thresholds with flankers of the same polarity to the target compared to flankers of opposite polarity. In line with the results from the statistical analyses, Figure 4.8.A shows that target- flanker similarity in polarity modulated crowding consistently in the typical periphery and fovea, as well as the amblyopic fovea.

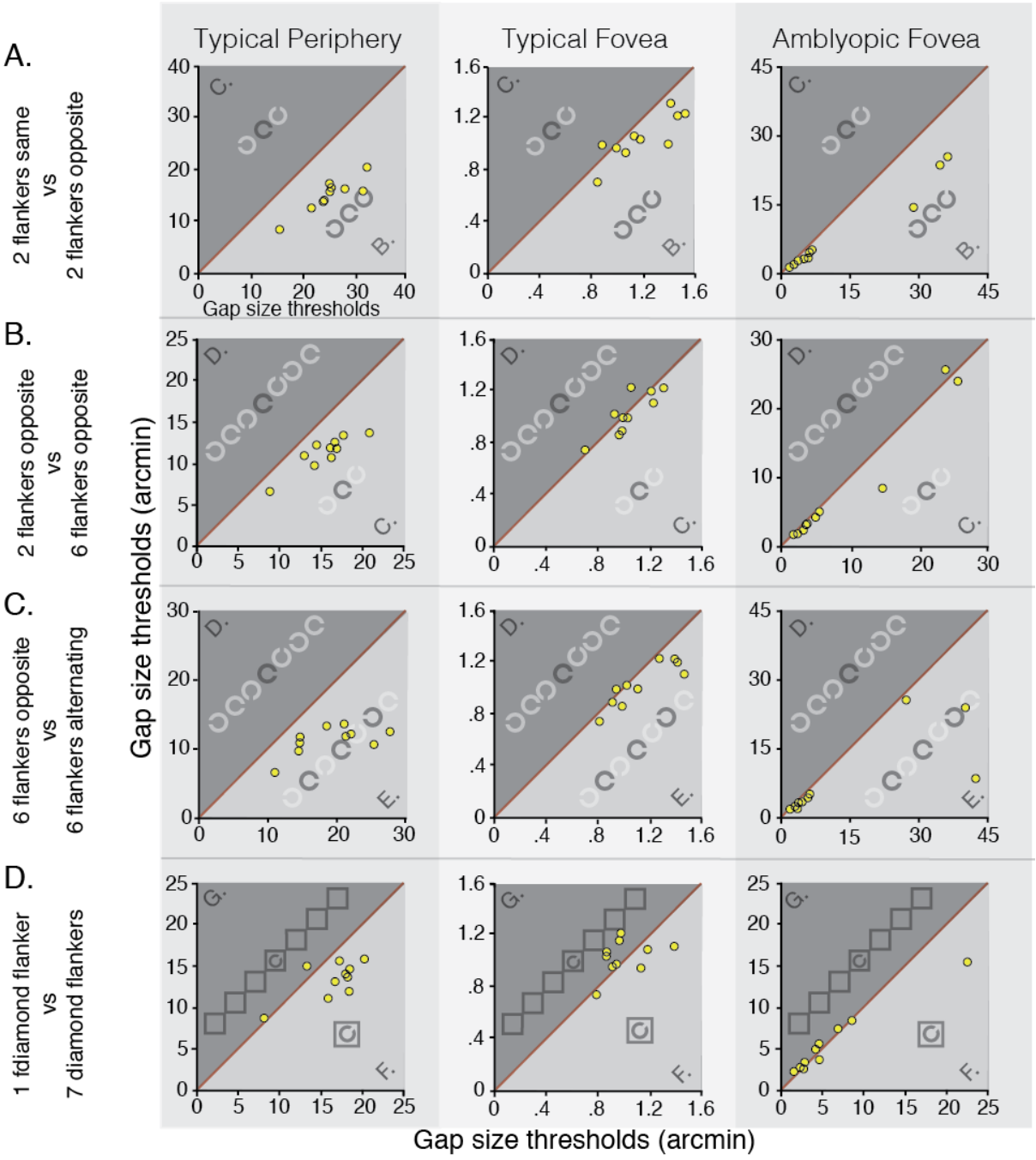

Figure 4-8 Gap size thresholds in one condition plotted against the other for each of the four crowding comparisons

A. Comparison of individual thresholds (in arcmins) with two flankers of the same polarity to the target and two flankers of opposite polarity. The left panel shows the individual thresholds from the visual periphery, the middle panel from the typical fovea, and the right panel from the amblyopic fovea. Each dot indicates the threshold for an individual observer. The stimuli on each triangle are indicative of the flanker condition. The triangle on which the dot is placed is indicative of a higher threshold and worse performance. The red line indicates equal performance.
B. Comparison of individual thresholds with two and six flankers of the opposite polarity to the target. Plotted with conventions as in A.
C. Comparison of individual thresholds with six flankers of opposite polarity and six flankers of alternating polarity. Plotted with conventions as in A.
D. Comparison of individual thresholds with a single diamond enclosing the target and seven diamond flankers. Plotted with conventions as in A.

Figure 4.8.B shows thresholds from the two-flanker opposite-polarity condition plotted against the six-flanker opposite-polarity condition. As expected based on the statistical analyses, all observers tested in the typical periphery had higher thresholds with two opposite polarity flankers compared to six, placing them in the lower half of the graph. However, although I find no significant difference between these conditions in neither the typical nor the amblyopic fovea, the pattern of results was different in these two instances (Figure 4.8.B). In the typical fovea, there was no consistency in observers' performance, with some observers having greater thresholds with two opposite polarity flankers, others with six, and one observer showing no difference between the conditions. In contrast, in the amblyopic fovea it appears from Figure 4.8.B that two opposite polarity flankers were more disruptive than six for 8 out of 10 of the observers. This suggests that although there was no consistent effect of opposite polarity flankers in the typical fovea, there was a weak trend of crowding being stronger with two opposite polarity flankers compared to six in the amblyopic fovea. As such, crowding in the amblyopic fovea may be relieved with additional opposite polarity flankers, but not to the extent observed in the typical periphery.

In Figure 4.8.C, thresholds from six opposite polarity flankers are plotted against thresholds from alternating polarity flankers. It is clear that in the typical periphery six alternating polarity flankers were more disruptive than six flankers of opposite polarity for all the observers. For one observer opposite polarity flankers were more disruptive, and another showed no difference between conditions. For the remaining observers in the control group gap size thresholds were greater with alternating polarity flankers, resulting in the significant difference between these two flanker conditions. Although the t-test did not yield significant results in the amblyopic fovea, thresholds were greater for all observers with alternating polarity flankers compared to opposite polarity flankers. However, as indicated by some observers placing very close to the line of unity, the difference between these two conditions was very small: for three observers it was less than 1 arcmin, and for five observers less than 2 arcmin. Nonetheless, this suggests that there was a consistent trend for crowding being stronger in the amblyopic fovea with six flankers of alternating polarity compared to six flankers of opposite polarity, indicative of a weak effect of the global flanker configuration.

Finally, in Figure 4.8.D, thresholds with a single diamond flanker are plotted against thresholds with an array of seven diamond flankers. In the typical periphery,

although two observers had lower thresholds with one diamond flanker, for the remaining observers seven diamonds aided target recognition relieving crowding. In the typical fovea, there is no consistent effect of the number of flanker diamonds, as three observers had almost identical thresholds, three had greater thresholds with one flanker diamond, and four had greater thresholds with seven diamonds. Similarly, for observers tested in the amblyopic fovea, one observer had almost equal thresholds in the two conditions, two observers had greater thresholds with the single diamond flanker and the remaining had greater thresholds with seven diamonds. As such, in line with the results from the statistical analyses, I observe no trend of the global configuration with diamond flankers in neither the typical nor the amblyopic fovea.

Overall, these results demonstrate that the similarity between local flankers and the target in contrast polarity modulates crowding in the typical and the amblyopic fovea, as crowding was strong with same-polarity flankers to the target and weak with opposite-polarity flankers in both instances of crowding. When additional opposite polarity flankers were placed, I did not find significant indications of flanker-flanker grouping. I only observed a weak trend for crowding to be further relieved with six opposite-polarity flankers relative to two for some observers with strabismic amblyopia, but no such trend in the typical fovea. In the typical fovea, I found that alternating polarity flankers were more disruptive than six opposite polarity flankers, indicative of target-flanker grouping. This effect was only significant in the typical fovea, whereas in the amblyopic fovea I only observed a trend. Our findings do not support any effects of target-flanker ungrouping and flanker-flanker grouping in neither the typical nor the amblyopic fovea, as crowding was similar with a single diamond flanker and multiple diamonds. Therefore, I find partial support for the findings by Rosen and Pelli (2015) in the fovea of observers with typical vision and amblyopia, but no evidence of the effects observed in Manassi et al. (2013).

4.3.2.2 Fixation Variability

Although trials in which eye movements exceeded the criterion distance were deemed invalid and did not contribute to QUEST estimates of gap size threshold (see Methods section), fixation variability could still be high for an observer whilst being within the criterion distance. If fixation variability was disproportionately high in one flanker condition over the other, this could have resulted in artificial

differences between the flanker conditions. This would mean that increased fixation variability, and not the differences between the target-flanker configurations, underlies the differences found in section 4.3.2.1. As amblyopia is characterised by poor fixation stability (Gonzalez et al., 2012; Subramanian et al., 2013), this increased fixation variability biasing the results is more likely for the group with amblyopia.

Figures 4.9.i and 4.9.ii show horizontal and vertical fixation variability, respectively, for the control and amblyopic groups. Although there were no marked differences between the different flanker conditions, it is clear that the amblyopic group shows greater fixation variability on average compared to the control group. The differences in fixation variability could have contributed to the main effect of group on gap size thresholds found in section 4.3.2.1, and as such it was important to examine them further.

In order to investigate these differences between the two groups, a 7 (flanker condition) × 2 (group) mixed effects ANOVA on horizontal and vertical fixation variability was conducted. For horizontal fixation variability, the analysis revealed that there was no significant difference between the different flanker conditions, $F(3.03, 51.44)= .83$, $P= .484$ (Greenhouse-Geisser corrected). Crucially, it showed that the increased horizontal fixation variability observed in Figure 5A for the amblyopic group was not significant, as the main effect of group was not significant, $F(1,17)= 3.41$, $P= .082$. The ANOVA on vertical fixation variability revealed that the main effect of flanker condition was not significant, $F(2.15, 36.63)= .645$, $P= .542$ (Greenhouse-Geisser corrected). Similarly to horizontal fixation variability, although the amblyopic group showed greater vertical fixation variability on average, I did not find a significant main effect of group, $F(1, 17)= 3.60$, $P= .075$. There results clearly show that there were no differences between the different flanker conditions in fixation variability that could have driven the differences in gap size thresholds discussed earlier. Additionally, these results also suggest that the magnitude of the differences between two groups in horizontal and vertical fixation variability was not significant, likely due to the majority of observers with amblyopia showing fixation variability within the range of controls.

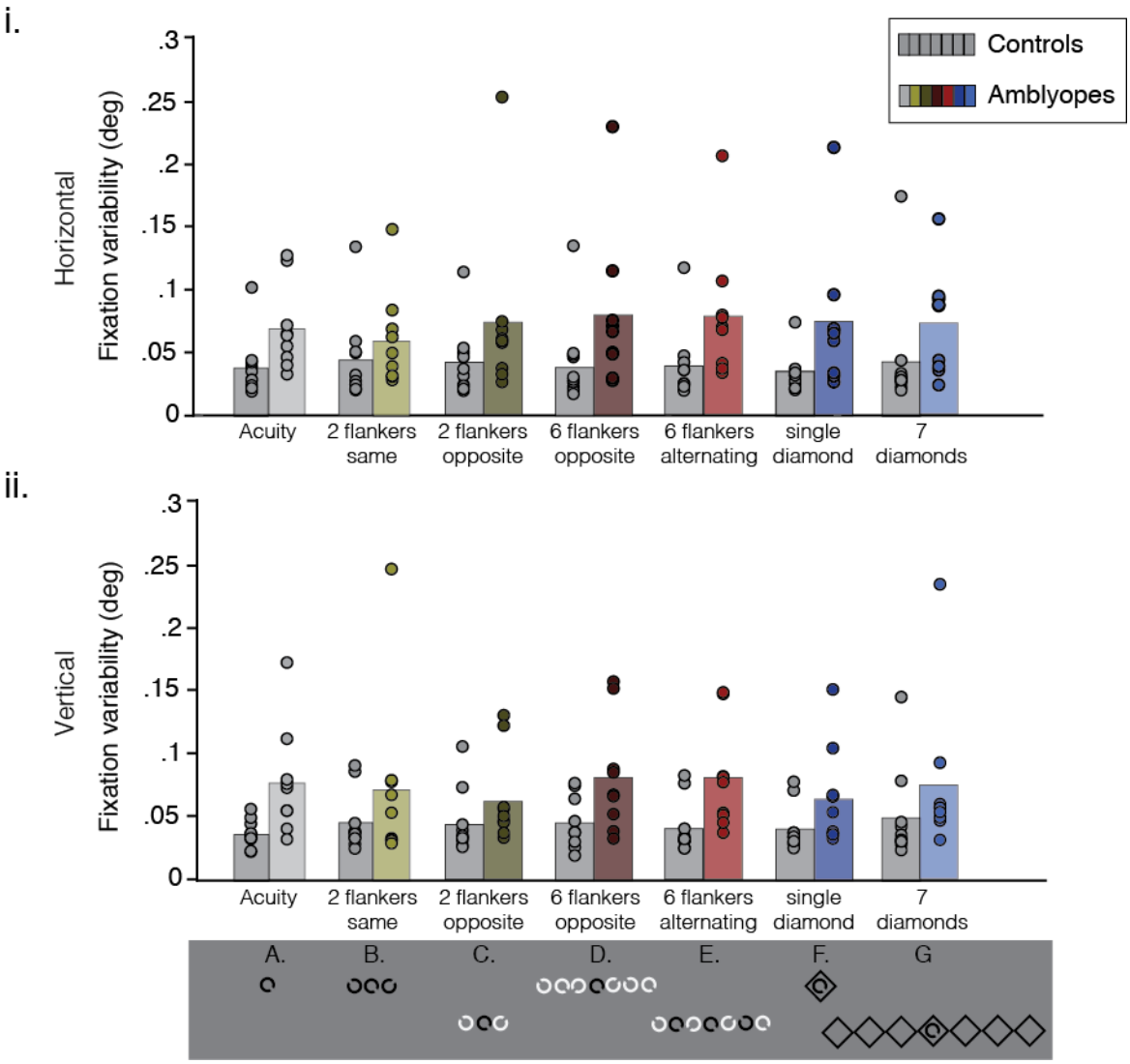

Figure 4-9 Horizontal and vertical fixation variability for the control and amblyopic groups

i. Horizontal fixation variability (in degrees of visual angle) for the control (N=10) and amblyopic (N=9) groups in the seven flanker conditions. Bars indicate the group mean for that flanker condition and dots indicate the mean fixation variability for each observer. Within each flanker condition, the control group is plotted with the light colour shade and the amblyopic group is plotted with the dark colour shade.
ii. Vertical fixation variability plotted with conventions as in A.

Despite the lack of statistical differences between the flanker conditions and between the two observer groups, it was important nonetheless to establish that the increased fixation variability observed for the amblyopic group did not have an effect on gap size thresholds. I thus investigated the relationship between gap size thresholds and an average measure of fixation variability (i.e. averaged across horizontal and vertical fixation variability) using Pearson correlations. If there increased fixation variability contributed to the differences between the flanker conditions in strabismic amblyopia, there should be a correlations between increased fixation variability and gap size thresholds. There was no significant

relationship between gap size thresholds and fixation variability for the acuity condition (r= .54, *P*= .136, n= 9), two flankers of the same polarity (r =.34, *P*= .371, n=9), two flankers of opposite polarity (r =.06, *P*= .870, n= 9), six flankers of opposite polarity (r= .28, *P*= .461, n= 9), six flankers of mixed polarity, (r= -.15, *P*= .696, n= 9), one diamond flankers (r= -.07, *P*= .655, n=9), or seven diamond flankers (r= .24, *P*= .539, n= 9). From the above results I can conclude that fixation variability did not significantly contribute to gap size thresholds in amblyopia. Therefore, fixation variability is likely not a confounding factor in either the presence or absence of differences between flanker conditions, nor in the larger gap size thresholds found for the amblyopic group.

4.3.3 **Discussion**

The aim of this study was to investigate the effects of the global configuration of the target and flankers on crowding in the amblyopic fovea. Amblyopic crowding was modulated by target-flanker similarity in polarity, with crowding being strong when two flankers were of the same polarity to the target and weak when two flankers were of opposite polarity. However, when flankers of opposite polarity to the target increased from two to six, there was no improvement in target recognition. Similarly, the addition of diamond flankers did not relieve crowding with one local diamond flanker. Taken together, these findings show that there was no flanker-flanker grouping in amblyopia. There was however a weak trend for alternating polarity flankers being more disruptive than opposite polarity flankers, suggestive of target-flanker grouping. Together, these results demonstrate that amblyopic crowding shows no evidence of flanker-flanker grouping, and only weak evidence for target-flanker grouping. Therefore, grouping cannot fully account for the effects of crowding in strabismic amblyopia.

In the typical periphery, crowding was modulated by the target-flanker similarity in contrast polarity, in line with previous findings (Chakravarthi & Cavanagh, 2007; Chung & Mansfield, 2009; Rosen & Pelli, 2015). This was also found for the amblyopic fovea: the extent of crowding was greater with flankers of the same polarity to the target and smaller with flankers of opposite polarity. These findings are inconsistent with previous studies that have found no difference in crowding with flankers of the same and opposite polarity to the target in strabismic amblyopia (Hess, Dakin, et al., 2001; Hariharan et al., 2005). This inconsistency could be related to the differences in the stimuli used between studies. Although

similarly to my paradigm, both previous studies investigating flanker polarity effects on amblyopic crowding have used Landolt-C targets, the flankers were either bars enclosing the target (Hess, Dakin, et al., 2001) or circular patches smaller than the target (Hariharan et al., 2005). If crowding in amblyopia also depends on the similarity between the target and flankers in shape and size similarly to the periphery (Kooi et al., 1994), this could have resulted in the same-polarity flankers not being as disruptive as in the experiment reported in this chapter. If flankers of the same polarity were not maximally disruptive, then the relieving effect of reversing flanker contrast polarity would be muted. The results here show that reversing the flanker polarity in amblyopia does not result in a complete release from crowding, but rather a decrease in the extent relative to flankers of the same polarity to the target. As such, the muted effect of target-flanker similarity in Hess, Dakin, et al. (2001) and Hariharan et al. (2005) would not be manifested as a ceiling effect, but as similar target recognition performance with same- and opposite- polarity flankers to the target, consistent with their results. Overall, the results demonstrate that amblyopic crowding is modulated by target-flanker similarity in contrast polarity.

In the typical fovea, the extent of crowding was also modulated by the similarity in polarity between the target and flankers. For the typical fovea, previous studies have shown that the effect of polarity relies on the size of the target: for near-threshold targets, reversing the polarity of the flankers abolishes crowding (Hess, Dakin, & Kapoor, 2000), whereas for suprathreshold targets there is no difference between same and opposite polarity flankers (Hess, Williams, & Chaundry, 2001; Ehrt, Hess, Williams, & Sher, 2003). The lack of modulatory effects with supratheshold targets could be indicative of a ceiling effect, pointing to the very small scale of these effects in the typical fovea. In this chapter, the differences in gap-size thresholds among different conditions in the typical fovea were tiny. Specifically, the difference between an isolated target and a target surrounded by flankers of opposite polarity was approximately a quarter of a minute of arc on average. This difference in size between isolated targets and targets of opposite polarity suggests that the target size was near-threshold, in line with the previous studies finding a target size-dependency of the polarity effect in the typical fovea. Note however that the elevation in thresholds with opposite polarity flankers relative to acuity shows that flankers still had an effect on target recognition, and performance was limited by crowding. Taken together, these findings demonstrate that the effect of target-flanker similarity in polarity with local flankers is common

across typical and amblyopic vision. However, as this effect would be predicted both by selectivity for target-flanker similarity along polarity, and the target-flanker grouping, it cannot differentiate between pooling and grouping accounts of crowding.

Although the findings clearly demonstrate that crowding in amblyopic and typical vision is modulated by target-flanker similarity in polarity, when flankers were multi-element arrays their effects were not the same across all instances of crowding tested. I first consider the findings from the typical periphery. When the number of flankers of opposite polarity to the target was increased from two to six, the extent of peripheral crowding was further reduced. Adding more flankers also had a relieving effect on crowding with diamond flankers, for which the extent of crowding was reduced with seven flanker diamonds compared to just one diamond enclosing the target. Taken together, these two results are line with flanker-flanker grouping relieving crowding in the visual periphery. On the other hand, when the flankers created an array of alternating polarity elements with the target, the extent of crowding was greater than with opposite polarity flankers, in line with target-flanker grouping. These findings replicate the pattern of results found in Rosen and Pelli (2015) and Manassi et al. (2013). They are also are consistent with previous studies showing that increasing the number of dissimilar flankers to the target has an increasingly relieving effect on crowding (Saarela et al., 2009), whereas flankers of alternating polarity and colour strongly disrupt target recognition (Manassi et al., 2012). The findings in the typical periphery are thus consistent with grouping accounts of crowding arguing that the global configuration of the target and flankers, and whether or not the target and flankers group, determine the strength (i.e. in this case the spatial extent) of crowding (Herzog & Manassi, 2015; Herzog et al., 2015; Francis et al., 2017).

Contrary to the typical periphery, manipulations of flanker polarity and shape in the amblyopic and typical fovea gave mixed results with regards to whether the global target-flanker configuration modulates crowding. When flankers of opposite polarity to the target were increased from two to six, the extent of crowding was similar in both instances of foveal crowding. Similarly, the extent of crowding did not differ between one and seven flanker diamonds. Taken together, these results show that additional flankers that match the ones most proximal to the target do not have a relieving effect, and thus do not reduce the extent of crowding in the amblyopic and typical fovea. However, when six flankers formed an array of alternating polarity

elements with the target, the extent of crowding was greater than with six opposite polarity flankers. The magnitude of this effect differed between the amblyopic and typical fovea. In the amblyopic fovea, only a weak trend was found, whereas in the typical fovea alternating polarity flankers substantially increased the extent of crowding relative to opposite polarity flankers. Although grouping by shapes as in Manassi et al. (2013) has not been tested previously in either the typical or the amblyopic fovea, the changes in polarity in the global target-flanker configuration have been examined in the typical fovea. Sayim et al. (2008) used a vernier offset discrimination task and found that adding more flanker lines of opposite polarity did not substantially improve offset discrimination relative to just two lines on each side of the vernier, but alternating flanker lines however strongly disrupted vernier discrimination. The findings on the typical fovea are thus consistent with Sayim et al. (2008). Therefore, in the amblyopic and typical fovea there was no evidence of flanker-flanker grouping, but evidence of target-flanker grouping.

A potential explanation for the lack of flanker-flanker grouping effects in amblyopic crowding relate to the perceptual distortions that have been found in some observers with amblyopia. These distortions are idiosyncratic to each observer with amblyopia, but can result in straight lines appearing “wavy” or jagged (Barrett et al., 2003). Based on the account by Francis et al. (2017) on how grouping could operate on crowding, these distortions could have resulted in the formation of boundaries between the target and flankers in conditions in which their boundaries are typically separated. Due to the variability of these distortions (Barrett et al., 2003; B. T. Barrett, A. Bradley, & P. McGraw, 2004), there are multiple ways that they could have affected the target and flanker percept. For example, if in the diamond flankers the lines did not appear straight but wavy, this might have created connected boundaries with the circular target. Although such distortions might account for the results in the diamond flanker conditions, it less clear how they would apply to manipulations of flanker polarity. The reduction in the extent of crowding with two opposite-polarity flankers compared to two same-polarity flankers suggests that boundary formation with regards to polarity was normal between the target and the flankers most proximal to it. Crucially, the typical fovea showed the exact same pattern of target-flanker configuration effects. This indicates that the lack of flanker-flanker grouping in the amblyopic fovea is not a consequence of a grouping deficit.

In this chapter, the only evidence for grouping effects determining crowding in the typical and amblyopic fovea was found in one condition with multi-element flanker arrays. If crowding were to be determined by grouping in the typical and amblyopic fovea, one would reason that grouping would modulate crowding across flanker configurations, not selectively only with alternating polarity flankers. As such, crowding in the two foveal instances cannot be fully accounted by a grouping framework (Herzog & Manassi, 2015; Herzog et al., 2015), and an interpretation of the results based on a grouping model such as the one by Francis et al. (2017) may not be necessary. Still, the trend for a larger extent of crowding with alternating polarity flankers relative to opposite polarity flankers poses a complication to the pooling framework followed in this thesis. Although pooling models have been designed primarily to account for the effects of crowding with only one flanker on each side of the target (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010; van den Berg et al., 2010; Harrison & Bex, 2015), these models typically assume that crowding occurs over a restricted spatial region and is determined by target-flanker similarity (e.g. Kooi et al., 1994). Consequently, if only the flankers most proximal to the target determined how disruptive crowding was on target recognition, then there should have been no difference in the extent of crowding between alternating and opposite polarity flankers. However, one explanation that would be consistent with pooling models concerns interference from the black flankers twice removed from the target. Despite not being the flankers most proximal to the target, they could still fall within the spatial extent of crowding. Due to their similarity to the target in polarity, and the demonstrated selectivity of both typical and amblyopic crowding for target-flanker similarity, they could still interfere with target recognition. Therefore, the interference from these farther black flankers would lead to an elevation in the extent of crowding relative to opposite polarity flankers in the typical and amblyopic fovea.

Although crowding in the amblyopic and typical fovea may not be well suited within a pooling framework, the effects of multi-element flanker arrays on peripheral crowding matched grouping predictions. If a grouping framework were to be abandoned, this raises the question of whether there could be a common process that could explain flanker configuration effects in the periphery and the typical and amblyopic fovea. It has been argued that the relieving effect of increasing the number of dissimilar flankers to the target may be due to positional cueing (Rosenholtz et al., 2017; Yu & Rosenholtz, 2018), instead of flanker-flanker grouping.

With multiple dissimilar flankers to the target, the flankers on each side of the target act as cues to the target position. In contrast, when target and flankers form a uniform texture, as with alternating polarity flankers, it is not clear which feature belongs to the target and which to the flankers. In these cases the flankers do not act as cues, as target position becomes more uncertain. As positional uncertainty is high in the typical periphery (Levi & Klein, 1986; Levi et al., 1987; Hess & Hayes, 1994), the cuing of target position could be alleviating positional uncertainty. Such a reduction in positional uncertainty, and not flanker-flanker grouping, could thus be underlying the improvement in target recognition when the number of opposite-polarity flankers and diamond flankers is increased. In contrast, in the fovea positional discrimination is highly accurate (Westheimer & Hauske, 1975; Westheimer, 1981), and these positional cues might not be required, resulting no additional decrease in the extent of crowding when more opposite-polarity or diamond flankers are added to the target. Quantitative differences in positional uncertainty between the fovea and the periphery could explain why additional dissimilar flankers to the target in shape and polarity may be utilised as position cues alleviating crowding in the periphery, but not the typical and amblyopic fovea.

A complication in this interpretation is that increased positional uncertainty has also been found in the amblyopic fovea (Levi et al., 1987; Demanins & Hess, 1996; Hess et al., 1997; Wang et al., 1998). Although positional uncertainty could be higher in the amblyopic than the typical fovea, it might still not have be as high as in the visual periphery. Levi et al. (1987) measured bisection acuity by having observers with amblyopia and typical vision judge whether a briefly flashed line bisected two horizontal reference lines, and found that thresholds of observers with amblyopia resembled those at 2.5° eccentricity in the typical periphery. Other studies have found similar position inaccuracy at the amblyopic fovea and higher eccentricities in the typical periphery. For example, Hussain et al. (2015) presented a reference probe at different locations in the visual field of observers with typical vision and amblyopia, and asked them to move a response probe at the opposite hemifield, so that the fixation bisected the two probes (reference and response). They found that the positional error of observers with amblyopia at 1° eccentricity closely matched the average of observers with typical vision at 7° eccentricity. In this chapter, observers with typical vision were tested at 15° eccentricity. Therefore, it is unlikely that positional uncertainty in the amblyopic fovea matched the adult periphery. As such, the quantitative difference in positional uncertainty between the

amblyopic fovea and the typical periphery could result in a better sense of the target position and thus less reliance on positional cues to determine the location of the target in the amblyopic fovea. This could thus account for the lack of an effect of the addition of more opposite polarity and diamond flankers in the amblyopic fovea.

The potential limitations of using a different stimulus in this chapter compared to most studies investigating grouping effects on crowding should also be addressed. Whereas here I used a Landolt-C target, the majority of research on grouping has used vernier targets (Sayim et al., 2008, 2010; Manassi et al., 2012, 2013; Manassi et al., 2015; Manassi et al., 2016). Although one would assume that if grouping were to modulate crowding it would do so regardless of the type of stimulus, it is still important to rule out whether the use of Landolt-Cs contributed to the lack of flanker-flanker grouping in the typical and amblyopic fovea. The advantage of verniers relative to Landolt-C stimuli is their smaller diameter that allows for smaller separations between the target and flankers, especially with flanker lines, and thus for the measurement of interactions at a finer spatial scale. The ability to measure interactions at a finer scale would have specifically benefited measurements of crowding in the typical fovea, for which crowding with Landolt-C targets showed differences among flanker conditions over a very small range (less than half a minute of arc). The use of verniers would be less beneficial in the amblyopic fovea that shows a disproportionate vernier acuity deficit relatively to letter acuity (Levi & Klein, 1985) that would result in elevated thresholds across conditions. Nonetheless, to rule out the involvement of differences between vernier and Landolt-C stimuli in the weak grouping effects found in amblyopia, vernier stimuli could be modificed for the purposes of measuring amblyopic crowding. In order to use vernier targets in amblyopia with the scaling method (Levi, Song, et al., 2007a; Song et al., 2014), the vernier could be modified to be comparable to Sloan letters (Sloan, 1959). Specifically, the size of the offset could be matched to the vernier line width, and the diameter of the whole element could be equal to five times the offset. These dimensions would be equivalent to the Landolt-C stimulus used in this thesis, were the gap of the C is equal to the width of the letter, and the diameter of the entire stimulus is equal to 5 times the gap. With these modifications, the role of the type of target stimulus on whether crowding in the typical and amblyopic fovea shows grouping modulations could be elucidated.

Overall, the findings from this chapter demonstrate that amblyopic crowding cannot be fully accounted for by a grouping mechanism. Although there was evidence of effects that could be described as target-flanker grouping on amblyopic crowding, there were no flanker effects suggestive of flanker-flanker grouping. This was also the case for crowding in the typical fovea, suggesting that these two instances of foveal crowding could not be fully accounted for within a grouping framework of crowding. Nonetheless, the effects of target-flanker grouping still posed a challenge for a traditional pooling framework. Finally, the dissociation between target-flanker grouping and flanker-flanker grouping suggested that mechanisms related to positional cueing and unrelated to grouping could be underlying effects that have been attributed to grouping.

5 **Chapter 5:** General Discussion

In the last two decades, the interest of vision scientists in better understanding the curious phenomenon of visual crowding has developed. However, this surge in research into visual crowding has primarily focused on the typical visual system, and particularly the visual periphery. Clinical instances of crowding, such as crowding in strabismic amblyopia, have not received the same amount of research interest. The primary focus of research in amblyopic crowding has been on understanding its relationship to acuity (Stuart & Burian, 1962; Levi & Klein, 1985; Stager, Everett, & Birch, 1990; Hess, Dakin, et al., 2001; Bonneh et al., 2004; Hariharan et al., 2005), and very few studies have focused on the underlying crowding mechanism. The aim of this thesis was to better understand the mechanism of amblyopic crowding, and investigate whether it can fit within a framework of crowding in typical vision.

In this thesis, I followed pooling models that argue that crowding in typical vision arises due to the integration (i.e. pooling) of the features of the target and flankers (Parkes et al., 2001; Pelli, 2008; Greenwood et al., 2009; Dakin et al., 2010; van den Berg et al., 2010; Harrison & Bex, 2015), in order to systematically promote perceptual similarity among adjacent regions of the visual field (Greenwood et al., 2010). The retinal emphasis to the fovea is carried through to the visual cortex, resulting in cortical undersampling of the visual periphery, and neurons with large receptive fields in order to ensure adequate coverage. According to pooling accounts, crowding arises from the integration of the target and flankers due to these large receptive fields. In this thesis, I explored whether crowding in the amblyopic fovea can be accounted for by pooling. In order to achieve this, in each experimental chapter I examined one question that has received research interest in peripheral vision, but had yet to be investigated with regards to amblyopic crowding. In this General Discussion, I will consider the findings from each experimental chapter in turn and discuss how they fit together within a pooling framework of crowding. I will then discuss alternative models of crowding and limitations of the framework I adopted, and suggest future directions.

First, in Chapter 2 I asked whether the effects amblyopic crowding has on the appearance of the target are the same as those in peripheral crowding. The perceptual effects of crowding, as indicated by the errors observers make, have been especially informative of the underlying mechanism in the typical visual periphery. When tasked with reporting the identity of a target in clutter in the

periphery, observers make systematic errors: they tend to report intermediate identities between the target and flankers (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010), or report the identity of the flankers (Strasburger et al., 1991; Ester et al., 2014; Ester et al., 2015). To explore the perceptual effects of amblyopic crowding, I used an orientation-matching task, and investigated whether the errors of children with amblyopia and typical vision were systematic and matched those in adult periphery.

In Chapter 3, I asked whether amblyopic and peripheral crowding have common neural correlates. In the periphery, the increase in the extent of crowding with eccentricity (Bouma, 1970; Toet & Levi, 1992) has been attributed to the insufficient sampling of the peripheral visual field (Parkes et al., 2001), with fewer neurons with large receptive fields responding to the peripheral stimulation (Dow et al., 1981; Van Essen et al., 1984). Amblyopia associated with strabismus has been shown to result in fewer neurons responding to the amblyopic eye in V1 and V2 (Crawford & von Noorden, 1979; H. Bi et al., 2011; Shooner et al., 2015). I hypothesised that crowding in amblyopia could result from a diversion of neural resources from the amblyopic eye to the fellow fixating eye, creating the need for neurons responding to the amblyopic eye to have large receptive fields. In order to investigate the relationship between amblyopic and peripheral crowding and receptive field size, I measured crowding and estimates of fMRI population receptive field (pRF) size (Dumoulin & Wandell, 2008) across the typical and amblyopic visual field.

In Chapter 4, I asked whether instead of pooling, grouping could account for amblyopic crowding. In the typical fovea and visual periphery, when more than one flankers are placed on each side of the target, crowding is said to be determined by Gestalt (Wertheimer, 1923) principles of grouping (Herzog & Manassi, 2015; Herzog et al., 2015). According to grouping accounts of crowding, crowding is strong when the flankers form a uniform configuration with the target and there is target-flanker grouping (Sayim et al., 2008; Manassi et al., 2012; Rosen & Pelli, 2015). In contrast, crowding is weak when the flankers form a uniform configuration with each other but not with the target and flanker-flanker grouping occurs (Manassi et al., 2013). To investigate whether grouping can determine crowding in amblyopia, I used different flanker arrangements that have been found to show target-flanker grouping and flanker-flanker grouping in the visual system.

5.1 The perceptual effects of amblyopic crowding

In Chapter 2, the aim was to investigate the perceptual effects of crowding in strabismic amblyopia, in order to examine whether they match those in the typical visual periphery. When tasked with reporting the identity of a crowded target, the reports observers make are indicative of the information available to them about the target. For crowded targets presented in the typical visual periphery of adults, the perceptual effects of crowding have been found to be systematic. Adult observers tend to report intermediate or average identities between the target and flankers (Parkes et al., 2001; Greenwood et al., 2009; Dakin et al., 2010) making what I referred to as assimilation errors. In other circumstances, adult observers have also been found to report the identity of the flankers, indicating that either parts of the flanker or the entire flanker element is substituted onto the target (Strasburger et al., 1991; Ester et al., 2014; Ester et al., 2015). These errors suggest that crowding systematically shifts the identity of the target to make it more similar to the identity of the flankers.

In order to examine the perceptual effects of amblyopic crowding, I tested children with strabismic amblyopia and typically developing vision, and adults in the visual periphery. I used an orientation-matching task, and asked observers to match the orientation of a reference Landolt-C to the orientation of crowded target Landolt-C when it was surrounded by flankers with a 30° orientation difference from the target, and flankers with a 90° orientation difference from the target. Children with amblyopia and typical vision made the same types of systematic assimilation and substitution errors as adults in the visual periphery: with flankers with a 30° orientation difference, children with amblyopia and typical vision primarily reported intermediate orientations between the target and flankers, indicative of assimilation errors, whereas with flankers with a 90° orientation difference, children with amblyopia and typical vision reported either the orientation of the target, or reported the orientation of the flankers, indicative of substitution errors. These findings thus demonstrated that crowding in amblyopic, developing and peripheral vision has the same systematic perceptual effects.

I then simulated these effects in the three instances of crowding using a population response pooling model. In line with previous models of crowding based on the principles of population coding (van den Berg et al., 2010; Harrison & Bex, 2015), a number of detectors selective for orientation were simulated, similarly to V1 neurons that respond preferentially to orientation (Schiller et al., 1976). Crowding

was simulated as the noisy weighted combination ('pooling') of the population responses to the target and flanker orientations. The model was consistent with traditional pooling accounts arguing that crowding involves two stages, feature detection and feature integration (Pelli et al., 2004). At the first stage of this common pooling mechanism, the population detected the target and flanker features by responding independently to their orientations. At the second stage, the responses of the population to the target and flankers were pooled, with weights determining the contribution of each to the combined response. Despite differences in the extent of crowding between developing, amblyopic, and peripheral crowding, the model assumed that once the target and flankers were within the extent of crowding, the pooling computation was the same.

However, within the extent of crowding there were differences in the contribution of the population response to the flankers to the pooled response depending on the flanker orientation. In the developing and amblyopic fovea and the near adult periphery, assimilation errors were more common with 30° flanker differences, whereas target responses were more frequent with 90° flanker differences. When simulating these instances of crowding, this response pattern resulted in generally higher flanker weights when the flankers differed by 30° from the target, and lower flanker weights when they differed by 90°. Therefore, flanker weights were modulated by the differences between the target and the flankers in orientation, with the population response to the flankers having a greater contribution to the pooled response when the flankers were more similar to the target. The modulation of the flanker weights in this model is similar to previous population response models that vary weights based on target-flanker distance (van den Berg et al., 2010; Harrison & Bex, 2015). In these models, weights are used to simulate the differential effects of the flankers depending on their separation from the centre of target - the closer the flankers were to the target, the more disruptive their effect on target recognition. As such, higher weights are attributed to flankers placed near the target and lower weights to distant flankers. van den Berg et al. (2010) also simulated the inner-outer asymmetry of crowding, by weighting an individual flanker positioned on the outer (more peripheral) side of the target higher than a flanker positioned in the inner (closer to the fovea) side of the target. With the use of weights, other spatial characteristics of peripheral crowding, such as the radial-tangential anisotropy (Toet & Levi, 1992) can be accounted for by radially positioned flankers being weighted higher than those tangentially positioned ones.

It is thus important to consider what determined the magnitude of the flanker weights in the population response pooling model presented in this thesis. I argued that cortical proximity between the target and flankers could be a determining factor in the magnitude of the flanker weights. Mareschal et al. (2010) have shown that when the separation between the target and flankers is kept constant, assimilation errors increase with eccentricity in the adult periphery. When eccentricity is kept constant and the separation between target and flankers is reduced, assimilation errors also increase. Increasing eccentricity and reducing target-flanker separation both reduce the cortical separation between the target and flankers, since less cortex is dedicated to the periphery than the fovea (Dow et al., 1981; Van Essen et al., 1984; Sereno et al., 1995; Engel et al., 1997) and the cortical representation of visual space is increasingly compressed with eccentricity. These results were thus taken to suggest that cortical and not physical proximity between the target and flankers determines the perceptual effects of crowding. Based on these findings, I argued that differences in the flanker weights based on target-flanker orientation similarity might also arise due to differences in the cortical proximity of the Landolt-C gaps. When mapped retinotopically, an orientation difference of 30° between the target and flankers places the gaps in closer cortical proximity than an orientation difference of 90°. Therefore, cortical proximity could be a modulatory factor in determining the magnitude of the flanker weights in the model. However, an exception to this weight modulation by the orientation of the flankers concerns the far adult periphery, where flanker weights were comparably high for both 30° and 90° flanker orientation differences. In this case, the cortical separation between the target and flanker gaps may be so small than any differences in orientation are not substantial enough to manifest as differences in the perceptual effects of crowding, resulting in similarly high weights for 30° and 90° target-flanker differences.

Despite the general trend for an increase in flanker weights with eccentricity regardless of flanker orientation, there was significant individual variation among adult observers in the rate of this increase. For some adult observers, the rate of increase in the flanker weights was rapid, with the flanker weights becoming equally high between the two flanker orientation conditions at 5° eccentricity. For other observers, this rate of increase was steeper, with the flanker weight for 90° differences becoming as high as that for 30° flanker differences at 15° eccentricity. These individual differences in the perceptual effects of crowding could be due to differences in the functional architecture of the visual cortex. Electrophysiological

studies in non-human primates and neuroimaging studies in humans have pointed to substantial individual variation in the topographical organisation of V1 and later visual areas (Dow et al., 1981; Van Essen et al., 1984; Benson & Winawer, 2018). Linking this variability to perceptual differences, Moutsiana et al. (2016) showed that the perceived size of an object is correlated with individual differences in V1 functional architecture: when pRFs are larger, as with increasing eccentricity, the size of target circles is perceived as smaller. Similarly, variation in the rate of reduction of the cortical magnification factor (Daniel & Whitteridge, 1961) -the amount of cortex dedicated to 1° of visual angle- may result in differences in the cortical separation between the target and flankers. Such variation in the cortical proximity of the target and flankers may in turn lead to differences in the perceptual effects of crowding and thus in the rate of increase in flanker weights with eccentricity. In fact, neuroimaging measures of cortical magnification factor (CMF) show that there is a close relationship between the reduction in CMF and the increase in pRF size (Harvey & Dumoulin, 2011). As such, individual differences in the general increase of pRF size with eccentricity found in Chapter 3 may also be related with differences in the perceptual effects of crowding in the adult periphery.

In addition to the adult periphery, there were also substantial individual differences in the magnitude of the weights in the groups of children with typical vision and amblyopia. More specifically, there was a subset of children in both groups who did not show an increase in target responses and thus a reduction in the magnitude of the flanker weights with 90° target-flanker differences. Instead, this subset of children made substitution errors with both 30° and 90° target-flanker differences, resulting in high flanker weights for both conditions. This pattern of response error thus resembled the general pattern in the adult periphery at higher eccentricities. Similarly to the adult periphery, one source for these individual differences in the developing and amblyopic fovea could involve idiosyncrasies in the functional architecture of visual areas. For example, Van Essen et al. (1984) reported significant differences between macaques' visual cortices in terms of the proportion of cortex dedicated to the fovea, with some animals showing increased emphasis on central relative to peripheral portions of the visual field compared to others. In Chapter 2, I speculated that idiosyncratic localised distortions of the retinotopic map corresponding to the fovea could result in misperceptions of the flanker feature for the target, thus increasing the flanker weights for both flaker orientation conditions in model. As such, idiosyncrasies in the functional

architecture of visual areas might be associated with individual differences in how crowded objects appear in the adult periphery, and the developing and amblyopic fovea.

The high incidence of substitution errors resulting in increased flaker weights for both 30° and 90° target-flanker differences could also be associated with increased positional uncertainty. Positional uncertainty could result in source confusion, with flanker features or whole flankers being mislocalised (Strasburger et al., 2011). Such mislocalisations may consequently lead to increased flanker reports, and thus an increase in the flanker weights when simulating these effects. This is consistent with evidence indicating that both peripheral and amblyopic vision are characterised by increased positional uncertainty (Rentschler & Treutwein, 1985; Hess & Hayes, 1994; Hess et al., 1997), typically manifested as a disproportionate reduction in vernier acuity relative to grating or letter acuity (Levi et al., 1987). Observers with amblyopia also show deficits in other tasks that rely on positional information, such as bisection and alignment tasks (Bedell & Flom, 1981; Levi & Klein, 1983). Additionally, during development vernier acuity reaches adult levels in the early teens (Carkeet et al., 1997; Skoczenski & Norcia, 2002), suggesting difficulties with positional accuracy at younger ages. As such, increased positional uncertainty could explain the high flanker weights for both 30° and 90° target-flanker differences at high eccentricities in the periphery, as well as in the subset of children with amblyopia and typical vision. This explanation would not be inconsistent with flanker weights being modulated by differences in cortical distance, as reductions in positional accuracy (as measured in vernier tasks) are correlated with reductions in V1 cortical magnification factor in primates (Levi et al., 1985).

Despite these instances in which flanker weights were comparably high for both 30° and 90° target-flanker differences, the findings generally indicated that target-flanker similarity in orientation modulated peripheral as well as developing and amblyopic crowding. Flankers were generally more disruptive when they were more similar to the target, as indicated by the greater frequency of target reports when the flankers had a 90° orientation difference from the target than when they differed by 30°. This was a significant finding for crowding in developing vision, for which selectivity for target-flanker similarity has not been previously investigated, and for crowding in amblyopic vision, which was thought to be largely immune to target-flanker differences (Hess, Dakin, et al., 2001; Hariharan et al., 2005). Although the model simulations in this thesis were limited to orientation, peripheral crowding

has been found to be selective for similarity along multiple visual features, including colour (Põder & Wagemans, 2007; Kennedy & Whitaker, 2010), motion (Bex et al., 2003), and contrast polarity (Chakravarthi & Cavanagh, 2007). In fact, the findings in Chapter 4 demonstrated that amblyopic crowding was also selective for contrast polarity: crowding on a black target was strong with black flankers and weak with white. As population coding is considered to be a general process through which the brain encodes sensory features (Pouget et al., 2000), the crowding of other visual features can also be accounted for by a population response pooling model, analogous to the one presented here for orientation. For example, recently it was shown that a population response pooling model can account for the selectivity of crowding for target-flanker similarity in colour and motion (Greenwood & Parsons, 2019). Therefore, population response pooling models provide a successful framework to account for the selectivity of crowding for similarity along multiple visual features and in different visual systems.

An implication of crowding resulting from the pooled responses of populations of detectors each selective for a particular visual feature is the involvement of multiple distinct populations of neurons, each selective for different visual dimensions, in crowding. The population response pooling model presented in this thesis simulates detectors selective for orientation similarly to neurons in V1 (Schiller et al., 1976), placing the neural site of orientation crowding as early as, but not restricted to, that visual area. Pooling of other features may involve populations of neurons both in V1 and later visual areas. For example, motion crowding could involve direction-selective populations of neurons in V1 (DeValois, Yund, & Hepler, 1982) and V5 (Albright, 1984), and colour crowding could also involve neurons in V1 and later visual areas, as neurons that respond to colour stimuli have been found in multiple visual areas (Solomon & Lennie, 2007). The involvement of multiple visual areas in crowding is consistent with the findings from Chapter 3 showing that fMRI measures of pRF size in V1, V2, and V3 follow the general pattern of crowding across the typical and amblyopic visual field, as well as previous neuroimaging studies reporting neural correlates in areas V1-V4 (J. Freeman et al., 2011b; Anderson et al., 2012). Therefore, crowding is likely to be distributed over multiple visual areas, with the principles of population response pooling being the same across different neural sites.

Pooling being a general process that occurs for various visual features and at multiple visual areas along the visual processing hierarchy is broadly consistent with

accounts arguing that crowding occurs over multiple visual processing stages (Whitney & Levi, 2011; Chaney et al., 2014). However, these accounts are based primarily on findings suggesting that crowding is selective for holistic similarity in faces. Louie et al. (2007) found that crowding for faces was selective for "holistic similarity", being strong only when both the target and the flankers were upright, and thus processed holistically. This was thus taken as evidence to suggest that in addition to being selective for fundamental visual features such as orientation in simple stimuli, crowding is also selective for holistic similarity in faces. However, once task difficulty for upright and inverted faces is matched, the effect indicative of holistic similarity disappears and crowding is modulated by the orientation of the flankers for both upright and inverted faces (Kalpadakis-Smith et al., 2018). In fact, Kalpadakis-Smith et al. (2018) simulated the crowding of faces by a population of detectors selective for the position of facial features (e.g. interocular separation) in line with theoretical proposals of a "face space" (Valentine, 1991) and neurophysiological recordings in the macaque IT lobe (Chang & Tsao, 2017). As such, population response pooling models can account for crowding along multiple visual features, and simple and complex objects. Therefore, there is no need to invoke independent crowding processes for simple and complex stimuli – population response pooling can adequately account for all instances of crowding whilst involving multiple visual areas.

Although the population response pooling model accounted for the responses of the majority of children with typical vision and amblyopia, there was a subset of children with amblyopia for which an alternative model performed best. For this subset of children, a population response noise model accounted for their response errors better than a population response pooling model. This model assumed a population of detectors that first responded independently to the target and flanker orientations, but instead of combining these responses at the crowding stage, it added noise to the target response. As such, no pooling of population responses was necessary to simulate the perceptual effects of crowding in these children. The population response to the flanker orientation was thus irrelevant in this crowding process, potentially pointing to a different underlying mechanism in these children with amblyopia. I speculated that in these cases that deviated from the children and adults with typical vision and the rest of the amblyopic group, errors may arise due to the perceptual distortions that characterise vision in the amblyopic eye (Fronius & Sireteanu, 1989; Lagreze & Sireteanu, 1991; Sireteanu, Lagreze, et al., 1993;

Sireteanu, Thiel, et al., 2008). Such distortions would alter the target appearance non-systematically, resulting in an increase of responses to random orientations. Not all amblyopes experience such distortions (Piano et al., 2015), and that might explain why only the errors of a subset of children with amblyopia were accounted for by a different noise model.

To conclude, in Chapter 2, I showed that the perceptual effects of crowding in amblyopic and developing vision largely matched those in the adult periphery. The variability in perceptual errors suggested that crowding is a multivariate and highly complex phenomenon (Agaoglu & Chung, 2016). Despite this variability, I showed that a population response pooling model can largely account for errors made by children with strabismic amblyopia and those with typical vision, and adults in the visual periphery, suggesting a common underlying mechanism. I speculated that the purpose of crowding is to regularise the appearance of the visual field promoting perceptual similarity among adjacent regions when neural resources are low and neurons with larger receptive fields are needed for adequate coverage. I then investigated the neural underpinnings of this common mechanism in Chapter 3.

5.2 Crowding and pRF size across the amblyopic visual field

In Chapter 2, a population response pooling account of crowding was found to account for the perceptual effects of crowding in amblyopia, suggesting a common mechanism with the visual periphery. In the typical visual system, pooling accounts (Parkes et al., 2001) have assumed that the minimal effect of crowding in the fovea and the increase in the extent of crowding with eccentricity (Bouma, 1970; Toet & Levi, 1992) may arise due to differences in cortical sampling of the visual field. In primates, a greater amount of cortex, and thus more neurons, are dedicated to the fovea than the periphery (Dow et al., 1981; Van Essen et al., 1984; Sereno et al., 1995; Engel et al., 1997). A traditional measure of this sampling process in V1 is the cortical magnification factor (Daniel & Whitteridge, 1961) that expresses the amount of cortex corresponding to 1° of visual angle. This measure of functional architecture has been associated with visual function: larger cortical magnification factor in V1 has been found to correlate with better acuity (Duncan & Boynton, 2003). This relationship opens the possibility of other measures of visual function, and in this case crowding, to also be associated with cortical sampling.

In the typical visual system, the undersampling of the periphery means that fewer neurons respond to larger regions of the visual field and must therefore have

larger receptive fields in order to ensure adequate coverage (Dow et al., 1981; Van Essen et al., 1984). Primate models of strabismic amblyopia point to a similar underlying basis – a shift in eye dominance with more neurons responding to the fellow fixating eye than the amblyopic eye has been found in primate V1 and V2 (H. Bi et al., 2011; Shooner et al., 2015). Neuroimaging studies in humans have shown that population receptive field (pRF) size in V1-V3 for the amblyopic eye is enlarged relative to the eyes of observers with unaffected vision (Clavagnier et al., 2015), consistent with an undersampling of the amblyopic eye resulting in increased receptive field size. In this chapter, I aimed to investigate whether an increase in receptive field size could be the underlying neural basis of the common crowding mechanism in typical and amblyopic vision. Based on the discussed findings on the functional architecture of the typical and amblyopic visual system (Dow et al., 1981; Van Essen et al., 1984; H. Bi et al., 2011; Clavagnier et al., 2015; Shooner et al., 2015), I first assumed that the increased extent of crowding with eccentricity arises due to increased pooling of peripheral visual input in populations of visual neurons with large receptive fields. Secondly, I assumed that the shift of neural resources from the amblyopic eye to the fellow fixating eye also results in enlarged receptive fields and increased extents of crowding for the amblyopic eye. To test the hypothesis that crowding increases with receptive field size in peripheral and amblyopic vision, I measured acuity and crowding across the typical and amblyopic visual field. I then used fMRI to estimate population receptive field (pRF) size (Dumoulin & Wandell, 2008), and explored the relationship between the individual patterns of variation in acuity, crowding, and pRF size.

I first found that for behavioural measures, in typical vision acuity became increasingly worse with eccentricity, and the extent of crowding increased. For the amblyopic eye, acuity was reduced and the extent of crowding was greater overall relative to typical vision across the visual field. Interestingly, the amblyopic deficit in both acuity and crowding was not uniform across the visual field. The greatest deficit for acuity and crowding was found in the fovea. This foveal deficit was greater for crowding than for acuity, with the extent of crowding being over eleven times the foveal average of that in typical vision, and acuity being four times the average in the typical fovea. Second, there was a naso-temporal difference: acuity was poorer and the extent of crowding was larger for the amblyopic eye in the temporal visual field. Although this naso-temporal difference in acuity and crowding was also found for observers with typical vision for the higher eccentricities tested, it

was more pronounced in amblyopia. Overall, the results demonstrated a reduction in acuity and increase in the extent of crowding with eccentricity for both typical and amblyopic vision. In amblyopia, the deficit persisted across the visual field, and was disproportionately large in the fovea.

When considering measures of fMRI pRF size, I found that in observers with typical vision, pRF size increased from 1 to 19° eccentricity for V1-V3, in line with previous findings (Dumoulin & Wandell, 2008; Clavagnier et al., 2015). For the amblyopic eye, pRF size was overall greater than in typical vision for V1-V3. This was consistent with a previous study showing greater pRF size between 1° and 6° eccentricity for V1-V3 in the amblyopic eye of strabismic amblyopes relative to both the fellow fixating eye and the eyes of observers with typical vision (Clavagnier et al., 2015). In addition to replicating this increase in pRF size for the same visual areas, I extended upon previous findings by better characterising the pattern of variation in pRF size across the visual field. By independently examining pRF size for the nasal and temporal visual fields, and extending measurements further into the visual periphery up to 19° eccentricity, it was possible to detect non-uniformities in the elevation of pRF for the amblyopic eye and thus the amblyopic deficit. Specifically, in V2 and V3, there was a naso-temporal difference, with the elevation in pRF size being greater in the temporal visual field for the amblyopic eye. The elevated pRF size in the temporal visual field was especially evident for parafoveal eccentricities, and was less pronounced for more eccentric visual field locations. This naso-temporal difference in V2 and V3 was similar to the observed pattern in the acuity and crowding measures. Overall, the findings from pRF mapping demonstrated that in both typical and amblyopic vision pRF size increased with eccentricity, with the enlargement in pRF size affecting the entire amblyopic visual field in a non-uniform manner.

At a group level, I found that behavioural measures of acuity and crowding largely matched the pattern of elevations in pRF size. As such, the findings were broadly consistent with pooling accounts of crowding (Parkes et al., 2001), that predict that regions of the visual field that show the greatest extent of crowding should also show increased pRF size. In the typical visual system the increase in pRF size could be associated with the cortical undersampling of the visual periphery (Dow et al., 1981; Van Essen et al., 1984; Sereno et al., 1995; Engel et al., 1997), resulting in neurons with large receptive fields. In turn, the increase of receptive field size with eccentricity could result in increasingly poorer resolution and thus reduced

acuity, and increasingly larger pooling regions and thus increased extents of crowding. In the amblyopic visual system, the elevation in pRF size could also arise due to a cortical undersampling of the amblyopic eye, with fewer neurons with larger receptive fields responding to the amblyopic relative to the fellow fixating eye. Such an undersampling of the amblyopic eye is consistent with animal studies showing that fewer neurons respond preferentially to the amblyopic eye in V1 (Crawford & von Noorden, 1979; Shooner et al., 2015) and in V2 (E. L. Smith et al., 1997; Kiorpes et al., 1998; H. Bi et al., 2011). Therefore, increased receptive field size due to cortical undersampling could account for the reduced acuity and increased crowding extents affecting the typical visual periphery and the amblyopic visual field. However, there were also important differences between the findings on acuity and crowding, and pRF size. First, across eccentricity the elevation in the temporal visual field was proportionally greater for pRF size than for acuity and crowding, suggesting that pronounced differences in pRF size may manifest as smaller differences in measures of visual function. Additionally, although on a group level the pattern of acuity and crowding across the visual field matched that of pRF size, on an individual level a significant relationship between the measures was not found. This is not necessarily indicative of a lack of relationship between receptive field size and crowding, but rather could be due to the relatively small sample numbers in the study. Nonetheless, the findings in this chapter suggested that increased population receptive field size may account for increases in the extent of crowding in both typical and amblyopic vision.

The commonalities between the overall pattern of elevations in the extent of crowding and pRF size in V1-V3 suggests the involvement of multiple visual areas in crowding. This is consistent with neuroimaging studies showing that changes in BOLD activation that are related to crowding are found in V1-V4 (J. Freeman et al., 2011b; Anderson et al., 2012). In Chapter 2, I simulated crowding as the pooled population response of detectors that are selective for orientation, similarly to neurons in V1 (Schiller et al., 1976). This model was specifically designed to test the pooling of orientation features between the target and flankers, and thus is not inconsistent with the involvement of multiple visual areas in crowding. Rather, the range of visual areas involved in pooling may be associated with how complex the features that are pooled are. Simple features such as orientation and spatial frequency of the target and flankers may be pooled in V1, for which V1 neurons have been shown to be selective (Movshon, Thompson, & Tolhurst, 1978; Tolhurst &

Thompson, 1981). More complex features such as the target-flanker contours (angles and curves) may be pooled in V2, where neurons have shown selectivity for more complex shape characteristics (Hedgé & Van Essen, 2000). V4 may process combine pooled features from previous sites, as V4 neurons have receptive field sizes that fit with Bouma's (1970) critical spacing (Motter, 2006, 2009) and show an inherent radial asymmetry in sensitivity (Motter, 2018) consistent with the radial-tangential anisotropy of crowding (Toet & Levi, 1992; Pelli et al., 2007; Greenwood et al., 2017). Additionally, a flanker placed at the outward side of the target (i.e. the side farther from the fovea) falls within the more sensitive part of the receptive field (Motter, 2018), consistent with the inward-outward asymmetry of crowding. As such, the involvement of distinct populations of neurons across multiple visual areas in crowding is consistent with the model presented in Chapter 2.

In addition to enlarged receptive fields responding to the amblyopic eye, it is important to consider how other neural correlates underlying pRF size could be related to crowding. As discussed in Chapter 3, in addition to an increase in receptive field size, increased positional scatter could have also contributed to an increase in estimates of pRF size for the amblyopic eye. Clavagnier et al. (2015) attributed the increase in pRF size found for the amblyopic eye to a combination of the increased size of component receptive fields and increased scatter, as they showed that there was more variability in the centres of the pRFs responding to the amblyopic eye than to the fellow fixating eye. A disorder in how the retina is mapped onto the cortex would result in an imprecise topographical map and provide a non-veridical representation of the visual field in amblyopia. Consequently, topographical disorder of cortical receptive fields would not allow the amblyopic visual system to maintain a precise spatial order. Topographical disorder has been linked to elevated position discrimination thresholds and thus increased positional uncertainty in amblyopia (Hess et al., 1978; Levi & Klein, 1985; Hess & Field, 1994), and perceptual distortions (Fronius & Sireteanu, 1989; Lagreze & Sireteanu, 1991; Sireteanu, Lagreze, et al., 1993; Sireteanu, Thiel, et al., 2008). The increased random errors made by children in amblyopia in Chapter 2 could also be linked with a topographical disorder in cortical maps. Matching the orientation of a Landolt-C target involves judging which side of a circle a gap is positioned. Gap positions could thus become more uncertain due to topographical disorder, resulting in the increased random responses found in the group with amblyopia. It could be the

case that children that made more random errors and showed greater variability in their responses have greater topographical disorder in their cortical receptive fields.

To conclude, in Chapter 3 I showed that in the typical visual system, acuity was reduced, the extent of crowding increased, and pRF size became increasingly larger in V1-V3 with eccentricity. In amblyopia, there was a reduction in acuity, an increase in the extent of crowding and an elevation in pRF size for the amblyopic eye relative to unaffected vision. The findings on a group level are thus broadly consistent with pooling models of crowding that predict that crowding arises due to the cortical undersampling of the visual field, with neurons with large receptive fields needed for sufficient coverage. Although a correlation between individual measures of visual function and pRF size was not found, the group results are in line with the population response pooling model described in Chapter 2 that assumes that pooling occurs when neurons have large receptive fields.

5.3 Grouping effects on amblyopic crowding

In the previous two experimental chapters, I adopted a pooling framework that views crowding as the outcome of the integration of features in close spatial proximity. In Chapter 3, the elevation in pRF size in V1-V3 for the amblyopic eye relative to unaffected vision was in line with crowding arising due to feature pooling within large receptive fields. However, the involvement of multiple visual areas in the amblyopic deficit also raises the question of whether the increased extent of crowding could be a higher-level deficit. To address this, in Chapter 4, I investigated whether grouping can modulate crowding in amblyopia, and thus whether amblyopic crowding can be accounted for by a different account of crowding. Grouping accounts propose that crowding is determined by how the target and the flankers look and whether or not they group (Herzog & Manassi, 2015; Herzog et al., 2015). According to grouping principles, when flankers form a uniform configuration with each other but segregate from the target, there is flanker-flanker grouping and the target is more easily recognised; thus crowding is weak (Manassi et al., 2012, 2013; Herzog, Sayim, Manassi, & Chicherov, 2016; Manassi et al., 2016). On the other hand, when flankers form a uniform configuration that includes the target, target-flanker grouping occurs and it is difficult to identify the target; thus crowding is strong (Manassi et al., 2012). Here I investigated whether modulations in amblyopic crowding follow flanker-flanker grouping and target-flanker grouping.

I first determined whether crowding in amblyopia was selective for the similarity between the target and the flankers in contrast polarity. In this condition, I presented a Landolt-C target surrounded by two identical Landolt-C flankers (one on each side horizontally) and reversed their contrast polarity. Similarly to the periphery and the typical fovea, with flankers of the same polarity, the spatial extent of amblyopic crowding was large; when the polarity of the flankers was reversed, the spatial extent of crowding was reduced, and crowding was thus relieved. This selectivity of amblyopic crowding for the similarity between the target and flankers was consistent with findings from Chapter 2, where I found that the target orientation was more readily reported when the flankers had a large orientation difference from the target, compared to when the flanker orientation difference was small. As such, the combination of the findings from these two chapters demonstrates that amblyopic crowding is selective for differences in the similarity between the target and the flankers.

Previous studies however have found mixed results on the selectivity of amblyopic crowding for target-flanker similarity: some have shown that flankers are less disruptive in the amblyopic fovea when they are dissimilar to the target in orientation (Levi, Hariharan, et al., 2002a) and for a subset of observers in contrast polarity (Hess, Dakin, et al., 2001). Other studies have found that the disruptive effects of flankers on a target are equal with similar and dissimilar flankers in polarity (Hariharan et al., 2005), orientation (Hariharan et al., 2005) and contrast (Levi, Hariharan, et al., 2002a). The discrepancy between these results and findings in this thesis are likely due to stimulus differences. Here I kept all other stimulus features (e.g. shape, size) the same between the target and flankers, and varied only one feature dimension: either orientation (Chapter 2) or contrast (Chapter 4). Other studies have used target and flankers of different shape and size, such as Gabor or Gaussian patches that form a Landolt-C or E target, surrounded by flanker bars (Levi, Hariharan, et al., 2002a; Hariharan et al., 2005). If crowding in amblyopia also shows a clear selectivity for shape and size like peripheral crowding (Kooi et al., 1994), flankers that were most similar to the target in these previous studies may not have been as disruptive as flankers used in this thesis that were identical to the target. This would in turn mute the release from crowding with dissimilar flankers along the dimensions tested. In fact, Song et al. (2014) have shown that flanker bars on a target letter are significantly less disruptive than flanker letters, in line with crowding in amblyopia being modulated by similarity in shape and size. This has

important implications for clinical measures of crowding that use bars instead of letters as flankers (Lalor, Formankiewicz, & Waugh, 2016), and may thus be underestimating the full extent of the crowding deficit.

When considering effects indicative of flanker-flanker grouping, the periphery was found to differ from the amblyopic and typical fovea. In the periphery, when the opposite polarity flankers were increased from two to six, the extent of crowding was reduced. This was in line with findings showing that adding more dissimilar to the target in polarity (Rosen & Pelli, 2015) or length (Manassi et al., 2012) alleviates crowding in the visual periphery. Similarly, when more diamonds were added to the sides of a diamond flanker enclosing the target, the extent of peripheral crowding was reduced, consistent with previous studies showing the same effect with vernier targets (Manassi et al., 2013; Manassi et al., 2016). However, in neither the typical nor the amblyopic fovea was there a relief from crowding by adding more flankers. As such, these effects that have been interpreted by grouping accounts as flanker-flanker grouping (Manassi et al., 2012, 2013; Herzog et al., 2016; Manassi et al., 2016) were only found in the typical periphery, but not the amblyopic or typical fovea.

Although findings on flanker-flanker grouping differentiated the periphery from the amblyopic and typical fovea, the three instances of crowding showed similar results when it came to conditions indicating target-flanker grouping. When six flankers were of alternating polarity, the extent of crowding was larger relative to when the six flankers were of opposite polarity to the target. This significant increase in the extent of crowding was found in both fovea and periphery in the typical visual system. As such, these findings were consistent with previous results in typical fovea and the periphery showing that when flankers form a regular pattern of alternating elements in polarity (Sayim et al., 2008; Rosen & Pelli, 2015) and colour (Sayim et al., 2008; Manassi et al., 2012), target recognition is worse compared to when they are dissimilar to the target. Crucially, a trend for this increase in the extent of crowding with alternating polarity flankers was also found in the amblyopic fovea, consistent with target-flanker grouping. Therefore, the findings in this chapter were consistent with grouping accounts *only* in cases in which the flankers were predicted to be most disruptive due to target-flanker grouping.

Why does then amblyopic crowding not show the relief expected because to flanker-flanker grouping? An account of grouping recently proposed by Francis et al. (2017) that describes a way in which grouping may operate on crowding,

assumes a first stage of pooling between the target and flankers (Francis et al., 2017). At the later grouping stage it assumes that boundaries between the target and flankers are formed. If these boundaries do not connect the target with the flankers but connect the flankers with each other, then the target is easily segmented from the flankers. In this case, flanker-flanker grouping occurs and crowding is weak. If these boundaries connect the target and flankers, target-flanker grouping occurs, as the target cannot be segmented from the flankers, and crowding is strong. However, if the earlier pooling stage were to be characterised by increased noise, as the findings from Chapter 2 indicate for strabismic amblyopia, then this noise would affect the subsequent grouping stage. If correct boundary formation and thus grouping depends to some extent on the pooling of features among many regions across the visual field, then the noise would accumulate when information is processed recurrently not allowing for the correct grouping of target and flankers. The increase in noise would particularly affect conditions in which flanker-flanker grouping is expected to occur, as there will be no relief from crowding. In contrast, in conditions in which target-flanker grouping is expected, crowding will still be strong as predicted by grouping accounts, as the increase in noise will disrupt target recognition. A complication to this interpretation is that I did find a relief from crowding with two opposite polarity flankers compared to two same polarity flankers, that based on grouping accounts (Francis et al., 2017) would suggest correct boundary formation in amblyopia. Crucially, an important limitation to this interpretation is that the typical fovea also showed a lack of flanker-flanker grouping effects. As children with typical vision did not show evidence of such increased noise for crowding conditions in Chapter 2, the lack of flanker-flanker grouping effects in typical adult fovea is unlikely to result from increased noise at the pooling stage.

In fact, a grouping account of crowding such as the one put forward by Francis et al. (2017) may not be necessary to account for the findings in the typical and amblyopic fovea. It would be reasonable to assume that if grouping were to modulate crowding in the typical and amblyopic fovea, it would do so across all multi-element flanker configurations – not selectively when flankers were of alternating polarity. If not grouping, then this raises the question of what is modulating crowding when flankers are of alternating polarity, for which condition the results matched grouping predictions. One explanation concerns interference from the black flankers twice removed from the black target. Despite not being the

flankers most proximal to the target, their addition might still interfere with its recognition, thus leading to an elevation in the extent of crowding relative to opposite polarity flankers. In fact, the combination of proximal white flankers and farther black flankers may lead to an additive crowding effect, similarly to supercrowding (Vickery, Shim, Chakravarthi, JIang, & Luedeman, 2009). In supercrowding, simultaneously masking the target and crowding it with flankers has a strong disruptive effect on target recognition, even though individually either the mask or the flankers have a weak effect. Likewise, alone either the proximal white flankers or the twice removed black flankers, due to dissimilarity in polarity or larger separation from the target respectively, would have a weak disruptive effect on target recognition. However, their combined presentation in the alternating polarity condition may lead to a stronger disruptive effect than either presented alone. A complication to this interpretation is that when opposite polarity flankers were increased from two to six, this did not lead to an increase in the extent of crowding. Rather, the extent of crowding remained constant in the typical and amblyopic fovea. However, in this case the additional flankers were both of opposite polarity *and* at a greater separation from the target, which could result in them having no effect on target recognition. Therefore, when added to the opposite polarity flankers proximal to the target that have a weak effect, this does not increase the extent of crowding, unlike black flankers in the alternating polarity condition.

Crucially, such an interpretation would allow for incorporating the findings from the amblyopic and typical fovea with flanker polarity manipulations in a population response pooling model. Similarly to the model presented for orientation in Chapter 2, the flanker weights could be determined first based on the similarity in polarity between the target and the flankers most proximal to it. If the target and flankers were of matching polarity, the flanker weights would be high, whereas if the flankers were of opposite polarity, the weights would be low. At a later additional stage, these weights could then be modified depending on the flankers twice removed from the target. If these farther flankers matched the proximal flankers, then the weights would remain the same. This would result in comparable effects of two and six flankers of opposite polarity on crowding, replicating the results for the typical and amblyopic fovea. On the other hand, if the farther flankers were of opposite polarity to the proximal flankers, the weights would be increased. As such, this later increase in flanker weights would reduce target recognition performance, simulating the increased extent of crowding found with alternating polarity flankers.

Note that this model would not simulate the effects of diamond flankers, as these may require populations of detectors sensitive to more complex features such as V2 neurons selective for contours (Hedgé & Van Essen, 2000). Nonetheless, a population response pooling model with an additional weight parameter may be able to account for the findings in the typical and amblyopic fovea with flanker polarity manipulations.

However, the findings from the visual periphery deviated from the typical and amblyopic fovea, showing a relief from crowding when additional opposite polarity or diamond flankers were added to the target. Therefore, it is important to consider what may be underlying these effects specific to the periphery. It has been proposed that unlike typical foveal vision where information is represented in high fidelity and crowding occurs over small distances, crowding in peripheral vision involves the representation of large spatial regions as texture (Balas et al., 2009; J. Freeman & Simoncelli, 2011). Modulations of peripheral crowding attributed to grouping may thus arise due to differences in the formation of texture boundaries between the target and flankers (Harrison & Bex, 2016). Specifically, in cases of target-flanker grouping, the target and flankers form a constant, uniform texture with no distinctive boundaries. In contrast, in cases of flanker-flanker grouping, clear texture boundaries are formed between the target and the flankers, segregating the flankers from the target. When boundaries between the target and flankers are distinctive, they might act as positional cues to the target, aiding target recognition and relieving crowding (Rosenholtz et al., 2017; Yu & Rosenholtz, 2018). The greater number of flankers, the more distinctive the texture boundary is, and the larger the cue to the target. As such, when the number of opposite polarity and diamond flankers is increased, the cue to the target is greater, and peripheral crowding is further relieved.

How could then the findings from the periphery, the typical and the amblyopic fovea be accounted for within a pooling framework of crowding? Following texture models (Balas et al., 2009) that are also discussed in the following section, texture boundaries could be incorporated in the population response pooling model. The contribution of the flankers to the combined population response could be weighted depending on whether they form a clear texture boundary segregating the target (low weight) or whether they form a constant texture with the target (high weight). An additional parameter would be necessary to account for the differences between the fovea and the periphery. Depending on the instance of crowding, this parameter

could influence the magnitude of the reduction in the flanker weights when a clear texture boundary is formed, segregating the target from the flankers. In the visual periphery this parameter would significantly lower the flanker weight, whereas in the typical and amblyopic fovea it would have little effect on the flanker weigh. As such, the relief from crowding when additional opposite polarity flankers or diamonds are added to the target will be simulated in the periphery, but not the typical or amblyopic fovea. This parameter could signify quantitative differences in positional uncertainty between the periphery and the typical and amblyopic fovea. In the typical and amblyopic fovea the two opposite polarity flankers or one diamond flanker may be sufficient positional cues to the target. In contrast, in the periphery where positional uncertainty is high (Levi et al., 1987; Hess & Field, 1993), additional flankers may provide even better cues to the location of the target, further relieving crowding. Although positional uncertainty is also high in the amblyopic fovea (Hess et al., 1978; Levi & Klein, 1985; Levi et al., 1987; Hess & Field, 1994), it may not be as high as at 15° eccentricity, explaining the differences between the two instances of crowding. Although the additional texture boundary stage and parameter results in a less parsimonious model, a simpler population response pooling model cannot account for the findings in Chapter 4. In fact, consistent with the increased pRF size in V1-V3 found in Chapter 3, an additional stage involving texture boundary formation could signify the involvement of multiple stages along the visual processing hierarchy in crowding.

To conclude, in Chapter 4 I found that the effects of differences in the global configuration of the target and flankers cannot be fully accounted for by a grouping account of crowding in amblyopia. However, the population response pooling model in its current form as described in Chapter 2 would also fail to account for the strong disruptive effects of flankers creating an alternating polarity pattern with the target. Therefore, I cannot fully exclude higher-level grouping modulations in amblyopic crowding.

5.4 Limitations & Alternative Frameworks

In this thesis, I have considered crowding in amblyopia within a pooling framework. However, it is important to also consider whether the findings in this thesis could fit within other frameworks of crowding. In this section, I briefly consider how my findings on crowding in amblyopia would fit first within a texture model of crowding (Balas et al., 2009; Keshvari & Rosenholtz, 2016), and second

within attentional accounts of crowding (Strasburger et al., 1991; He et al., 1996; Intriligator & Cavanagh, 2001; Tripathy & Cavanagh, 2002; Strasburger, 2005).

5.4.1 Texture Models

Texture models propose that crowding is a mechanism that statistically summarises the target and flankers, and view the crowded percept as "forced texture perception" (Balas et al., 2009; Keshvari & Rosenholtz, 2016). According to texture models, the visual system computes summary statistics of the crowded visual input over local pooling regions that overlap and tile the entire visual field. These models are thus similar to pooling accounts of crowding in that they assume pooling regions equivalent to integration fields (Pelli et al., 2004; Pelli & Tillman, 2008), that could have an underlying basis in visual neurons with receptive field sizes that scale with eccentricity. However, contrary to traditional pooling (Parkes et al., 2001) and population response pooling models (van den Berg et al., 2010; Harrison & Bex, 2015), instead of taking as the unit of the crowding mechanism the target and flanker features such as their orientation or position (Greenwood et al., 2009; Dakin et al., 2010; van den Berg et al., 2010; Harrison & Bex, 2015), texture models take the image statistics.

Texture accounts of crowding (Balas et al., 2009; Keshvari & Rosenholtz, 2016) suggest that the observer might be treating the crowded signal as texture. Within this texture, the identity of the flankers could also be available, resulting in substitution errors. Taking the stimuli from an original experiment that showed an increase in substitution errors when the flankers were more similar to the target (J. Freeman et al., 2012), Keshvari and Rosenholtz (2016) generated synthesised images of the texture statistics (mongrels) of the letter stimuli used in that experiment. They then presented the mongrels in the visual periphery, and asked observers to identify the letters in the synthesised images. They showed that based on the mongrel images, observers made the same type of substitution errors found in the original experiment (Keshvari & Rosenholtz, 2016). This demonstrates that when the crowded percept is represented as texture, observers have access to the flanker features, and can make substitution errors. Similarly, the average identity of the target and flankers may still be available to the observer when the crowded percept is represented as texture, resulting in assimilation errors when reporting the identity of the target. As such, texture models could account for the substitution and assimilation errors made by children with amblyopia and typical vision, and adults in

the visual periphery, reported in Chapter 2. Although it is unclear how texture models would predict the increased in random responses found in children with amblyopia, the information loss associated with representing the crowded percept as a mongrel may also result in random errors.

Texture models are also broadly consistent with crowding resulting from the cortical undersampling of the peripheral visual field. They assume that in the typical visual system, the pooling regions over which summary statistics are computed are smaller in the fovea and increase with eccentricity. These models also take the size of the pooling region over which summary statistics are computed to be equal to Bouma's (1970) critical spacing (Balas et al., 2009). Although they do not make direct predictions on whether this pooling computation arises from the increase of receptive field size with eccentricity, the increase of the pooling region with eccentricity is consistent with crowding arising due to cortical undersampling of the visual periphery, with fewer neurons with large receptive fields responding to peripheral stimulation (Dow et al., 1981; Van Essen et al., 1984; Sereno et al., 1995; Engel et al., 1997). The increase of texture pooling regions with eccentricity would thus be in line with the findings of Chapter 3, where the general pattern of crowding matched the pattern of V1-V3 pRF size across the typical visual field. Although texture models have not been applied to amblyopic crowding, the increased extent of crowding and pRF size in V1-V3 for observers with amblyopia could be accommodated by larger pooling regions computing summary statistics than those in the typical visual field. To incorporate the non-uniformity of the amblyopic deficit across the visual field, these pooling regions should be disproportionately large in the fovea relative to the periphery, and larger in the temporal than the nasal visual field. Finally, the involvement of multiple visual areas in crowding would be consistent with texture models that assume that features like orientation are first detected by V1-like detectors, with summary statistics then computed at a later visual processing stage.

With regards to the results of Chapter 4, texture accounts of crowding could also account for the effects of target-flanker grouping found in amblyopic crowding. As previously discussed in section 5.3, effects attributed to target-flanker grouping and flanker-flanker grouping could be associated with differences in the formation of texture. Briefly, in cases of flanker-flanker grouping texture boundaries between the target and the flankers could be disconnected, segregating the target from the flankers. In contrast, in cases of target-flanker grouping, texture boundaries

between the target and flankers may be less clear, leading to the target and flankers having a connected boundary and difficulties with target segregation. These texture boundaries could be acting as position cues (Rosenholtz et al., 2017; Yu & Rosenholtz, 2018), alleviating the effects of the flankers when clear texture boundaries are formed, and disrupting target recognition when the texture boundaries are unclear. However, as texture models have not been tested with stimulus configurations alike those used in grouping studies. Crucially, texture models in their current form would need some modification to account for the differences found in Chapter 4 between the periphery and the typical and amblyopic fovea.

Overall, texture models (Balas et al., 2009; Keshvari & Rosenholtz, 2016) are broadly consistent with the findings in this thesis. This is not surprising, as similarly to the pooling framework adopted, texture models also assume that crowding in the typical visual system arises due to pooling regions that increase with eccentricity, which could arise due to the increase of receptive field size with eccentricity. A benefit of texture models relative to population response pooling models is that they represent all visual features simultaneously as summary statistics, and may thus be able to account for the effects of crowding along all stimulus features in one instance. This would negate the need of multiple distinct populations of detectors, each selective for a visual feature, in order to account for crowding along different visual dimensions. Therefore, the effects of crowding on orientation (Chapter 2) and polarity and shape (Chapter 4) could be reproduced within one single texture model. However, when it comes to crowding in amblyopia, it still remains untested whether the computed summary statistics of texture models provide a reliable representation of the perceptual experience of crowding. Texture models in the future could focus on whether synthesised images of the texture statistics could represent crowded vision in amblyopia.

5.4.2 Attentional Models

Amblyopia is considered a disorder of spatial vision (McKee et al., 2003), and as such I reasoned that amblyopic crowding is likely a spatial deficit and thus adopted a pooling framework. If instead of pooling, attention was involved in the elevation in crowding across the amblyopic visual field, then it should follow that there is some general attentional deficit specific to the amblyopic eye. Recently, it was proposed that due to the association between eye-movements and attention

(Goldberg & Wurtz, 1972), the poor fixation stability in strabismic amblyopia reported in this thesis and elsewhere (Subramanian et al., 2013) could result in an attentional deficit specific to the amblyopic eye (Verghese, McKee, & Levi, 2019). It is thus important to consider whether and how the findings presented in this thesis could be explained by attentional accounts of peripheral crowding (Strasburger et al., 1991; He et al., 1996; Intriligator & Cavanagh, 2001; Strasburger, 2005).

Although attentional accounts generally argue that crowding arises due to limitations of attention (Strasburger et al., 1991; He et al., 1996; Intriligator & Cavanagh, 2001; Strasburger, 2005), they deviate with regards to the mechanism they attribute these limitations to. The attentional account of crowding by Strasburger et al. (1991) and Strasburger (2005) is similar to substitution accounts, in that it argues that crowding in the periphery arises due to an imprecise attentional spotlight, zooming in and selecting the wrong object (a flanker instead of the target). The account of crowding by He et al. (1996), Intriligator and Cavanagh (2001), and Tripathy and Cavanagh (2002) is different, in that it proposes that when multiple objects are close together, limitations in attentional resolution lead to a failure in individuating the target from the flankers, and consequently to their binding. In this subsection, I briefly discuss how both these accounts may explain the findings presented in this thesis.

In order to be able to account for the perceptual effects of crowding in amblyopic, developing and peripheral vision found in Chapter 2, limitations in attentional resolution should result both in systematic assimilation and substitution errors, as well as random errors. The account by Strasburger (2005) and Strasburger et al. (1991) predicts substitution errors due to the inability to accurately focus spatial attention in light of the increased positional uncertainty reported in developing, amblyopic and peripheral vision. Assimilation errors are difficult to account for within this framework, as they require some combination or pooling of the target and flanker features. However, if the limited attentional resolution of attention arises due to a failure to individuate the target and flankers (He et al., 1996; Intriligator & Cavanagh, 2001; Tripathy & Cavanagh, 2002), this may result in a coarser representation of the crowded target. If this coarser representation of the target that includes bound features of the flankers, this may result in reports of intermediate target and flanker identities, taken as indicative of assimilation errors. It should be noted that this prediction is purely speculative, as the attentional account of crowding put forward by He et al. (1996), Intriligator and Cavanagh (2001), and

Tripathy and Cavanagh (2002) is agnostic with regards to perceptual effects of crowding. It is unclear how both attentional accounts of crowding would predict the increased random errors observed in a subset of children with amblyopia.

As I cannot fully exclude attention as a contributing factor for the for the perceptual errors reported in Chapter 2, it is also important to consider whether attention may also be a contributing factor to the increased pRF size in V1, V2, and V3 found for the amblyopic eye. Unfocused attention from a visual scene due to perceptual load has been shown to modulate overall levels of activity in the visual cortex of observers with typical vision (Schwartz et al., 2005). de Haas et al. (2014) showed that when perceptual load at fixation was high, pRF size significantly increased in V1, V2, and V3, compared when perceptual load was low. This suggested that withdrawing attention from the visual periphery due to high perceptual load in the fovea resulted in larger pRFs. This is broadly consistent with attention narrowing the tuning of population responses of V1 neurons (Martinez-Trujillo & Treue, 2004; Murray & Wojciulik, 2004). If there is a general spatial attention deficit for the amblyopic eye in strabismic amblyopia, this could lead to coarser positional tuning at the level of the neural populations, and an increase in pRF size for the amblyopic eye. Note however that in Chapter 3 I reported substantial variations in the enlargement of pRF size across the visual field for observers with strabismic amblyopia. As such an attentional deficit would necessarily have to be non-uniform across the visual field to account for the pRF deficit in the amblyopic eye.

In Chapter 4, I found differences between the amblyopic and typical fovea and the periphery with regards to whether crowding was modulated by grouping. In all three instances of crowding, crowding was strong when the target formed a uniform configuration with the flankers (target-flanker grouping). However, in contrast to the periphery, crowding in the typical and amblyopic fovea was not relieved when the flankers formed a uniform configuration with each other excluding the target (flanker-flanker grouping). In order for attentional accounts of crowding to account for these findings, they must incorporate differences between the amblyopic and typical fovea and the periphery that could explain the lack of so-called flanker-flanker grouping effects in the former. Within the context of the attentional account by Strasburger (2005) and Strasburger et al. (1991), these findings could be explained by differences in the precision of the attentional spotlight due to variations in positional uncertainty between the typical and

amblyopic fovea, and the visual periphery. When positional uncertainty is high as in the periphery (Levi et al., 1987; Hess & Field, 1993), the uniform independent configuration of the flankers in cases of flanker-flanker grouping may act as a positional cue to the target, aiding the zooming in of the attentional spotlight. Although this is broadly consistent with my interpretation of these findings in Chapter 4, it is purely speculative. In fact, neither attentional accounts of crowding (Strasburger et al., 1991; He et al., 1996; Intriligator & Cavanagh, 2001; Tripathy & Cavanagh, 2002; Strasburger, 2005) have used stimulus configurations similar to those used in grouping studies (Manassi et al., 2012; Herzog et al., 2015; Herzog et al., 2016; Manassi et al., 2016), nor attempted to account for similar results.

Overall, an attentional framework would have difficulties explaining the enlargement of pRF size across the amblyopic visual field found in Chapter 3, and would need modifications to account for the findings in Chapters 2 and 4. Additionally, an attentional framework of amblyopic crowding would necessarily imply an attentional deficit in amblyopia. Evidence suggests that observers with amblyopia show difficulties in counting sequentially presented objects in object enumeration tasks (Sharma et al., 2000), abnormalities in attentional blink tasks (Popple & Levi, 2008), and can track fewer moving objects in multiple object tracking tasks (Ho et al., 2006; Tripathy & Levi, 2008). However, due to the multiple deficits in amblyopic vision, it is not clear whether difficulties in these tasks represent a genuine attentional deficit, or whether they are associated with the increased extent of peripheral crowding in amblyopia found in Chapter 3, positional uncertainty (Levi et al., 1987; Hess et al., 1997; Wang et al., 1998), perceptual distortions (Pugh, 1958; Sireteanu, Wolf-Dietrich, et al., 1993; Sireteanu, Thiel, et al., 2008), and especially deficits in motion perception (Simmers et al., 2003; Simmers et al., 2005; Aaen-Stockdale & Hess, 2008).

To conclude, the findings in this thesis are better accounted for by a pooling framework of crowding that views amblyopia as a disorder of spatial vision (McKee et al., 2003) than an attentional account of crowding. However, the modulations of pRF size by attention (de Haas et al., 2014) raise the issue of the involvement of feedback processing not specifically in crowding, but in vision in general. The pooling framework adopted in this thesis does not take into account the generally accepted notion that both feed-forward and feedback processing play a significant role in object recognition (Wyatte, Jilk, & O'Reilly, 2014), but views crowding as a bottom-up feed-forward mechanism. Although amblyopic crowding may indeed

better fit within a pooling than an attentional framework of crowding, this does not rule out the involvement of feedback processes in the amblyopic deficit. Consistent with the pRF results from Chapter 3, the emerging consensus on amblyopia (Kiorpes & Daw, 2018) based on recent evidence (Li et al., 2007; H. Bi et al., 2011; Clavagnier et al., 2015; Shooner et al., 2015) is that both V1 and later visual areas are involved in the deficit. However, it remains unclear whether both feed-forward and feedback interactions among these visual areas are affected. As such, it is still important to elucidate whether both feedback and feed-forward interactions are affected in amblyopia.

5.5 Future Directions

A pooling framework of crowding on the whole provides a successful framework for considering crowding in strabismic amblyopia. This opens the door to consider whether other clinical instances of crowding which have not received the same amount of research interest as crowding in amblyopia, could also be explained within a pooling framework. Studies on crowding in the fovea of observers with posterior cortical atrophy have shown that it depends on the separation between the target and flankers (Crutch & Warrington, 2007, 2009; Yong et al., 2014), is selective for similarity between the target and flankers in contrast polarity (Yong et al., 2014), and perceptual errors when reporting a crowded target are best described by target-flanker averaging (Yong et al., 2014). On a cortical level, crowding effects in PCA were associated with lower grey matter volume within the right collateral sulcus (Yong et al., 2014), an area that if retinotopically mapped could correspond to V3 (Yeatman, Rauschecker, & Wandell, 2013), V3a (Grill-Spector & Malach, 2004) or V4 (Sereno et al., 1995). However, it is unclear whether this reduction in grey matter volume is associated with an increase in receptive field size for the neurons in this area, an assumption central to the pooling framework. Nonetheless, these findings suggest that crowding in posterior cortical atrophy shares important characteristics with crowding in peripheral vision and amblyopia, indicating that pooling may provide a successful framework in this instance. Additionally, the high occurrence of simultanagnosia – a selective impairment in recognising complex visual arrays despite preserved recognition of individual parts (Coslett & Saffran, 1991)– in observers with posterior cortical atrophy (Tang-Wai et al., 2004{McMonagle, 2006 #535)} could potentially disrupt grouping modulations of

crowding with multi-element flanker arrays. As such, a grouping framework might not be suitable for crowding in this instance.

However, less is known about crowding in the fovea of children and adults with dyslexia (Spinelli, De Luca, Judica, & Zoccolotti, 2002; Martelli et al., 2009; Moores et al., 2011; Zorzi et al., 2012). The few studies on dyslexic crowding have focused primarily in improving reading in observers with dyslexia, and have shown that increasing the separation between letters improves reading rates (Moores et al., 2011; Zorzi et al., 2012). This suggests that dyslexic crowding depends on the separation between the target and flankers. Although the perceptual effects of dyslexic crowding have not been investigated, the frequent letter confusions and feature reversals reported in dyslexia (Perri, Bartolomeo, & Silveri, 1996) could be in part due to feature integration and positional uncertainty, in line with the pooling framework presented in this thesis. However, a potential candidate neural site for such a pooling mechanism remains unclear, as research on the neural basis of dyslexia has focused primarily on temporal, parietal, and fusiform regions (Norton, Beach, & Gabrieli, 2015) and the neural underpinnings of dyslexic crowding have not been investigated. It is similarly unclear whether an alternative grouping framework could account for dyslexic crowding. Although there is evidence to suggest that observers with dyslexia have difficulties in grouping identical shapes in a Gestalt figure into rows or columns (Lewis & Frick, 1999), a different study has found that they do not experience any difficulties in perceiving groups in contrast to parts (Keen & Lovegrove, 2000). As such, it is difficult to predict from the limited evidence whether any difficulties in grouping shapes may also disrupt grouping modulations of dyslexic crowding. Overall, more research is required on determining the crowding mechanism in dyslexia and its neural underpinnings.

Finally, even less is known about the elevation in the extent of crowding in the fovea of observers with nystagmus (Chung & Bedell, 1995; Pascal & Abadi, 1995). Congenital nystagmus is characterised by involuntary oscillations of the eyes primarily along the horizontal meridian (Abadi, 2002). The limited evidence available on this instance of crowding suggests that it is not merely a consequence of the involuntary eye movements. Specifically, it has been shown that the increased extent of crowding in the fovea of observers with congenital nystagmus cannot be replicated in the fovea of observers with typical vision simply by imposing image motion to simulate involuntary oscillations (Chung & Bedell, 1995). Similarly to crowding in peripheral vision (Feng et al., 2007), crowding in nystagmus was

recently shown to have a horizontal-vertical asymmetry, with the extent of crowding being greater along the horizontal dimension (Tailor, Dahlmann-Noor, Theodorou, & Greenwood, 2018). This asymmetry was not associated with eye movement parameters such as velocity, fixational offsets, and degree of foveation of the stimuli (Tailor, 2019), suggesting it might not result from the primarily horizontal eye movements. Together, these findings suggest that similarly to amblyopic crowding, crowding in nystagmus is likely a sensory deficit reflecting long-term neural changes. However, the underlying neural basis of crowding in nystagmus and whether it is susceptible to grouping modulations remains unclear. Therefore, more research is required to determine whether instances of crowding in developmental dyslexia and nystagmus could be accounted for successfully within a pooling framework.

Posterior cortical atrophy, dyslexia, and nystagmus are complex disorders with vastly different clinical characteristics. One might thus reason that this variability makes it unlikely that a single framework would account for crowding in all three instances. Central to pooling accounts of crowding (Parkes et al., 2001) is that crowding in the typical visual system arises due to the undersampling of the visual periphery, with neurons responding to peripheral stimulation having large receptive fields to ensure adequate coverage. Based on this assumption of pooling accounts, a first step towards tackling the complex issue of a unified crowding framework would be to investigate whether these three clinical instances of increased foveal crowding also show increased receptive field size.

5.6 Conclusion

In this thesis, crowding in strabismic amblyopia was investigated to better understand whether it matches crowding in the typical visual system. A pooling framework of crowding was adopted that posits that crowding results from the increased integration of the target and flanker features as a mechanism to promote perceptual similarity among adjacent regions of the visual field when neurons have large receptive fields. Here, I considered whether such an account could describe crowding in amblyopic vision. It was demonstrated that the perceptual effects of crowding in amblyopia are largely systematic and resemble those in the adult periphery. Although a minority of observers with amblyopia showed non-systematic perceptual effects, a model simulating pooled responses of populations of visual neurons characterised the perceptual effects of amblyopic crowding in the majority

of observers with amblyopia. This suggested a common underlying mechanism with crowding in the typical periphery. I then showed that overall both amblyopic and peripheral vision are characterised by reduced acuity, elevated extents of crowding, and an increase in fMRI estimates of population receptive field size in V1-V3. Together, these results suggested that the common pooling mechanism in peripheral and amblyopic vision mat be associated with cortical undersampling in the periphery and the amblyopic visual field. When considering whether an alternative grouping account can explain crowding in amblyopia, I found that amblyopic crowding followed grouping predictions with some flanker configurations but not others. The pattern of flanker configuration effects in the amblyopic fovea matched the one observed in typical fovea, suggesting grouping cannot fully account for these instances of crowding. As a whole, I conclude that pooling generally provides a successful framework to account for crowding in strabismic amblyopia.

6 Appendices

6.1 Appendix A: Supplementary Information for Chapter 2

6.1.1 Clinical Details of Children with Amblyopia and Unaffected Vision

Clinical details of the children with typical vision (N=20) tested in Chapter 2 are presented in Table 6.1. below.

Initials	Age (months)	Sex	Refractive Error	logMAR acuity	TNO stereo
IK	69	M	L: Plano R: Plano	L: 0.00 R: 0.00	60”
SR	78	F	L: Plano R: Plano	L: 0.00 R: 0.00	60”
JN	82	F	L: Plano R: Plano	L: 0.00 R: 0.00	60”
RN	82	F	L: Plano R: Plano	L: 0.06 R: 0.00	60”
DM	94	M	L: Plano R: Plano	L: 0.00 R: 0.00	60”
MK	88	M	L: Plano R: Plano	L: 0.00 R: 0.00	60”
AS	73	M	L: Plano R: Plano	L: 0.10 R: 0.04	120”
TD	60	M	L: Plano R: Plano	L: 0.10 R: 0.04	60”
AT	54	M	L: +0.50 DS R: +1.00 DS	L: 0.16 R: 0.18	120”
AO	73	M	L: +0.50 DS R: +1.00 DS	L: 0.10 R: 0.20	60”
AP	84	M	L: +0.25 DS R: +0.25 DS	L: 0.00 R: 0.00	60”
LD	36	F	L: Plano R: Plano	L: 0.00 R: 0.10	60”
BK	93	M	L: -0.75/-1.00×50° R: -3.00 DS	L: 0.12 R: 0.20	120”
BG	70	M	L: Plano R: Plano	L: -0.10 R: -0.10	60”
FH	73	M	L: Plano R: Plano	L: 0.00 R: 0.00	60”
MH	73	M	L: Plano R: Plano	L: 0.00 R: 0.00	60”
JP	89	M	L: Plano R: Plano	L: 0.00 R: 0.00	60”
LK	61	F	L: Plano R: Plano	L: -0.10 R: -0.10	60”
MP	74	F	L: Plano R: Plano	L: 0.02 R: 0.10	60”
JS	56	F	L: Plano R: Plano	L: 0.00 R: 0.00	60”

Table 6-1 Clinical details of children with typical vision included in the group with typical vision (N=20).

Age is reported in months. Optical correction includes cylindrical and spherical values with the appropriate axes for each eye (L = left eye, R = right eye). logMAR acuity is also reported for each eye. Results from the TNO stereo-acuity test and are reported in seconds of arc.

Clinical details of the children with strabismic amblyopia (N=20) tested in Chapter 2 are presented in Table 6.2 below.

Initials	Age (months)	Sex	Ocular alignment	Refractive Error	logMAR acuity	TNO Stereo.
AG	89	M	L SOT 20Δ	L: +5.25 DS R: +1.75 DS	L: 0.32 R: -0.06	Nil
NK	63	M	L SOT 30Δ	L: +6.5/-1.25 ×175° R: +4.5/-0.5 ×5°	L: 0.36 R: 0.1	Nil
EL	100	F	R SOT 40Δ	L: +2.25/-0.25 ×180° R: +1.25/-2.00 ×15°	L: 0.04 R: 0.32	Nil
RA	62	M	R SOT 35Δ	L: +2.00 DS R: +2.50/-5.00 ×180°	L: 0.02 R: 0.22	Nil
KM	67	F	n: L SOT 8Δ d: L SOT 6Δ	L: +6.25/-1.00 ×180° R: +4.25/-0.25 ×180°	L: 0.46 R: 0.00	Nil
OB	55	M	n: L SOT 18Δ d: L SOT 9Δ	L: +2.00/-1.00 ×180° R: +0.50/-0.50 ×180°	L: 0.36 R: 0.00	360”
AS	70	F	L XOT 4Δ	L: +7.00/-0.50 ×180° R: +6.75 DS	L: 0.76 R: 0.00	Nil
EB	37	F	R SOT 30Δ	L: +7.25/-1.50 ×5° R: +8.00/-1.50 ×70°	L: 0.12 R: 1.10	Nil
AA	89	F	n: R SOT 25Δ d: R SOT 18Δ	L: +7.25/-2.00 ×170° R: +7.75/-1.00 ×20°	L: 0.00 R: 0.24	Nil
SP	76	F	n: R SOT 16Δ/ L 14Δ d: R SOT 10Δ/ L 0Δ	L: +4.00/ +0.5 ×180° R: +4.50/ -0.5 ×175°	L: 0.00 R: 0.38	Nil
OS	77	M	L XOT 35Δ	L: +5.50/-0.50 ×180° R: +3.00/-0.50 ×180°	L: 0.50 R: 0.00	Nil
HC	86	F	n: R SOT 10Δ d: R SOT 4Δ	L: +3.00/-0.50 ×180° R: +5.00/-1.00 ×180°	L: 0.10 R: 0.36	Nil
LC	64	F	L XOT 20Δ	L: +6.25/-1.25 ×180° R: Plano	L: 0.50 R: 0.00	Nil
FG	79	M	R SOT 85Δ	L: 0.50/-0.25×180° R: 5.50/-0.75 ×180°	L: 0.00 R: 0.54	480”
FH	83	M	L SOT 45Δ	L: +6.50/ -1.00 ×180° R: +4.75/ -1.25 ×175°	L: 0.92 R: 0.00	Nil
EM	70	F	L XOT 10Δ	L: +6.00 DS R: +6.00 DS	L: 0.22 R: 0.00	Nil
RM	85	M	L XOR 25Δ	L: +7.00/-1.00 ×180° R: +1.25/-0.25 ×180°	L: 0.86 R: 0.00	Nil
AM	73	M	L SOT 40Δ	L: +8.00/-1.50 ×70° R: +7.50/ -1.00 ×20°	L: 0.72 R: 0.18	Nil
MH	45	F	L SOT 10Δ	L: +7.50/-1.00 ×70° R: +6.50/-0.50 ×180°	L: 0.36 R: 0.12	NIl
JF	44	M	L SOT 20Δ	L: +7.00/-0.5 ×20° R: +6.00/-1.00 ×120°	L: 0.44 R: 0.02	Nil

Table 6-2 Clinical details of children with strabismic amblyopia included in the amblyopic group (N=20).

The “ocular alignment” column indicates the outcome of near (n) and distance (d) prism tests. SOT = esotropia, XOT = exotropia. The degree of deviation is shown in prism dioptres and the amblyopic eye is denoted. Remaining columns are in the same format as Table 6.1 above.

6.1.2 Population Response Noise Model

6.1.2.1 Model Description

In Chapter 2 section 2.3.3, I presented a population response pooling model of crowding. This model successfully captured the errors made by children with unaffected vision and amblyopia and adults in the typical periphery, when tasked with reporting the orientation of a crowded target. However, in addition to this model, which will be referred to as the *final* model, two alternative models were also constructed in the attempt to simulate the perceptual effects of crowding.

The first model rested on the assumption that crowding does not have a systematic effect on target appearance, but rather distorts the target orientation randomly. This assumption is based on the evidence suggesting that amblyopic vision is characterised by perceptual distortions (Pugh, 1958; Sireteanu, Wolf-Dietrich, et al., 1993; Sireteanu, Thiel, et al., 2008). As such, this model was tested only on the data from the children with typical vision and amblyopia. I will refer to this model as the *noise model*.

Similarly to the final model, I assumed a population of detectors selective to orientation, based on the well-documented orientation selectivity of visual neurons in V1 (Schiller et al., 1976). Each detector was sensitive to a range of orientations, with a Gaussian tuning function and peak on a particular orientation from -180° to 180° (the full range of possible response error in our orientation-matching task) and gradually lesser sensitivity to nearby orientations. Based on the principles of population coding, the population activity distribution is similarly Gaussian for uncrowded targets, with a peak on or near the orientation of the Landolt-C target. The perceived orientation is read out from this peak. As in the final model, the population response random model had three stages.

I used one free parameter for the bandwidth of the detectors in the population. The first stage of the model was identical to the final model:

$$y = \theta + \alpha\sigma_e$$

Where θ represented the target orientation θ_t or the flanker orientation θ_f, σ_e represented the error and α was the magnitude of this error and the second free parameter in the model. The population response to the target was a Gaussian function centred on 0°, whereas for the flankers it centred on either 30° or 90°.

The second stage was different from the final model. Crowding was depicted as a process that distorts the response to the target in a non-systematic way, resulting in a noisy estimate of the target orientation. Contrary to the final model,

there was no pooling stage where the population response to the flanker orientation was taken into account and pooled with the population response to the target orientation. In contrast, in the noise model the addition of flankers to the target adds random noise. Therefore, the third and final free parameter for this model was crowding noise. The population response to the target and flankers thus was:

$$y_c = y_t + \beta\sigma_n$$

Where σ_n represented the added random noise to the population response to the target and β was the third and final free parameter in this model– the magnitude of this noise. This resulted in three free parameters for the noisy estimate of the target orientation: the bandwidth of the underlying detectors, the Gaussian noise (σ_e), and the crowding noise (σ_n).

In the third stage of the noise model, the perceived orientation of the target was read from the peak of the population response to the target and flankers. In this model, errors arise due to the addition of crowding noise, resulting in peak population responses at orientations that did not correspond to the target, that are then read out as the perceived orientation. As such, the decision stage was identical to the final model. I ran 1000 trials of this 3-stage model for each flanker orientation condition. I determined the best fitting parameters for the control group and group with amblyopia using a coarse-to-fine fitting procedure. The best fitting parameters were determined by minimising the least squared error (LSE) between the response error distributions and the simulated population response distributions.

6.1.2.2 Model Comparison

I ran 1000 iterations of the final model presented in Chapter 2 section 2.3.3 and the noise model with the best fitting parameters for each determined by the coarse-to-fine fitting procedure. Note that each iteration consisted of 1000 trials per flanker condition, as in the model description above. To determine which of the two models best fitted the response distributions of the children with unaffected vision and amblyopia, the Akaike Information Criterion (AIC) was used (Akaike, 1973, 1974). The AIC considers the least squared error (LSE), and corrects for the number of parameters included in each model. The LSE was computed by comparing the response error distributions with the mean population response distributions from the 1000 simulations for each flanker condition (uncrowded, 30° flanker difference from the target, 90° flanker difference). I then took the mean LSE value across flanker conditions to compute the AIC. Note that the final model had five

parameters, whereas the noise model had three. Figure 6.1 shows group data of the children with unaffected vision (Fig. 6.1 A) and amblyopia (Fig. 6.1 B) along with the mean and the range of the population response distributions from the 1000 simulations. AIC values for fits to the individual data are presented in section 6.1.4 of this Appendix.

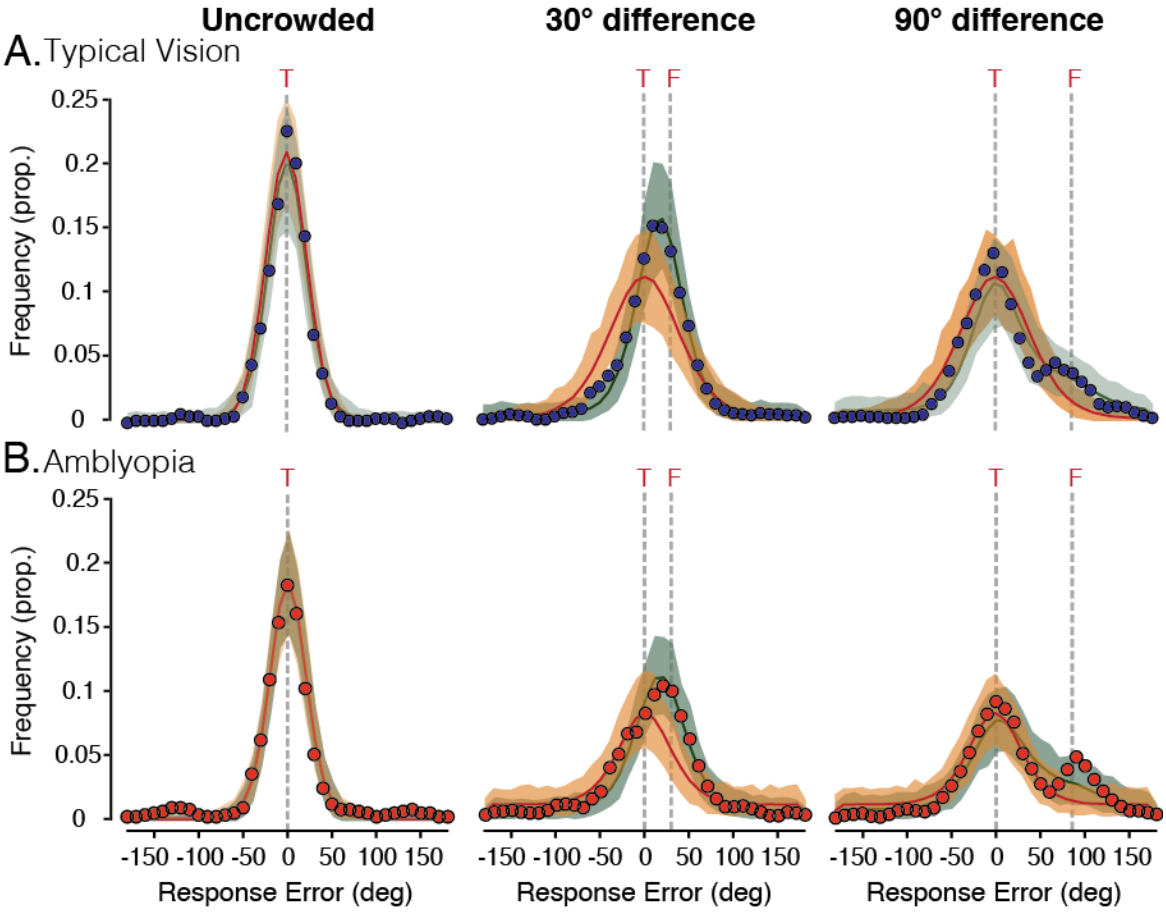

Figure 6-1 Comparison between the final model and the noise model on the response error distributions of children in the group with typical vision and amblyopia

A. *Model comparison for the group with typical vision:* The orange line indicates the mean distribution of 1000 iterations of the population response random model, and the green line the mean distribution of the population response pooling model. The colour-coded shaded areas represent the range of distributions for each model in 1000 iterations. The group response error distribution from the orientation-matching task is presented by the dots. The grey line indicates the target location ('T'), and for the two flanker conditions in which flankers were present, the flanker location ('F').
B. *Model comparison for the group with amblyopia*: Plotted with the same conventions as in A.

It is clear from Figures 6.1 A and 6.1 B that both the noise model (orange) and the final population response pooling model (green) captured well the uncrowded response error distribution. This is not surprising as they contained the same parameters for simulating the response error distribution to uncrowded targets (i.e. the bandwidth and Gaussian noise). Clear differences between the two

models however arise for the two conditions that included flankers, as the models differed significantly in their crowding stage. When the flankers differed by 30° from the target, the noise model resulted in a simulated response distribution centred at 0°, and thus failed to capture the shift of the response error distributions towards the flankers in both the group data of the children with typical vision and amblyopia. This is due to the lack of the combination of the population responses to the target and flanker orientation. In contrast, the final model that includes this combination and weights the contribution of the population response to the target. As such, it captures the shift towards the flankers and clearly outperforms the noise model in this condition. When the flankers differed by 90° from the target, the response error distributions were bimodal for both children with typical vision and amblyopia, with the first peak centred at 0° and the second centred at 90°. The noise model did not capture this bimodality. The failure of the noise model was due to the lack of inclusion of the population response to the flankers at the crowding stage, resulting in a unimodal distribution centred on 0°. Although the difference between the models is not as pronounced as in the condition in which the flankers differed by 30° from the target, it is clear that the final model outperformed the noise model in this condition too.

The better performance of the final model was clear from the AIC values. Note that the lower the AIC value, the better the model fits the data, regardless of the number of parameters. For the response error distributions of the control group, the noise model had an AIC value of -187.34, whereas the final model had a lower value of -228.52. Similarly, for the data from the group with amblyopia, the noise model had an AIC value of -194.05, whereas the final model had a lower value of -228.68. As such, I conclude that the perceptual effects of crowding in developing and amblyopic vision at the group level do not merely add random noise to the target orientation. Rather, the perceived orientation of the target is systematically shifted to resemble the flanker orientation.

6.1.3 Noise-less Population Response Pooling Model

6.1.3.1 Model Description

The simulations with the noise model demonstrated that a noisy estimate of the target orientation did not adequately characterise the perceptual effects of crowding in children with unaffected vision and amblyopia at a group level. Rather, the simulations presented in section 6.1.2.2 of this Appendix suggested that a

combinational pooling stage where the population response to the flankers is taken into account was necessary. With such a stage, it was possible to simulate the shift of the distribution towards the flankers in the condition in which the flankers differ by 30° from the target, as well as the bimodality of the responses when flankers differ by 90°. However, in addition to the flanker weight parameters that simulate these effects seen in the data, in the final model I also included a crowding noise parameter at the pooling stage. To establish whether crowding noise was required to characterise the data, I also ran a preliminary model without it. I call this alternative model the *noise-less* population pooling response model.

The noise-less population response pooling model had three stages. The details of the first and third stages were identical to the final model, and are described above in section 6.1.2.1. As such, here I focus on the second stage that differentiated this model from the final one. At the second stage, I simulated the effects of crowding on perceived target orientation. I took a weighted average of the population response to the target and flankers, but in contrast to the final model, I did not add noise to the combined population response. The combination of the population responses to target and flanker was:

$$y_c = y_t w_t + y_f w_f$$

Where w_t and w_f were the weights for the target and flankers, respectively. The flanker weight could be any value from 0-1, with the weight of the target being equal to one minus the value of the flanker weight. As the data showed that flankers had different effects on the perceived orientation of the target depending on the flanker orientation difference, different weights were used for the condition in which the flankers differed 30° from the target and for the condition in which they differed 90° from the target. This resulted in four free parameters: the bandwidth of the underlying detectors, the Gaussian noise (σ_e), the flanker weight for the 30° flanker difference condition ($w_{f_{30°}}$) and the flanker weight for the 90° flanker difference condition ($w_{f_{90°}}$). These parameters are identical to the final model, with exception of the crowding noise that was not included in this alternative model.

In the final stage of the noise-less population response pooling model, the perceived orientation of the target was read from the peak of the population response to the target and flankers. Errors arise due to the weighted combination of the population response to the target and flankers. If the flanker weight was low, the combined population response would be centred near the target, resulting in a peak

close to the target and a small error. If the flanker weight was high, the combined population response would be centred near the flankers, resulting in a peak away from the target and an error of a larger magnitude.

I ran 1000 trials of the 3-stage noise-less population response pooling model for each flanker orientation condition (uncrowded, 30° flanker difference, 90° flanker difference). I determined the best fitting parameters for the data using a coarse-to-fine fitting procedure and using the least squared error (LSE) between the response error distributions and the simulated population response distributions. To determine whether the crowding noise parameter was required for all the groups, I simulated the response error distributions of children with typical vision and amblyopia and those of adults tested in the visual periphery.

6.1.3.2 Model Comparison

I ran 1000 iterations of the noise-less population response pooling model and the final model presented in Chapter 2 section 2.3.3 with the best fitting parameters from the coarse-to-fine fitting procedure. Each iteration consisted of 1000 trials per flanker condition, as in the model description above. Similarly to the model comparison described in section 6.1.2.2 of this Appendix, the Akaike Information Criterion (AIC) was used (Akaike, 1973, 1974) to determine which model best characterised the data. I discuss the model fits below for the group data of the children with typical vision and amblyopia (section 2.3.2.2.1) and the adults (section 2.3.2.1.1).

6.1.3.2.1 Children

Figure 6.2 shows group data of the children with unaffected vision (Fig. S1.2 A) and amblyopia (Fig. S1.2 B) along with the mean and the range of population response distributions from 1000 iterations of the final model and the noise-less population response pooling model. Note that AIC values for fits to the individual data are presented in section 6.1.4.

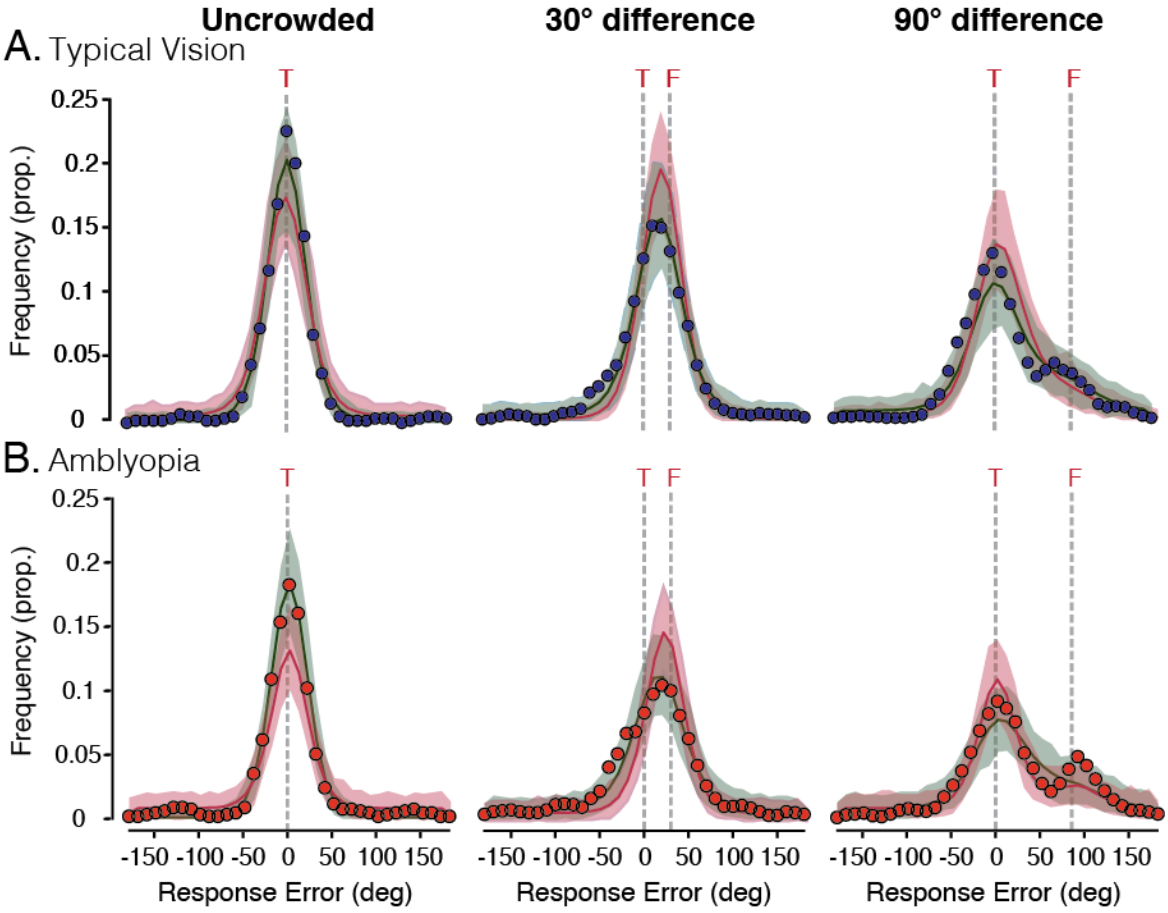

Figure 6-2 Comparison between the final model and the noise-less model on the response error distributions of children in the control and amblyopic groups

A. *Model comparison for the group with typical vision:* The pink line indicates the mean distribution of 1000 iterations of the noise-less population response pooling model, and the green line the mean distribution of the final population response pooling model. The colour-coded shaded areas represent the range of distributions for each model in 1000 iterations. The group response error distribution from the orientation-matching task is presented by the dots. The grey line indicates the target location ('T'), and for the two flanker conditions in which flankers were present, the flanker location ('F').
B. *Model comparison for the group with amblyopia*: Plotted with the same conventions as in A.

It is clear from Figure 6.2A and 6.2B that the noise-less population response pooling model (pink) overestimated the height of the peak of the uncrowded response error distribution, and as indicated by the increased height of the base of the simulated response distribution, overestimated the frequency of responses at the tails of the unimodal response error distribution. Figure 6.2A also indicates that for the group with typical vision, the noise-less population response pooling model overestimated the bandwidth of the response error distribution. In contrast, the final model (green) captured the distribution of response errors to uncrowded targets well for the group with typical vision, and almost perfectly for the amblyopic group. In the condition in which the flankers differed 30° from the target, the noise-less model captured the shift of the response error distribution from the uncrowded condition

towards the flankers. However, it overestimated the peak and underestimated the bandwidth of the response error distribution for both groups. The final model followed the data in this flanker condition, matching the peak and bandwidth of the unimodal response error distribution. In the condition in which the flankers differed 90° from the target, for the group with typical vision, the noise-less model captured the first peak of the response error distribution centred on the target, but underestimated the height of the second peak located at the flanker orientation. For the group with amblyopia, the noise-less model overestimated the height of the first peak, but similarly with the data from the group with typical vision, underestimated the height of the second peak. In contrast, the final model underestimated the first peak of the bimodal response error distribution for both groups, but captured well the second peak for the control group and less so for the group with amblyopia. Although in the condition in which the flankers differed by 90° from the target it is clear which model best fitted the data, in the uncrowded and 30° flanker difference condition it is clear that the final model outperformed the noise-less population response pooling model.

In line with these observations, the AIC values for the noise-less population response pooling model were higher than those for the final model. The noise-less population response pooling model had an AIC value of -190.51 for the response distributions of the group with typical vision, and a -179.65 for the group with amblyopia. As reported in Appendix A section 6.1.2.2 above, the final population response pooling model had AIC values of -228.52 for the group with typical vision, and -230.62 for the amblyopic group. Note that the only difference between the two models was the crowding noise parameter that was included in the final model, but not the noise-less population response pooling model. With a single noise parameter in both uncrowded and crowded cases, the noise-less population response pooling model had to compromise by adding more noise in the first stage than required to successfully capture the uncrowded data but less noise in the conditions in which there is crowding. These model simulations suggest that when the population responses to the target and flankers are pooled, additional noise is necessary in order to capture the perceptual effects of crowding in developing and amblyopic vision.

6.1.3.2.2 Adults

When examining the individual response error distributions in sections 2.3.2.1.2 and 2.3.2.2.2 of Chapter 2, it was clear that random responses under crowding conditions were more common in children than adults tested in the visual periphery. It was thus essential to establish that the crowding noise parameter was necessary to simulate the perceptual effects of crowding not only in the typically developing and amblyopic fovea, but also in the typical periphery. To achieve this, I compared the final model with the noise-less population response pooling model. Figure 6.3 shows group response error distributions at 2.5° eccentricity (Fig. S1.3 A) and 15° eccentricity (Fig. S1.3 B) along with the mean and the range of 1000 iterations of the final model (green) and the noise-less population response pooling model (pink). These eccentricities were selected to illustrate the performance of the models in the adult periphery as they were maximally different

At 2.5° eccentricity, when the target was uncrowded, the noise-less model underestimated the height of the peak of the response error distribution, while overestimating the bandwidth. It had the opposite effect when flankers differed by 30° from the target, overestimating the height of the peak of the unimodal response distribution while underestimating its bandwidth. The final model clearly followed the data in both these flanker conditions. When the flankers differed by 90° from the target, the noise-less population response pooling model followed the data, capturing both peaks of the bimodal distribution, and outperformed the final model that muted the second peak. As such, at 2.5° eccentricity the noise-less population response pooling model performed similarly to the children's data described above in section 6.1.3.2.1.

With the group response error distributions from 15° eccentricity, the noise-less population response model captured well the height of the peak in the uncrowded condition, with a small underestimation of the bandwidth. This suggests that the noise in the conditions in which the target was uncrowded and crowded was more balanced than at 2.5° eccentricity. When the flankers differed by 30° from the target orientation, the noise-less population response pooling model overestimated the height of the peak of the response error distribution, and underestimated the bandwidth. The final model followed the data in both these conditions, capturing the height of the peak and the bandwidth better than the noise-less population response pooling model. When the flankers differed by 90° from the target, the noise-less population response pooling model overestimated

the height of the both peaks of the bimodal response error distribution, and underestimated their bandwidth. The final model followed the height of the two peaks, while overestimating the bandwidth. Overall, at 15° eccentricity, the differences between the two models are less clear at the uncrowded and 90° flanker difference conditions, but the final model clearly outperforms the noise-less population response pooling model when the flankers differ by 30° from the target.

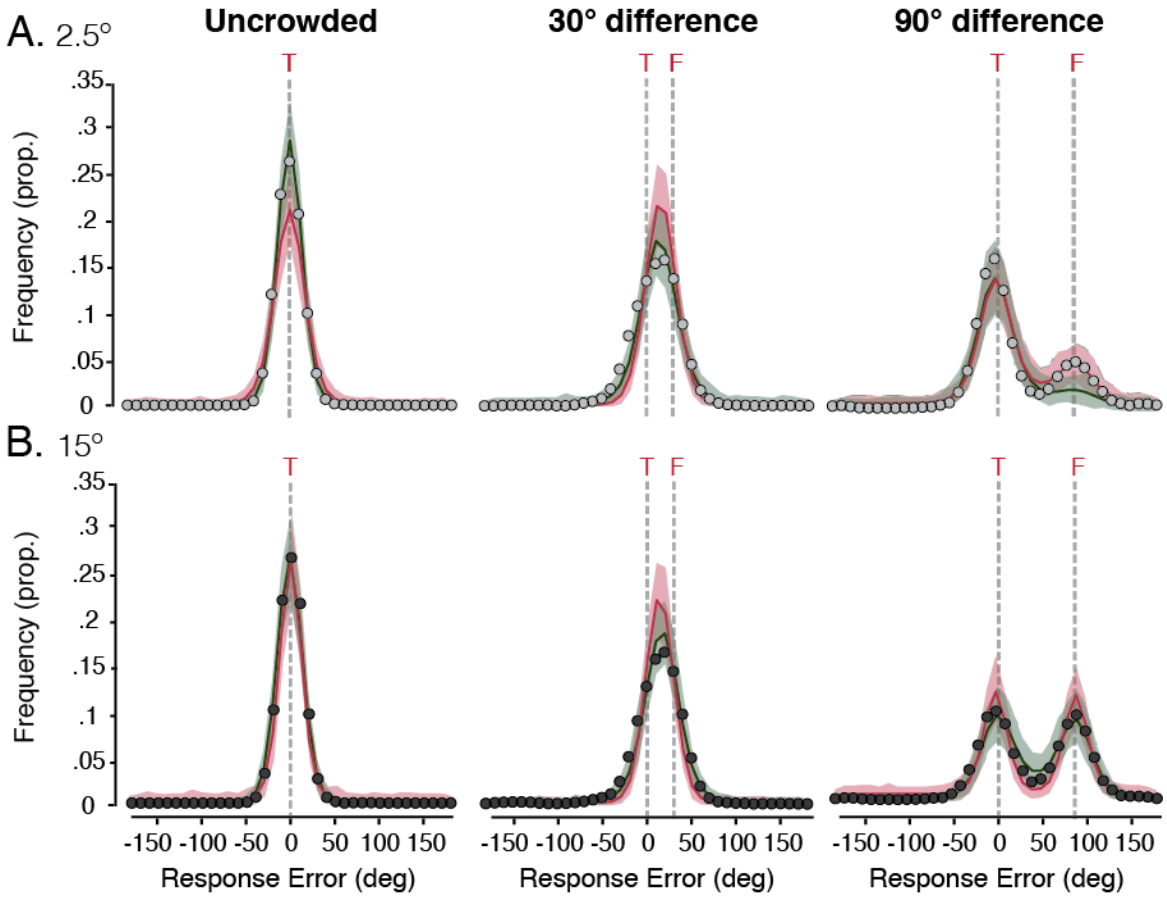

Figure 6-3 Comparison between the final model and the noise-less model on the response error distributions at 2.5° and 15° in the adult periphery

A. *Model comparison for 2.5° eccentricity:* The pink line indicates the mean distribution of 1000 iterations of the simple population response pooling model, and the green line the mean distribution of the final population response pooling model. The colour-coded shaded areas represent the range of distributions for each model in 1000 iterations. The group response error distribution from the orientation-matching task is presented by the dots. The grey line indicates the target location ('T'), and for the two flanker conditions in which flankers were present, the flanker location ('F').
B. *Model comparison for 15° eccentricity*: Plotted with the same conventions as in A.

The AIC values clearly demonstrate that the final model that included the crowding noise parameter clearly outperforms the noise-less population response pooling model. For the response error distributions at 2.5° eccentricity, the AIC value for the noise-less population response pooling model was -175.30 while for

the final model it was -208.24. Although the difference between the two models was smaller for the response error distributions at 15°, the AIC value for the noise-less population response pooling model was –182.64 and for the final model -210.77. Therefore, a crowding noise parameter is required to simulate the perceptual effects of crowding not only in the developing and amblyopic fovea, but in the adult periphery as well.

6.1.4 Model Comparison: Individual Fits

6.1.4.1 Typically Developing Children

Table 6.3 below shows the AIC values from 1000 simulations of the two alternative models and the final model on the individual response error distributions of children with typical vision. Note that the lower the AIC value, the better the fit of the model to response error distributions. The model with the best fit to the data has the AIC value corresponding to each observer in the group of children with typical vision in bold.

Observer ID	Noise	Noise-less	Final Model
1	-156.15	-143.53	**-161.38**
2	**-190.82**	-164.73	-186.49
3	-148.98	-157.76	**-160.36**
4	-152.09	-150.90	**-154.74**
5	-146.62	-157.38	**-163.38**
6	-120.61	-135.84	**-136.05**
7	-134.12	-154.54	**-184.43**
8	-168.84	**-183.02**	-178.94
9	-156.39	-183.62	**-203.99**
10	-132.82	-179.19	**-180.60**
11	-171.99	-163.82	**-188.58**
12	-176.27	-175.47	**-177.03**
13	-127.25	-152.67	**-155.72**
14	-134.68	**-135.70**	-130.08
15	-128.91	-183.63	**-185.05**
16	-132.28	**-165.71**	-154.66
17	-176.76	-175.46	**-176.84**
18	-185.49	-176.38	**-185.67**
19	-169.80	-182.25	**-183.96**
20	-168.67	-143.46	**-187.98**

Table 6-3 AIC values for the three models of the individual response error distributions of children with unaffected vision.

Note that lower values are indicative of a better fit. The best-fitting model is selected with bold.

From the twenty children tested, for the response error distributions of sixteen children the final model performed best, in line with the group data. For three children, the population response pooling model without the noise parameter performed better than the final model. For this subset of children, response errors did not become substantially more variable in the crowding conditions relative to the uncrowded condition. Therefore, the addition of crowding noise to the pooled population response was not required, as the variability induced in the pooled population response from the combination of the responses to the target and flanker orientations was sufficient. However, their response error distributions did shift towards the flankers when the target was crowded, and flanker weights were necessary in order to capture this shift. As such, the perceptual errors of this subset

of children were still best captured by a model that included the pooling of population responses to the target and flankers. The mechanism underlying the perceptual errors in this subset of children is thus consistent with the sixteen children, for which crowding was depicted as the pooling of population responses to the target and flanker orientations.

There was also one child with typical vision for which the noisy model captured their response errors best. As can be seen from the AIC values, two models that included crowding noise, the noisy model and the final model, significantly outperformed the noise-less population response pooling model. As such, crowding noise was necessary to capture the response error distributions for this observer. Crowding substantially increased the variability in this observer's responses, resulting in responses with increased variability to crowded targets. However, as the majority responses were to the target orientation, there was no shift in the distribution of the response errors towards the flankers in neither of the two crowding conditions. Consequently, flanker weights were not required to capture this response error pattern. Therefore, this observer deviates from the group with typical vision, in that the model that best describes their perceptual errors did not simulate crowding as the pooling of population responses to the target and flanker orientations.

6.1.4.2 Children with Amblyopia

Table 6.4 below shows the AIC values from 1000 simulations of the three models (two alternatives and the final) on the individual response error distributions of children in the group with amblyopia. The model with the best fit to the data has the value corresponding to each individual in the group of children with amblyopia marked in bold.

Observer ID	Noise	Noise-less	Final Model
1	**-168.99**	-156.79	-165.60
2	-123.42	-171.72	**-177.43**
3	**-175.23**	-163.58	-166.03
4	**-175.19**	-162.08	-170.19
5	-111.34	-176.05	**-178.83**
6	-156.44	-170.18	**-187.98**
7	**-172.15**	-167.51	-163.14
8	-169.83	-170.06	**-177.01**
9	-189.91	-170.39	**-190.87**
10	-112.99	-128.66	**-136.81**
11	-148.47	-163.80	**-176.71**
12	-177.26	-138.87	**-181.17**
13	**-177.40**	-158.41	-163.81
14	-144.84	-177.23	**-177.26**
15	-161.10	-165.09	**-165.16**
16	-131.32	-159.20	**-159.80**
17	-193.41	-161.83	**-195.02**
18	-161.19	-164.44	**-167.63**
19	-156.24	-154.96	**-156.57**
20	-156.50	-162.80	**-183.19**

Table 6-4 AIC values for the three models of the individual response error distributions of children with amblyopia.

Note that lower values are indicative of a better fit. The best-fitting model is selected with bold.

For fifteen out of the twenty children with amblyopia tested, the final population response pooling model best characterised their data. However, for the remaining five, the noisy model outperformed the final model. Similarly to the child with typical vision for which the noisy model best captured their data, crowding significantly increased response variability for these five observers with amblyopia. However, the majority of responses were of 0 error, indicative of reports of the target orientation. As such, relative to the uncrowded condition, crowding did not result in a shift of the response error distribution towards the flankers. The response errors for one of these individual observers is discussed in section 2.3.2.2.2, as an

example of the subset of children with amblyopia who showed increased response variability. In these cases, the model that best described the perceptual errors did not simulate crowding as the pooling of population responses to the target and flanker orientations, but a noisy population response to the target orientation. Therefore, these observers deviated from the rest of the group with amblyopia, as well the majority of children with typical vision and adults, for which a pooling mechanism best captured the perceptual errors under crowded conditions. As discussed in Chapter 2 (section 2.4), the vision in the amblyopic eye of these children could be characterised by increased perceptual distortions (Fronius & Sireteanu, 1989; Lagreze & Sireteanu, 1991; Sireteanu, Lagreze, et al., 1993; Sireteanu, Baumer, et al., 2008), resulting in increased random errors.

6.2 Appendix B: Supplementary Information for Chapter 3

6.2.1 Stepwise Regression Analyses Model Progression

In section 3.3.2.1, the aim was to understand the sources of variation in pRF size between the DE and the AME, and the DE and the FFE. To achieve this, stepwise linear regression analyses were conducted. pRF size in V1, V2, and V3 was included as the dependent variable, and eccentricity (1°-19°), visual field (coded 0 for nasal and 1 for temporal), and eye (DE and FFE, DE and AME) as predictors, as well as the interactions between the predictor variables were also included in the models. Note that for the predictor variable for eye, the DE was always coded as 0, whereas either the AME or the FFE (depending on the analysis) were coded as 1. A forward-selection procedure was applied, with predictor being added at each step of the model.

6.2.1.1 pRF size: Dominant Eye of Controls and Amblyopic Eye

In section 3.3.2.1.1. differences in pRF size between the dominant eye (DE) of controls and the amblyopic eye (AME) of observers with amblyopia were investigated using stepwise multiple linear regression analyses. Here, I present the tables with the full progression of the models discussed in Chapter 3.

6.2.1.1.1 pRF size in V1

Table 6.5 shows the variance in pRF size in V1 explained by the model, as indicated by the R^2, as well as the unstandardized B coefficient values for each predictor at each step with the addition of an extra variable. For pRF size in V1, the best model was reached in three steps, with no variables removed. In the first step, eccentricity was added. The model with just the eccentricity accounted approximately 31% of the total variance in V1 pRF size. At the second step, eye (DE and AME) was added as a significant predictor. Together with eccentricity, they accounted for approximately 38% of the total variance in pRF size. In the third and final step, the interaction between eccentricity and eye was added. Although the addition of the interaction between eccentricity and eye did not substantially increase the variance explained by the model, as indicated by the very small change in R^2, it accounted for a significant amount of variance in pRF size.

V1	**Model 1**		**Model 2**		**Model 3**	
R	.558		.615		.618	
R^2	.312		.378		.382	
R^2 adjusted	.311		.376		.376	
R^2 change	.312		.066		.004	
Predictors	B	*P*	B	*P*	B	*P*
Eccentricity	.131	**>.001**	.131	**>.001**	.117	**>.001**
VF	-.042	.178	-.042	.157	-.042	.156
Eye	.257	>.001	.660	**>.001**	.377	**.017**
Eccentricity × VF	-.056	.100	-.056	.084	-.056	.083
Eccentricity × Eye	280	>.001	.138	.040	.028	**.040**
VF × Eye	.181	>.001	.046	.201	.046	.200
Eccentricity × VF × Eye	.176	>.001	.053	.132	.033	.380

Table 6-5 Results of stepwise regression analysis for pRF size in V1 (DE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model was statistically significant in predicting pRF size in V1, $F(3,721) = 147.74$, $P < .001$, and accounted for approximately 38% of the variance ($R^2 = .382$, adjusted $R^2 = .379$). The B values in Table 6.5 for the significant coefficients included in the final model show that pRF size in V1 increased with eccentricity by 117° for every 1° eccentricity, and the AME had larger pRF sizes by .377° on average. Eccentricity increased at a greater rate of .028° per 1° eccentricity for the AME relative to the DE, as indicated by the B value for this predictor.

6.2.1.1.2 pRF size in V2

For pRF size in V2, the best model was reached in 5 steps, which can be seen in Figure 6.6 below. Similarly to pRF size in V1, in the first step eccentricity was added as a predictor. The model including only eccentricity accounted for approximately 42% of the total variance in pRF size in V2 for the DE and the AME. In the second step, eye was added as a predictor. The model with eccentricity and eye accounted for 50% of the total variance in pRF size. In the third step, the interaction between eccentricity and eye was added to the model. Up to this point the model was identical to the one described above for pRF size in V1. However, the interactions between eccentricity and eye and visual field and eye were also added in this model in the fourth and fifth steps, respectively. The addition of the final three

predictors did not lead in as a substantial increase in the amount of variance explained by the model, as eccentricity and eye.

V2	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
R	.646		.714		.716		**.**718		.722	
R^2	.417		.509		.512		.515		.522	
R^2adjusted	.417		.508		.510		.512		.518	
R^2change	.417		.092		.003		.003		.007	
Predictors	B	*P*	B	*P*	B	*P*	B	*P*	B	*P*
Eccentricity	.173	**>.001**	.173	**>.001**	.187	**>.001**	.194	**>.001**	.202	**>.001**
VF	-.029	.312	-.029	.271	-.029	.270	.079	.145	.011	.859
Eye	.303	.397	.890	**>.001**	1.186	**>.001**	1.186	**>.001**	.968	**>.001**
Eccentricity × VF	-.059	.059	-.059	.040	-.059	.040	-.014	**.040**	-.030	**>.001**
Eccentricity × Eye	.263	>.001	-126.	.035	-.030	**.035**	.030	**.034**	.030	**.033**
VF × Eye	.204	>.001	.040	.213	.040	.212	.126	.002	.435	**.002**
Eccen × VF × Eye	.147	>.001	-.020	.529	.002	.946	.094	.040	-.057	.454

Table 6-6 Results of stepwise regression analysis for pRF size in V2 (DE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

Overall, pRF size in V2 significantly predicted by the final model, F(5, 721)= 156.26, *P*< .001, with the model accounting for approximately 52% of the variance in V2 estimates of pRF size, (R^2= .522, adjusted R^2= .518). The B values in Table 6.6 show that pRF size in V2 increased by .202° for every 1° eccentricity, with this increase accounting for the largest proportion in variance in the model. pRF size also differed between the DE and the AME, with the AME having larger pRF size on by .968° average, and eccentricity increasing by a greater rate of .03° for the AME. The model also showed that the effect of eccentricity was also dependent on the visual field, with the B coefficient indicating that the temporal visual field showing smaller pRF sizes by a factor of .03° for every 1° eccentricity. The effect of eye was also dependent on the visual field, with the AME showing larger pRF sizes by .435° in the temporal visual field compared to the DE of controls.

6.2.1.1.3 pRF size in V3

For pRF size in V3, the final model was reached in five steps and followed the same progression and included the same predictor variables as the final model

for pRF size in V2. As shown in Table 6.7, in the first step, eccentricity was added, and this model accounted for approximately 44% of the variance in pRF size in V3. In the second stage, eye was added. The model with eccentricity and eye accounted for approximately 48% of the total variance in pRF size in V3. Then the interaction between eccentricity and eye, the interaction between eccentricity and visual field, and finally the interaction between visual field and eye were added, in steps three, four, and five. Similarly to the models for pRF size in V1 and V2, the majority of the variance accounted for by the model was captured in the first two steps by the eccentricity and eye predictors.

V3	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
R	.662		.690		.705		.708		.713	
R^2	.439		.476		.498		.501		.509	
R^2adjusted	.438		.474		.495		.498		.505	
R^2change	.439		.037		.022		.004		.008	
Predictors	B	*P*	B	*P*	B	*P*	B	*P*	B	*P*
Eccen	.272	**>.001**	.272	**>.001**	.329	**>.001**	.341	**>.001**	.354	**>.001**
VF	-.039	.160	-.039	.147	-.039	.138	.057	.301	-.022	.712
Eye	.193	>.001	. 867	**>.001**	2.080	**>.001**	2.080	**>.001**	1.724	**>.001**
Eccen × VF	-.066	.033	-.066	.027	-.066	.024	-.024	**.024**	-.049	**>.001**
Eccen × Eye	.108	>.001	-.338	>.001	-.121	**>.001**	-.121	**>.001**	-.121	**>.001**
VF × Eye	.140	>.001	.041	.216	.041	.206	.135	.001	.713	**.001**
Eccen × VF × Eye	.063	.030	-.055	.094	.109	.891	.109	.019	-.039	.615

Table 6-7 Results of stepwise regression analysis for pRF size in V3 (DE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model with the five predictors overall explained a significant proportion of the variance in V3 pRF size, $F(5, 721)= 148.29$, $P< .0001$, accounting for approximately 51% of the variance ($R^2= .509$, adjusted $R^2= .505$). The B values in Table 6.7 show that in V3, there was a .354° increase in pRF size for every 1° increase in eccentricity. The B value for eye signified that the AME had larger pRF sizes by 1.724° on average compared to the DE. The temporal visual field showed a smaller increase with 1° eccentricity compared to the nasal visual field by .049°, and the AME showed a smaller increase in pRF size by .121° for every 1° eccentricity compared to the DE. The coefficient value for the interaction between visual field

and eye showed that in the temporal visual field the AME had greater pRF sizes by .713° compared to the DE.

6.2.1.2 pRF size: Dominant Eye of Controls and Fellow Fixating Eye

In section 3.3.2.1.2. pRF size in the regions of interest were investigated using stepwise multiple linear regression analyses, to compare the dominant eye (DE) of controls and the fellow fixating eye (FFE) of observers with amblyopia. Here, I present the tables with the full progression of the models outlined in that section.

6.2.1.2.1 pRF size in V1

Table 6.8 shows the full model progression for pRF size in V1 for the DE and FFE. The final model was reached in three steps, with no variables removed. In the first step, eccentricity was added. The model just with eccentricity as a predictor accounted for almost 30% of the variance. In the second step, the interaction of eye with visual field was added. Together with eccentricity, this model explained approximately 32% of the variance. The final model also included visual field as a predictor, that was added in the third step.

V1	**Model 1**		**Model 2**		**Model 3**	
R	.544		.569		.581	
R^2	.295		.324		.337	
R^2 adj.	.294		.322		.334	
R^2 change	.295		.028		.013	
Predictors	B	*P*	B	*P*	B	*P*
Eccentricity	.116	**>.001**	.116	**>.001**	.116	**>.001**
VF	-.003	.936	-.139	>.001	-.325	**>.001**
Eye	.145	>.001	.071	.062	.002	**.**972
Eccentricity × VF	.015	.660	-.099	.011	.077	.274
Eccentricity × Eye	.138	>.001	.058	.141	-.002	.957
VF × Eye	.168	>.001	.461	**>.001**	.673	**>.001**
Eccen. × VF × Eye	-.015	.635	-.038	.218	-.038	.214

Table 6-8 Results of stepwise regression analysis for pRF size in V1 (DE and FFE).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model was significant in predicting pRF size in V1, $F(3, 721)= 121.65$, $P< .001$, and accounted for approximately 34% of the total variance ($R^2= .337$, adjusted $R^2= .334$). As indicated by the B coefficient values in Table 6.8, for every 1° increase in eccentricity there was .116° increase in V1 pRF size. The B coefficient indicated that the nasal visual field had larger pRFs by .325° compared to the temporal. The interaction between eye and visual field in the model indicated that the effect of visual field was different for the two eyes, and that the FFE had larger pRFs by .673° in the temporal visual field relative to the DE.

6.2.1.2.2 pRF size in V2

For pRF size in V2, the final model was reached in five steps, with the removal of one predictor variable. In the first step, eccentricity was added. The model in the first step accounted for approximately 45% of the variance. In the second step, eye was added as a predictor. This model explained approximately 2.5% more variance than the model including only eccentricity. In the third step, the interaction between eccentricity and eye was added, that led to a small improvement in the variance explained by the model, as shown by the change in R^2. In the fourth step, visual field was included. With the inclusion of visual field the predictor for eye was no longer significantly, likely due to shared variance. In fact,

the effect of visual field was primarily driven by the DE of controls that had larger pRFs in the nasal visual field, and to a much lesser degree by the FFE, that only had larger pRFs in parafoveal eccentricities in the nasal visual field (see Figure 3.9 in Chapter 3). The FFE showed larger pRFs for the temporal visual field, as captured by the interaction between visual field and eye, and indicated by the coefficient value in Table 6.9. As such, the combination of these two predictors, visual field and the interaction between visual field and eye, were likely sufficient to account for differences in variation in pRF size between the two eyes, resulting in eye not being a significant predictor in the final model.

V2

	Model 1		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
R	.670		.690		.695		.699		.698	
R^2	.449		.476		.483		.488		.487	
R^2adj.	.448		.475		.481.		.485		.484	
R^2change	.449		.027		.007		.005		-.002	
Predictors	B	*P*	B	*P*	B	*P*	B	*P*	B	*P*
Eccentricity	.180	**>.001**	.180	**>.001**	180	**>.001**	.180	**>.001**	.180	**>.001**
VF	.006	.820	.006	.816	-.097	.009	-.285	**.009**	-.362	**>.001**
Eye	.164	>.001	.484	**>.001**	.306	**>.001**	.163	**.143**	.055	.143
Eccentricity × VF	.026	.396	.026	384	-.047	.211	.086	.162	.086	.162
Eccentricity × Eye	.147	>.001	-.048	.440	-.048	.437	-.048	.436	.026	.496
VF × Eye	.164	>.001	.103	.002	.356	**>.001**	.641	**>.001**	.804	**>.001**
Eccen. × VF × Eye	.056	.045	.043	.115	.034	.209	.034	.208	.034	.208

Table 6-9 Results of stepwise regression analysis for pRF size in V2 (DE and FFE).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model with the four predictors was significant in predicting pRF size in V2, $F(3, 721)= 226.79$, $P< .001$, and accounted for approximately 49% of the total variance ($R^2= .487$, adjusted $R^2= .484$). As the B coefficients in Table 6.9 show, for every 1° increase in eccentricity there was an increase of .180° in pRF size in V2. The B coefficient for visual field shows that the nasal visual field had larger pRFs by .362° compared to the temporal visual field. In fact, similarly to V1, the inclusion of the interaction between eye and visual field in the model demonstrated that the effect of visual field was different for the DE and the FFE.

1.1.1.1.1 pRF size in V3

For pRF size in V3, the best model was reached in four steps, with no variables removed at any step. In the first step, eccentricity was added. This model accounted for approximately 51% of the variance in pRF size in V3 across the DE and the FFE. At the second step, the interaction between visual field and eye was added as a predictor. This increased the total variance explained by the model to 53%. At the third step, visual field was added, that led to only a small improvement in the total variance explained by the model of less than 1%. In the fourth and final step, the interaction between eccentricity and eye was added as a significant predictor, that also only improved the total variance explained by the model by less than 1%.

V3	**Model 1**		**Model 2**		**Model 3**		**Model 4**	
R	.718		.729		.733		.737	
R^2	.515		.531		.538		.543	
R^2, adj.	.514		.530		.536		.541	
R^2change	.515		.016		.007		.006	
Predictors	B	*P*	B	*P*	B	*P*	B	*P*
Eccentricity	.293	**>.001**	.293	**>.001**	.293	**>.001**	.311	**>.001**
VF	.003	.919	-.098	.001	-.438	**.001**	-.618	**>.001**
Eye	.084	.001	.016	.622	-.044	.224	.110	.059
Eccentricity × VF	.023	.416	-.059	.069	.090	.123	.090	.121
Eccentricity × Eye	.037	.197	-.047	.151	-.106	.003	-.038	**.003**
VF × Eye	.126	>.001	.663	**>.001**	.950	**>.001**	1.331	**>.001**
Eccen. × VF × Eye	.123	>.001	.037	.500	.037	.498	.110	.059

Table 6-10 Results of stepwise regression analysis for pRF size in V3 (DE and FFE).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model was significant in predicting pRF size in V3, $F(4, 721)= 213.22$, $P<.001$, and explained approximately 54% of the variance ($R^2= .543$, adjusted $R^2= .541$). The inclusion of eccentricity as a predictor in the model indicated that pRF size in V3 significantly increased with eccentricity. The B coefficient value for eccentricity in Table 6.10 indicates that for every 1° increase in eccentricity there was an increase of .311° in pRF size in V3. Visual field was also a

significant predictor in the final model, with the B coefficient showing that nasal visual field having larger pRFs than the temporal visual field by .618°. Interestingly, the B coefficient for the interaction between eccentricity and eye shows that it was the DE that had larger pRFs by .038° for every 1° increase in eccentricity, compared to the FFE. However, the coefficient for the significant predictor of the interaction between visual field and eye indicates that the FFE has larger pRFs by 1.331° compared to the DE.

6.2.2 Comparison of pRF Size Between the Eyes of Observers with Amblyopia

In Chapter 3 my aim was to establish whether there were differences in population receptive field (pRF) in amblyopia relative to unaffected vision. In order to achieve this, I compared pRF size in V1, V2, and V3 between the dominant eye of controls (DE) and the amblyopic eye (AME) or the fellow fixating eye (FFE) of observers with amblyopia. I found that pRF size was overall greater for the AME relative to the DE across the visual field for all regions of interest. In V2 and V3, the enlargement in pRF size relative to the DE was especially large in the temporal visual field. There was no overall enlargement of pRF size for the FFE relative to the DE. However, there was a localised enlargement specific to the temporal visual field in all regions of interest. Although these results suggested that amblyopia affects pRF size in both eyes relative to unaffected vision, they also indicated the magnitude of the effect was different between the AME and the FFE. In order to investigate these interocular differences in pRF size for observers with amblyopia, in this section I compared pRF size between the AME and the FFE.

Figure 6.4 shows pRF size in areas V1, V2, and V3 for the AME and the FFE across eccentricities in the nasal and temporal visual fields. It is clear that pRF size was larger for the AME compared to the FFE in all regions of interest. However, this difference was not uniform across the visual field or of the same magnitude for all regions of interest. In V1, pRF size for the AME is overall greater than pRF size for the FFE. In V2, pRF size for the AME is also overall greater, but this elevation is especially large for eccentricities up to 10° in the temporal visual field. This elevation in pRF size for the AME relative to the FFE in the temporal visual field is maintained in V3, but there is also a more pronounced elevation for the AME at higher eccentricities in the nasal visual field.

In order to investigate whether the AME has significantly larger pRFs compared to the FFE, stepwise regression analyses were conducted comparing the FFE to the

AME in V1, V2, and V3. As in the analyses outlined in section 2.1 of Chapter 3, the following predictors were included: eccentricity (1-19), visual field (nasal and temporal, coded as 0 and 1, respectively), eye (FFE and AME, coded as 0 and 1 respectively), and the interactions between them. Observer was also included to account for potential between-observer variability. Results from the analyses are outlined below.

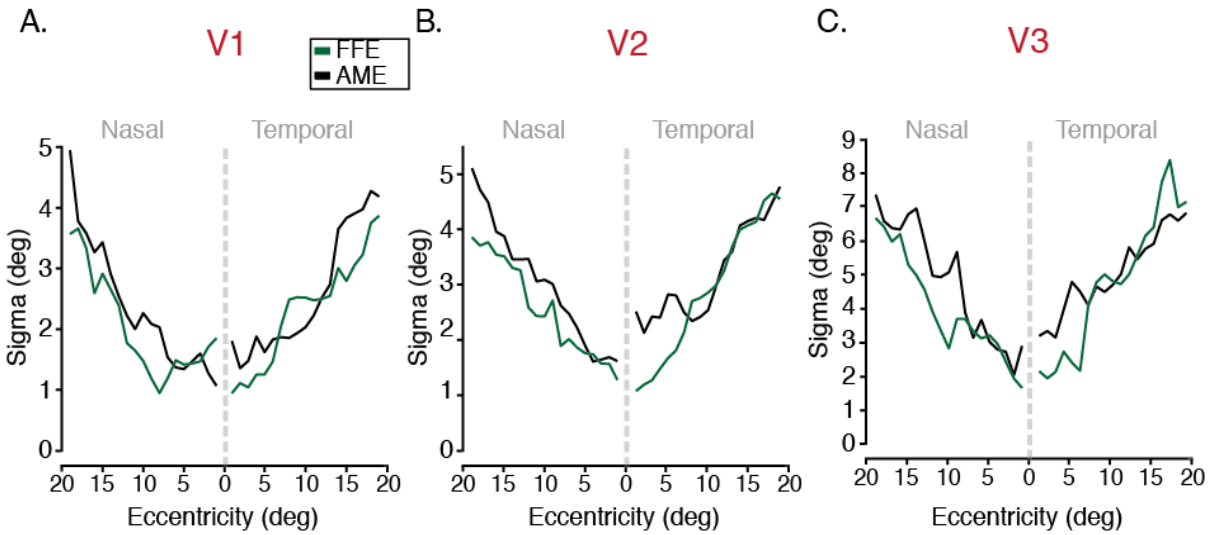

Figure 6-4 pRF size across eccentricity in the nasal and temporal visual fields for V1, V2 and V3

A. Average pRF size (sigma) in V1 across eccentricities (1°-19°) in the nasal and temporal visual fields. Solid lines indicate the mean pRF size in degrees of visual angle for the fellow-fixating eye (FFE) and the amblyopic eye (AME) of observers with amblyopia (N=9). The dashed grey line separates the nasal and temporal visual fields.
B. Average pRF size (sigma) in V2. Plotted with the same conventions as in A.
C. Average pRF size (sigma) in V3. Plotted with the same conventions as in A.

6.2.2.1 pRF size in V1

The best fitting model from the stepwise regression analysis was reached in two steps with no variables removed, and included eccentricity and the interaction between eccentricity and visual field as predictors. Table 6.11 shows the progression of the model, with the addition of the predictor at each step.

In the first step, eccentricity was included as the predictor in the model. This model accounted for approximately 31% of the total variance in pRF size in V1. In the second and final step, the interaction between eccentricity and visual was added as a significant predictor to the model.

V1	**Model 1**		**Model 2**	
R	.557		.581	
R^2	.310		.338	
R^2adj.	.309		.336	
R^2change	.310		.027	
Predictors	B	*P*	B	*P*
Eccentricity	.131	**>.001**	.099	**>.001**
VF	.149	>.001	.063	.157
Eye	.121	>.001	.050	.166
Eccentricity × VF	.214	>.001	.043	**>.001**
Eccentricity × Eye	.148	>.001	.063	.138
VF × Eye	.126	>.001	.064	.070
Eccen. × VF × Eye	.144	>.001	.069	.082

Table 6-11 Model progression for stepwise regression on pRF size in V1 (FFE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model was significant, $F(2, 683) = 173.55$, $P<0.0001$, and accounted for approximately 34% of the variance in pRF size in V1 ($R^2= 0.338$ adjusted $R^2= 0.336$). Table 6.12 shows the regression coefficients of the predictors in the final model together with their correlations with pRF size (Pearson R), and their squared semi-partial correlations (sr^2).

Model	B	SE B	β	Pearson R	sr^2
Intercept***	.990	.085			
Eccentricity ***	.099	.010	.420	.557	.104
Eccen. × VF***	.043	.008	.214	.483	.027

Table 6-12 Stepwise regression results for variables significantly predicting pRF size in V1 (FFE & AME)

The dependent variable was pRF size in V1 in the FFE and AME. R^2=0.338, adjusted R^2=0.336. *sr²* is the squared semi-partial correlation (***$P<0.001$).

For eccentricity, the unstandardized coefficient indicates that for every 1° increase in eccentricity there was a .099° increase in pRF size in V1. The interaction between eccentricity and visual field shows that for every 1° eccentricity there is an increase of .043° in V1 pRF size in the temporal visual field. This is especially clear for the FFE in Figure 6.4, where eccentricities beyond 6° in the temporal visual field appear to have larger pRFs than the corresponding ones in the nasal visual field.

Overall in the model eccentricity explained the largest percentage of unique variance, accounting for 10.4% in pRF size, whereas the interaction between eccentricity and visual field only accounted for 2.7%.

In Chapter 3 section 3.3.2.1, when comparing pRF size between the DE of controls and the AME, and the DE and the FFE, the results showed a naso-temporal difference in the deficit in both eyes of observers with amblyopia relative to the DE: the temporal visual field had larger pRFs than the nasal. This difference relative to unaffected vision was found in V2 and V3 for the AME, and across regions of interest for the FFE. This difference is also captured in the results of this analysis, as the rate by which pRF size increased with eccentricity was greater for the temporal visual field compared to the nasal visual field in both the AME and the FFE. Overall, the results from this analysis indicate that the AME and the FFE did not significantly differ in pRF size in V1.

6.2.2.2 pRF size in V2

For V2, the best fitting model for pRF size was reached in two steps, with no variables removed, and included eccentricity and visual field as predictors. Table 6.13 shows the two steps in the model progression, with the addition of one predictor at each step.

V2	**Model 1**		**Model 2**	
R	.595		.615	
R^2	.354		.378	
R^2 adj.	.353		.376	
R^2 change	.354		.024	
Predictors	B	*P*	B	*P*
Eccentricity	.164	**>.001**	.164	**>.001**
VF	.155	>.001	.330	**>.001**
Eye	.137	>.001	.056	.191
Eccentricity × VF	.194	.001	.102	.071
Eccentricity × Eye	.114	>.001	.012	.771
VF × Eye	.138	>.001	.024	.683
Eccen. × VF × Eye	.102	.002	-.054	.277

Table 6-13 Model progression for stepwise regression on pRF size in V2 (FFE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

In the first step, eccentricity was added. The model including just eccentricity as a predictor explained approximately 35% of the total variance in pRF size in V2. In the second step visual field was included, that increased the total variance in pRF size accounted for by the model by more than 2%.

The final model with the two predictors was significant, F(2, 683) = 206.97, $P<.001$, and accounted for approximately 38% of the variance in pRF size in V1 (R^2= .378, adjusted R^2= .376). Table 6.14 below shows the information of the predictors included in the final model.

Model	B	SE B	β	Pearson R	sr^2
Intercept***	.970	.115			
Eccen***	.164	.008	.595	.595	.354
VF***	.330	.065	.155	.155	.024

Table 6-14 Stepwise regression results for variables significantly predicting pRF size in V2 (FFE vs AME)

The dependent variable was pRF size in V2 in the FFE and AME. R^2=0.378, adjusted R^2=0.376. sr^2 is the squared semi-partial correlation (***P<0.001).

pRF in V2 size significantly increased with eccentricity. The B value for eccentricity presented in Table 6.14 indicates that for every 1° increase in eccentricity there was a .164° increase in pRF size in V2. The inclusion of visual field in the model suggests that pRF size in V2 differs between the nasal and temporal visual fields for the eyes of observers with amblyopia, with the temporal visual field showing larger pRF size on average by .330° according to the B value. This difference is especially clear in Figure 6.4 for the AME that shows larger pRFs especially for eccentricities up to 8° in the temporal visual field, whereas in the FFE this naso-temporal difference becomes apparent beyond 8°. As in the previous models described above, eccentricity had the highest unique variance accounted, explaining 35.4% of the variance in V2 pRF size, whereas visual field only accounted for 2.4%. These results suggest that the AME does not significantly differ from the FFE in pRF size in V2, similar to the results comparing pRF size in V1 between the two eyes. However, a significant difference between the nasal and temporal visual fields, with both eyes showing larger pRFs responding to the temporal visual field. This enlargement in pRF size for the temporal visual field is of a greater magnitude than the one observed in V1 above.

6.2.2.3 pRF size in V3

For V3, the best model from the stepwise regression was reached after two steps, without the exclusion of any predictors. Similarly to the previous analyses reported above for V1 and V2, the final model included eccentricity and visual field as predictors. Table 6.15 below shows the progression of the model over the two steps.

V3	**Model 1**		**Model 2**	
R	.589		.607	
R^2	.347		.369	
R^2 adjusted	.346		.367	
R^2 change	.347		.021	
Predictors	B	*P*	B	*P*
Eccentricity	.232	**>.001**	.232	**>.001**
VF	.147	>.001	.447	**>.001**
Eye	.113	>.001	.019	.652
Eccentricity × VF	.191	>.001	.111	.050
Eccentricity × Eye	.077	0.25	-.039	.369
VF × Eye	.177	>.001	-.031	.597
Eccen. × VF × Eye	.024	.442	-.009	.767

Table 6-15 Model progression for stepwise regression on pRF size in V3 (FFE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model significantly predicted pRF size in V3, F(2,683)= 198.72, *P*< .0001, and accounted for approximately 37% of the variance. Information on the predictors that were included in the final model is shown below in Figure 6.16

Model	B	SE B	β	Pearson R	sr^2
Intercept***	1.900	.165			
Eccentricity***	.232	.012	.589	.589	.347
VF***	.447	.093	.147	.147	.022

Table 6-16 Stepwise regression results for variables significantly predicting pRF size in V3 (FFE and AME)

The dependent variable was pRF size in V3 in the FFE and AME. R^2=0.369, adjusted R^2=0.367. *sr²* is the squared semi-partial correlation (***P<0.001).

The model showed that pRF size increases with eccentricity in V3. The B value for eccentricity presented in Table 6.16 indicates that for every 1° increase in eccentricity there was a .164° increase in pRF size in V3. Visual field was also a significant predictor (coded as 0 for the nasal and 1 for the temporal visual field), showing that pRF size in V3 differs between the nasal and temporal visual fields, with the temporal visual field showing larger pRF size on average by .447°. Similarly to V2, this difference can be seen in Figure 6.4 for the AME up to ~8° eccentricity, whereas for the FFE the temporal visual field has larger pRFs than the nasal visual field from ~10° eccentricity. As in the previous models described above, eccentricity had the highest unique variance accounted, explaining 34.7% of the variance in V2 pRF size, whereas visual field only accounted for 2.2%. The results from this analysis suggest that pRF size in V3 is not significantly different between the two eyes of observers with amblyopia. However, the results are in line with the results of analyses in section 3.3.2.1 of Chapter 3, that demonstrate that both the AME and the FFE show enlarged pRFs responding to the temporal visual field compared to unaffected vision.

6.2.2.4 Summary

In this section, differences in pRF size in V1, V2, and V3 between the two eyes of observers with amblyopia were investigated. The results of the analyses showed that in all regions of interest, pRF size increased with eccentricity. In V1, pRF size increased at a greater rate with eccentricity for the temporal visual field. In V2 and V3, there was a naso-temporal asymmetry in average pRF size, with the temporal visual field having larger pRF size than the nasal visual field. The analyses thus point to greater pRF size for both eyes in the temporal visual field relative to nasal, and show that the differences in pRF size between the two eyes of observers with amblyopia in Figure 6.4 are not significant.

These results are in line with the findings from Chapter 3 (section 3.3.2.1) investigating differences in pRF size between the dominant eye of controls (DE) and the AME, and the DE and the FFE. The analyses comparing the DE to the AME pointed to a significant elevation in pRF size in V2 and V3 for the AME in the temporal visual field. For the FFE, the analyses showed that this elevation in pRF size in the temporal visual field relative to the DE was clear across regions of interest. As such, the greater pRF size found here for the temporal visual filed in

both eyes of observers with amblyopia is consistent with the naso-temporal difference in the elevation in pRF size relative to the DE in Chapter 3.

The results reported here also indicate that the two eyes of observers with amblyopia do not differ substantially in pRF size. When compared the DE in Chapter 3, the AME showed larger pRF size on average in all regions of interest. In contrast, the FFE did not show this overall elevation in pRF size relative to the DE. Taken together with the results reported in this supplementary section, the findings point to an elevation in pRF size for the FFE, but not to the same extent as the AME. As discussed in the Discussion in Chapter 3 (section 3.4), this suggests that although the FFE is clinically considered “normal”, as optotype acuity is similar to that of observers with unaffected vision, amblyopia might affect both eyes.

6.2.3 Additional Eye Movement Analyses

In Chapter 3, the results in section 3.3.2.2 show that there were no significant differences in monocular fixation variability during the fMRI scans between the eyes of observers with amblyopia and controls. However, there was a small difference in average fixation variability between the AME and the DE. Poor fixation stability for the AME could have affected estimates of pRF size and consequently the results of the analyses reported in the Results section 3.3.2.1. Fixation variability was computed as eye positions deviating from the fixation cross. In Chapter 3 (Figure 3.10), I show that average fixation variability was approximately 2° for the AME. Due to this magnitude of positional variance in fixation, the pRFs that would have been most affected are the ones at parafoveal eccentricities (i.e. the ones extending approximately 2° from fixation). Figures 6.5 shows average fixation variability for the AME of all observers in the amblyopia group plotted against pRF size at parafoveal eccentricities in the regions of interest.

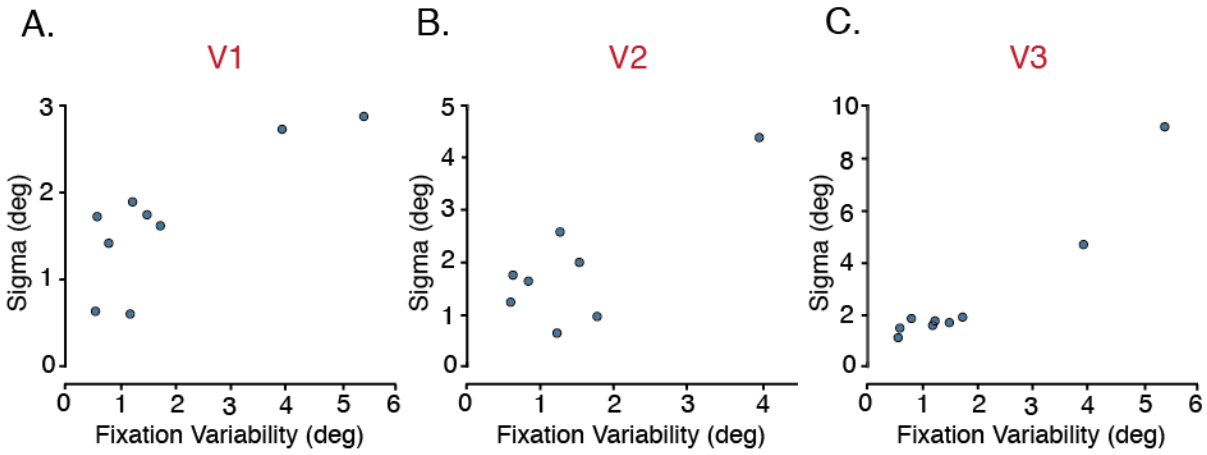

Figure 6-5 Average fixation variability for the AME against parafoveal pRF size (sigma) in the regions of interest

A. Individual fixation variability (averaged across the X- and Y-axis) for the AME (N=9), plotted against average pRF size in V1 for 1-3° eccentricity. Values are in degrees of visual angle.
B. Individual fixation variability (averaged across the X- and Y-axis) for the AME (N=8), plotted against average pRF size in V2 for 1-3° eccentricity. Note that one observer is not included, as it was not possible to estimate pRF size due to insufficient voxels responding to 1-3° eccentricity. Values are in degrees of visual angle.
C. Individual fixation variability (averaged across the X- and Y-axis) for the AME (N=9), plotted against average pRF size in V3 for 1-3° eccentricity. Values are in degrees of visual angle.

In order to investigate whether increased fixation variability could be underlying the increased pRF size for parafoveal eccentricities, Pearson correlations were conducted between an average estimate of horizontal and vertical fixation variability and the average pRF size estimates for eccentricities 1°-3° for the AME (averaged for visual field). The analyses revealed that there were significant correlations between average fixation variability and pRF size for parafoveal eccentricities in V1 ($r(9)= .82$, $n= 9$, $P= .007$), V2 ($r(8)= 0.78$, $n= 8$, $P= .023$), and V3 ($r(9)= .96$, $n= 9$, $P<.001$).

However, as mentioned in section 3.3.2.2. of Chapter 3, the elevation in the mean fixation variability for the AME was driven primarily by two observers with amblyopia, who had greater fixation variability than the rest of the group with amblyopia and the controls. From Figures 6.5 A-C, it is clear that two observers with amblyopia deviate from the group. The increased fixation variability for these observers is likely driving the significant correlations reported above. Note that for one of these observers, it was not possible to calculate average pRF size in V2 due to insufficient number of voxels responding to 1-3° eccentricity. Figure 6.5 shows that these two observers also had the greatest pRF size for the parafoveal eccentricities. Removing them from the correlation analyses changes the outcome the results in V1 ($r= .35$, $n=7$, $P= .441$), V2 ($r= -.63$, $n= 7$, $P= .894$), and V3 ($r= .69$, $n=$

7, *P*= .089). As the increased fixation variability could be contributing to the increased pRF size in these observers, this could have biased the results reported in section 3.3.2.1.1, resulting in the significant differences in pRF size between the DE and the AME. Therefore, the observers with the highest AME fixation variability were removed from the regression analyses comparing pRF size estimates between the DE and the AME, in order to investigate whether the outcome changes. Stepwise regression analyses without the two observers are reported below.

6.2.3.1 pRF size in V1

For pRF size in V1, in the original analyses with the full sample of observers with amblyopia, the final model included eccentricity, eye, and the interaction between eccentricity and eye. The final model from the stepwise regression analysis with the reduced amblyopic sample contained the following predictors: eccentricity (1°-19°), visual field the interaction between eccentricity and eye, and the interaction between eccentricity and eye.

The progression of the model on the data from the reduced sample of observers with amblyopia with the addition of one variable at each step can be seen in Table 6.17 below. By removing the two observers with the highest fixation variability, the new model was reached in four steps, with no variables removed. In the first step eccentricity was added, and this model accounted for approximately 36% of the total variance in pRF size in V1. In the second step, the interaction between eccentricity and eye was added, and the model with these two predictors accounted for a total variance of approximately 40% in pRF size. In the third and fourth steps, the interaction between visual field and eye were added, and the visual field as predictors.

The exclusion of the two observers with amblyopia with the highest fixation variability resulted in the exclusion of eye as a predictor in the regression model, and the inclusion of visual field and the interaction between visual field and eye, which were not significant predictors in the original model.

	Model 1		Model 2		Model 3		Model 4	
R	.596		.631		.634		.641	
R^2	.356		.398		.402		.411	
R^2 adjusted	.355		.396		.399		.408	
R^2 change	.356		.042		.004		.010	
Predictors	B	*P*	B	*P*	B	*P*	B	*P*
Eccentricity	.132	**>.001**	.114	**>.001**	.117	**>.001**	.121	**>.001**
VF	-.041	.190	-.041	.176	-.123	.001	-.298	**.001**
Eye	.189	>.001	.041	.641	-.004	.995	-.070	.319
Eccentricity × VF	-.053	.132	-.053	.119	-.123	.002	-.054	.438
Eccentricity × Eye	.221	>.001	.044	**>.001**	.036	**>.001**	.026	**.002**
VF × Eye	.162	>.001	.073	.045	.217	**.045**	.480	**>.001**
Eccen. × VF × Eye	.169	>.001	.064	.105	-.004	.957	.030	.666

Table 6-17 Results of stepwise regression analysis for pRF size in V1 (DE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model was statistically significant in predicting pRF size in V1, $F(4,645) = 111.99$, $P < .00001$, and accounted for approximately 41% of the variance ($R^2 = 0.411$, adjusted $R^2 = 0.408$). Table 6.18 shows the regression coefficients of the predictors in the final model together with their correlations with pRF size (Pearson R), and their squared semi-partial correlations (sr^2). For the interpretation of the regression coefficients, note that visual field was coded as 0 for nasal and 1 for temporal, and eye was coded as 0 for the DE and 1 for the AME.

Model	B	SE B	β	Pearson R	sr^2
Intercept***	.694	.085			
Eccentricity***	.121	.008	.548	.596	.238
VF**	-.298	.092	-.123	-.041	.010
Ecc × Eye**	.026	.008	.129	.413	.009
VF × Eye***	.480	.134	.160	162	.012

Table 6-18 Stepwise regression results for variables significantly predicting pRF size in V1 (DE and AME) .

The dependent variable was pRF size in V1 in the DE and AME. R^2=0.411, adjusted R^2=0.408. sr^2 is the squared semi-partial correlation (** *P*<.01, ***P<0.001). Note that for this model the two observers with amblyopia with the highest fixation variability were removed from the analysis.

Similarly to the analyses with the full sample of observers with amblyopia, pRF size in V1 was predicted by eccentricity, with pRF size increasing by .121° for every 1° eccentricity. Excluding the two observers revealed visual field as a significant predictor, with the temporal visual field having smaller pRF sizes by .298° on average. Contrary to the original analyses in which eye was a significant predictor in pRF size in V2, in this model ran on the reduced sample eye was not included as a significant predictor. Differences between the DE and the AME however still emerged, as indicated by the significant interactions of eccentricity and visual field with eye. The B value for this interaction indicated that for the AME pRF size in V1 increased by an additional .026° for every 1° increase in eccentricity compared to the DE. The interaction between visual field and eye showed that in the temporal visual field the AME showed pRF size increased by .480° compared to the DE, demonstrating that the significant naso-temporal asymmetry was not due to the excluded observers. The greatest unique variance was explained by eccentricity that accounted for 23.8%, followed by the interaction between visual field and eye that accounted for 1.2%, whereas visual field and the interaction between visual field and eye accounted for less than 2% in total. Overall these results suggest that, although eye was not a significant predictor, the final regression model still revealed differences between the DE and the AME, as the effects of eccentricity and visual field were eye-dependent. We can thus conclude that removing the two observers with amblyopia with high fixation variability does not substantially alter the results of the analyses on pRF size in V1 between the DE and the AME.

6.2.3.2 pRF size in V2

For pRF size in V2, in the original analyses with the full sample of observers with amblyopia, the final model had the following five predictors: eccentricity, eye, the interaction between eccentricity and eye, the interaction between eccentricity and visual field, and the interaction between visual field and eye.

The progression of the model with reduced sample of observers with amblyopia can be seen in Table 6.19. In the first step, eccentricity was added, and at this stage the model accounted for over 50% of the variance in pRF size in V2. At the second stage, eye was added, and the total variance accounted by the model increased to 57%. In third step the interaction between eccentricity and visual field was added, that only slightly improved the model, and in the fourth the interaction

between visual field and eye, that also led to a small improvement of the model as indicated by the R^2 change.

The removal of the two observers with amblyopia with the highest fixation variability resulted in the removal of interaction between eccentricity and eye as a predictor.

	Model 1		**Model 2**		**Model 3**		**Model 4**	
R	.713		.756		.759		.764	
R^2	.508		.572		.576		.584	
R^2 adjusted	.507		.570		.574		.581	
R^2 change	.508		.064		.005		.008	
Predictors	B	*P*	B	*P*	B	*P*	B	*P*
Eccentricity	.186	**>.001**	.195	**>.001**	.202	**>.001**	.202	**>.001**
VF	-.037	.179	-.037	.150	.099	.065	.036	.524
Eye	.252	>.001	.734	**>.001**	.734	**>.001**	.500	**>.001**
Eccentricity × VF	-.076	.013	-.076	.013	-.017	**.008**	-.032	**>.001**
Eccentricity × Eye	.240	>.001	.004	.939	.004	.939	.004	.938
VF × Eye	.180	>.001	.042	.196	.133	.001	.469	**.001**
Eccen. × VF × Eye	.132	>.001	-.009	.775	.064	.100	-.071	.215

Table 6-19 Results of stepwise regression analysis for pRF size in V2 (DE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The final model was statistically significant in predicting pRF size in V2, $F(4,645) = 224.75$, $P < .0001$, and accounted for approximately 58% of the variance ($R^2 = 0.584$, adjusted $R^2 = 0.581$). Table 6.20 below presents information for the significant predictors included in the final regression model.

Model	B	SE B	β	Pearson R	sr^2
Intercept ***	.372	.082			
Eccentricity***	.202	.008	.774	.713	.448
Eye***	.500	.101	.172	.252	.016
Eccen × VF ***	-.032	.008	-.141	.247	.011
VF × Eye**	.469	.138	.133	.180	.008

Table 6-20 Stepwise regression results for the variables significantly predicting pRF size in V2 (DE and AME)

The dependent variable was pRF size in V2 in the DE and AME. R^2=0.584, adjusted R^2=0.581. sr^2 is the squared semi-partial correlation (** P<0.01, ***P<0.001). Note that for this model the two observers with amblyopia with the highest fixation variability were removed from the analysis.

pRF size in V2 was predicted by eccentricity, with pRF size increasing by .202° for every 1° eccentricity. Eye was also a significant predictor, with the AME showing increased pRF size in V2 by .500° compared to the DE. The B value is smaller here than in the original regression model. This reduction in the B coefficient for eye from the analyses with the full sample reported in Chapter 3 can be attributed to the removal of the two observers with the greatest fixation variability. These observers also had large pRF size estimates in V2, leading to increased differences between the DE and AME, as indicated by the B coefficient. Similarly to the analysis with the full amblyopic sample, the effect of eccentricity was dependent on the visual field, with the temporal visual field showing smaller pRF sizes by a factor of .032° for every 1° eccentricity compared to the nasal visual field. The effect of eye was also dependent on the visual field, with the AME showing larger pRF sizes by .469° in the temporal visual field compared to the DE of controls. The B value for this predictor was slightly increased in this model with the adjusted sample, indicating that removing the two observers lead to a greater naso-temporal difference for the AME. Eccentricity explained the largest unique variance, accounting for almost 45% of the variance in pRF size in V2, whereas eye accounted for 1.6% and the remaining two predictors for less than 2% in total. The results of this analysis indicate that the removal of the two observers with the highest fixation variability has little effect on the analyses on pRF size in V2 between the DE and the AME.

6.2.3.3 pRF size in V3

For V3, the final model with the adjusted sample size for the amblyopia group contained the same predictors as the model with the full amblyopic sample

reported in Chapter 3: eccentricity, eye, the interaction between eccentricity and eye, the interaction between eccentricity and visual field, and the interaction between visual field and eye. The progression for this model can be seen in Table 6.21.

At the first step, eccentricity was added that accounted for more than 53% of the total variance in pRF size in V3. At the second stage, eye was added as a predictor, and the total variance explained by the model increased to more than 56%. In the remaining three steps, the interaction between visual field and eye, the interaction between eccentricity and eye, and the interaction between eccentricity and visual field were added. These increased the total variance explained by the model, but not to the same magnitude as the first two predictors added.

	Model 1		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
R	.733		.752		.755		.761		.764	
R^2	.538		.566		.571		.579		.583	
R^2 adjusted	.537		.564		.569		.577		.580	
R^2 change	.538		.028		.005		.009		.004	
Predictors	B	*P*	B	*P*	B	*P*	B	*P*	B	*P*
Eccentricity	.306	**>.001**	.306	**>.001**	.306	**>.001**	.328	**>.001**	.350	**>.001**
VF	-.024	.380	-.024	.365	-.116	.001	-.051	.375	-.051	.373
Eye	.167	>.001	.773	**.001**	.524	**.001**	.298	**.066**	.836	**.002**
Ecc. × VF	-.042	.158	-.042	.145	-.125	>.001	-.045	**>.001**	-.045	**>.001**
Ecc. × Eye	.125	>.001	-.143	.014	-.143	.014	-.143	.013	-.054	**.013**
VF × Eye	.157	>.001	.088	.007	.498	**.007**	.950	**>.001**	.950	**>.001**
Ecc. × VF × Eye	.112	>.001	.028	.374	-.130	.019	-.078	.178	0.22	.763

Table 6-21 Results of stepwise regression analysis for pRF size in V3 (DE and AME).

B indicates the unstandardized regression coefficient and *P* the significance of the predictor in that model. Predictors included in each step have their *P*-value in bold.

The model was reached in 5 steps and accounted for approximately 58% of the variance in pRF size in V2 (R^2= 0.583, adjusted R^2= 0.580). Table 6.22 below presents information for the predictors included in the final regression model.

Model	B	SE B	β	Pearson R	sr^2
Intercept***	.711	.158			
Eccentricity***	.350	.015	.841	.773	.347
Eye**	.836	.268	-.383	.193	.006
Eccen × VF ***	-.045	.013	-.139	.233	.009
Eccen × Eye*	-.054	.022	-.338	.364	.004
VF × Eye***	.950	.214	.135	.140	.012

Table 6-22 Stepwise regression results for the variables significantly predicting pRF size in V3 (DE and AME)

The dependent variable was pRF size in V3 in the AME and DE. R^2=0.583, adjusted R^2=0.580. sr^2 is the squared semi-partial correlation (* P<.05, ** P<0.01, ***P<0.001). Note that for this model the two observers with amblyopia with the highest fixation variability were removed from the analysis.

Eccentricity was a significant predictor of pRF size in V3, with a .350° increase in pRF size for every 1° increase in eccentricity. Eye was significant predictor, but removing the two observers with the highest fixation variability reduced the B estimate from 1.724° to .836°, indicating that the difference between the DE and the AME was smaller with the reduced amblyopic sample. The interaction between eccentricity and visual field indicated that the temporal visual field showed a smaller increase with 1° eccentricity compared to the nasal visual field by .045°. The significant interaction between eccentricity and eye indicated that the AME showed a smaller increase in pRF size by .054° for every 1° eccentricity compared to the DE. Finally, the interaction between visual field and eye suggested that in the temporal visual field the AME had greater pRF sizes by .950°. In the case of this predictor the B value increased with the removal of the two amblyopic observers, leading to a greater naso-temporal difference in V3 pRF size for the AME. These results indicate that the exclusion of the two observers with the highest fixation variability had no effect on the results of the analyses on pRF size in V3 between the DE and the AME.

6.2.3.4 Summary

In this section, I explored whether the differences in pRF size between the DE and the AME found in Chapter 3 (section 3.3.2.1.1) can be explained by differences in fixation variability. I found that the small difference in average fixation variability observed in Chapter 3 (section 3.3.2.2) that was not statistically significant

was driven by two observers that had disproportionally poor fixation stability relative to the rest of the group with amblyopia and the controls. These observers also had large pRF sizes, and thus I excluded them from the regression analyses comparing pRF size in the regions of interest to examine whether the outcomes change.

The outcomes of the regression analyses with the reduced sample of observers with amblyopia were very similar to those reported in Chapter 3 with the full amblyopic sample. For V1 pRF size in V1, the analyses with the reduced sample did not reveal an overall difference in pRF size between the AME and the DE, as the analyses with the full sample. However, it did show differences between the two eyes, indicative of elevations in pRF size for the AME. Particularly, pRF size increased with eccentricity at a greater rate for the AME. Additionally, the characteristic naso-temporal difference in the elevation for the amblyopic eye was also clear, with the temporal visual field having increased pRF size relative to the DE.

For pRF size in V2, the analysis with the reduced sample were almost identical to the analyses with the full sample of observers with amblyopia. The analyses showed that pRF size was overall greater for the AME relative to the DE, and that the temporal visual field was more affected in the AME. The analysis with the full sample had shown that pRF in V2 size increases at a greater rate with eccentricity for the AME relative to the DE, but with the exclusion of the two observers with high fixation variability, this effect was not found.

For pRF size in V3, the exclusion of the two observers with amblyopia with high fixation variability led to identical results as the original analyses with the full sample. In line with the original analysis, the supplemental analysis showed that here was an overall difference in pRF size between the AME and the DE, and the rate at which pRF size increased with eccentricity was greater for the AME relative to the DE. The naso-temporal difference in the elevation in pRF size for the AME was also maintained, with the temporal visual field showing greater pRF sizes than the nasal for the AME relative to the DE.

Taken together, the results from these supplementary analyses demonstrate that the increased pRF size found for the AME relative to the DE in Chapter 3 cannot be explained by the increased fixation variability shown by a minority of observers with amblyopia.

References

Aaen-Stockdale, C., & Hess, R. F. (2008). The amblyopic deficit for global motion is spatial scale invariant. *Vision Research, 48*, 1965-1971.

Aaen-Stockdale, C., Ledgeway, T., & Hess, R. F. (2007). Second-order optic flow deficits in amblyopia. *Investigative Ophthalmology & Visual Science, 48*, 5532-5538.

Abadi, R. V. (2002). Mechanisms underlying nystagmus. *Journal of the Royal Society of Medicine, 95*, 231-234.

Agaoglu, M. N., & Chung, S. T. L. (2016). Can (should) theories of crowding be unified? *Journal of Vision, 16*(15), 10.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & B. F. Csaki (Eds.), *Second International Symposium on Information Theory* (pp. 267-281). Budapest: Academiai Kiado.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transations on Automatic Control AC, 19*, 716-723.

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology, 52*, 1106-1130.

Algaze, A., Roberts, C., Leguire, L., Schmalbrock, P., & Rogers, G. (2002). Functional magnetic resonance imaging as a tool for investigating amblyopia in the human visual cortex: A pilot study. *Journal of American Association of Pediatric Ophthalmology and Strabismus, 6*(5), 300-308.

Altmann, L., & Singer, W. (1986). Temporal integration in amblyopic vision. *Vision Research, 26*, 1959-1968.

Alvarez, I., De Haas, B., Clark, C., Rees, G., & Schwarzkopf, D. (2015). Comparing different stimulus configurations for population receptive field mapping in human fMRI. *Frontiers in Human Neuroscience, 9*(96). doi:10.3389/fnhum.2015.00096

Amano, K., Wandell, B. A., & Dumoulin, S. O. (2009). Visual field maps, population receptive field sizes, and visual field coverage in the human MT + complex. *Journal of Neurophysiology, 102*, 2704-2718.

Anderson, E., Dakin, S. C., Schwarzkopf, D. S., Rees, G., & Greenwood, J. A. (2012). The neural correlates of crowding-induced changes in appearance. *Current Biology, 22*(13), 1199-1206.

Andriessen, J. J., & Bouma, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research, 16*, 71-78.

Ashburner, J., Barnes, G., Chen, C. C., Daunizeau, J., Flandin, G., Friston, K., . . . Phillips, C. (2012). *SPM12 Manual*. London: Functional Imaging Laboratory, Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL.

Atkinson, J., Anker, S., Evans, C., & McIntyre, A. (1987). *The Cambridge Crowding Cards for preschool visual acuity testing.* Paper presented at the Transactions of the 6th International Orthoptic Congress, Harrogate, UK.

Atkinson, J., & Braddick, O. J. (1983). Assessment of visual acuity in infancy and early childhood. *Acta Ophthalmologica Scandinavica, 157*, 18-26.

Avetisov, E. S. (1979). Visual acuity and contrast sensitivity of the amblyopic eye as a function of the stimulated region of the retina. *American Journal of Optometry and Physiological Optics, 56*, 465-469.

Baker, F. H., Grigg, P., & von Noorden, G. K. (1974). Effects of visual deprivation and strabismus on the response of neurons in the visual cortex of the monkey, including studies on the striate and prestriate cortex in the normal animal. *Brain Research, 66*, 185-208.

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision, 9*(12), 1-13.
Banks, W. P., Larson, D. W., & Prinzmetal, W. (1979). Asymmetry of visual interference. *Perception & Psychophysics, 25*(6), 447-456.
Barnes, G. R., Hess, R. F., Dumoulin, S. O., Achtman, R. L., & Pike, G. B. (2001). The cortical deficit in humans with strabismic amblyopia. *Journal of Physiology, 533*, 281-297.
Barrett, B. T., Bradley, A., & McGraw, P. (2004). Understanding the neural basis of amblyopia. *The Neuroscientist, 10*(2), 106-117.
Barrett, B. T., Bradley, A., & McGraw, P. V. (2004). Understanding the neural basis of amblyopia. *The Neuroscientist, 10*(2), 106-117.
Barrett, B. T., Pacey, I. E., Bradley, A., Thibos, L. N., & Morrill, P. (2003). Nonveridical visual perception in human amblyopia. *Investigative Ophthalmology & Visual Science, 44*, 1555-1567.
Bedell, H. E., & Flom, M. C. (1981). Monocular spatial distortion in strabismic amblyopia. *Investigative Ophthalmology & Visual Science, 20*, 263-268.
Bedell, H. E., & Flom, M. C. (1983). Normal and abnormal space perception *American Journal of Optometry and Physiological Optics, 60*(6), 426-435.
Bedell, H. E., Siderov, J., Waugh, S. J., Zemanova, R., Pluháček, F., & Musilová, L. (2013). Contour interaction for foveal acuity targets at different luminances. *Vision Research, 89*, 90-95.
Benson, N. C., & Winawer, J. (2018). Bayesian analysis of retinotopic maps. *eLife, 7*, e40224.
Berman, N., & Murphy, E. H. (1981). The critical period for alteration in cortical binocularity resulting from divergent and convergent strabismus. *Developmental Brain Research, 2*(2), 181-202.
Bex, P. J., & Dakin, S. C. (2005). Spatial interference among moving targets. *Vision Research, 45*, 14.
Bex, P. J., Dakin, S. C., & Simmers, A. J. (2003). The shape and size of crowding for moving targets. *Vision Research, 43*, 2895-2904.
Bi, H., Zhang, B., Tao, X., Harwerth, R. S., Smith, E. L., & Chino, Y. M. (2011). Neuronal responses in visual area V2 (V2) of macaque monkeys with strabismic amblyopia. *Cerebral Cortex, 21*(9), 2033-2045.
Bi, T., Cai, P., Zhou, T., & Fang, F. (2009). The effect of crowding on orientation-selective adaptation in human early visual cortex. *Journal of Vision, 9*(11), 13-13.
Blake, R., Tadin, D., Sobel, K. V., Raissian, T. A., & Chong, S. C. (2006). Strength of early visual adaptation depends on visual awareness. *Procedings of the National Academy of Sciences of the United States of America, 103*(4783-4788).
Bonneh, Y. S., Sagi, D., & Polat, U. (2004). Local and non-local deficits in amblyopia: Acuity and spatial interactions. *Vision Research, 44*(27), 3099-3110.
Bonneh, Y. S., Sagi, D., & Polat, U. (2007). Spatial and temporal crowding in amblyopia. *Vision Research, 47*, 1950-1962.
Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(241), 177-178.
Braddick, O., & Atkinson, J. (2011). Development of humn visual function. *Vision Research, 51*(13), 1588-1609.
Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436.
Breuer, F. A., Blaimer, M., Heidemann, R. M., Mueller, M. F., Griswold, M. A., & Jakob, P. M. (2005). Controlled aliasing in parallel imaging results in higher acceleraton (CAIPIRINHA) for multi-slice imaging. *Magnetic Resonance in Medicine, 53*, 684-691.

Caldara, R., & Seghier, M. L. (2009). The fusiform face area responds automatically to statistical regularities optimal for face categorization. *Human Brain Mapping, 30*, 1615-1625.
Carkeet, A., Levi, D. M., & Manny, R. E. (1997). Development of Vernier acuity in childhood. *Optometry and Vision Science, 74*(9), 741-750.
Cattaneo, Z., Vecchi, T., Monegato, M., Pece, A., Merabet, L. B., & Carbon, C. C. (2013). Strabismic amblyopia affects relational but not featural and Gestalt processing of faces. *Vision Research, 80*, 1-12.
Celebrini, S., Thorpe, S., Trotter, Y., & Imbert, M. (1993). Dynamics of orientation coding in area V1 of the awake primate. *Visual Neuroscience, 10*(5), 811-825.
Chakravarthi, R., & Cavanagh, P. (2007). Temporal properties of the polarity advantage effect in crowding. *Journal of Vision, 7*, 11.
Chaney, W., Fisher, J., & Whitney, D. (2014). The hierarchical sparse selection model of visual crowding. *Frontiers in Integrative Neuroscience, 8*, 73.
Chang, L., & Tsao, D. Y. (2017). The code for facial identity in the primate brain. *Cell, 169*(6), 1013-1028.
Chastain, G. (1982). Confusability and interference between members of parafoveal letter pairs. *Perception & Psychophysics, 32*, 576-580.
Chen, J., He, Y., Zhu, Z., Zhou, T., Peng, Y., Zhang, X., & Fang, F. (2014). Attention-dependent early cortical suppression contributes to crowding. *The Journal of Neuroscience, 34*(32), 10465-10474.
Chui, T. Y., Song, H., & Burns, S. A. (2008). Adaptive-optics imaging of human cone photoreceptor distribution. *Journal of the Optical Society of America A, 25*, 3021-3029.
Chung, S. T., & Bedell, H. E. (1995). Effect of retinal image motion on visual acuity and contour interaction in congenital nystagmus. *Vision Research, 35*(21), 3071-3082.
Chung, S. T., Levi, D. M., & Legge, G. E. (2001). Spatial frequency and contrast properties of crowding. *Vision Research, 41*(14), 1833 - 1850.
Chung, S. T., & Mansfield, J. S. (2009). Contrast polarity differences reduce crowding but do not benefit reading performance in peripheral vision. *Vision Research, 49*(23), 2782-2789.
Ciuffreda, K. J., Kenyon, R. V., & Stark, L. (1979). Fixational eye movements in amblyopia and strabismus. *Journal of the American Optometric Association, 50*, 1251-1258.
Ciuffreda, K. J., Kenyon, R. V., & Stark, L. (1980). Increased drift in amblyopic eyes. *British Journal of Ophthalmology, 64*, 7-14.
Clavagnier, S., Dumoulin, S. O., & Hess, R. F. (2015). Is the cortical deficit in amblyopia due to reduced cortical magnification, loss of neural resolution, or neural disorganisation? *The Journal of Neuroscience, 35*(44), 14740-14755.
Coates, D. R., Levi, D. M., Touch, P., & Sabesan, R. (2018). Foveal Crowding Resolved. *Scientific Reports, 8*(1), 9177. doi:10.1038/s41598-018-27480-4
Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends in Cognitive Sciences, 20*(5), 324-335.
Connolly, M., & Van Essen, D. (1984). The representation of the visual field in parvicellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *Journal of Comparative Neurology, 226*, 544-564.
Coslett, H. B., & Saffran, E. (1991). Simultanagnosia: To see but not two see. *Brain, 114*(4), 1523-1545.
Crawford, M. L., & von Noorden, G. K. (1979). The effects of short-term experimental strabismus on the visual system in Macaca mulatta. *Investigative Ophthalmology & Visual Science, 18*, 496-505.

Crutch, S. J., & Warrington, E. K. (2007). Foveal crowding in posterior cortical atrophy: A specific early-visual-processing deficit affecting word reading. *Cognitive Neuropsychology, 24*, 843-866.
Crutch, S. J., & Warrington, E. K. (2009). The relationship between visual crowding and letter confusability: Towards an understanding of dyslexia in posterior cortical atrophy. *Cognitive Neuropsychology, 26*, 471-498.
Curcio, C. A., & Allen, K. A. (1990). Topography of ganglion cells in human retina. *Journal of Comparative Neurology, 300*, 5-25.
Curcio, C. A., Sloan, K. R., Kalina, R. E., & Hendrickson, A. E. (1990). Human photoreceptor topography. *Journal of Comparative Neurology, 292*, 497-523.
Dakin, S. C., Cass, J., Greenwood, J. A., & Bex, J. P. (2010). Probabilistic, positional averaging predicts object-level crowding effects with letter-like stimuli. *Journal of Vision, 10*(10), 1-16. doi:10.1167/10.10.14
Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortica surface-based analysis. I. Segmentation and surface reconstruction. *NeuroImage, 9*, 179-194.
Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field on the cebral cortex in monkeys. *Journal of Physiology, 159*, 203-221.
Danilova, M. V., & Bondarko, V. M. (2007). Foveal contour interaction and crowding effects at the resolution limit of the visual system. *Journal of Vision, 7*(2).
de Haas, B., Schwarzkopf, D., Anderson, E., & Rees, G. (2014). Perceptual load affects spatial tuning of neuronal populations in human early visual cortex. *Current Biology, 24*(2), 66-67.
de Vries, J. P., Hooge, I. T. C., Wiering, M. A., & Verstraten, F. A. J. (2011). Saccadic selection and crowding in visual search: Stronger lateral masking leads to shorter search times. *Experimental Brain Research, 211*(1), 119-131.
Dekker, T. M., Schwarzkopf, D. S., de Haas, B., Nardini, M., & Sereno, M. I. (2017). Population receptive field tuning properties of visual cortex during childhood. *bioRxiv, 213108*.
Demanins, R., & Hess, R. F. (1996). Effect of exposure duration on spatial uncertainty in normal and amblyopic eyes. *Vision Research, 36*, 1189-1193.
DeValois, R. L., Yund, E. W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research, 22*, 531-544.
DeYoe, E. A., Bandettini, P., Neitz, J., Miller, D., & Winans, P. (1994). Functional magnetic resonance imaging (FMRI) of the human brain. *Journal of Neuroscience Methods, 54*(2), 171-187.
Dow, B. M., Snyder, A. Z., Vautin, R. G., & Bauer, R. (1981). Magnfication factor and receptive field size in foveal striate cortex of the monkey. *Experimental Brain Research, 44*, 213-228.
Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *NeuroImage, 39*(2), 647-660.
Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron, 38*, 659-671.
Ehrt, O., Hess, R. F., Williams, C. B., & Sher, K. (2003). Foveal contrast thresholds exhibit spatial frequency- and polarity-specific contour interactions. *Journal of the Optical Society of America A, 20*, 11-17.
Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organisation in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex, 7*, 181-192.
Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 583-597.

Ester, E. F., Klee, D., & Awh, E. (2014). Visual crowding cannot be wholly explained by feature pooling. *Journal of Experimental Psychology: Human Perception & Performance, 40*(3), 1022.
Ester, E. F., Zilber, E., & Serences, J. T. (2015). Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *Journal of Vision, 15*(1), 4-4. doi:10.1167/15.1.4
Estes, W., Allmeyer, D. H., & Reder, S. M. (1976). Serial position functions for letter identification at brief and extended exposure durations. *Perception & Psychophysics, 19*, 1-15.
Fang, F., & He, S. (2008). Crowding alters the spatial distribution of attention modulation in human primary visual cortex. *Journal of Vis§, 8*, 1-9.
Farivar, R., Thompson, B., Mansouri, B., & Hess, R. F. (2011). Interocular suppression in strabismic amblyopia results in an attenuated and delayed hemodynamic response function in early visual cortex. *Journal of Vision, 11*(16).
Farzin, F., Rivera, S. M., & Whitney, D. (2009). Holistic crowding of Mooney faces. *Journal of Vision, 9*(6), 1-15.
Feng, C., Jiang, Y., & He, S. (2007). Horizontal and vertical asymmetry in visual spatial crowding effects. *Journal of Vision, 7*(2), 1-10.
Ferrera, V. P., Nealy, T. A., & Maunsell, J. H. (1994). Responses in macaque visual area V4 following inactivation of the parvocellular and magnocellular LGN pathways. *Journal of Neuroscience, 14*, 2080-2088.
Fischl, B. (2012). FreeSurfer. *NeuroImage, 62*, 774-781.
Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordiante system. *NeuroImage, 9*, 195-207.
Flom, M. C., & Bedell, H. E. (1985). Identifying amblyopia using associated functions, acuity and non-acuity features. *American Journal of Optometry and Physiological Optics, 62*, 153-160.
Flom, M. C., Heath, G. G., & Takahashi, E. (1963). Contour interaction and visual resolution: Contralateral effect. *Science, 142*, 979-980.
Flom, M. C., Weymouth, F. W., & Kahneman, D. (1963). Visual resolution and contour interaction. *Journal of the Optical Society of America, 53*(9), 1026-1032.
Fortenbaugh, F. C., Silver, M. A., & Robertson, L. C. (2015). Individual differences in visual field shape modulate the effects of attention on the lower visual field advantage in crowding. *Journal of Vision, 15*(2), 19.
Francis, G., Manassi, M., & Herzog, M. H. (2017). Neural dynamics of grouping and segmentation explain properties of visual crowding. *Psychological Review, 124*(4), 483-504.
Freeman, J., Chakravarthi, R., & Pelli, G. D. (2012). Substitution and pooling in crowding. *Attention, Perception, & Psychophysics, 74*(2), 379-396.
Freeman, J., Donner, T. H., & Heeger, D. J. (2011a). Inter-area correlations in the ventral visual pathway reflect feature integration. *Journal of Vision, 11*(4), 1-23.
Freeman, J., Donner, T. H., & Heeger, D. J. (2011b). Inter-area correlations in the ventral visual pathway reflect feature integration. *Journal of Vision, 11*(4), 15.
Freeman, J., & Simoncelli, E. (2011). Metamers of the ventral stream. *Nature Neuroscience, 14* (9), 1195-1201.
Freeman, R. D., & Tsumoto, T. (1983). An electrophysiological comparison of convergent and divergent strabismus in the cat: Electrical and visual activation of single cortical cells. *Journal of Neurophysiology, 49*(1), 238-253.

Fronius, M., & Sireteanu, R. (1989). Monocular geometry is selectively distorted in the central field of strabismic amblyopes. *Investigative Ophthalmology & Visual Science, 30*(9), 2034-2044.
Geiger, G., & Lettvin, J. Y. (1987). Peripheral vision in persons with dyslexia. *The New England Journal of Medicine, 316*, 1238-1243.
Giaschi, E. D., Regan, D., Kraft, S. P., & Kothe, A. C. (1993). Crowding and contrast in amblyopia. *Optometry and Vision Science, 70*, 192-197.
Goffaux, V., & Rossion, B. (2007). Face inversion disproportionately impairs the perception of vertical but not horizontal relations between features. *Journal of Experimental Psychology: Human Perception & Performance 33*, 995-1001.
Goldberg, M. E., & Wurtz, R. H. (1972). Activity of superior colliculus in behaving monkey. II. Effect of attention on neural responses. *Journal of Neurophysiology, 35*, 560-574.
Gonzalez, E. G., Wong, A. M., Niechwiej-Szwedo, E., Tarita-Nistor, L., & Steinbach, M. J. (2012). Eye position stability in amblyopia and in normal binocular vision. *Investigative Ophthalmology & Visual Science, 53*, 5386-5394.
Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2009). Positional averaging explains crowding with letter-like stimuli. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 12130-13135.
Greenwood, J. A., Bex, P. J., & Dakin, S. C. (2010). Crowding changes appearance. *Current Biology, 20*(6), 496-501.
Greenwood, J. A., & Parsons, M. J. (2019). Dissociable effects of visual crowding on the perception of colour and motion. *bioRxiv*. doi:https://doi.org/10.1101/639450
Greenwood, J. A., Szinte, M., Sayim, B., & Cavanagh, P. (2017). Variations in crowding, saccadic precision, and spatial localisation reveal the shared topology of spatial vision. *Procedings of the National Academy of Sciences of the United States of America, 114*(17), E3573- E3582.
Greenwood, J. A., Tailor, V. K., Sloper, J. J., Simmers, A., Bex, P. J., & Dakin, S. C. (2012). Visual acuity, crowding, and stereo-vision are linked in children with and without amblyopia. *Visual Psychophysics and Physiological Optics, 53*(12), 7655-7665.
Grill-Spector, K., & Malach, R. (2004). THE HUMAN VISUAL CORTEX. *Annual Review of Neuroscience, 27*(1), 649-677. doi:10.1146/annurev.neuro.27.070203.144220
Grönlund, M. A., Andersson, S., Aring, E., Hård, A.-L., & Hellström, A. (2006). Ophthalmological findings in a sample of Swedish children aged 4-15. *Acta Ophthalmologica Scandinavica, 84*, 169-176.
Gunton, K. B. (2013). Advances in amblyopia: What have we learned from PEDIG trials? *Pediatrics, 131*, 540-547.
Gur, M., Kagan, I., & Snodderly, D. M. (2005). Orientation and direction selectivity of neurons in V1 of alert monkeys: Functional relationships and laminar distributions. *Cerebral Cortex, 15*(8), 1207-1221.
Gur, M., & Snodderly, D. M. (2007). Direction selectivity in V1 of alert monkeys: Evidence for paralllel pathways for motion processing. *Journal of Physiology, 585*, 383-400.
Hariharan, S., Levi, D. M., & Klein, S. A. (2005). "Crowding" in normal and amblyopic vision assessed with Gaussian and Gabor C's. *Vision Research, 45*(5), 617-633.
Harrard, R. (1996). Psychophysics of suppression. *Eye, 10*, 270-273.
Harrison, W. J., & Bex, P. J. (2015). A unifying model of orientation crowding in peripheral vision. *Current Biology, 25*(24), 3213-3219.

Harrison, W. J., & Bex, P. J. (2016). Reply to Pachai *et al. Current Biology, 26*(9), R353-R354.
Harrison, W. J., Mattingley, J. B., & Remington, R. W. (2013). Eye movement targets are released from crowding. *Journal of Neuroscience, 33*(7), 2927-2933.
Harrison, W. J., Retell, J. D., Remington, R. W., & Mattingley, J. B. (2013). Visual crowding at a distance during predictive remapping. *Current Biology, 23*(9), 793-798.
Harvey, B. M., & Dumoulin, S. O. (2011). The relationship between cortical magnification factor and population receptive field size in human visual cortex: Constancies in cortical architecture. *The Journal of Neuroscience, 31*(38), 13604-13612.
Harwerth, R. S., Smith, E. L., Boltz, R. L., Crawford, M. L., & von Noorden, G. K. (1983). Behavioural studies on the effects of abnormal early visual experience in monkeys: Spatial modulation sensitivity. *Vision Research, 23*, 1501-1510.
He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*, 334-337.
Hedgé, J., & Van Essen, D. (2000). Selectivity for complex shapes in primate visual area V2. *The Journal of Neuroscience, 20*, 1-6.
Hendrickson, A. E., Movshon, J. A., Eggers, H. M., Gizzi, M. S., Boothe, R. G., & Kiorpes, L. (1987). Effects of early unilateral blur on the macaques visual system. 2. Anatomical observations. *Journal of Neuroscience, 7*, 1327-1339.
Hensch, T. K. (2004). Critical period regulation. *Annual Review of Neuroscience, 27*, 549-579.
Herzog, M. H., & Manassi, M. (2015). Uncorking the bottleneck of crowding: A fresh look at object recognition. *Current Opinion in Behavioural Sciences, 1*, 86-93.
Herzog, M. H., Sayim, B., Chicherov, V., & Manassi, M. (2015). Crowding, grouping, and object recognition: A matter of appearance. *Journal of Vision, 15*, 1-18.
Herzog, M. H., Sayim, B., Manassi, M., & Chicherov, V. (2016). What crowds in crowding? *Journal of Vision, 16*(11), 25.
Hess, R. F., Campbell, F. W., & Greenhalgh, T. (1978). On the nature of the neural abnormality in human amblyopia; neural aberrations and neural sensitivity loss *Pflügers Archives: European Journal of Physiology, 377*(3), 201-207.
Hess, R. F., Dakin, S. C., & Kapoor, N. (2000). The foveal “crowding” effect: physics or physiology. *Vision Research, 49*, 365-370.
Hess, R. F., Dakin, S. C., Tewfik, M., & Brown, B. (2001). Contour interaction in amblyopia: Scale selection. *Vision Research, 41*(17), 2285-2296.
Hess, R. F., & Field, D. (1993). Is the increased spatial uncertainty in the normal periphery due to spatial undersampling or uncalibrated disarray? *Vision Research, 33*(18), 2663-2670.
Hess, R. F., & Field, D. (1994). Is the spatial deficit in strabismic amblyopia due to loss of cells or an uncalibrated disarray of cells? . *Vision Research, 34*, 3397-3406.
Hess, R. F., & Hayes, A. (1994). The coding of spatial position by the human visual system: Effects of spatial scale and retinal eccentricity. *Vision Research, 34*(5), 625-643.
Hess, R. F., & Howell, E. R. (1977). The threshold contrast sensitivity function in strabismic amblyopia: Evidence for a two type classification. *Vision Research, 17*, 1049-1055.
Hess, R. F., & Jacobs, R. J. (1979). A preliminary report of acuity and contour interactions across the amblyope's visual field. *Vision Research, 19*, 1403-1408.

Hess, R. F., McIlhagga, W., & Field, D. J. (1997). Contour integration in strabismic amblyopia: The sufficiency of an explanation based on positional uncertainty. *Vision Research, 37*(22), 3145-3161.
Hess, R. F., Williams, C. B., & Chaundry, A. (2001). Contour interaction for easily resolvable stimulus. *Journal of the Optical Society of America A, 18*, 2414-2418.
Ho, C. S., Paul, P. S., Asirvatham, A., Cavanagh, P., Cline, P., & Giaschi, D. (2006). Abnormal spatial selection and tracking in children with amblyopia. *Vision Research, 46*(19), 3174-3283.
Howell, E. R., Mitchell, D. E., & Keith, C. G. (1983). Contrast thresholds for sine gratings of children with amblyopia. *Investigative Ophthalmology & Visual Science, 24*(6), 782-787.
Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148*(574-591).
Hubel, D. H., & Wiesel, T. N. (1965). Binocular interaction in striate cortex of kittens reared with artificial squint. *Journal of Neurophysiology, 28*, 1060-1072.
Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. *Visual Cognition, 9*(7), 889-910.
Hussain, Z., Svensson, C.-M., Besle, J., Webb, B. S., Barrett, B. T., & McGraw, P. (2015). Estimation of corticla magnification from positional error in normally sighted and amblyopic subjects. *Journal of Vision, 15*(2), 25.
Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *The Journal of Comparative Neurology, 387*(167-178).
Huttenlocher, P. R., de Courten, C., Garey, L. J., & Van der Loos, H. (1982). Synaptogenesis in human visual cortex - Evidence for synapse elimination during normal development. *Neuroscience Letters, 33*(3), 247-252.
Intriligator, J. M., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43*(3), 171-216.
Jeon, S. T., Hamid, J., Mauer, D., & Lewis, T. L. (2010). Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *Journal of Experimental Child Psychology, 107*(4), 423-437.
Kalpadakis-Smith, A. V., Goffaux, V., & Greenwood, J. A. (2018). Crowding for faces is determined by visual (not holistic) similarity: Evidence from judgements of eye position. *Scientific Reports, 8*(1), 12556.
Keen, A. G., & Lovegrove, W. J. (2000). Transient deficit hypothesis and dyslexia: Examination of whole–parts relationship, retinal sensitivity, and spatial and temporal frequencies. *Vision Research, 40*(6), 705-715.
Kennedy, G. J., & Whitaker, D. (2010). The chromatic selectivity of visual crowding. *Journal of Vision, 10*(6), 15.
Keshvari, S., & Rosenholtz, R. (2016). Pooling of continuous features provides a unifying account of crowding. *Journal of Vision, 16*(3), 39.
Kiorpes, L., & Daw, N. (2018). Cortical correlates of amblyopia. *Visual Neuroscience, 35*(e016), 1-5.
Kiorpes, L., Kiper, D. C., O'Keefe, L. P., Cavanaugh, J. R., & Movshon, J. A. (1998). Neuronal correlates of amblyopia in the visual cortex of macaque monkeys with experimental strabismus and anisometropia. *Journal of Neuroscience, 18*, 6411-6424.
Kiorpes, L., & McKee, S. P. (2006). Visual processing in amblyopia: Animal studies. *Strabismus, 14*(1), 3-10.
Kiorpes, L., Movshon, J. A., Chalupa, L. M., & Werner, J. S. (2003). Neural limitations on visual development in primates *The visual neurosciences* (pp. 159-173). Cambridge, MA: MIT Press.

Kirschen, D. G., & Flom, M. C. (1978). Visual acuity at different retinal loci of eccentrically fixating functional amblyopes. *American Journal of Optometry and Physiological Optics, 55*, 144-150.
Koffka, K. (1935). *Principles of Gestalt Psychology*. New York: Harcourt, Brace.
Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision, 8*, 255-279.
Kravitz, D. J., & Behrmann, M. (2011). Space-, object-, and feature- based attention interact to organise visual scenes. *Attention, Perception, & Psychophysics, 73*(8), 2434-2447.
Kravitz, D. J., Saleem, K. S., Baker, C. I., Ungerleider, L. G., & Mishkin, M. (2013). The ventral visual pathway An expanded neural framework for the proessing of object quality. *Trends in Cognitive Sciences, 17*(1), 26-49.
Krumhansl, C. L., & Thomas, E. A. (1977). Effect of level of consfusability on reporting letters from briefly presented visual displays. *Perception & Psychophysics, 21*, 269-279.
Lagarias, J. C., Reeds, J. A., Wright, M. H., & Wright, P. E. (1998). Convergence properties of the Nedler-Mead simplex method in low dimensions. *SIAM Journal on Optimization, 9*, 112-147.
Lagreze, W.-D., & Sireteanu, R. (1991). Two-dimensional spatial distortions in human strabismic amblyopia. *Vision Research, 31*(1271-1288).
Lalor, S. J. H., Formankiewicz, M. A., & Waugh, S. J. (2016). Crowding and visual acuity measured in adults usign paediatric test letters, pictures and symbols. *Vision Research, 121*, 31-38.
Leat, S. J., Li, W., & Epp, K. (1999). Crowding in central and eccentric vision: The effects of contour interaction and attention. *Investigative Ophthalmology & Visual Science, 40*, 504-512.
Lettvin, J. Y. (1976). On seeing sidelong. *Sciences, 16*(4), 10-20.
Lev, M., Yehezkel, O., & Polat, U. (2014). Uncovering foveal crowding? *Scientific Reports, 4*, 4067.
Levi, D. M. (2000). Crowding is size invariant in foveal visino, but not in peripheral or amblyopic vision. *Investigative Ophthalmology & Visual Science (Supplement), 41*, S534.
Levi, D. M. (2006). Visual processing in amblyopia: Human studies. *Strabismus, 14*, 11-19.
Levi, D. M. (2008). Crowding - An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635-654.
Levi, D. M. (2013). Linking assumptions in amblyopia. *Visual Neuroscience, 30*, 277-287.
Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology, 19*(23), 1988-1993.
Levi, D. M., & Carney, T. (2011). The effect of flankers on three tasksin central, peripheral, and amblyopic vision. *Journal of Vision, 11*(1), 1-10.
Levi, D. M., Hariharan, S., & Klein, S. A. (2002a). Suppressive and facilitatory spatial interactions in amblyopic vision. *Vision Research, 42*, 1379-1394.
Levi, D. M., Hariharan, S., & Klein, S. A. (2002b). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision, 2*, 3.
Levi, D. M., & Harwerth, R. S. (1978). Contrast evoked potentials in strabismic and anisometropic amblyopia. *Investigative Ophthalmology & Visual Science, 17*, 571-575.
Levi, D. M., & Klein, S. A. (1982). Hyperacuity and amblyopia. *Nature, 298*, 268-270.
Levi, D. M., & Klein, S. A. (1983). Spatial localization in normal and amblyopic vision. *Vision Research, 23*(10), 1005-1017.

Levi, D. M., & Klein, S. A. (1985). Vernier acuity, crowding and amblyopia. *Vision Research, 25*(7), 979-991.
Levi, D. M., & Klein, S. A. (1986). Sampling in spatial vision. *Nature, 320*, 360-362.
Levi, D. M., & Klein, S. A. (1990). Equivalent intrinsic blur in spatial vision. *Vision Research, 30*(12), 1971-1973.
Levi, D. M., Klein, S. A., & Aitsebaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification. *1985, 25*(7), 963-977.
Levi, D. M., Klein, S. A., & Hariharan, S. (2002). Suppressive and facilitatory interactions in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision, 2*(2), 140-166.
Levi, D. M., Klein, S. A., & Yap, Y. L. (1987). Positional uncertainty in peripheral and amblyopic vision. *Vision Research, 27*(4), 581-597.
Levi, D. M., Song, S., & Pelli, D. G. (2007a). Amblyopic reading is crowded. *Journal of Vision, 7*(2), 21.
Levi, D. M., Song, S., & Pelli, D. G. (2007b). Amblyopic reading is crowded *Journal of Vision, 7*(2), 1-17.
Levi, D. M., Yu, C., Kuai, S. G., & Rislove, E. (2007). Global controur processing in amblyopia. *Vision Research, 47*, 512-524.
Lewis, J. P., & Frick, R. W. (1999). Row blindness in Gestalt grouping and developmental dyslexia. *Neuropsychologia, 37*(3), 385-393.
Li, X., Dumoulin, S. O., Mansouri, B., & Hess, R. F. (2007). Cortical deficits in human amblyopia: Their regional distribution and their relationship to the contrast detection deficit. *Investigative Ophthalmology & Visual Science, 48*(4), 1575-1591.
Li, X., Mullen, K. T., Thoompson, B., & Hess, R. F. (2011). Effective connectivity anomalies in human amblyopia. *NeuroImage, 54*(1), 505-516.
Lim, H. T., Yu, Y. S., Park, S. H., Ahn, H., Kim, S., Lee, M., . . . Koo, B. S. (2004). The Seoul Metropolitan Preschool Vision Screening Program: results from South Korea. *British Journal of Ophthalmology, 88*(7), 929-233.
Liu, J., Harris, A., & Kanwisher, N. (2010). Perception of face parts and face configurations: An fMRI study. *Journal of Cognitive Neuroscience, 22*(1), 203-211.
Liu, L., & Arditi, A. (2000). Apparent string shortening concomitant with letter crowding. *Vision Research, 40*, 1059-1067.
Livne, T., & Sagi, D. (2007). Configuration influence on crowding. *Journal of Vision, 7*(2), 1-12.
Logothetis, N. K., & Charles, E. R. (1990). The minimum motion technique applied to determine isoluminance in psychophysical experiments with monkeys. *Vision Research, 30*, 829-838.
Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature, 412*, 150-157.
Louie, E. G., Bressler, D. W., & Whitney, D. (2007). Holistic crowding: Selective interference between configural representations of faces in crowded scenes. *Journal of Vision, 7*(2), 24-24.
Maione, M., Berardi, D., & Cerimele, D. (1983). The horse and zebras test in visual acuity and contrast sensitivity assessment. *Ophthalmic Paediatrics & Genetics, 2*(3), 149-152.
Malania, M., Herzog, M. H., & Westheimer, G. (2007). Grouping of contextual elements that affect Vernier thresholds. *Journal of Vision, 7*(2), 1-7.
Manassi, M., Hermens, F., Francis, G., & Herzog, M. H. (2015). Release of crowding from pattern completion. *Journal of Vision, 15*(8), 16.

Manassi, M., Lonchampt, S., Clarke, A., & Herzog, M. H. (2016). What crowding can tell us about object representations. *Journal of Vision, 16*(3), 35.
Manassi, M., Sayim, B., & Herzog, M. H. (2012). Grouping, pooling and whe bigger is better in visual crowding *Journal of Vision, 12*(10), 1-13. doi:10.1167/12.10.13
Manassi, M., Sayim, B., & Herzog, M. H. (2013). When crowding of crowding leads to uncrowding. *Journal of Vision, 13*(13), 1-10.
Manassi, M., & Whitney, D. (2018). Multi-level crowding and the paradox of object recognition in clutter. *Current Biology, 28*(3), R127-R133.
Manning, C., Jones, P. R., Dekker, T. M., & Pellicano, E. (2018). Psychophysics with children: Investigating the effects of attentional lapses on threshold estimates. *Attention, Perception, & Psychophysics, 80*, 1311-1324.
Mareschal, I., Morgan, M., & Solomon, J. A. (2010). Cortical distance determines whether flankers cause crowding or the tilt illusion. *Journal of Vision, 10*(8), 1-14.
Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowidng, reading, and developmental dyslexia. *Journal of Vision, 9* (4), 1-18.
Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision, 5*(1), 58-70.
Martinez-Trujillo, J. C., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology, 744-751*, 744-751.
Matsuo, T., Matsuo, C., Matsuoka, H., & Kio, K. (2007). Detection of strabismus and amblyopia in 1.5- and 3- year-old children by a preschool vision-screening program in Japan. *Acta Medica Okayama, 61*(1), 9-16.
Mazer, J. A., Vinje, W. E., McDermott, J., Schiller, P. H., & Gallant, J. L. (2002). Spatial frequency and orientation tuning dynamics in area V1. *Procedings of the National Academy of Sciences of the United States of America, 99*, 1645-1650.
McKean-Cowdin, R., Cotter, S. A., Tarczy-Hornoch, K., Wen, G., Kim, J., Borchert, M., & Varma, R. (2013). Prevalence of amblyopia or strabismus in Asian and non-Hispanic white preschool children: Multi-Ethnic Pediatric Eye Disease Study. *Ophthalmology, 120*(10), 2117-2124.
McKee, S. P., Levi, D. M., & Movshon, J. A. (2003). The pattern of visual deficits in amblyopia. *Journal of Vision, 3* (5), 5-5.
Merigan, W. H. (2000). Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Visual Neuroscience, 17*, 949-958.
Miles, W. (1928). Ocular dominance: Methods and results. *Psychological Bulletin, 25*, 155-156.
Millin, R., Arman, A. C., Chung, S. T. L., & Tjan, B. S. (2014). Visual crowding in V1. *Cerebral Cortex, 24*(12), 3107-3115.
Mooney, C. (1957). Age in the development of closure ability in children. *Canadian Journal of Psychology, 11*(4), 219-226.
Moores, E., Cassim, R., & Talcott, J. B. (2011). Adults with dyslexia exhibit large effects of crowding, increased dependence on cues, and detrmimental effects of distractors in visual search tasks. *Neuropsychologia, 49*(14), 3881-3890.
Motter, B. C. (2002). Crowding and object integration within the receptive field of V4 neurons. *Journal of Vision, 2*(7), 274a.
Motter, B. C. (2006). Modulation of transient and sustained response components of V4 neurons by temporal crowding in flashed stimulus sequences. *Journal of Neuroscience, 26*, 9683-9694.

Motter, B. C. (2009). Central V4 receptive fields are scaled by the V1 cortical magnification and correspond to a constant-sized sampling of the V1 surface. *Journal of Neuroscience, 29*, 5749-5757.
Motter, B. C. (2018). Stimulus conflation and tuning selectivity in V4 neurons: A model of visual crowding. *Journal of Vision, 18*(1), 15.
Motter, B. C., & Simoni, D. A. (2007). The roles of cortical image separation and size in active visual search performance. *Journal of Vision, 7* (2), 1-15.
Moutsiana, C., de Haas, B., Papageorgiou, A., van Dijk, J. A., Balraj, A., Greenwood, J. A., & Schwarzkopf, D. S. (2016). Cortical idiosyncrasies predict the perception of object size. *Nature Communications, 7*(12110), 1-12.
Movshon, J. A., Eggers, H. M., Gizzi, M. S., Hendrickson, A. E., Kiorpes, L., & Boothe, R. G. (1987). Effects of early unilateral blur on the macaques visual system. 3. Physiological observations. *Journal of Neuroscience, 7*, 1340-1351.
Movshon, J. A., & Lennie, P. (1979). Pattern-selective adaptation n visual cortical neurones. *Nature, 26*(278), 850-852.
Movshon, J. A., & Simoncelli, E. (2014). Representation of naturalistic image structure in the primate visual cortex. *Cold Spring Harbor Symposia on Quantitative Biology, 79*, 115-122.
Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology, 283*, 53-77.
Muckli, L., Kiess, S., Tonhausen, N., Singer, W., Goebel, R., & Sireteanu, R. (2006). Cerebral correlates of impaired grating perception in individual, psychophysically assessed human amblyopes. *Vision Research, 46*, 506-526.
Murray, S. O., & Wojciulik, E. (2004). Attention increases selectivity in the human lateral occipital complex. *Nature Neuroscience, 7*, 70-74.
Nandy, A. S., & Tjan, B. S. (2007). The nature of letter crowding as revealed by first- and second-order classification images. *Journal of Vision, 7*(5), 1-26.
Nandy, A. S., & Tjan, B. S. (2012). Saccade-confounded image statistics explain visual crowding. *Nature Neuroscience, 15*, 463-469.
Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal, 7*, 308-313.
Neri, P., Luu, J. Y., & Levi, D. M. (2007). Sensitivity to biological motion drops by ~ 1/2 log-unit with inversion, and is unaffected by amblyopia. *Vision Research, 47*, 1209-1214.
Niederhauser, S., & Mojon, D. S. (2002). Normal isopter position in the peripheral visual field in goldmann kinetic perimetry. *Ophthalmologica, 216*(6), 406-408.
Norton, E. S., Beach, S. D., & Gabrieli, J. D. E. (2015). Neurobiology of dyslexia. *Current Opinion in Neurobiology, 30*, 73-78.
Parkes, L., Lund, J., Angelluci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience, 4*(7), 6.
Pascal, E., & Abadi, R. V. (1995). Contour interaction in the presence of congenital nystagmus. *Vision Research, 35*(12), 1785-1789.
Pelli, D. G. (1997). The VideoToolBox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.
Pelli, D. G. (2008). Crowding: A cortical constraint in object recognition. *Current Opinion in Neurobiology, 18*(4), 445-451.
Pelli, D. G., Burns, C. W., Farell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research, 46*, 4646-4674.

Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision, 4*(12), 12-12.
Pelli, D. G., Robson, J. G., & Wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clinical Vision Sciences, 2*, 187-199.
Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience, 11*(10), 1129-1135.
Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, N. J., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision, 7*, 1-36.
Perri, R., Bartolomeo, P., & Silveri, M. C. (1996). Letter dyslexia in a letter-by-letter reader. *Brain and Language, 53*(55), 390-407.
Petrov, Y., & Meleshkevich, O. (2011a). Asymmetries and idiosyncratic hot spots in crowding. *Vision Research, 51*(10), 1117-1123.
Petrov, Y., & Meleshkevich, O. (2011b). Locus of spatial attention determines inward–outward anisotropy in crowding. *Journal of Vision, 11*(4), 1-1. doi:10.1167/11.4.1
Petrov, Y., & Popple, A. V. (2007). Crowding is directed to the fovea and preserves only feature contrast. *Journal of Vision, 7*(2), 8.
Petrov, Y., Popple, A. V., & McKee, S. P. (2007). Crowding and surround suppression: Not to be confused. *Journal of Vision, 7* (2), 1-9.
Piano, M. E. F., Bex, P. J., & Simmers, A. J. (2015). Perceptual visual distortions in adult amblyopia and their relationship to clinical features. *Investigative Ophthalmology & Visual Science, 56*(9), 5533-5542.
Pinon, M. C., Gattass, R., & Sousa, A. P. (1998). Area V4 in Cebus monkey: Extent and visuotopic organization. *Cerebral Cortex, 8*, 685-701.
Põder, E., & Wagemans, J. (2007). Crowding with conjunctions of simple features. *Journal of Vision, 7*(2), 1-12.
Polat, U., Bonneh, Y., Ma-Naim, T., Belkin, M., & Sagi, D. (2005). Spatial interactions in amblyopia: Effects of stimulus parameters and amblyopia type. *Vision Research, 45*(11), 1471-1479.
Polat, U., Sagi, D., & Norcia, A. M. (1997). Abnormal long-range spatial interactions in amblyopia. *Vision Research, 37*, 737-774.
Popple, A. V., & Levi, D. M. (2008). The attentional blink in amblyopia. *Journal of Vision, 8*(13), 1-12.
Portilla, J., & Simoncelli, E. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients *International Journal of Computer Vision, 40*, 49-71.
Posner, I. P., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25-42.
Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience, 1*(125-132).
Preslan, M. W., & Novak, A. (1996). Baltimore Vision Screening Project. *Ophthalmology, 103*(1), 105-109.
Preslan, M. W., & Novak, A. (1998). Baltimore Vision Screening Project: Phase 2. *Ophthalmology, 105*(1), 150-153.
Pugh, M. (1958). Visual distortions in amblyopia. *British Journal of Ophthalmology, 42*, 449-460.
Rentschler, I., & Treutwein, B. (1985). Loss of spatial phase relationships in extrafoveal vision. *Nature, 313*, 308-310.
Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience, 2*(1019-1025).
Riesenhuber, M., & Poggio, T. (2000). Models of object recognition. *Nature Neuroscience, 3*, 1199-1204.

Ringach, D. L., Shapley, R. M., & Hawken, M. J. (2002). Orientation selectivity in macaque V1: Diversity and laminar dependence. *Journal of Neuroscience, 22*, 5639-5651.
Rislove, E., Hall, E. C., Stavros, K. A., & Kiorpes, L. (2010). Scale-dependent loss of global form perception in strabismic amblyopia. *Journal of Vision, 10*(1-13).
Robaei, D., Kifley, A., Rose, K. A., Ojaimi, E., Martin, F. J., & Mitchell, P. (2008). Causes and associations of amblyopia in a population-based sample of 6-year old Australian children. *Archives of Ophthalmology, 124*(6), 878-884.
Robaei, D., Rose, K. A., Ojaimi, E., Kifley, A., Huynh, S., & Mitchell, P. (2005). Visual acuity and the causes of vision loss in a population-based sample of 6-year-old Australian children. *Ophthalmology, 112*, 1275-1282.
Rolfs, M. (2009). Microsaccades: Small steps on a long way. *Vision Research, 49*(20), 2415-2441.
Rosa, M. G., & Tweedale, R. (2005). Brain maps, great and small: Lessons from comparative studies of primate visual cortical organization. *Philosophical transactions of the Royal Society B: Biological Sciences, 360*(1456), 665-691.
Rosen, S., & Pelli, D. G. (2015). Crowding by a repeating pattern. *Journal of Vision, 15*(6), 1-9.
Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science, 2*, 437-457.
Rosenholtz, R., Yu, D., & Keshvari, S. (2017). *Challenges to pooling models of crowding reconsidered*. Paper presented at the 40th European Conference for Visual Perception, Berlin, Germany.
Rossion, B. (2008). Picture-plane inversion leads to qualitative changes of face perception. *Acta Psychologica, 128*(2), 274-289.
Rovamo, J., & Raninen, A. (1990). Cortical acuity and the luminous flux collected by retinal ganglion cells at various eccentricities in human rod and cone vision. *Vision Research, 30*(1), 11-21.
Saarela, T., & Herzog, M. H. (2009). Crowding in multi-element arrays: Regularity of spacing. *Journal of Vision, 9*(8), 1017.
Saarela, T., Sayim, B., Westheimer, G., & Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *Journal of Vision, 9*(2), 5.
Saarela, T., Westheimer, G., & Herzog, M. H. (2010). The effect of spacing regularity on visual crowding. *Journal of Vision, 10*(10), 1-7.
Sayim, B., Westheimer, G., & Herzog, M. H. (2008). Contrast polarity, chromaticity, and stereoscopic depth modulate contextual interactions in vernier acuity. *Journal of Vision, 8*(8), 12.
Sayim, B., Westheimer, G., & Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychological Science, 21*(5), 641-644.
Schiller, P. H., Finlay, B. L., & Volman, S. F. (1976). Quantitative studies of single-cell properties in monkey striate cortex ||. Orientation specificity and ocular dominance. *Journal of Neurophysiology, 38*(6), 1320-1333.
Schor, C., & Hallmark, W. (1978). Slow control of eye position in strabismic amblyopia. *Investigative Ophthalmology & Visual Science, 17*, 577-581.
Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J., & Driver, J. (2005). Attentional load an sensory competition in human vision: Modulation of fMRI responses by load at fixation during task-irrelevant stimulation in the peripheral visual field. *Cerebral Cortex, 15*, 770-786.
Schwarzkopf, D. S., Anderson, E. J., de Haas, B., White, S. J., & Rees, G. (2014). Larger extrastriate population receptive fields in autism spectrum disorders. *Journal of Neuroscience, 34*(7), 2713-2724.
Schwarzkopf, D. S., de Haas, B., & Alvarez, I. (2018). SamSrf 6 - Toolbox for pRF modelling. Retrieved from https://osf.io/2rgsm/

Secen, J., Culham, J., Ho, C., & Giaschi, D. (2011). Neural correlates of the mutliple-object tracking deficit in amblyopia. *Vision Research, 51*(23-24), 2517-2527.
Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., . . . Tootel, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*, 889-893.
Seth, A. K. (2014). A predictive processing theory of sensorimotor contigencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive Neuroscience, 5*(2), 97-118.
Sharma, V., Levi, D. M., & Klein, S. A. (2000). Undercounting features and missing features: Evidence for a high-level deficit in strabismic amblyopia. *Nature Neuroscience, 3*, 496-501.
Shooner, C., Hallum, L. E., Kumbhani, R. D., Ziemba, C. M., Garcia-Marin, V., Kelly, J. G., . . . Kiorpes, L. (2015). Population representation of visual information in areas V1 and V2 of amblyopic macaques. *Vision Research, 114*(56-67).
Siderov, J., Waugh, S. J., & Bedell, H. E. (2013). Foveal contour interaction for low contrast acuity targets. *Vision Research, 77*, 10-13.
Simmers, A. J., Gray, L. S., McGraw, P. V., & Winn, B. (1999). Contour interaction for high and low contrast optotypes in normal an damblyopic observers. *Ophthalmic and Physiological Optics, 19*(3), 253-260.
Simmers, A. J., Ledgeway, T., & Hess, R. F. (2005). The influences of visibility and anomalous integration processes on the perception of global spatial form versus motion in human amblyopia. *Vision Research, 45*(4), 449-460.
Simmers, A. J., Ledgeway, T., Hess, R. F., & McGraw, P. V. (2003). Deficits to global motion processing in human amblyopia. *Vision Research, 43*, 729-738.
Simmers, A. J., Ledgeway, T., Mansouri, B., Hutchinson, C. V., & Hess, R. F. (2006). The extent of the dorsal extra-striate deficit in amblyopia. *Vision Research, 46*, 2571-2580.
Sireteanu, R., Baumer, C. C., & Iftime, A. (2008). Temporal instability in amblyopic vision: Relationship to a displacement map of visual space. *Investigative Ophthalmology & Visual Science, 49*, 3940-3954.
Sireteanu, R., & Best, J. (1992). Squint-induced modification of visual receptive fields in the lateral suprasylvian cortex of the cat: Binocular interaction, vertical effect and anomalous correspondence. *European Journal of Neuroscience, 4*, 235-242.
Sireteanu, R., & Fronius, M. (1981). Naso-temporal asymmetries in human amblyopia: Consequence of long-term interocular suppression. *Vision Research, 21*, 1055-1063.
Sireteanu, R., Lagreze, W.-D., & Constantinescu, D. H. (1993). Distortions in two-dimensional visual space perception in strabismic observers. *Vision Research, 33*(5-6), 677-690.
Sireteanu, R., Thiel, A., Fikus, S., & Iftime, A. (2008). Patterns of spatial distortions in human amblyopia are invariant to stimulus duration and instruction modality. *Vision Research, 48*(9), 1150-1163.
Sireteanu, R., Wolf-Dietrich, L., & Constantinescu, D. H. (1993). Distortions in two-dimensional visual space perception in strabismic observers. *Vision Research, 5*, 677-690.
Sjöstrand, J. (1981). Contrast sensitivity in children with strabismic and anisometropic amblyopia. A study of the effect of treatment. *Acta Ophthalmologica, 59*, 25-34.
Skoczenski, A. M., & Norcia, A. M. (2002). Late maturation of visual hyperacuity. *Psychological Science, 13*(6), 537-541.
Sloan, L. L. (1959). New test charts for the measurement of visual acuity at far and near distances. *American Journal of Ophthalmology, 48*(6), 807-813.

Sloan, L. L. (1968). The photopic acuity-luminance function with special reference to parafoveal vision. *Vision Research, 8*(7), 901-911.
Smith, A. (2015). Binocular vision: Joining up the eyes. *Current Biology, 25*(15), R661-R663.
Smith, A. T., Singh, K. D., Williams, A. L., & Greenlee, M. W. (2001). Estimating receptive field size from fMRI data in human striate and extrastriate visual cortex. *Cerebral Cortex, 11*(12), 1182-1190.
Smith, E. L., Chino, Y. M., Cheng, H., Crawford, M. L., & Harwerth, R. S. (1997). Residual binocular interactions in the striate cortex of monkeys reared with abnormal binocular vision. *Journal of Neurophysiology, 78*, 1353-1362.
Solomon, S. G., & Lennie, P. (2007). The machinery of colour vision. *Nature Review Neuroscience, 8*, 276-286.
Song, H., Levi, D. M., & Pelli, D. G. (2014). A double dissociation of the acuity and crowding limits to letter identification, and the promise of improved visual screening. *Journal of Vision, 14*(5), 1-37.
Spinelli, D., De Luca, M., Judica, A., & Zoccolotti, P. (2002). Crowding effects on word identification in developmental dyslexia. *Cortex, 38*, 179-200.
Stager, D. R., Everett, M. E., & Birch, E. E. (1990). Comparison of crowding bar and linear optotype acuity in amblyopia. *American Orthoptic Journal, 40*(1), 51-56.
Strasburger, H. (2005). Unfocused spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision, 5*(11), 1024-1037.
Strasburger, H., Harvey, L. O. J., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters. *Perception & Psychophysics, 49*(6), 495-508.
Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision, 11*(5), 1-82.
Stuart, J. A., & Burian, H. M. (1962). A study of separation difficulty and its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology, 53*, 471-477.
Subramanian, V., Jost, R. M., & Birch, E. E. (2013). A quantitative study of fixation stability in amblyopia. *Investigative Ophthalmology & Visual Science, 54*, 1998-2003.
Tailor, V. (2019, 10th February 2019). [Visual crowding in congenital idiopathic nystagmus].
Tailor, V., Dahlmann-Noor, A., Theodorou, M., & Greenwood, J. A. (2018). Visual crowding in congenital nystagmus, sensory deficit or image motion? *Investigative Ophthalmology & Visual Science, 59*(9), 1080-1080.
Tanaka, J. W., & Farah, M. J. (1993). Parts and wholes in face recognition. *Quarterly Journal of Experimental Psychology: A Human Experimental Psychology, 46*, 225-245.
Tang-Wai, D. F., Graff-Radford, N. R., Boeve, B. F., Dickson, D. W., Parisi, J. E., Crook, R., . . . Petersen, R. C. (2004). Clinical, genetic, and neuropathologic characteristics of posterior cortical atrophy. *Neurology, 63*(7), 1168-1174. doi:10.1212/01.wnl.0000140289.18472.15
Thomas, J. (1978). Normal and amblyopic contrast sensitivity functions in central and peripheral retinas. *Investigative Ophthalmology & Visual Science, 17*, 746-753.
Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research, 32*(7), 1349-1357.
Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23*, 775-785.

Tolhurst, D. J., & Thompson, I. D. (1981). On the variety of spatial frequency selectivities shown by neurons in area 17 of the cat. *Proceedings of the Royal Society of London. Series B. Biological Sciences, 213*(1191), 183-199.
Tootel, R. B., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., . . . Rosen, B. R. (1995). Visual motion after-effect in human cortical area MT revelead by functional magnetic resonance imaging. *Nature, 11*(375), 139-141.
Treisman, A. M., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*(97-136).
Treutwein, B. (1995). Adaptive psychophysical procedures *Vision Research, 35*(17), 2503-2522.
Tripathy, S. P., & Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size. *Vision Research, 42*, 2357-2369.
Tripathy, S. P., & Levi, D. M. (2008). On the effective number of tracked trajectories in amblyopic human vision. *Journal of Vision, 8*(4), 8.
Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology, 43*(2), 161-204.
van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2007). On the generality of crowding: Visual crowding in size, saturation, and hue compared to orientation. *Journal of Vision, 7*(2), 14.
van den Berg, R., Roerdink, J. B. T. M., & Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Computational Biology, 6*(1), e1000646.
van Dijk, J. A., de Haas, B., Moutsiana, C., & Schwarzkopf, D. S. (2016). Intersession reliability of population receptive field estimates. *NeuroImage, 143*, 293-303.
Van Essen, D. C., Newsome, W. T., & Maunsell, J. H. (1984). The visual field representation in the striate cortex of the macaque monkey: Asymmetries, anisotropies, and individual variability. *Vision Research, 24*, 429-448.
Verghese, P., McKee, S. P., & Levi, D. M. (2019). Attention deficits in Amblyopia. *Current Opinion in Psychology, 29*, 199-204.
Vickery, T. J., Shim, W. M., Chakravarthi, R., JIang, Y. V., & Luedeman, R. (2009). Supercrowding: Weakly measking a target expands the range of crowding. *Journal of Vision, 9*(2), 12.
Vlaskamp, B. N., & Hooge, I. T. (2006). Crowding degrades saccadic search performance *Vision Research, 46*(3), 417-425.
Vul, E., Hanus, D., & Kanwisher, N. (2009). Attention as inference: Selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General, 138*(546-560).
Wallis, T. S. A., Bethge, M., & Wichmann, F. A. (2016). Testing models of peripheral encoding using metamerism in an oddity paradigm. *Journal of Vision, 16*(2), 4-4.
Wallis, T. S. A., Funke, C. M., Ecker, A. S., Gatys, L. A., Wichmann, F. A., & Bethge, M. (2018). Image content is more important than Bouma's Law for scene metamers. *bioRxiv*.
Wandell, B. A., Brewer, A. A., & Dougherty, R. F. (2005). Visual field map clusters in human cortex. *Philosophical transactions of the Royal Society B: Biological Sciences, 360*(1456), 693-707.
Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. *Neuron, 56*(2), 366-383.
Wang, H., Levi, D. M., & Klein, S. A. (1998). Spatial uncertainty and sampling efficiency in amblyopic position acuity. *Vision Research, 38*(9), 1239-1251.

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2), 113-120.
Webber, A. L., & Wood, J. (2005). Amblyopia: Prevalence, natural history, functional effects and treatment. *Clinical and Experimental Optometry, 88*, 365-375.
Wertheim, T. (1894). Über die indirekte Sehschärfe. *Zeitschrift für Psychololgie & Physiologie der Sinnesorgane, 7*(172-187).
Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychologische Forschung, 4*(1), 301-350.
Westheimer, G. (1975). Visual acuity and hyperacuity. *Investigative Ophthalmology & Visual Science, 14*, 570-572.
Westheimer, G. (1981). Visual hyperacuity. *Progress in Sensory Physiology, 1*, 1-30.
Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research, 15*, 1137-1141.
Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160-168.
Wichmann, F. A., & Hill, J. N. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics, 63*(8), 1314-1329.
Wilkinson, F., Wilson, H. R., & Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays. *Journal of the Optical Society of America A, 14*(9), 2057-2068.
Witton, C., Talcott, J. B., & Henning, B. (2017). Psychophysical measurements in children: Challenges, pitfalls, and considerations. *PeerJ, 5*, e3231.
Wolford, G. (1975). Perturbation model for letter identification. *Psychological Review, 82*, 184-199.
Wyatte, D., Jilk, D. J., & O'Reilly, R. C. (2014). Early recurrent feedback facilitates visual object recognition under challenging conditions. *Frontiers in Psychology, 5*(674).
Yeatman, J. D., Rauschecker, A. M., & Wandell, B. A. (2013). Anatomy of the visual word form area: Adjacent cortical circuits and long-range white matter connections. *Brain and Language, 125*(2), 146-155.
Yehezklel, O., Sterkin, A., Lev, M., & Polat, U. (2015). Crowding is proportional to visual acuity in young ang aging eyes. *Journal of Vision, 15*(8), 23-23.
Yildirim, F., Meyer, V., & Cornelissen, F. W. (2015). Eyes on crowding: Crowding is preserved when responding by eye and similarly affected identity and position accuracy. *Journal of Vision, 15*(2), 21.
Yong, K. X. X., Shakespeare, T. J., Cash, D., Henley, S. M. D., Nicholas, J. M., Ridgway, G. R., . . . Crutch, S. J. (2014). Prominent effects and neural correlates of visual crowding in a neurodegenerative disease population. *Brain, 137*(12), 3284-3299. doi:10.1093/brain/awu293
Yu, C., & Rosenholtz, R. (2018). Similarity effects in peripheral vision: Improved representation of cuing? *Journal of Vision, 19*(10), 20-20.
Zeki, S. M. (1969). Representation of central visual fields in prestriate cortex of monkey. *Brain Research, 14*, 271-291.
Zeki, S. M. (1978). Uniformity and diversity of structure and function in the monkey prestriate visual cortex. *Journal of Physiology, 277*, 273-290.
Zeki, S. M., Watson, J. D. G., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. J. (1991). A direct demonstration of functional specialisation in human visual cortex. *The Journal of Neuroscience, 11*(3), 641-649.
Zhang, B., Zheng, J., Watanabe, I., Maruko, I., Bi, H., Smith, E. L., & Chino, Y. M. (2005). Delayed maturation of receptive field center/surround mechanisms in

V2. *Procedings of the National Academy of Sciences of the United States of America, 102*(16), 5862-5867.
Zhen, Z., Fang, H., & Liu, J. (2013). The hierarchical brain network for face recognition *PLoS One, 8*(3), e59886.
Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., . . . Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. *Procedings of the National Academy of Sciences of the United States of America, 109*(28), 11455-11459.