7. General Discussion

In this thesis, I strived to understand human visuospatial perception. Using psychophysical methods, eye-tracking and neuro-imaging techniques, I investigated how human observers localize and identify objects in order to understand the role of a number of factors in human spatial perception.

7.1. Summary of the Findings

In chapter 2, I made a comparison of perception and visual saccadic decision making with respect to localization performance. I placed two objects at similar eccentricities but in opposite hemifields. When the similarity of a shared feature (i.e. eccentricity in this case) increases, one should expect poorer decision-making (Freeman, Chakravarthi, & Pelli, 2012), which is what we found. Intriguingly, we observed that the eyes follow the perceived position of an object (as determined on the basis of the selected directions) rather than its physical position. This indicates a shared neural representation for position for decision-making and eye-movement planning. Despite a substantial number of studies and many detailed reviews devoted to the topic of the similarities and differences between perception and oculomotor planning and action (Schütz, Braun, & Gegenfurtner, 2011; Spering & Carrasco, 2015) – to my knowledge – prior to my study a comparison of perception and visual saccadic decision making with respect to localization had not been made in a straightforward task.

In chapter 3, I checked whether the manner of response in a target localization task had any effect on localization performance. My findings again support the notion that eye movements and decision-making have shared neural representations. One way to interpret this finding is that position information is interpreted and determined before saccades and manual responses are programmed (also evident from Chapter 2). This is also an important finding given the ongoing debate on whether eye movements reveal different mechanisms in crowding tasks than would manual (motor) movements (Deubel & Schneider, 1996; Harrison, Mattingley, & Remington, 2013; Harrison, Retell, Remington, & Mattingley, 2013; Lisi & Cavanagh, 2015). It could also be considered as a supporting finding for many existing studies in the literature. Most of those studies have used manual responses to unravel crowding mechanisms, whereas most natural visual behavior in which it plays an important role (such as visual search and reading) relies on making appropriate eye movements.
In chapter 4, I found that recognition and saccadic localization performance for a target object are similarly affected by crowding. Furthermore, also the similarity level of the target and the flankers has a similar effect on the magnitude of crowding, and on the perceived localization error. This is an important finding as it provides a way out of the division present in the literature between crowding being either the result of confusion of object positions or of a breakdown in feature integration (Freeman et al., 2012). This work also confirms the ideas suggested in recent modeling work that crowding is the result of uncertainty on both identity and location (van den Berg, Johnson, Martinez Anton, Schepers, & Cornelissen, 2012; van den Berg, Roerdink, & Cornelissen, 2010) and provides a guide for the further refinement of neurocognitive models of crowding.

Having investigated the effect of shared features of objects and response modes in various target localization tasks, I proceeded to examine the neural correlates of some aspects of spatial processing using fMRI.

In chapter 5, I mapped the visual field using a retinotopic mapping stimulus created using orientation contrast instead of luminance contrast (which is the conventional way used in the literature). In the conventional stimulation method used in fMRI, only a moving aperture is stimulated by means of luminance contrast while leaving the background blank. In my new method, both the aperture and the background are filled with oriented patches thus creating a continuously present stimulus. The visual field maps obtained using the new stimulus were consistent with those determined with the conventional retinotopy stimulus. However, the population receptive field sizes that were estimated were smaller and the locations of their receptive centers differed in such a way from those obtained using the conventional method that the foveal representation appeared to be more extended, in particular in higher order visual areas. I propose that the differences between the two mapping techniques could be explained by a combination of lateral inhibition and leakage of BOLD signals between voxels.

In chapter 6, the last experimental chapter of this thesis, I examined whether it is possible to use a shape illusion – that arises from manipulating low-level orientation information – to determine where visual shape information is processed in the brain. I had two possible expectations for the outcome: if the illusion is evoked by illusory contours, I expected to find traces of the illusion in the BOLD signal in early visual areas. If – on the other hand – the illusion arises at a later stage in the visual hierarchy, the illusion should not leave a trace in early visual cortex. In that case, the illusion-related signals would most likely only show up in higher order visual areas. I approached this by first determining the population receptive field properties of voxels and then analyzing specifically those voxels in the regions where we expected that the illusion might arise. Indeed, I found that the visual shape illusion did not leave a clear trace in the early visual cortex. However, a trend for an illusory shape-related effect was present in LO2, a visual area previously implicated in object recognition.
7.2. Discussion

7.2.1. Using eye movements to indicate perception

In my thesis, eye movements were used as an indicator of object identity and location (e.g. chapter 2-4). First, I examined the saccadic landing direction and used it as a replacement for the more common manual responses as an indication of 'which' object was chosen as a target. Second, the saccadic landing point was used as an indication of 'where' the observers thought that the target was located. The first parameter was treated as an indication of perception. This might appear confusing for the fact that both measurements represent perception, in a sense. However, due to the nature of the task (i.e. making an eye movement to the closer object), observers were expected to carry out two tasks. First to decide on which object is the target object and, second, to plan an eye movement towards this object's position. The former action is an indication for 'what' is selected and the latter one for 'where' the object is. In our study, the former task was considered to be a more direct indication of perception.

Although both manual and eye movement responses could be considered as indicators of perception, either of these could still be taken as a priority response. For instance in chapter 2, while the absolute landing location of the eyes was an important factor, picking the right side was more crucial for scoring a correct response. As the task goes by, subjects can modulate their attention so that each task could contain a prior task, as well as other sub-tasks. It might be of use if the subjects have a main focus during a task, as this might be considered as the most direct indication of perception. The tasks could thus be defined accordingly for studying conscious perception.

7.2.2. On the Context Dependency of Spatial Location Information

It is possible to elaborate on the results of this thesis in different ways in terms of the context dependency of spatial location information.

Based on the results of chapter 4, I conclude that the derivation of spatial information is a type of perceptual process that is also highly dependent on context, similar to the way that the derivation of identity information is (e.g. as in crowding or lightness and brightness illusions).

In contrast, in chapter 6 – albeit in a different context, namely that of shape recognition – I found no evidence for a context dependency of spatial location information at the level of the early visual areas. Since there were no differential signals, there was also no indication for feedback from higher-level areas.

Is there a way to explain this difference? In crowding, the distance between target and surrounding objects influence the recognition performance. This distance is known as the interference zone or “critical spacing”. Bouma (1970) found that the distance between objects should roughly be half the target eccentricity for interference to occur. Objects within this region will affect target identification negatively due to distraction. Although not all objects that are surrounded by other objects are crowded as it – by definition – requires objects to be similar in feature characteristics. Also the distortion of
the position information of individual objects or parts of objects depends on the strength of crowding or masking interactions. In other words, it will depend on the distances between objects or object parts. The approximate spacing between the objects in the studies in which I did find evidence for interactions in position information was \( \sim 1.05 \) deg of visual angle or less (chapter 4) while for the study in which it was not present this was \( \sim 2.3 \) deg of visual angle (chapter 6). The effect of objects’ distances on perceived position may require further studied. I propose further studies on this subject in section 7.4 below.

A further note is that the flankers presented in chapter 2 were not discussed in detail within the scope of localization performance. This is because these flankers preserved the same target-flanker orientation difference throughout the experiment. A crucial aspect of crowding is to be able modify the difference between target and flankers to control the ambiguity of the target object. Although the localization performance mildly differed between flanked and isolated conditions, the results did not depend on this contextual difference provided by the flankers. However, I observed that in case of crowding the context (created through the flankers) had a significant effect on the perceived position, evident from the findings in chapter 4.

### 7.2.3. On the Methods of Choice and on the Interpretation of Results

It is as important to decide how to answer a research question as it is to choose which question to ask. I will comment on what the implications of my findings could be for the future choice of methods?

Visual crowding concerns a decreased discrimination performance and has been demonstrated to be present for various stimulus aspects such as object size, color, shapes or letters (e.g. van den Berg, Roerdink, & Cornelissen, 2007). To this day, there has been a trend to explain crowding using a single model of the underlying mechanisms (Whitney & Levi, 2012). In this thesis, I limited my experiments to study crowding in just two domains, namely position and orientation. If a crowding model concerning relatively complex objects such as letters and faces were to be developed, my findings on perceived position and orientation should nevertheless be taken into account. In my view, their potential effect on higher-level cognitive processes should not be underestimated. To my knowledge, it is best to compare crowding models only if they use objects of similar complexity levels (in other words, similar properties).

The chapter 3 findings – which show that crowding is preserved when observers make an eye movement response rather than a manual keyboard response – might be considered good news for the use of eye movements in localization experiments. It indicates that eye movements are a valid measure to assess crowding. Although both manual and eye movement responses give consistent outcomes in terms of perceived position, eye movements provide the researcher more information in a similar amount of time. For this reason, I recommend eye movements as a response tool and not only as a method to control for fixation in psychophysics studies.
Based to my findings in chapter 5, I recommend that using "matched" mapping stimuli for different experiments might be an important consideration. In chapter 5, I found that orientation- and luminance-contrasted retinotopy provided different population receptive field (pRF) properties. For the same voxel, the estimated receptive field sizes and positions were significantly different between these two mapping techniques. This implies that depending on the stimulus, the same neural population (the neurons in a voxel) was assigned to a different part of the visual field. This implies that it may be important to use comparable stimuli for mapping as for addressing an experimental question. Thus far, selecting the mapping stimulus from a comparable stimulus type as used for the main experiment has been used only in a very few occasions, e.g. faces (Rajimehr, Bilenko, Vanduffel, & Tootell, 2014) and biological motion (Saygin & Sereno, 2008). The results of chapter 5 stress the importance of this approach once again. Using similar stimulus features for mapping and main experiment should better guarantee matched (i.e. similar) pRF properties. Moreover, it is important to verify this correspondence when elaborating on the results of any given experiment and when comparing results between studies.

7.3. Conclusions

I conclude that the perception of the spatial location and the identity of objects are highly intertwined processes that similarly affect human eye movements and perception. I found no evidence that saccadic decision-making is different from decisions that were revealed via manual responses only while observers had to fixate.

Moreover, in my neuroimaging studies, I found that using orientation rather than luminance contrast provides a better option for spatial mapping of the human visual cortex in particular for higher order areas, such as those involved in global shape perception. The results inform about the importance of using accurate visual field mapping techniques. My new and flexible method can have a wide application in future experiments.

I also used fMRI to localize the physiological mechanism underlying a visual illusion. In this illusion, the perceived position of the elements that make up a global shape is strongly altered by varying their orientation only. I found that, the responses of early visual areas V1 and V2 are driven by physical position. Additionally, my results tentatively suggested that the responses of a higher level visual area (LO2) correlated with the illusory change in shape perception.

7.4. Future Research & Applications

I found no difference between perception and eye movements and concluded that they share the same neural substrate for decision-making. This opens up eye movements as a valid way to probe a large number of aspects of visual perception, in particular in ophthalmic, neurological and psychiatric patients.

Since Chapter 5 proposes a new and complementary visual field mapping technique, future studies could use integrated stimuli that combine the properties of the current orientation-contrast and luminance-contrast modulated stimuli. More detailed – and
therefore potentially more informative – models could be created that take the actual orientation of each Gabor in each frame into account when modeling the BOLD responses. In that case, also periods of blank backgrounds (i.e. Gabors only present in the bar aperture or no Gabors present at all) could be included to enable modeling this aspect of the response as well.

I concluded that different mapping techniques could stress different properties of the visual system and that they might be used to assess different aspects of observers’ or patients’ brains. In the long term, having more accurate mapping techniques could also have a use in diagnosing patients. For instance, it was found that in autism spectrum disorder, patients have larger receptive field sizes for the extrastriate regions (Schwarzkopf, Anderson, Haas, White, & Rees, 2014). Other studies suggest reduced population receptive field sizes in schizophrenic individuals (Anderson, Tibber, Schwarzkopf, Rees, & Dakin, 2012). It would be important to verify these results using additional mapping techniques, to verify their extend and generalizability.

This could be useful since in psychiatric conditions such as schizophrenia, where patients appear to be less affected by context (see an effect either one way or another) (Dakin et al., 2005; Robol et al., 2013) one could also use various types of mapping techniques to probe higher order visual perception. These patients’ brains might reveal more or less differences depending on the mapping techniques used, which could lead us to a better understanding of their perception.

The retinotopic mapping stimulus in chapter 5 consists of a large amount of small gabor patches to render the percept of a moving bar aperture. This bar was an example of an illusory shape, in that there were no actual edges defined. The visual system inferred edges on the basis of the orientation differences in the stimulus. From these – in turn – the presence of an object may be inferred. But this might also occur on the basis of orientation-based grouping of the gabor patches. Although it was possible to infer a perceptually visible bar aperture by reading-out the activity in the cortex, I did not find a significant difference between the two modeling procedures, where one predicted the bar aperture to be perceived as a closed shape with only edges and the other as a filled-in shape (see edge and field models in chapter 5). Hence, reading out the details of the perceived shape is still not a fully performed task. What could be helpful is to define more distant border locations (by defining a thicker bar aperture) while taking the cortical magnification and spatial resolution of fMRI into account in future studies.

Irrespective of perceiving the bar aperture with edges only or as a filled-in shape, the visual field mapping results showed consistency of the perceived edge eccentricity with the luminance-contrast defined bar aperture. However, the presence of an illusory visible object cannot always be read-out from the cortex, as I also found in chapter 6. The spacing between object-parts – by affecting the degree of crowding and perceptual grouping – might be a factor that determines whether or not a shape is perceived. This could be further tested to further our understanding of shape perception.

Shape stimuli similar to the ones used in chapter 6 might also be used to find out where conscious perception occurs, both in healthy control and in neurological or neuropsychiatric patients. In cases of neurological lesion-based visual recognition impairments, it is often hard to determine what the origin of the impairment is. Stimuli
such as used in chapter 6 might provide a useful combination of clear low-level visual features (orientation) while also bridging the gap to higher-level visual information (objects). Using such stimuli, it might be possible to more selectively and precisely probe for deficiencies in object perception.

I investigated spatial perception using relatively simple, low-level visual stimuli and shapes and by varying relatively simple features such as orientation and location. Future studies might also consider studying spatial perception in visual scenes of higher complexity. Understanding the mechanisms behind spatial perception might be possible by developing biologically inspired models that can explain the results of experiments in which the stimulus complexity increases gradually. These gradual modifications could be manipulated in several ways such as by increasing object numbers, changing stimulus intensity per field of view, modification of positions and different degrees of jitter in contour integration. As the stimulus is made more complex in such ways, involvement of higher-order areas would be increased contingently and controlling the involvement of these areas could be studied systematically.

Finally, the 3T fMRI scans we have carried out in this thesis have given an idea of the functional organization of the cortex during recognition of spatial features. However, studying these aspects at 7T or higher and at a columnar resolution and at the level of different laminae and laminar connections could lead to a even better understanding of spatial processing in the human brain.
7.5. References


