Identifying the Function of Peripheral Vision in Early and Late Scene Processing

by

Jatheesh Srikantharajah

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2019

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

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

Abstract

The vast majority of information available in the surrounding environment to be processed by the visual system is located in the peripheral visual field. Despite this, our understanding of the functionality of peripheral vision is limited in regards to its influence on the affective and cognitive experiences we have in different settings and in response to different scenes. Thus, this thesis investigates the function of peripheral vision during the time course of scene viewing, and contrasts its role with that of the central visual field. We first establish that peripheral vision is sufficient for affective impressions of stimuli presented for brief periods of time (Experiment 1). In that study, participants perceive natural scenes presented in the peripheral vision field for as little as 50 ms to be more pleasant and interesting than urban scenes. We also contrast performance on evaluating scene gist and forming affective impressions during peripheral presentation of scenes and find that with a peripheral scene presentation of 50 ms, performance in identifying scene gist is reliably above chance. Since in the real world, any stimulus is likely to be located in the peripheral visual field before any fixation is made, our work suggests that the initial impression people form to real-world visual stimuli is the result of peripheral processing. In a follow-up study, we contrast performance on scene gist between central and peripheral vision through presenting conflicting scenes simultaneously to the two visual fields for brief (67 ms) periods of time (Experiment 2). That experiment revealed that despite the advantage in visual acuity and cortical magnification for the central visual field, it is peripheral vision that people rely on for the rapid identification of scene gist. Thus, we argue that

peripheral vision is specialized to rapidly identify the category to which a scene belongs. Finally, in our third experiment, we contrast eye movements during the time course of viewing a scene presented for 20 s during full vision and when vision is limited to either the central or peripheral visual field. In this experiment, we compare whether central and peripheral vision are associated with focal and ambient visual processes, as argued in the literature (Trevarthen, 1968). Focal vision involves detailed processing of specific objects or features in the environment, while ambient vision involves processing the locations of stimuli, detecting movement, and facilitating navigation. Prior work investigating changes in eye movements as a function of scene viewing time suggest that ambient visual processing occur primarily during early scene viewing, while focal visual processing occurs during late scene viewing (Pannasch et al., 2008). In our experiment, we find that while central vision is associated with focal eye movements and peripheral vision is associated with ambient eye movements, the peripheral visual field also plays a role in focal processing. This work establishes that not only does peripheral vision play a substantial role in early affective processing of scenes, it has some importance in late scene processing as well.

Acknowledgements

I would like to thank my supervisor, Dr. Colin Ellard, for his guidance and support. I would also like to thank my lab mates, along with my fellow CNS graduate students for their insight. I would also like to thank my friends and family, for providing constant positive energy throughout this process.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

At any point in time, the vast majority of visual information available in the environment is presented to the peripheral visual field. Central vision includes the fovea and the parafovea, which in total comprise the centremost 10° of the visual field, while peripheral vision includes everything outside of that (Larson & Loschky, 2009; Rayner, 1998). Historically, the function of peripheral vision has been associated with vision under low levels of illumination (Zele & Cao, 2015), processing of movement (McColgin, 1960; Findlay, 1982), and the visual guidance of posture (Howard, 1980) and eye movements (Matin, 1986). Moreover, information from the visual periphery influences behaviour, whether it is through guiding navigation (Barton et al., 2014) or providing affordances that facilitate grasping and other motor movements (Makris et al., 2013). Yet despite the vast disadvantage in terms of available visual information, it is central vision that is most prominent in processing colour (Mullen, 1991), object recognition (Nelson & Loftus, 1980), and resolving fine details.

During early work in vision science, one of the major differences posited to exist between central and peripheral vision that vision in the peripheral field often “produces little impression in consciousness” (Trevarthen, 1968). This claim is associated with research that shows that at any moment, there is a limited amount of visual information that can be perceived or recalled. For example, work on change blindness demonstrates that people often do not perceive changes to objects, scenes, and real-world environments (Simons & Levin, 1997), suggesting that the extent of visual information that is processed may be rather low. However, recent work (Cohen et al., 2016) has questioned this line of reasoning, arguing that in subjective impressions of the world, peripheral information from the surrounding environment is very much available in consciousness. Instead, while processing of information from the central visual field is advantaged in terms of the richness of detail and colour, peripheral information is also represented, but at a lower resolution.

1.1 Neural Mechanisms Underlying Central and Peripheral Vision

The differences between central and peripheral vision in terms of visual resolution are the result of their underlying neural mechanisms. Information from the central visual field is processed by the fovea and parafovea, which have a substantially higher number and density of cone photoreceptors, when compared to more peripheral parts of the retina (Rolls & Cowey, 1970). In fact, the density of cones decreases substantially with eccentricity (Curcio & Sloan, 1992), and the fovea in particular is specialized for high-acuity visual processing (Provis et al., 2013). Since cones are responsible for processing colour and other high spatial frequency information (Roorda et al., 1999), the decrease in the prevalence of cones in the peripheral visual field makes colour and high spatial frequency processing more challenging. This is further compounded by cortical magnification, as areas of the lateral geniculate nucleus (LGN) and striate cortex that receive projections from the fovea are disproportionately large (Connolly & van Essen, 1984), considering the limited amount of space in the visual field from which they process information. In fact, cortical magnification is directly related to visual acuity (Cowey & Rolls, 1970; Duncan & Boynton, 2003). One of the most substantial consequences of reduced cone density and cortical resources for the peripheral visual field is the phenomenon of crowding (Pelli et al., 2004), in which people struggle to identify objects in peripheral locations within an environment where there is distracting information nearby.

However, some researchers argue (Hansen et al., 2009, Rosenholtz et al., 2016) that while colour processing is worse in the periphery, it is not entirely absent. Essentially, it is argued that colour processing is a function of the distance between cones and not the density of cones, so while there is a dramatic decrease in cone density in the peripheral visual field, the distance between cones does not decrease to as great a degree (Rosenholtz et al., 2016). Thus, when people are able to recall the colour of their peripheral environment, it may indeed be the result of some form of peripheral colour processing, and not necessarily the result of prior foveal processing from a series of preceding fixations. Another possibility is that degraded processing of colour information in the peripheral visual field is in turn affected by higher-priority information from foveal vision, colouring perception of the peripheral scene.

1.2 Low Spatial Frequency Processing in Peripheral Vision

Peripheral vision finds itself in a position where it is tasked with processing information covering a substantially larger area of the visual field than central vision, while having a fraction of the neural resources to do so. In order to carry out this task, while peripheral vision is not highly adept at resolving high spatial frequencies, its ability to process low spatial frequency information is substantially better. Some evidence for the division of central and peripheral information by spatial frequency information comes from work in which scenes had either central or peripheral information filtered for low or high spatial frequencies (Cajar et al., 2016). Central vision was more strongly affected by the absence of high spatial frequencies than the absence of low spatial frequencies, while in peripheral vision, the absence of low spatial frequencies was more problematic than the absence of high spatial frequency information.

Peripheral vision exploits its processing capability for low spatial frequencies by condensing large amounts of similar information pooled over regions of the visual field in order to represent visual phenomena in terms of summary statistics (Rosenholtz et al., 2012). These summary statistics are comprised of a series of visual dimensions that can be processed from low spatial frequency information. The resulting visual ensembles include dimensions such as the average orientation, speed of motion, size, and brightness within certain regions of the peripheral visual field. Many of these ensemble statistics can be processed within 50 ms, and moreover, they can be calculated even when attention is reduced (Alvarez & Oliva, 2009). The low-level visual characteristics that can be computed in visual ensembles are associated with processing openness, symmetry, complexity, and depth. These characteristics are highly involved in the process of recognizing scenes, as they are diagnostic of the affordances that are present (Oliva & Torralba, 2001). Moreover, there is work showing that masking low spatial frequency information has a more disruptive effect on scene recognition when compared to high spatial frequency information (Loschky et al., 2007). In addition to facilitating recognition of scenes, there is evidence that when cognitive load is elicited by images of scenes, low spatial frequency information is required (Valtchanov & Ellard, 2015). In this study, participants experienced higher levels of cognitive load in response to urban scenes when compared to natural scenes only when low spatial frequency information was available in the images.

Furthermore, the neural underpinnings of peripheral vision are highly associated with areas involved in processing the structure of scenes. Information from the peripheral visual field projects to magnocellular ganglion cells, which are part of the dorsal visual stream. Importantly, dorsal stream activity is associated with processing the locations of stimuli in the environment (Mishkin et al., 1983), and is specifically involved in scene processing. In other work, peripheral stimuli activate neural regions responsible for scene processing substantially more when compared to foveal stimuli (Baldassano et al., 2016). In this study, the near periphery was particularly associated with scene processing regions, as the occipital place area (OPA) was maximally active at 10° of visual angle, while both the retrosplenial cortex and parahippocampal place area (PPA) were maximally active at 15° of visual angle.

1.3 Thesis Objectives

This thesis investigates the function of peripheral vision during the time course of scene processing. Since the literature on peripheral vision has heavily emphasized work on scene and object recognition, we aim to identify the function of peripheral vision in the formation of preferences. During early visual processing, we contrast the role of peripheral vision in judgments of gist and affective evaluations of scenes. Specifically, we aim to identify how rapidly affective and gist processing occurs in the peripheral visual field. Moreover, we aim to identify whether processing of gist is not just possible, but particularly reliant on the peripheral visual field by contrasting performance on this task between central and peripheral vision.

In addition, we aim to identify whether the function of central and peripheral vision changes during the time course of visual processing. With the exception of visual search, the majority of research on the functionality of the peripheral visual field has focused on its performance on tasks that involve rapid visual processing. As a result, there is a lesser amount of clarity about the role of peripheral vision during later scene processing. Therefore, one of the priorities of this thesis will be to compare peripheral visual processing between early and late scene viewing, and furthermore, to compare peripheral vision with both central and full vision at each of those intervals.

Chapter 2: Affective Processing of Scenes in Peripheral Vision

Introduction

In the visual perception literature, there is an extensive amount of research demonstrating that tasks such as object recognition (Thorpe et al., 1996), scene perception (Rousselet et al., 2005; Greene & Oliva, 2009), and emotion perception (Haberman & Whitney, 2007) can be completed rapidly within a single fixation. For any perceptual task completed within a single fixation, the majority of information that is processed will be in the peripheral visual field, since there is not sufficient time for a fixation to be made. Recent work has demonstrated that across a number of these dimensions, peripheral vision is sufficient to accurately carry out perceptual recognition tasks. Processing of scene gist in the peripheral visual field (Larson & Loschky, 2009) is especially accurate, as people can accurately identify the basic level category of a scene (e.g. ‘beach’, ‘street’) presented for 106 ms at close to a 95% rate. This ability holds true even in the far peripheral field, as global scene properties (e.g. natural/artificial) can be recognized at above chance levels at 70° eccentricity with a 80 ms presentation (Boucart et al., 2013). In terms of object recognition, the far periphery (70.5° eccentricity) can effectively recognize whether a natural scene presented for 28 ms contains an animal (Thorpe et al., 2001).

Moreover, people are capable of identifying the emotional gist of images of faces presented briefly in the peripheral visual field (Rigoulot et al., 2011). In this experiment, while performance on this task did decline with eccentricity, reaction times to fearful faces were lower than reaction times to neutral faces, suggesting that there is a bias towards preferentially processing fearful stimuli in the far peripheral visual field. Other work in the literature demonstrates that amygdala activation occurs during rapid processing of emotional gist in the periphery, but there is mixed evidence in regards to whether amygdala activation is higher for emotional stimuli in the peripheral (Bayle et al., 2011) or the central visual field (Almeida et al., 2015). Rapid face processing in the peripheral visual field is not merely limited to identifying facial emotion, as it has been demonstrated to be effective at identifying beauty (Guo et al., 2011), and moderately effective at judging cuteness (Kuraguchi & Ashida, 2015).

In contrast to the literature on face perception, there is a paucity of scene perception research investigating whether affective processing occurs rapidly in the peripheral visual field. Given that in the visual environment, the majority of information is present in the peripheral visual field, it is possible that initial impressions of scenes will be associated with peripheral visual functioning. While there is some evidence that people can categorize scenes in terms of whether they are pleasant or unpleasant in the periphery (Calvo et al., 2011), this work is limited to a 150 ms presentation interval. In that study, processing of gist was superior to affective categorization in terms of reaction times and in accuracy. However, it remains to be seen whether affective processing in the periphery is possible at more rapid time intervals where gist identification has been shown to occur. Similarly, it is unknown whether identifying the interestingness of stimuli with some accuracy can occur in the peripheral visual field. Thus, in the first experiment of this thesis, we set out to test whether peripheral vision was sufficient for the formation of affective impressions and evaluation of levels of interest for briefly presented scenes. Processing of low-level frequency information in the peripheral visual field was tested by measuring levels of openness. Moreover, this experiment contrasts whether processing of gist information would occur more rapidly than processing of affective information in peripheral vision.

2.1 Stimulus Validation

Prior to running Experiment 1, we conducted a pilot study in which we validated whether under normal viewing conditions, participants would identify the natural images in our dataset as different from urban images on the measures of pleasantness, interestingness, and openness.

2.1.1 Hypothesis

We predicted that under normal viewing conditions, participants would identify natural scenes to be more pleasant, interesting, and open than urban scenes. This would follow previous work by Kaplan and colleagues (Kaplan et al., 1989), in which under normal visual conditions, openness and pleasantness were among a host of dimensions in which natural and urban scenes differed.

2.1.2 Methods

2.1.2.1 Design

This was a within-subjects study in which stimulus type was manipulated (2: Urban, Natural).

2.1.2.2 Stimuli

A total of 104 images were gathered for this study from the Scene Understanding (SUN) Database (Xiao et al., 2010) and Flickr. The stimulus set included 52 natural scenes and 52 urban scenes. Natural scenes consisted of beaches, bodies of water (i.e. lakes, oceans) or forests. Urban images largely consisted of city streets and buildings. Scenes were stripped of their colour information in MATLAB, and through the use of the SHINE toolbox (Willenbock et al., 2010), these scenes were equalized on low-level visual characteristics such as luminance and spatial frequency. All scene images were kept at a size of 1920 x 1200 pixels and presented on a 60 Hz LCD monitor, extending to a maximum of 22.1° eccentricity at a viewing distance of 82 cm. A forehead and chinrest was employed in this study to maintain consistency with subsequent experiments in which it was employed to control viewing distance. The experiment was constructed through the use of the psychopy library in python (Peirce, 2007).

2.1.2.3 Procedure

All participants viewed the 104 natural and urban scenes, with the order of scene presentation randomized. Trials proceeded as follows: participants first viewed a central fixation cross for a period of 500 ms, and then viewed a scene for a total of 10 s. After that time interval, participants would answer a series of questions about the scenes, rating the images on pleasantness, interestingness, and openness. Measures of pleasantness, interestingness, and openness consisted of a Likert scale ranging from 1 to 7, ranging from low to high on those categories. There were two breaks scheduled during the study after the thirty-sixth and seventy-second trial.

2.1.2.4 Participants

A total of 8 undergraduates from the University of Waterloo participated in this study in exchange for course credit. All participants in this study had normal or corrected-to-normal vision. This study had 80% power to identify a very large effect size of $d = 1.5$. Power calculations were computed through the use of the pwr package in R (Champely et al., 2018).

2.1.3 Results and Discussion

As predicted, participants evaluated natural scenes differently than the urban scenes on all three measures. Natural scenes were considered to be significantly more interesting ($t(1,7) = 4.44$, $p = .003$), pleasant ($t(1,7) = 6.96$, $p < .001$), and open ($t(1,7) = 10.92$, $p < .001$), than the urban scenes under normal visual conditions. Our pilot testing confirms that under normal visual conditions, the natural images in our stimulus set were reliably different than the urban images on the three dependent measures.

2.2 Rapid Gist Processing

As indicated previously, the objective of the first experiment was to identify the time course of gist and affective processing in the peripheral visual field. Prior work in the literature indicates that the speed of gist processing is affected by the type of category that is being evaluated. Under normal visual conditions, processing of broader categories (i.e. man-made instead of ‘indoor’) occurs more rapidly than more specific categories (Joubert et al., 2007). In that study, processing of scene contexts was contrasted with processing of specific objects and beings in scenes. One of the striking results was that object processing and scene processing occur simultaneously. Moreover, delayed reaction times to scenes was associated with the presence of salient objects, indicating that there was a distractor effect. Supporting evidence for differences in the time course of gist processing by category type comes from work on age-related macular degeneration (AMD) (Tran et al., 2010). Since people with AMD have a reduced central visual field, object recognition is impaired, which in turn impairs processing of categories that involve information from objects. In that study, participants were worse at categorizing scenes presented for 300 ms when they were asked to identify whether the gist of the scene was indoor or outdoor, when compared to identifying whether it was natural or urban. Similarly, in studies where stimuli are presented to the far peripheral field at an eccentricity of 70°, classification of scenes by naturalness is superior to categorizing by indoor/outdoor (Boucart et al., 2013).

2.3 Hypotheses

We employed presentation lengths during this study that were brief enough such that people would not be able to initiate a saccade while a scene was being presented. Time intervals in this study ranged from 50 ms to 150 ms, which are brief enough to lie within the range of a single fixation. Since scene gist processing is particularly rapid in early perceptual processing, we anticipated that at the earliest time intervals for scene presentation, the ability to recognize gist would remain intact. In contrast, we anticipated that affective judgments and evaluations of interest would be accurate for the three longer time intervals (83.3 ms, 116.6 ms, 150 ms), but did not make a prediction for the shortest time interval (50 ms). The rationale for this prediction is that prior work suggests that peripheral gist processing is more rapid than affective evaluation (Calvo et al., 2011). This implies that at the earliest time intervals after scene presentation, identifying the gist of a scene will be possible, but processing of the affective quality of a scene will have yet to occur. That stated, since there is a lack of evidence that 50 ms is too rapid of a time interval for affective impressions to be possible in the peripheral visual field, we chose not to make a prediction for affective judgments and evaluations of interest for that time interval. Since judgments of facial beauty can occur accurately with a 100 ms prediction (Guo et al., 2011), we adopt the prediction that at a similar time interval in our study (83.3 ms), affective impressions of peripheral stimuli should occur. Finally, since judgments of openness within scenes occur extremely rapidly due to a reliance on rapidly processed low spatial frequency information (Oliva & Torralba, 2001), we anticipated that accurate judgments of openness would also be formed at the most rapid scene presentation interval. This line of reasoning resulted in the following hypotheses:

H1. Natural scenes will be rated as more pleasant than urban scenes when presented for at least 83.3 ms.

H2. Natural scenes will be rated as more interesting urban scenes when presented for at least 83.3 ms.

H3. Natural scenes will be rated as more open than urban scenes at all time intervals.

H4. Gist evaluation will be accurate at all time intervals.

Moreover, we were also interested in evaluating whether the accuracy of gist identification in peripheral vision was contingent on the type of scene. Under normal visual conditions, tasks involving rapid processing of scene gist demonstrate an advantage for natural scenes (Rousselet et al., 2005; Greene & Oliva, 2009). However, when scenes are presented for a relatively larger amount of time, the advantage for natural scene processing decreases. This resulted in the following prediction:

H5. Gist evaluation will be superior for natural scenes than urban scenes at the earliest two time intervals for this study.

2.4 Methods

2.4.1 Design

This study employed a 2 x 4 within-subjects design in which scene type (2: Natural, Urban) and scene duration (4: 50 ms, 83.3 ms, 116.6 ms, 150 ms) were manipulated.

2.4.2 Stimuli

The stimuli for this study consisted of the 104 natural and urban scenes that were validated in the previous experiment under normal visual conditions. Scenes had a size of 1920 x 1200 pixels, extending to a maximum of 22.1° eccentricity. However, the central 10° of the visual field was occluded through the presentation of a gaze-contingent black circle on the screen to ensure that only peripheral information from each scene was available for viewing. All images were presented on a 60 Hz monitor and viewed with a forehead and chinrest from a viewing distance of 82 cm. The duration of scene presentation was manipulated between trials. Images were presented for 50 ms, 83.3 ms, 116.6 ms, or 150 ms, with 26 trials per scene duration. Within each time interval, there were 13 trials in which natural scenes were displayed, and another 13 trials in which urban scenes displayed.

Scene gist was measured by asking participants to identify the basic category of the scene they had viewed as either ‘natural’ or ‘urban’. For each trial, we were thus able to evaluate the accuracy of gist identification. As in the validation study, interestingness, openness, and pleasantness of the scenes were evaluated through Likert scales from 1 to 7, where 7 represented the highest rating (i.e. most interesting).

2.4.3 Procedure

We employed a Mirametrix S2 eye-tracker recording gaze position at 60 Hz. Participants viewed 104 natural and urban scenes presented in a random order. Each trial began with the display of a fixation cross at the centre of the screen, which remained on the screen for a minimum of 500 ms. At this point, if participants were fixating the centre of the screen, the fixation cross would disappear and an image of the scene was presented along with the central 10° occluder. If participants were not fixating the centre, the fixation cross would remain on the screen until a central fixation was made. Scenes would be displayed for either 50 ms, 83.3 ms, 116.6 ms, or 150 ms. Following the image, a mask would be presented for a duration of 50 ms. After this point, participants would respond to the measures about scene gist and the three scales about interestingness, openness, and pleasantness. Participants were not limited in the period of time they were allotted to answer each question. There were also two scheduled breaks in the study to try to alleviate any fatigue participants might experience from continually looking at a monitor.

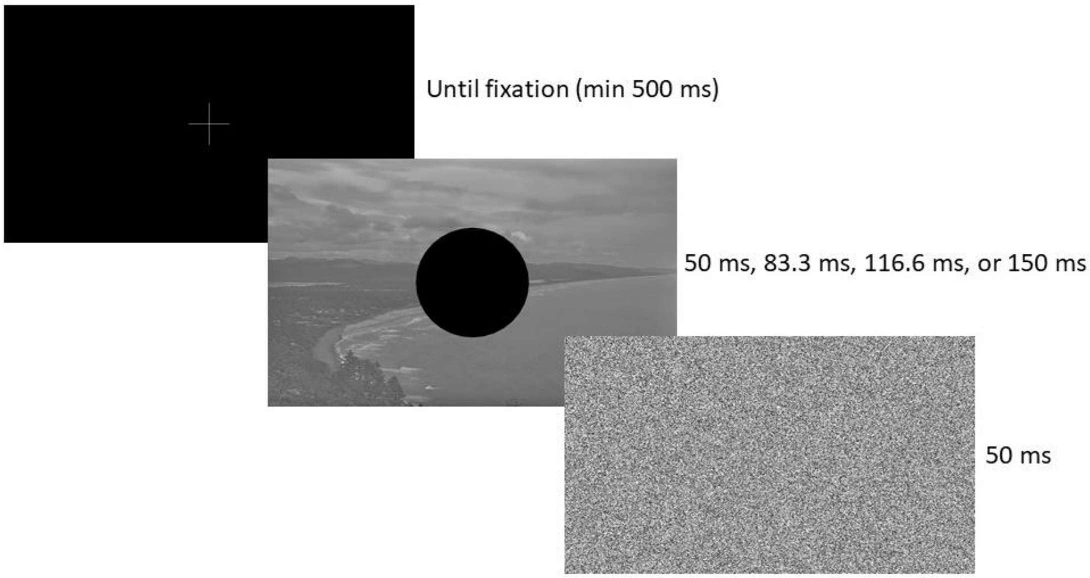

Figure 1. Experiment paradigm

2.4.4 Participants

A total of 50 undergraduates from the University of Waterloo participated for course credit. All either had normal or corrected-to-normal vision. This study had 90% power to identify a moderate effect size of $d = 0.5$.

2.5 Results

Since scenes in this study were presented for different time intervals, we tested whether the duration of presentation influenced performance on gist identification or ratings of pleasantness, interestingness, and openness.

2.5.1 Gist

We employed a 2x4 repeated measures ANOVA to identify whether the factors of scene type and time interval had a main effect on gist identification, and if there was any interaction between the two. A Greenhouse-Geisser correction was applied to the ANOVA due to violations of sphericity. The resulting ANOVA revealed that there was a main effect of scene type ($F(1,49) = 13.21$, $p < .001$) and time interval ($F(2.56,125.39) = 63.55$, $p < .001$). There was also a significant interaction between scene type and time interval ($F(2.41,118.13) = 8.44$, $p < .001$).

On average, there was a 7% difference in accuracy rate between gist identification of urban scenes ($M = 89.1\%$) and natural scenes ($M = 82.0\%$). A series of follow-up post-hoc t-tests analyzed whether gist identification for urban scenes was superior to natural scenes at each time interval, with a Bonferroni-corrected alpha level of $p < .0125$. Paired t-tests revealed that there were no significant differences between natural and urban gist identification at the 150 ms interval ($t(49) = 0.10$, $p = .092$), but there was a marginally significant effect at 116.6 ms ($t(49) = 2.58$, $p = .013$). Performance at the 150 ms period was very close to ceiling, with both urban and natural scenes identified at a 93% rate. At the 83.3 ms ($t(49) = 2.86$, $p = .006$) and 50 ms ($t(49) = 3.98$, $p < .001$) intervals, there were significant differences between natural and urban gist processing.

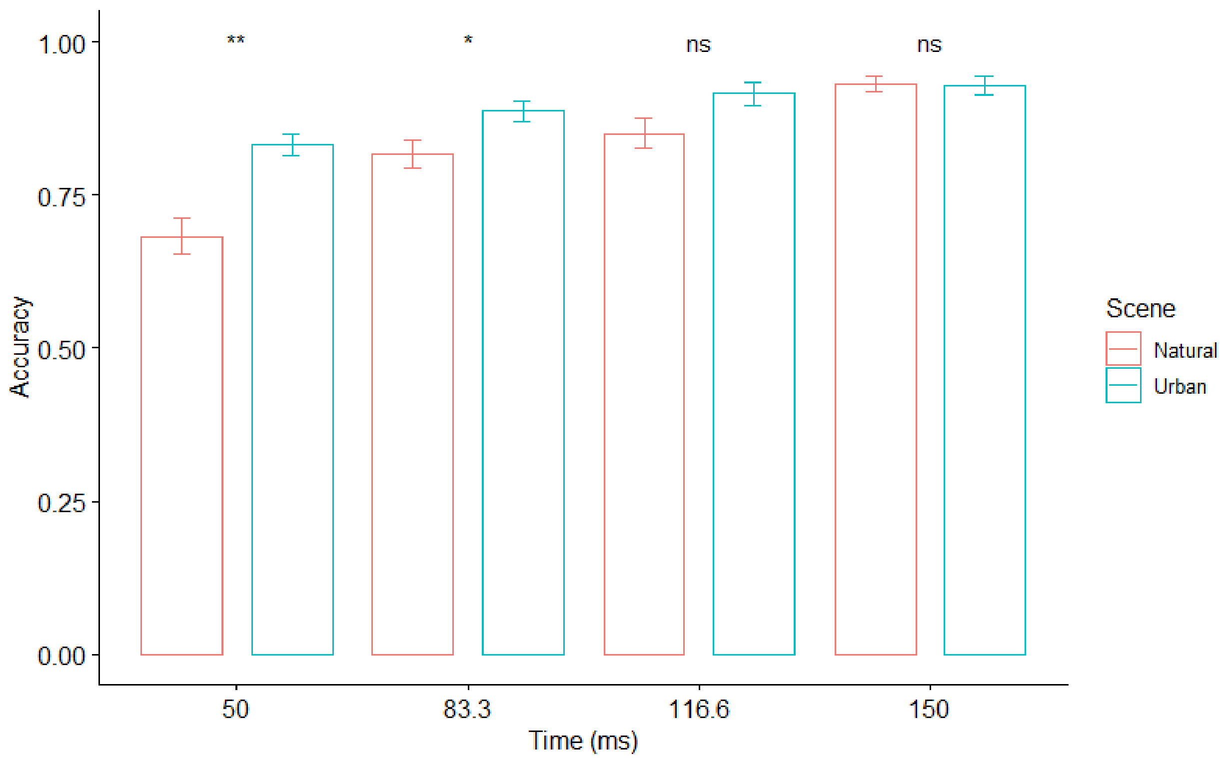

Figure 2. Accuracy of scene gist identification by stimulus type and presentation length. Red bars refer to performance on natural scenes, while blue bars indicate performance on urban scenes. Error bars indicate one standard error. * indicates p < .0125

A series of one-sample t-tests was carried out to evaluate whether gist identification was significantly higher than chance at each of the time intervals for scene presentation. For the 150 ms time interval, gist identification was significantly better than chance ($t(49) = 36.54$, $p < .001$).. At the 116.6, 83.3, and 50 ms intervals, gist performance on natural or urban scenes was compared separately with chance. A Bonferroni correction was applied to these comparisons, resulting in an alpha level of $p < .0125$. At the 116.6 ms interval, gist identification was significantly higher than chance for both natural ($t(49) = 14.73$, $p < .001$) and urban scenes ($t(49) = 22.55$, $p < .001$). At the 83.3 ms interval, gist identification was significantly better than chance for both natural ($t(49) = 14.43$, $p < .001$) and urban scenes ($t(49) = 22.93$, $p < .001$). At the 50 ms interval, gist identification was also significantly better than chance for both natural ($t(49) = 6.03$, $p < .001$) and urban ($t(49) = 19.10$, $p < .001$) scenes.

2.5.2 Pleasantness

A series of planned comparisons was carried out at each time interval comparing ratings of natural scenes with urban scenes on pleasantness. A Bonferroni correction was applied to these comparisons, resulting in an alpha level of $p < 0.0125$. Paired t-tests revealed that ratings of pleasantness were higher for natural scenes than urban scenes when presented for 50 ms ($t(49) = 6.43$, $p < .001$), 83.3 ms ($t(49) = 7.37$, $p < .001$), 116.6 ms ($t(49) = 9.19$, $p < .001$), and 150 ms ($t(49) = 9.98$, $p < .001$).

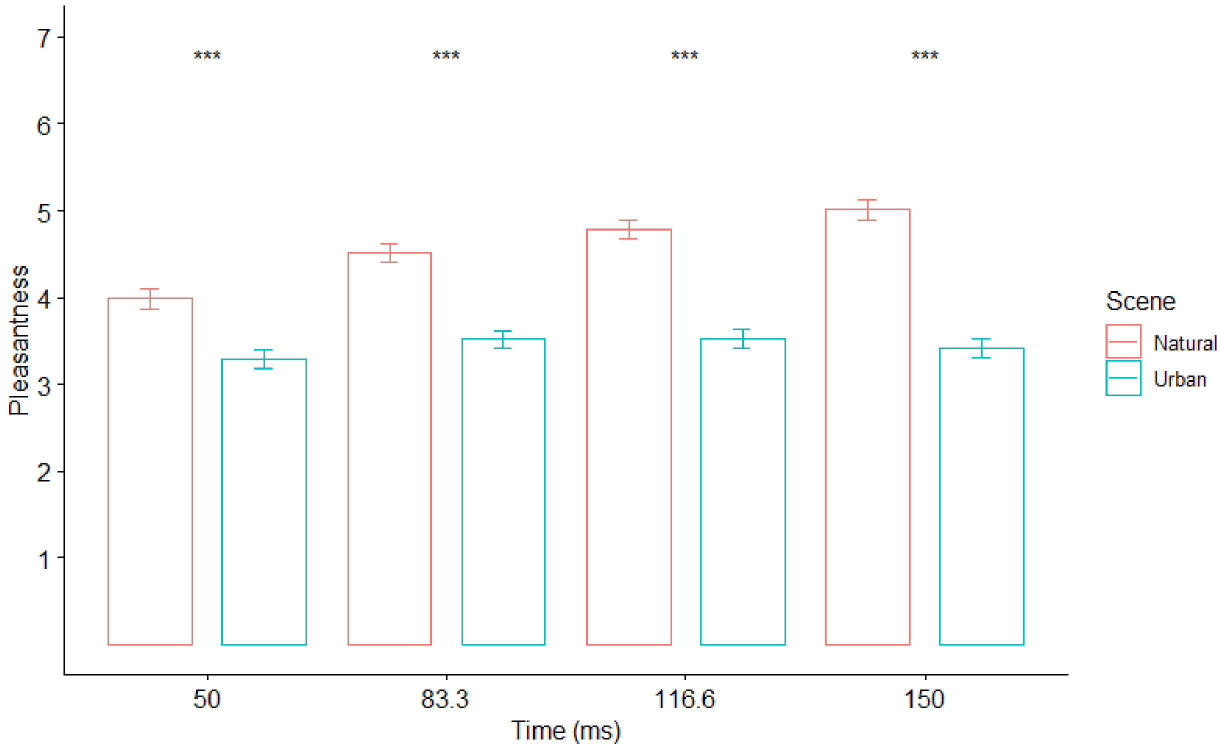

Figure 3. Ratings of pleasantness by scene type and presentation length. Red bars indicate natural scenes, while blue bars indicate urban scenes. Error bars indicate one standard error. * indicates p < .0125. ** represents p < .0025. *** represents p < .00025.

2.5.3 Interestingness

A Bonferroni-corrected series of planned contrasts were carried out at each time interval comparing ratings of interestingness on natural and urban scenes. The alpha level used for these comparisons was p < 0.0125. Paired t-tests revealed that ratings of interestingness were higher for natural scenes than urban scenes when presented for 50 ms (t(49) = 4.52, p < .001), 83.3 ms (t(49) = 6.68, p < .001), 116.6 ms (t(49) = 9.83, p < .001), and 150 ms (t(49) = 9.44, p < .001).

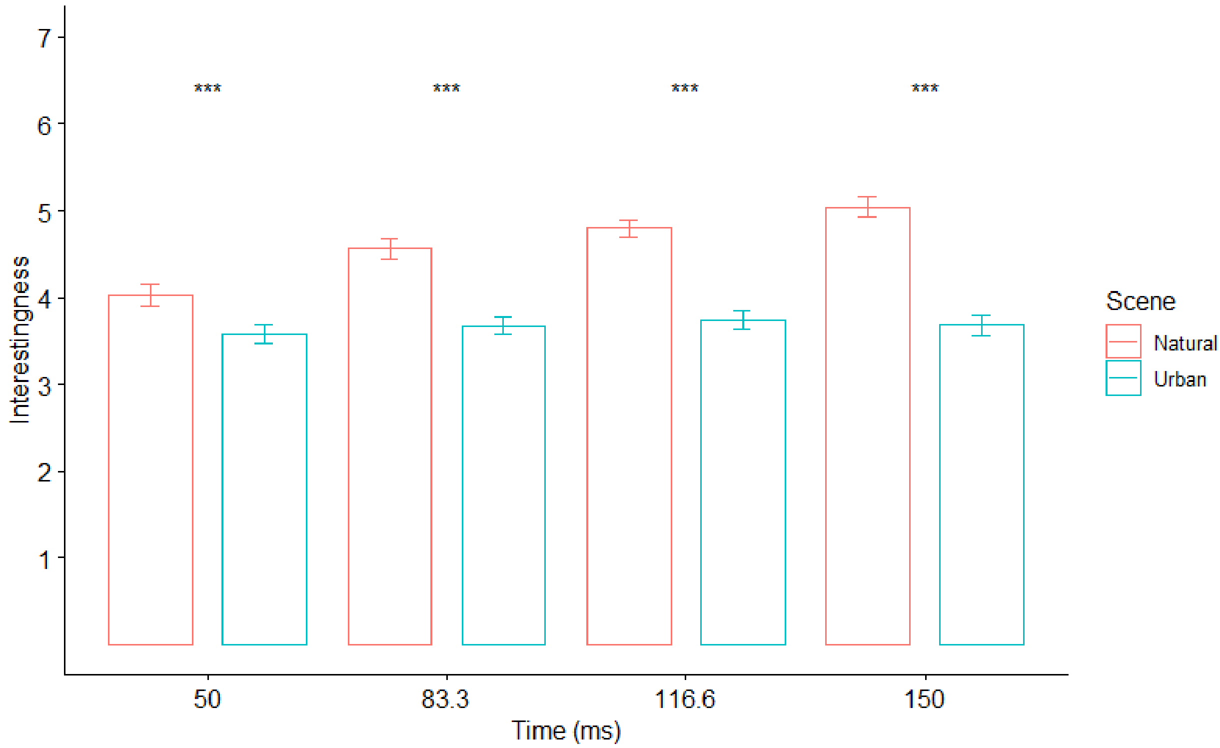

Figure 4. Ratings of interestingness by scene type and presentation length. Red bars indicate natural scenes, while blue bars indicate urban scenes. Error bars represent one standard error. * indicates p < .0125. ** represents p < .0025. *** represents p < .00025.

2.5.4 Openness

A series of Bonferroni-corrected planned comparisons were completed at each time interval, employing an alpha level of $p < 0.0125$ to contrast ratings of openness for natural and urban scenes. Paired t-tests revealed that ratings of openness were higher for natural scenes than urban scenes when presented for 50 ms ($t(49) = 8.28$, $p < .001$), 83.3 ms ($t(49) = 9.05$, $p < .001$), 116.6 ms ($t(49) = 12.50$, $p < .001$), and 150 ms ($t(49) = 12.66$, $p < .001$).

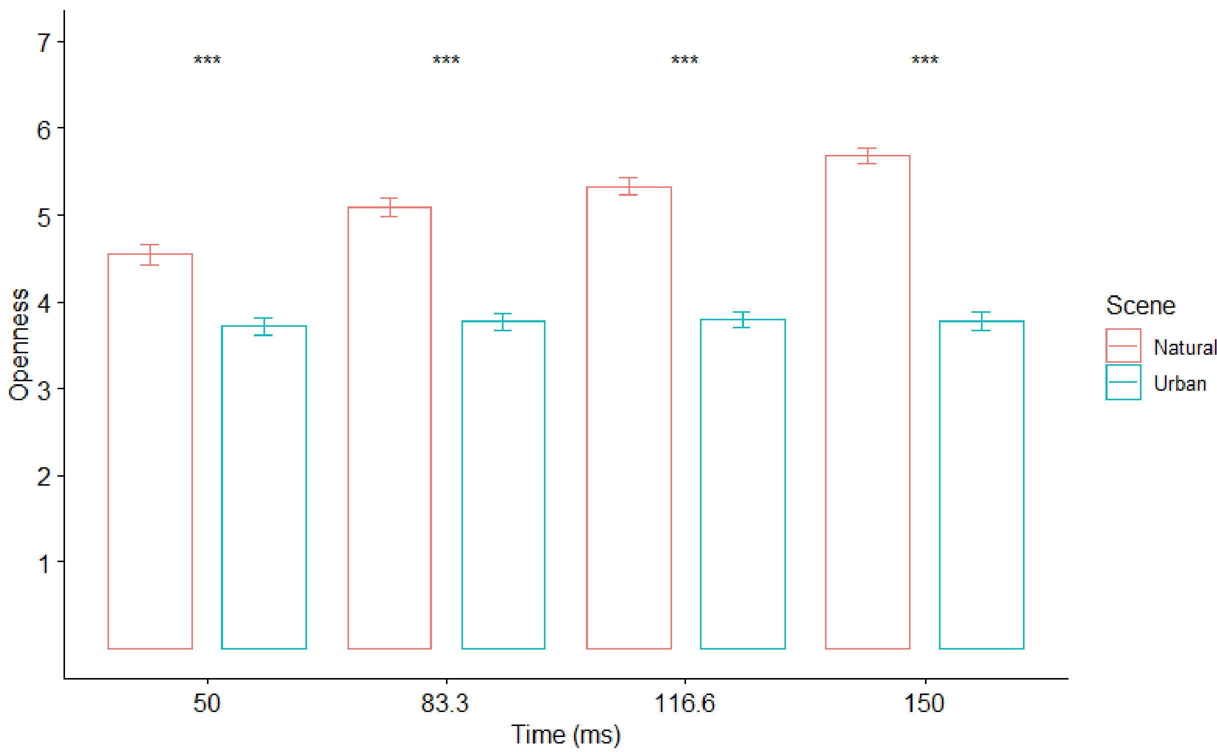

Figure 5. Ratings of openness by scene type and presentation length. Red bars indicate natural scenes, while blue bars indicate urban scenes. Error bars indicate one standard error. * indicates p < .0125. ** represents p < .0025. *** represents p < .00025.

2.6 Discussion

In this study, we demonstrated that people were capable of forming affective judgments about scenes that they viewed in the peripheral visual field for very brief periods of time. People rated natural scenes to be more interesting, pleasant, and open when compared to urban scenes, even when limited to a viewing period of only 50 ms. Similarly, people were capable of identifying the gist of a peripheral scene shown for the same time interval at reliable levels.

While previous research indicates that gist processing occurs extremely rapidly (Joubert et al., 2007), at the shortest time intervals in this study, accurate judgments of affective quality and evaluations of interest were just as possible as judgments of gist and openness for natural and urban scenes. That stated, it remains a possibility that at time intervals smaller than 50 ms, processing of gist is superior to affective processing. This study is especially interesting because it suggests that peripheral vision may play a prominent role in the formation of initial impressions to stimuli in our surrounding environment. It clarifies that the role of peripheral vision is not limited to simply providing information to identify the gist of peripheral stimuli. Nor is peripheral vision limited to identifying what parts of the visual environment are the most salient and determining where attention should be allocated, whether it is through guiding visual fixations (Matin, 1986) or influencing the direction of microsaccades (i.e. Laubrock et al., 2005). One explanation for this phenomenon comes from the fact that peripheral vision is associated with low spatial frequency processing, which is particularly important during rapid scene processing (Schnys & Oliva, 1994). Not only is low spatial frequency information from scenes associated with gist processing, it also plays a substantial role in affective processing (Valtchanov & Ellard, 2015). Low-level visual features are important in guiding subjective perceptions of naturalness in scenes (Berman et al., 2014), which in turn influence affective impressions (Coburn et al., 2019). Importantly, perceptions of naturalness predicted by low-level features have a larger impact on visual preference when judgments are faster (Kardan et al., 2015), suggesting that rapid affective impressions are especially reliant on processing low-level visual features.

Moreover, research using ERPs suggests that the P1 amplitude (100 – 150 ms) when viewing different scenes is larger for pleasant scenes when compared to neutral scenes when only low spatial frequency information was available (Alorda et al., 2007), while there is no effect when only high spatial frequencies are available. Since the majority of visual information available to be processed in the environment is located in the peripheral visual field, which is in turn specialized at processing low spatial frequency information, this suggests that peripheral vision may play a substantial role in not only determining which stimuli are the most important or worthy of attention, but also in the most immediate judgments of those stimuli.

2.6.1 Affective Impressions as a Cascading Process

Some supporting evidence for this result comes from work in object recognition, in which the orbitofrontal cortex (OFC) is activated within 80 to 130 ms of object presentation, processing low spatial frequency information from magnocellular inputs (Barrett & Bar, 2009). Activity within the first 150 ms of stimulus presentation in the OFC is associated with affective and associative processing (Shenhav et al., 2012). In addition, OFC activity spikes a second time between 200 and 450 ms, which is interpreted as indicating modification of initial affective impressions (Barrett & Bar, 2009). This may be particularly relevant for information in the peripheral visual field in which there is an immediate peripheral evaluation and then a subsequent foveation after a saccade. In fact, early OFC activity (135 ms) increases when foveal stimuli are congruent with previously presented unpleasant or pleasant peripheral scenes (D’Hondt et al., 2013), suggesting that the OFC is responsible for integrating peripheral affective information. We propose that for peripheral stimuli, there is an initial activation of OFC indicating an initial affective impression, and then as foveation provides more rich detail, there is a second spike in OFC activity integrating this information to modify the initial impression.

Due to the sheer magnitude of the peripheral visual field, every stimulus is likely to be viewed in the peripheral visual field before a central fixation is ever made. This suggests that any response to a visual stimulus may involve a cascading process that begins with peripheral processing. Salient information is processed in the peripheral visual field, which may lead to an immediate affective impression. Afterwards, there is a central fixation made on the salient region of space, which brings to bear focal processing that moderates or strengthens the initial impression. In conjunction, processing of contextual information from areas of visual space to the periphery of the item of interest further affects that impression (i.e. Joubert et al., 2007). Through such a pattern, peripheral vision may play a substantial role in guiding judgments about the world, but because salient information elicits fixations and visual resolution is particularly high in the central visual field, people associate affective judgments with central, focal viewing, instead of the preceding preconscious peripheral process.

2.6.2 Natural and Urban Scene Gist

Contrary to our prediction, we found that across a series of time intervals from 50 to 150 ms, gist processing in the peripheral visual field was more accurate in evaluating urban scenes than natural scenes. This was a particularly surprising result, as under normal visual conditions, there is an advantage for processing natural scenes, especially when scenes are presented for very brief periods of time (Rousselet et al., 2005; Greene & Oliva, 2009). This result also contrasted with prior studies on gist perception in the peripheral visual field, in which there is no difference in processing gist between natural and man-made scenes (Larson & Loschky, 2009). The differences between the results in our experiment and this previous peripheral gist processing study may come down to the duration of scene presentation. In that experiment, images were presented for 106 ms, indicating that by that point, processing of natural and urban scenes occurs at equal levels of competence in the peripheral visual field. This is not dissimilar to data from our experiment, as in the closest time interval (116.6 ms), the advantage of natural scenes when compared to urban scenes was only marginally significant.

There are a number of possible interpretations of this result. First, it may be the case that the advantage in processing urban scenes during our experiment was due to the fact that colour information was absent in our stimuli. In experiments where people have access to their entire visual fields, colour has been shown to improve processing of natural scenes (Rousselet et al., 2005). However, the magnitude of the effect of colour information in facilitating judgments of scene gist in their study wassubstantially lower than the magnitude of the difference in our experiment between natural and urban gist processing, which suggests that there is at least another factor at work.

A second possibility originates from the argument that object information is particularly important during scene processing (i.e. Biederman, 1987). Urban scenes typically include not only more objects than natural scenes, but also contain a greater diversity of objects. The argument for this account rests on the claim that because processing of objects within urban scenes occurs rapidly, that in turn facilitates identification of the broader scene category. In contrast, since natural scenes have a lesser number of objects, processing may be slower. In our experiment, many of the natural scenes consisted of lakes and oceans, and as a result would not contain a large number of objects. While there is some appeal to this line of reasoning, there are several possible objections to be raised. Object processing is facilitated by the presence of colour in scenes and is particularly prominent when colour is especially diagnostic of the object (Oliva & Schyns, 2000; Rossion & Pourtois, 2004). However, all scenes in this experiment were greyscale, so any facilitation effect of colour on object processing would not affect the current results. Moreover, object processing is highly associated with the central visual field (Nelson & Loftus, 1980), as the peripheral visual field is significantly worse at this activity. For example, crowding of objects (Pelli et al., 2004) would be more likely to occur in urban scenes than in natural scenes, due to the difference in object prevalence by scene type. Since there was no central information available whatsoever in this experiment, any posited enhancement of urban scene gist processing due to object prevalence would have to be the result of peripheral processing. Ultimately, any account invoking object processing as an explanation for improved gist processing of urban scenes in the peripheral visual field must describe how peripheral processing of objects facilitates peripheral processing of scenes, which will be challenging considering the flaws of peripheral vision in object recognition.

Comparing performance under normal conditions with performance with only peripheral vision present yields an interesting pattern. When viewing scenes with colour under normal viewing conditions, people are capable of identifying natural and man-made images over 90% of the time, even when scenes are only presented for 26 ms (Rousselet et al., 2005). By contrast, in our experiment, performance was substantially lower in the peripheral visual field at 50 ms (M = 75.8%) and 83.3 ms (M = 85.2%), though back to a similar level at 150 ms (M = 93%). What this indicates is not that peripheral vision is incapable of processing gist information from natural or urban scenes. Instead, processing of gist in the peripheral visual field simply takes longer to complete when compared to full vision. Our data suggests that in particular, natural scenes require more time for gist recognition in the peripheral visual field, even though they are processed the most rapidly with the full visual field. As a result, it may be the case that the central visual field plays a substantial role in processing information from natural scenes. This would provide an excellent explanation as to why colour information particularly facilitates processing of natural scenes, since visual processing of colour is especially reliant on the central visual field (Mullen, 1991).

One major limitation of this work was that the length of stimulus presentation was not short enough in any time interval to prevent either gist or affective processing to occur. Thus, any future experiment aiming to identify which of these processes occurs most rapidly in the peripheral visual field must use shorter time intervals. Similarly, given the influence of colour in facilitating object recognition and its potential impact on scene gist recognition, future work should investigate the importance of colour on the time course of affective and gist judgments in the peripheral visual field.

Chapter 3: Contrasting Gist Processing in the Central and Peripheral Visual Fields

Introduction

The preceding experiment demonstrated that peripheral vision is sufficient to identify scene gist and form affective impressions of briefly presented scenes. Having identified that peripheral vision is capable to carry out these tasks, we wanted to compare whether peripheral vision would be more prominent than central vision in the identification of scene gist. In previous work contrasting the two (Larson & Loschky, 2009), people were shown images of scenes to either their central or peripheral visual field and were then asked to identify the basic-level gist of those scenes. In their experiments, performance on scene gist was equal when a scene was presented to the peripheral visual field, when compared to presentation to the full visual field. In contrast, foveal vision (1°) was significantly worse than both peripheral vision and full vision, and in fact was only a few percentage points better than chance at this task. Central vision, which includes both foveal vision and parafoveal vision (1-5° radius), was also significantly worse than both peripheral and full vision. In a later study, when the extent of the available peripheral visual field in an experiment was extended to 180° of visual angle, the identification of scene gist in peripheral vision was not adversely affected even with a central scotoma covering 20° of the visual field (Loschky et al., 2019).

As in our first two experiments, the methodology for this line of research along these lines involves testing central and peripheral vision in isolation, through presenting scenes either to the centre or to the periphery. One exception comes from a study (Lukavsky, 2019) where compound scenes were created in which scene information was presented simultaneously for a brief period of time (33 ms) to the central (5.54° radius) and peripheral (5.54 - 11.54° radius) visual fields. In this study, information in the central and peripheral visual fields were either part of the same scene, or instead were from different scenes. People were instructed to attend to either the central or peripheral visual field, and then had to identify the gist of the scene to which they were attending. The results from this study indicated that judgments of scene gist were less accurate when people were attending to the peripheral visual field when compared to the central visual field. Moreover, scene gist judgments were also less accurate when a contrasting scene (as opposed to a similar scene) was displayed in the part of the visual field to which they were not attending.

We argue that in order to identify whether the peripheral visual field is more important for identifying scene gist than the central visual field, one should employ a methodology that directly contrasts central and peripheral vision. The experiments by Loschky and colleagues (Larson & Loschky, 2009; Loschky et al., 2019) show that in isolation, peripheral vision performs better than central vision, but they do not directly compare central and peripheral visual processing within one stimulus. Our objective was to identify how people make judgments of scene gist when presented with different information simultaneously in the central and peripheral visual fields, and unlike the work by Lukavsky, in the absence of any attentional instructions. Moreover, since we were interested in evaluating the absolute differences in gist processing between central and peripheral vision, we did not equate the amount of the available visual field between the two. We contend that by definition, the peripheral visual field possesses an advantage in terms of size when compared to central vision, which the latter compensates for by virtue of its advantage in cortical magnification (Larson & Loschky, 2009). Thus, in this experiment, we directly contrasted gist recognition in the central and peripheral visual field by presenting participants with compound scenes including different scene information in the central and periphery, and then measuring the probability at which people evaluate the gist of the scenes with either the central or peripheral scene.

3.1 Hypothesis

Judgments of scene gist are processed rapidly and are associated with magnocellular inputs to the dorsal visual stream (Barrett & Bar, 2009). Although the amount of cortical resources available to process information in central vision is much higher than peripheral vision due to cortical magnification, we argue that peripheral vision is specialized for the rapid identification of scene gist, due to its importance in the magnocellular visual pathway. When people are shown different scenes concurrently in the central and the peripheral visual field, peripheral information should be more strongly associated with judgments of gist. Thus, we have the following prediction for this experiment:

> H1. When people are shown compound scenes with different information simultaneously in the centre and periphery, judgments of gist will be more likely to indicate the peripheral scene.

3.2 Methods

3.2.1 Design

This experiment employed a within-subjects design in which gist responses (2: central scene vs. peripheral scene) were measured.

3.2.2 Stimuli

For this experiment, we gathered 240 images from the SUN database and from Flickr. These images belonged to four different categories: lakes, forests, home interiors, and streets. As in the first two experiments, scenes were stripped of colour information and equalized on low-level visual characteristics. Afterwards, a series of 120 compound scenes were created, in which information from one scene category comprised the central 10° of the image, while information from another scene category comprised the peripheral information in the image. Each scene category appeared in the central visual field for 30 scenes, and in the peripheral visual field for 30 scenes.

Central Scene	Peripheral Scene	# of Trials
Lake	Forest	10
Lake	Home Interior	10
Lake	Street	10
Forest	Lake	10
Forest	Home Interior	10
Forest	Street	10
Home Interior	Lake	10
Home Interior	Forest	10
Home Interior	Street	10
Street	Lake	10
Street	Forest	10
Street	Home Interior	10

Table 1. Number of trials in which each scene category was displayed in central or peripheral vision.

As in the first experiment, scene images were 1920 x 1200 pixels and viewed at a distance of 82 cm, extending to a maximum of 22.1° eccentricity. All scenes were presented on a 60 Hz LCD monitor.

3.2.3 Procedure

During the study, participants were informed that a scene would appear on the screen during each trial and that they would need to answer questions about it. They were instructed to fixate a central fixation cross and try to identify a scene displayed briefly afterwards. As in experiment 1, trials consisted of a central fixation cross, followed by the compound scene, and then followed by a mask. A Mirametrix S2 eye-tracker was employed to measure gaze position. The fixation cross was presented for the screen for a total of 500 ms. After that, the compound scene would be presented once gaze position was identified as being at the centre of the screen. Each scene was presented for a total of 66.6 ms, equal to four frames. Following each scene, a mask was presented for a total of 50 ms, equal to three frames. After this, participants would be asked to identify the gist of the scene and were instructed to guess if they were not sure. Each gist identification question consisted of a two-alternative forced choice question in which participants had to select either the basic-level scene category (e.g. lake, forest) present in the central visual field or the basic-level category present in the peripheral visual field. The order of scenes and the order of options for the gist question (whether the central or peripheral scene was the first option) was randomized for this study. To avoid fatiguing participants, there were two scheduled breaks during the study after the fortieth and eightieth trials.

3.2.4 Participants

A total of 20 undergraduates from the University of Waterloo participated in this study. All either had normal or corrected-to-normal vision. In experiment 1, the effect size of gist processing when compared to chance in the peripheral visual field was $d = 2.26$ at the shortest time interval (50 ms). We anticipated that the effect size of the difference between central and peripheral vision would be smaller, and as a result designed this study to have 85% power to identify an effect size of $d = 0.7$.

3.3 Results

For each trial, we recorded whether participants evaluated the gist of the compound scene as being the peripheral or the central scene. Then, for each participant, we calculated the average number of trials where the gist was identified as being the peripheral scene (herein described as peripheral trials) or the central scene (herein described as central trials). We conducted a 2x4 repeated measures ANOVA to identify visual location (2: central, peripheral) or scene type (4: lake, forest, street, home interior) had a main effect on gist accuracy. Due to violations of sphericity, a Greenhouse-Geisser correction was applied to this ANOVA. This analysis revealed a significant main effect of visual location ($F(1,19) = 283.41$, $p < .001$) and scene type ($F(2.31,43.88) = 14.36$, $p < .001$), but there was no interaction ($F(2.77, 52.61) = 0.75$, $p = .52$). Further comparisons were conducted to identify differences in gist identification between each scene type, with a Bonferroni-corrected alpha level of $p < .0083$. Gist identification of lake scenes was significantly less frequent than streets ($t(1,19) = 3.26$, $p = .0041$) and home interiors ($t(1,19) = 6.97$, $p < .001$), but not different than forests ($t(1,19) = 1.09$, $p = .29$). Performance for forest scenes was significantly less frequent than home interiors ($t(1,19) = 4.14$, $p < .001$), but not different from streets ($t(1,19) = 1.47$, $p = .16$). Performance for street scenes was also significantly less frequent than home interiors ($t(1,19) = 4.32$, $p < .001$).

For each scene type, we compared the frequency of identifying the gist of the scene as belonging to that category between trials in which it appeared in the centre and trials in which it appeared in the periphery. These comparisons were Bonferroni-corrected, resulting in an alpha level of $p < .0125$. Across lakes ($t(1,19) = 12.04$, $p < . 001$), forests ($t(1,19) = 11.28$, $p < . 001$), streets ($t(1,19) = 12.25$, $p < . 001$), and home interiors ($t(1,19) = 12.39$, $p < . 001$), participants were more likely to identify a scene as belonging to a certain scene category when it appeared in the peripheral visual field when compared to appearing in the central visual field.

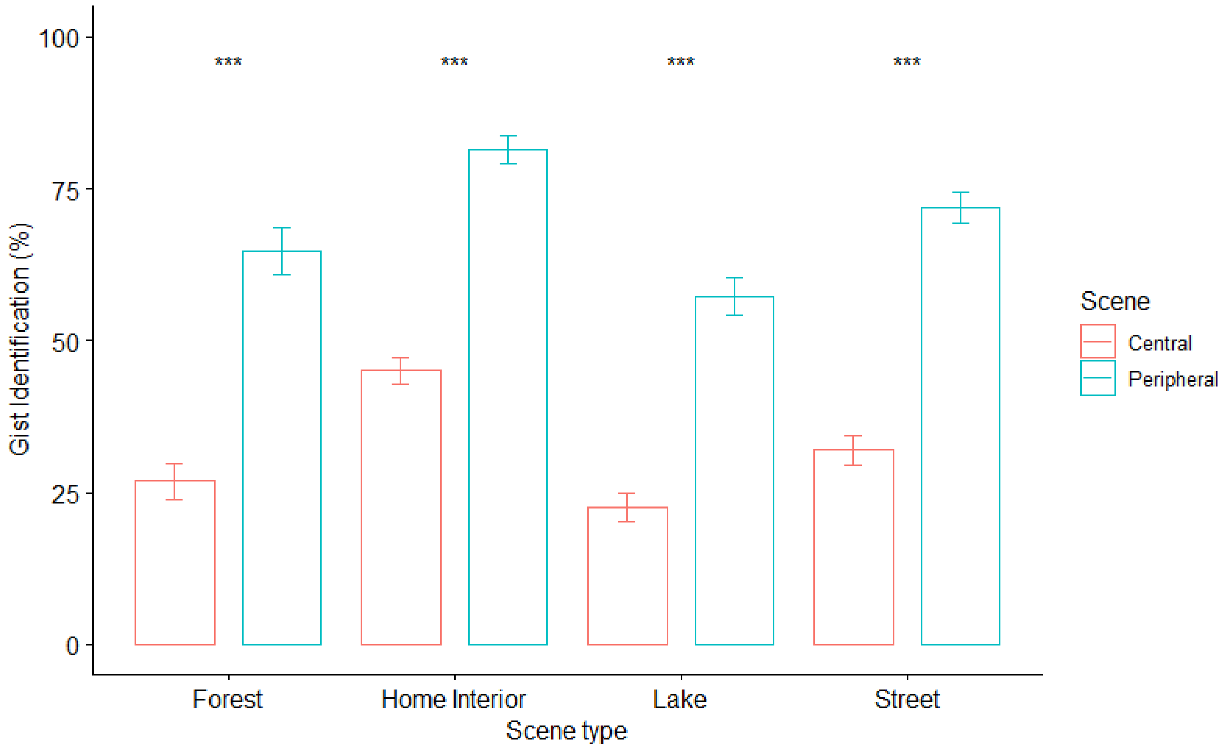

Figure 6. Identification of scene gist by scene type and visual field. Red bars indicate the percentage of trials in which participants identify the gist of a compound scene as being the central scene category. Blue bars indicate the percentage of trials in which participants identify the gist of a compound scene as being the peripheral scene category. *** indicates $p < .00025$.

3.4 Discussion

The results from this experiment showed that when presented with different scene information concurrently to the central and peripheral visual fields, participants were significantly more likely to identify the gist of the scene as being the peripheral scene. Thus, when given a choice between scene information in the central and peripheral visual fields, people rely on peripheral vision significantly more than central vision when making judgments about scene gist. Although there were differences between scene categories in overall gist identification, the advantage for identifying peripheral scenes held true across all scene types in the experiment. However, while we were able to demonstrate that there is an absolute peripheral advantage for gist processing, our experiment does not provide any evidence as to why this advantage exists. One possible explanation is that early scene processing is associated with a greater bias towards low spatial frequency information (Schnys & Oliva, 1994). Since peripheral vision is associated with low spatial frequency processing, this may be the catalyst for its advantage in gist recognition.

A competing explanation argues that the amount of information available to the peripheral visual field explains differences in the ability to categorize scene gist (Larson & Loschky, 2009). In their experiments, a marginal increase in the amount of information in the central visual field resulted in a greater increase in scene gist accuracy when compared to a marginal increase in the amount of information in the peripheral visual field. In addition, central vision requires substantially less information to be available in order to perform equally well at gist recognition when compared to the peripheral visual field. Due to the effects of eccentricity on cone density within the retina and the amplifying power of cortical magnification (Curcio & Sloan, 1992), a marginal increase in available image area within the central visual field increases visual acuity to a much greater degree than a similar increase in the peripheral visual field. However, when controlling for differences in cortical magnification, central vision is in fact less efficient than peripheral vision that what one would predict given its resource advantage (Larson & Loschky, 2009). An interesting direction for future work would be to identify whether this peripheral advantage in efficiency holds true not just when stimuli are presented individually to either the central or peripheral visual field, but when competing information is available in scenes in the centre and periphery. It would be particularly interesting to identify the amount of visual information required in the centre and periphery for judgments of gist to be equally likely between the two options. One could then control for cortical magnification and then identify whether central or peripheral vision is more efficient at processing gist.

Chapter 4: Do Eye Movements in Central and Peripheral Vision Reflect Focal and Ambient Processing?

Introduction

The previous experiments investigated the role of peripheral vision during various scene recognition and evaluation tasks during early perceptual processing. While there is a substantial amount of work investigating the functionality of the peripheral vision field in early visual processing, there is less work identifying the function of the periphery during later visual processing. One resulting question is whether the functionality of peripheral vision emphasizes early scene processing and whether central vision is more prominent during late scene processing. The literature on the time course of scene processing suggests that ambient visual processes may be associated with early scene processing, while late scene processing involves focal visual processes (Eisenberg & Zacks, 2016).

An interesting formulation of the differences between central and peripheral vision during extended visual processing is provided by early work identifying the existence of two separate visual systems – focal and ambient vision (Trevarthen, 1968). Trevarthen identifies focal vision with the central visual field, while he associates ambient vision with the peripheral visual field. Under this theory, the ambient visual field provides information about the organization of space, and the motion of any objects or beings that are located with that region of space. This information in turn organizes focal activity occurring within the central visual field, which is responsible for detailed visual processing within small regions of space. Trevarthen contends that the relative importance of focal and ambient visual processing is contingent on both visual illumination and task demands (Trevarthen, 1968). Not only is focal processing diminished during low levels of visual illumination as in scotopic vision (Zele & Cao, 2015), he argues that large, continuous motion inherently involves a decrease in central acuity, with more resources available for ambient, peripheral processing. While peripheral vision is particularly sensitive to the movement of surrounding stimuli (Finlay, 1982), it may be also especially active during the movement of the observer when compared to activities that involve a more static experience.

Ambient and focal processes are associated with different eye movement patterns (Unema et al., 2005). Ambient visual scanning involves larger saccade amplitudes and shorter fixation lengths, while focal processing involves shorter saccade amplitudes and longer fixation lengths. As people view scenes for longer periods of time, the average length of fixations increases, while the amplitude of saccades decreases. Essentially, when people first view a scene, the initial viewing pattern involves scanning large portions of the scene for shorter intervals, and as time progresses, people engage in more intense fixation of smaller, salient regions (Pannasch et al., 2008). A similar conception of ambient and focal processing defines fixations as ambient or focal as a function of preceding eye movements. Ambient fixations are those where the preceding saccade amplitude is below 5°, while focal fixations are those where the preceding saccade amplitude is greater than 5° (Pannasch & Velichkovsky, 2009). One reason for this distinction is because 5° represents the extent of parafoveal vision, with saccades into the periphery landing onto information represented in at a lower visual resolution (Trevarthen, 1968).

In other words, the prevalence of ambient and focal fixations is affected by the amount of time people look at a scene. During early scene processing, people engage primarily in ambient visual processing (Pannasch et al., 2008), and as time passes, they engage in more focal visual processing. A similar pattern occurs when people are exposed to novel visual events, as they engage in more exploratory eye movements, but after a period of time, they return to a focal mode of processing by fixating items of interest for a longer period of time to gather more detailed information (Eisenberg & Zacks, 2016).

While central and peripheral vision are in theory associated with focal and ambient processing, there have yet to be any studies testing if manipulating whether people have access to central or peripheral vision results in changes to ambient or focal visual processing. Thus, in this experiment, we aim to identify whether central and peripheral vision map onto focal and ambient visual processes, as measured through eye movements. In addition, we investigate how eye movements under central and peripheral visual conditions change as a function of the length of time that people view scenes.

4.1 Hypotheses

This experiment will include a peripheral condition when vision is limited to the peripheral visual field, a central condition in which vision is limited to the central vision field, and a control condition where information from the full visual field is available. We claim that peripheral vision should map onto ambient visual processing, while central vision should map onto focal visual processing. As a result, ambient and focal fixations should be highly associated with peripheral and central processing, respectively. Due to the association between peripheral vision and ambient processing, there should be differences in saccade amplitudes and fixation lengths between peripheral vision and both the control and central vision conditions. Moreover, we expect to replicate findings from previous work indicating that ambient eye movements are biased towards early scene processing, while focal eye movements are biased towards late scene processing. Thus, we propose the following hypotheses:

H1. In the control condition, fixation durations will increase and saccade amplitudes will decrease during the time course of scene viewing.

H2. Fixation lengths will be shorter in peripheral vision when compared to both central vision and the control condition for both early and late scene viewing intervals.

H3. Saccade amplitudes will be smaller in central vision when compared to the control condition for both early and late scene viewing.

H4. Saccade amplitudes will be larger in peripheral vision than either central vision or the control condition for both early and late scene viewing.

H5. Saccade amplitudes will be smaller at late scene viewing intervals than for early scene viewing intervals for peripheral vision.

H6. The prevalence of ambient fixations will be higher in peripheral vision than central or full vision.

H7. The prevalence of focal fixations will be higher in central vision than peripheral or full vision.

H8. In the control condition, the prevalence of ambient fixations should decrease between early and late scene processing.

4.2 Methods

4.2.1 Design

This study employed a 2 x 3 within-subjects design in which stimulus type (2: Natural, Man-made) and visual condition (3: Central vision, Peripheral vision, Full vision) were manipulated.

4.2.2 Stimuli

For this experiment, we gathered 90 images of natural and man-made scenes from the SUN database and from Flickr. Natural scenes consisted primarily of forests and bodies of water (i.e. lakes, rivers), while man-made scenes consisted primarily of streets and housing interiors. Scenes were stripped of colour information and equalized on low-level visual characteristics such as spatial frequency and luminance via the SHINE toolbox in MATLAB. All images in this study were 1024 x 768 pixels, extending to a maximum 18.5° eccentricity. Scenes were viewed on a CRT monitor with a refresh rate of 60 Hz at a 71.5 cm distance, while participants rested their head on a forehead and chinrest.

4.2.3 Procedure

An SR Eyelink 1000 remote eye-tracking system recording gaze monocularly at 1000 Hz was used to record eye movements during the experiment. Fixations and saccades were detected online, with a spatial accuracy greater than 0.5°. Saccades were identified by deflections in eye position greater than 0.1° with a minimum acceleration of $8000°/s^2$ and a minimum velocity of 30°/s, maintained for at least 4 ms. Eye position was gathered for the dominant eye of each participant. At the start of each trial, a fixation cross was presented at the centre of the screen. When participant gaze was identified as being at the centre of the screen, the scene was presented. Each scene was displayed on the screen for a duration of 20 seconds. Participants were instructed to freely explore the scene and ensure that their gaze remained within the area of the monitor and not outside of the scene. For the central and peripheral vision conditions, we employed a gaze-contingent window and scotoma. Thus, in central vision, only the area of the screen within the central 10° was visible, with the rest of the screen blacked out. Similarly, for peripheral vision, a black circle covered the central 10° of the screen. When participants' gaze was outside of the boundaries of the scene, the entire scene was occluded, serving as a cue for participants to return their gaze back onto the scene. In total, each participant viewed 30 scenes with central vision, 30 scenes with peripheral vision, and 30 scenes with the full visual field. Visual condition was subdivided by the manipulation of scene type, such that within each visual condition, 15 natural scenes and 15 urban scenes were presented. The order of scenes and the visual condition under which each scene was presented was randomized. There were two scheduled breaks during the study after the thirtieth and sixtieth trials.

4.2.4 Participants

A total of 16 undergraduates (11 female, 5 male) with normal or corrected-to-normal vision from the University of Waterloo participated in this study. This study had 80% power to identify an effect size of $d = 0.75$.

4.3 Data Analysis

All data analysis was carried out using R. Blinks were identified as periods in which pupil information was missing. Saccades or fixations occurring within a 100 ms interval of a blink were discarded. Fixations longer than 1000 ms and shorter than 80 ms were also removed from the dataset. Trials in which the total blink time was equal to or greater than 5 s (equal to 25% of the trial length) were excluded. Moreover, since previous studies indicate that fixations and saccades are right-skewed (Pannasch et al., 2008), we conducted statistical analyses with median fixation and saccade data for the intervals of interest, instead of using the means.

4.4 Results

4.4.1 Fixation Lengths

We employed a 2x3 repeated measures ANOVA to identify whether time interval (2: early, late) or visual condition (3: control, central, and peripheral) had a main effect on the duration of fixations. Early time intervals were defined as the first four seconds of scene viewing, while late time intervals were defined as the last four seconds of scene viewing. A Greenhouse-Geisser correction was applied due to a violation of sphericity. There were significant main effects of time ($F(1,15) = 64.13$, $p < .001$) and visual condition ($F(1.68,25.2) = 9.84$, $p = .001$), along with a significant interaction between time and visual condition ($F(1.69,25.41) = 5.94$, $p = .01$). Planned follow-up contrasts revealed that during early scene viewing, fixation lengths in the control condition were significantly longer than in the peripheral condition ($t(1,15) = 4.08$, $p < .001$). In contrast, there were no significant differences in fixation length between central and peripheral vision during early ($t(1,15) = 0.77$, $p = .45$) processing. During late scene viewing, there were no differences in fixation length between the control and peripheral conditions ($t(1,15) = .89$, $p = .39$). There was, however, a significant difference between central and peripheral fixation lengths during late scene viewing ($t(1,15) = 3.61$, $p = .003$), but it was in the opposite direction to what we predicted. Surprisingly, exploratory comparisons revealed that fixation lengths were shorter under central vision than the control condition during both early ($t(1,15) = 3.38$, $p = .004$) and late ($t(1,15) = 4.25$, $p < .001$) scene viewing.

As predicted, there was a significant increase in fixation length during the control condition between the early and late time intervals ($t(1,15) = 3.64$, $p = .002$). Exploratory comparisons revealed that in both central ($t(1,15) = 6.60$, $p < .001$), and peripheral ($t(1,15) = 8.60$, $p < .001$) visual conditions, fixation lengths were shorter during early scene viewing than late scene viewing.

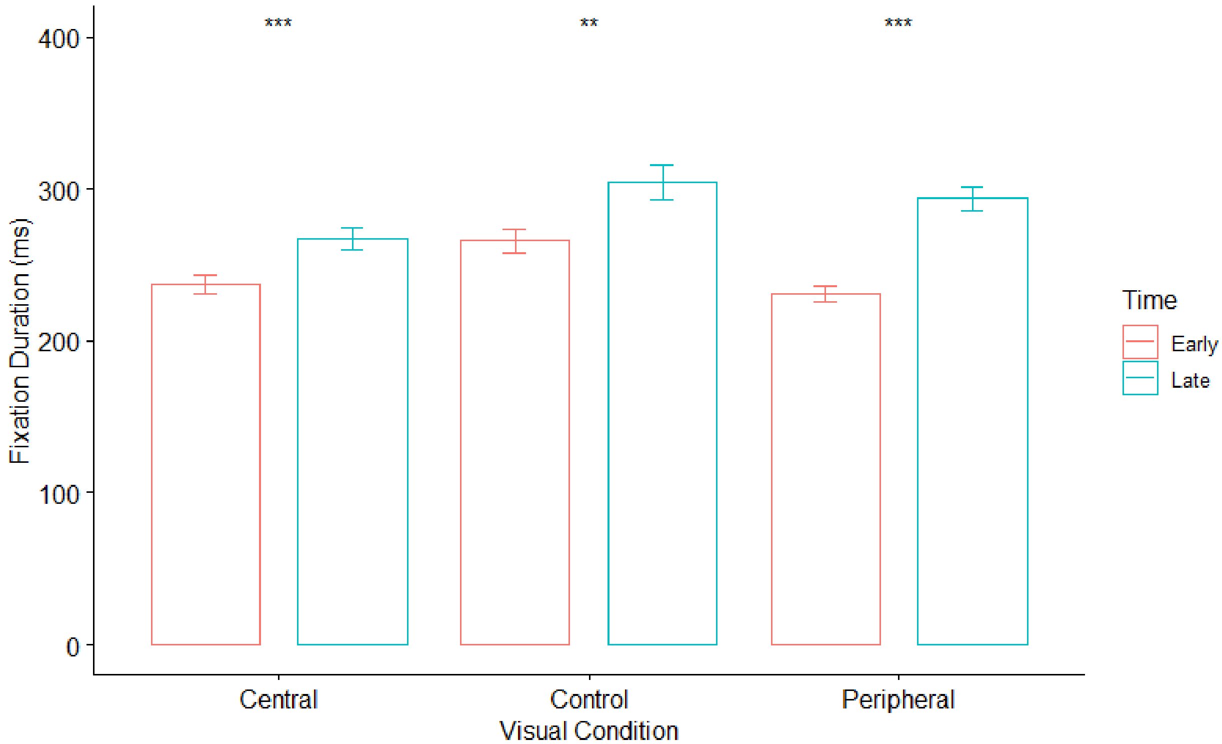

Figure 7. Fixation length by visual condition and early/late scene processing. Red bars indicate early time intervals (0-4 s), while blue bars late time intervals (16-20 s) when people view scenes. Error bars indicate one standard error. * indicates p < .05. ** indicates p < .01. *** indicates $p < .001$.

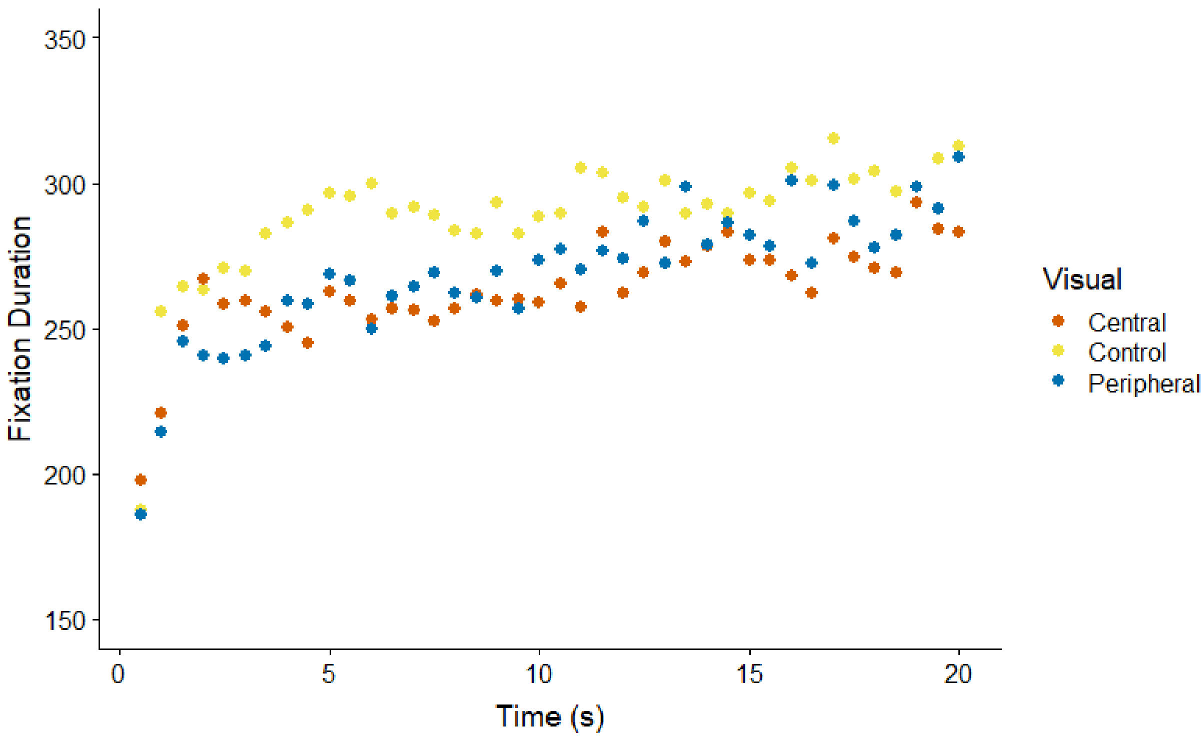

Figure 8. Median fixation lengths across central vision, peripheral vision, and the control condition over a series of 500 ms time intervals while viewing a scene.

4.4.2 Saccade Amplitudes

A 2x3 repeated measures ANOVA was employed to identify the effects of time interval (2: early, late) or visual condition (3: control, central, and peripheral) on saccade amplitudes. Early time intervals were defined as the first four seconds of scene viewing, while late time intervals were defined as the last four seconds of scene viewing. A Greenhouse-Geisser correction was applied due to a violation of sphericity. There were main effects of time ($F(1,15) = 15.60$, $p = .001$) and visual condition ($F(1.27,19.01) = 146.15$, $p < .001$) on saccade amplitude, along with a significant interaction ($F(1.58,23.69) = 19.92$, $p < .001$). Planned contrasts revealed that there were significant differences between the control condition and the peripheral condition on saccade amplitude at both early ($t(1,15) = 13.73$, $p < .001$) and late ($t(1,15) = 7.36$, $p < .001$) scene viewing intervals. Peripheral vision also differed significantly from central vision in saccade amplitude during early ($t(1,15) = 16.34$, $p < .001$) and late ($t(1,15) = 8.61$, $p < .001$) intervals. Moreover, central vision differed from the control condition in saccade amplitude during both early ($t(1,15) = 8.49$, $p < .001$) and late ($t(1,15) = 5.36$, $p < .001$) time intervals.

As predicted, saccade amplitudes decreased significantly between early and late scene processing in peripheral vision ($t(1,15) = 5.79$, $p < .001$). However, in both central vision ($t(1,15) = 0.94$, $p = .36$) and the control condition ($t(1,15) = 1.24$, $p = .23$), there were no differences between saccade amplitudes when the first four seconds of scene viewing were compared with the last four seconds of scene viewing.

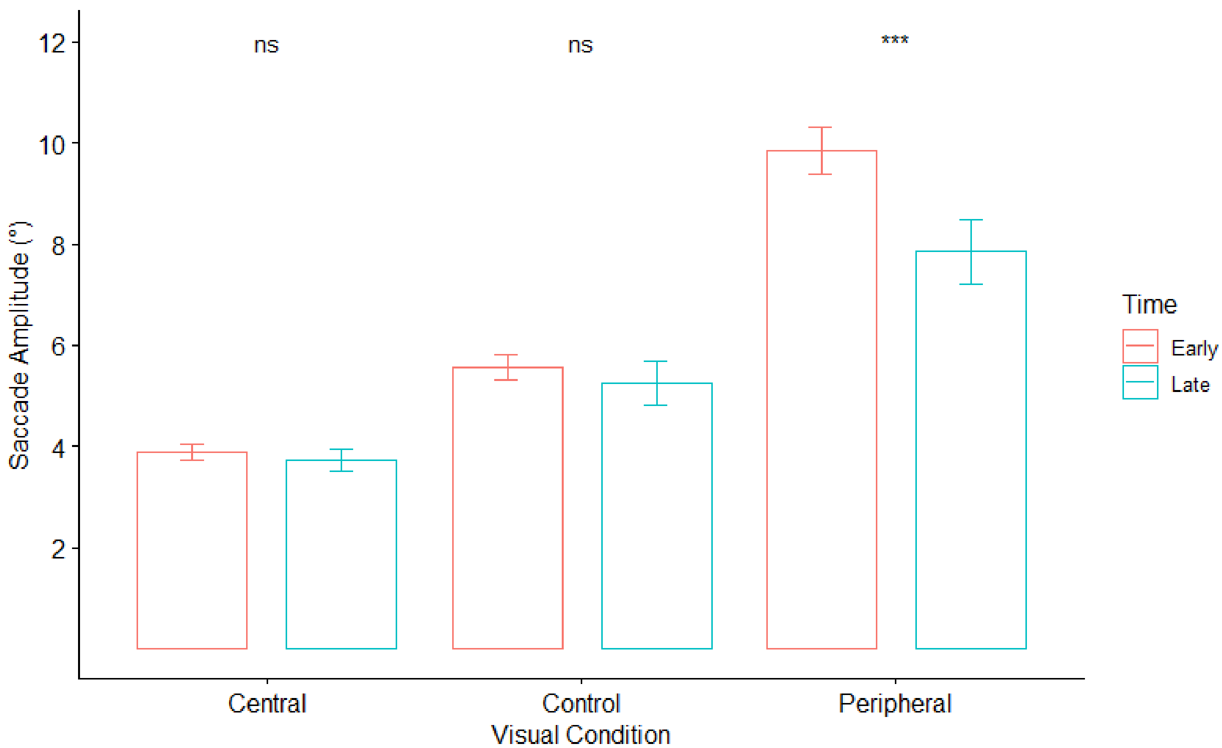

Figure 9. Saccade amplitudes by visual condition and early/late scene processing. Red bars indicate saccade amplitudes during early scene processing, while blue bars indicate saccade amplitudes during late scene processing. Error bars indicate one standard error. *** represents $p < .001$.

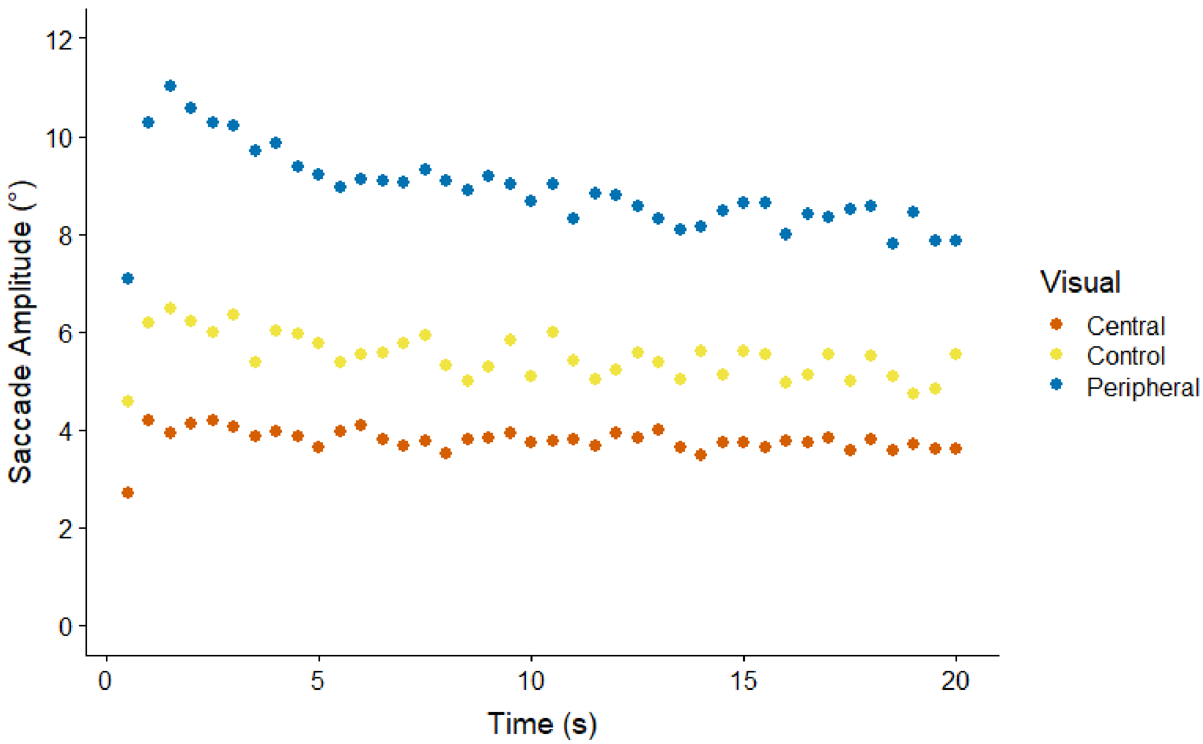

Figure 10. Median saccade amplitudes for central vision, peripheral vision, and full vision over 500 ms time intervals during scene viewing.

4.4.3 Focal and Ambient Fixations

Ambient fixations were defined as fixations in which the preceding saccade had an amplitude greater than 5°, while focal fixations consisted of fixations in which the preceding saccades was smaller than 5°. Thus, any fixation had to be either a focal fixation or an ambient fixation. For each trial, we calculated the prevalence of these two types of fixations as a percentage by dividing the number of ambient fixations by the total number of fixations.

We employed a 2x3 repeated measures ANOVA to identify whether time interval (early vs. late) or visual condition (control vs. central vs. peripheral) had a main effect on the prevalence of focal and ambient fixations. A Greenhouse-Geisser correction was applied to the ANOVA due to violations of sphericity. Early time intervals consisted of the first four seconds of scene presentation (Unema et al., 2005), while late time intervals consisted of the last four seconds. There was a significant effect of time ($F(1,15) = 10.13$, $p = .006$) and visual condition ($F(1.32,19.84) = 181.57$, $p < .001$), and also a significant interaction ($F(1.69,25.3) = 15.27$, $p < .001$). Planned comparisons revealed that in the control condition, the prevalence of ambient fixations was significantly higher than in central vision ($t(1,15) = 8.27$, $p < .001$). In peripheral vision, ambient fixations were significantly more prevalent than the control ($t(1,15) = 12.91$, $p < .001$) and the central conditions ($t(1,15) = 13.29$, $p < .001$). In the control condition, there was a significant decrease in ambient fixations between early ($M = 52.8\%$) and late ($M = 47.5\%$) scene viewing ($t(1,15) = 2.25$, $p = .040$).

We conducted two further exploratory comparisons about the relationship between time interval and visual condition on the prevalence of ambient and focal fixations. The first comparison demonstrated that in peripheral condition, there was a significant decrease in the prevalence of ambient fixations between early ($M = 87.0\%$) and late ($M = 73.2\%$) scene viewing ($t(1,15) = 4.66$, $p < .001$). In the second comparison, there was no difference in the control condition in the prevalence of early ($M = 28.6\%$) and late ($M = 28.5\%$) ambient fixations ($t(1,15) = 0.09$, $p = .93$).

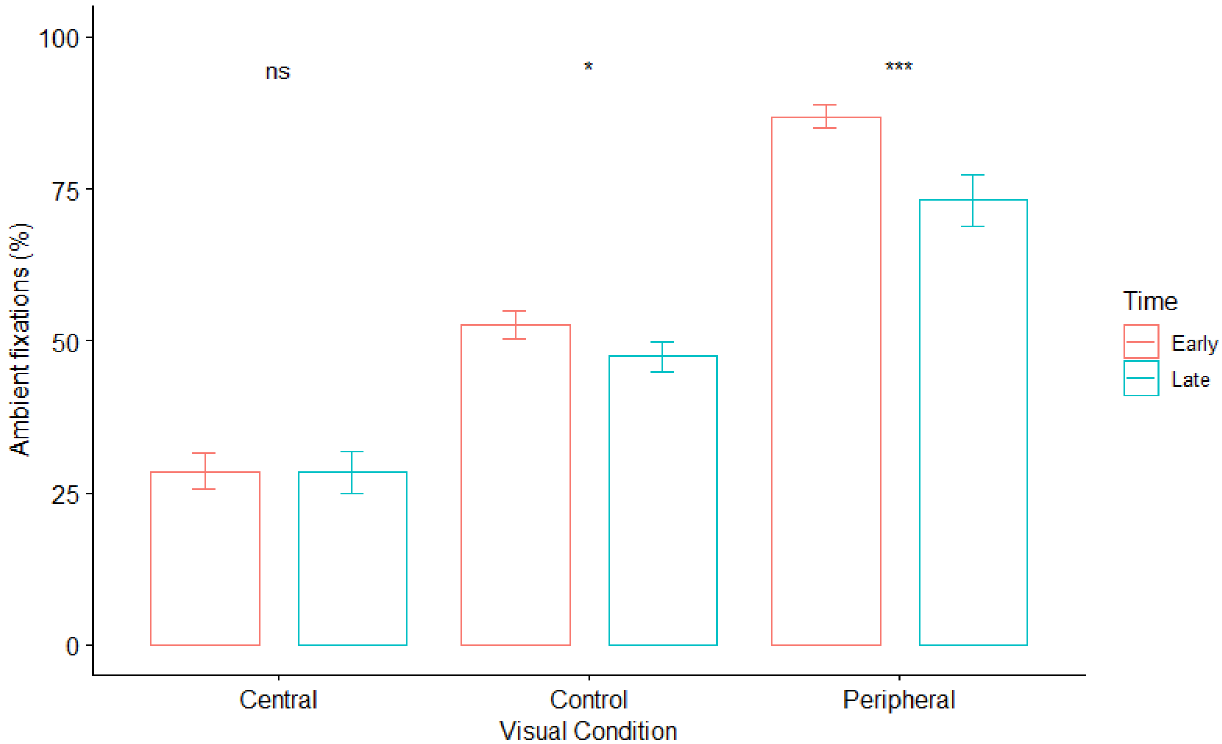

Figure 11. Prevalence of ambient fixations by visual condition and early/late scene processing. Red bars indicate ambient fixation prevalence during early scene processing, while blue bars indicate ambient fixation prevalence during late scene processing. Error bars indicate one standard error. * indicates $p < .05$. *** represents $p < .001$.

4.5 Discussion

As predicted, we found that the visual field influenced the size of the saccades people make when viewing a scene. At both early and late time intervals, saccades in the peripheral condition were larger than under full vision, which in turn involved larger saccades than central vision. Similarly, fixation durations were shorter during early scene processing for peripheral vision in comparison with full vision. However, during late scene processing, there were no differences in fixation lengths between peripheral and full vision. In fact, during this time interval, fixations were longer in peripheral vision when compared to central vision. Moreover, we found that there was an increase in fixation lengths and a decrease in saccade amplitudes between early and late time intervals in peripheral vision. Since increased fixation lengths during late scene processing reflect focal processing (Eisenberg & Zacks, 2016), these results suggest that peripheral vision is not limited to ambient visual processing and plays some role during focal visual processing. This result is in line with other research in which central or peripheral information was masked during the first 20-120 ms of fixations on scenes (van Diepen & d'Ydewalle, 2003). In that study, absence of central vision during the earliest part of fixations resulted in increased mean fixation lengths, when compared to the absence of peripheral vision.

For the central and full vision conditions, our data revealed mixed results. While fixation durations increased during both of these conditions between early and late processing, there was no change in saccade amplitudes. Our data suggests that the lack of difference in saccade amplitudes between the earliest and latest time intervals was due to the fact that saccade amplitudes were especially low during the first 500 ms bin of scene viewing (see Figures 12-14). The saccadic results conflict with previous work under normal visual conditions, in which the earliest saccades were the largest in a trial (e.g. Unema et al., 2005; Pannasch et al., 2008; Pannasch & Velichkovsky, 2009). It is unclear why the amplitude of initial saccades was smaller in our experiment than in previous literature, especially because initial saccade amplitude did differ between the visual conditions, suggesting that there is not a common cause to the phenomenon.

As predicted, under normal viewing conditions, the type of fixations that people make changed during the time course of viewing the scene. During early scene viewing, the prevalence of ambient fixations was higher, while during later scene viewing, people engaged in more focal fixations. Our results showed that peripheral vision involved a greater prevalence of ambient fixations when compared to central and full vision, reflecting its importance in ambient processing. However, the prevalence of ambient fixations did decrease between early and late time intervals, providing further evidence for the claim that while peripheral vision is predominantly associated with ambient processing, it may also serve some function in terms of late, focal processing. In contrast, central vision involved a higher prevalence of focal fixations than either peripheral or full vision, with no change by time interval, providing support for the association between the central visual field and focal visual processing (Trevarthen, 1968).

Chapter 5

General Discussion

The objective of this thesis was to identify and compare the function of peripheral vision with central vision during early and late scene processing. We were particularly interested in analyzing early perceptual processes, as previous work has demonstrated that processing of scene qualities occurs very quickly under normal visual conditions (Rousselet et al., 2005; Greene & Oliva, 2009). Moreover, research on peripheral vision has emphasized early perceptual processing, as prior experiments have shown that peripheral vision is sufficient to identify the gist of scenes (Larson & Loschky, 2009) and evaluate the emotion of briefly presented faces (Rigoulot et al., 2011). Recent theory suggests that peripheral vision is not only involved in identifying gist, but it is also particularly important for identifying various affective qualities of the surrounding environment (Rooney et al., 2017). However, the relationship between peripheral vision and the formation of affective impressions had not been tested empirically. Thus, we set out to identify whether peripheral vision was involved in affective processing.

In our first experiment, we demonstrated that peripheral vision was sufficient not only in processing information about the category of a briefly presented scene, but to form affective impressions. Moreover, participants' ratings of their level of interest in peripherally-presented natural and urban scenes differed in the same manner as under normal viewing conditions. This is particularly interesting because our work implies that not only does processing of information guide future eye movements (Yamamoto & Philbeck, 2012; Wastlund et al., 2018), we are to some degree aware of our impressions of peripheral information, even without a direct fixation. What these experiments suggest is that the peripheral visual field may play a profound role in the formation of initial impressions to stimuli. Due to the configuration of the visual field, the vast majority of visual information is located in the visual periphery. Any stimuli that a person fixates on is extremely likely to have been viewed previously in the periphery. Given that the choice of what information to foveate is influenced by peripheral processes, peripheral vision is involved in preconscious processing of the stimuli that become subjects of fixations. Thus, we argue that the formation of affective impressions involves a cascading process starting with the peripheral visual field. In this process, people initially view a stimulus in the periphery, they form an affective evaluation, and then they foveate the stimulus. Information from the central visual field then amplifies or modifies that initial evaluation, as the stimulus is processed with a substantially higher visual resolution.

In addition to evaluating the function of peripheral vision in isolation, we directly contrasted the role of central and peripheral vision in early scene gist processing (Experiment 2). In previous work (Larson & Loschky, 2009; Loschky et al., 2019), studies comparing the efficacy of gist processing between the two visual fields would separate central and peripheral scenes into different trials, with people exposed to only central or peripheral information on each scene. In contrast, we directly compared central and peripheral vision by showing people scenes in which the two visual fields contained conflicting information. Our study demonstrated that when people were briefly (67 ms) shown scenes with different information in the centre and periphery, they were much more likely to identify the gist of that scene as being the peripheral scene category, when compared to the central scene. This result is particularly interesting, since not only does the central visual field have a substantial advantage in the number of cone photoreceptors, but it is also advantaged by the amount of cortical magnification that occurs. Thus, our work provides evidence arguing that peripheral vision is especially specialized to process information important for scene gist, considering its efficacy at this task despite its limitations in cortical processing.

Having established some of the tasks that peripheral visual processes support during early scene processing, we set out to investigate the function of the peripheral visual field during late scene processing. In previous work, peripheral vision has been associated with ambient visual processes, while central vision has been associated with focal visual processes (Trevarthen, 1968). Through investigating eye movement patterns, we tested whether peripheral vision was associated with ambient processing, and whether central vision was associated with focal processing. Amongst other things, ambient vision provides information about the organization of space along with the relative location of objects and other salient items within the environment. In contrast, focal vision is responsible for gathering detailed information from very small regions of space through fixations. In the eye-movement literature, the ambient/focal divide is associated with a divide between early and late scene processing (Eisenberg & Zacks, 2016). During early scene viewing, people engage in ambient processing through exploratory eye movements and make shorter fixations, with the goal of gathering information about scene layout and gist. In contrast, during late scene viewing, people engage in focal processing through making shorter saccades and longer fixations on salient information.

Therefore, in our final experiment (experiment 3), we investigated whether central and peripheral visual processes map onto ambient and focal eye movements. We found that people engaged in significantly more ambient fixations during peripheral vision, while in central vision, people completed more focal fixations. Moreover, saccadic data revealed that in peripheral vision, saccades were significantly larger than in full vision, while in central vision, saccades were significantly smaller than in full vision. However, our work showed that fixation lengths increased and saccade amplitudes decreased during the duration of scene viewing in the peripheral visual field, which is associated with a transition between focal to ambient visual processing. Further supporting evidence comes from the fact that the prevalence of ambient fixations in peripheral vision between early and late viewing. Thus, this experiment provides evidence that central vision is associated with focal visual processing. However, peripheral vision is associated not only with ambient visual processing, it also plays a role in focal processing.

The three studies in this thesis demonstrate that peripheral vision plays a substantial role in early visual processing, whether it is through processing scene gist or facilitating the formation of affective impressions. However, our eye-tracking work suggests that peripheral vision is important not just during early scene processing, but during later time intervals as well. Future studies should investigate further what the function of peripheral vision is during late scene processing, and how it interacts with central vision to moderate the perception and evaluation of scenes over time.

Bibliography

Almeida, I., Soares, S. C., & Castelo-Branco, M. (2015). The distinct role of the amygdala, superior colliculus and pulvinar in processing of central and peripheral snakes. *PLoS One, 10*(6), e0129949.

Alorda, C., Serrano-Pedraza, I., Campos-Bueno, J. J., Sierra-Vázquez, V., & Montoya, P. (2007). Low spatial frequency filtering modulates early brain processing of affective complex pictures. *Neuropsychologia, 45*(14), 3223-3233.

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences of the United States of America, 106*(18), 7345-7350.

Baldassano, C., Esteva, A., Fei-Fei, L., & Beck, D. M. (2016). Two distinct scene-processing networks connecting vision and memory. *Eneuro, 3*(5), 10.1523/ENEURO.0178-16.2016. eCollection 2016 Sep-Oct.

Barrett, L. F., & Bar, M. (2009). See it with feeling: Affective predictions during object perception. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(1521), 1325-1334.

Barton, K. R., Valtchanov, D., & Ellard, C. (2014). Seeing beyond your visual field: The influence of spatial topology and visual field on navigation performance. *Environment and Behavior, 46*(4), 507-529.

Bayle, D. J., Henaff, M., & Krolak-Salmon, P. (2009). Unconsciously perceived fear in peripheral vision alerts the limbic system: A MEG study. *PLoS One, 4*(12), e8207.

Berman, M. G., Hout, M. C., Kardan, O., Hunter, M. R., Yourganov, G., Henderson, J. M., et al. (2014). The perception of naturalness correlates with low-level visual features of environmental scenes. *PloS One, 9*(12), e114572.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115.

Boucart, M., Moroni, C., Thibaut, M., Szaffarczyk, S., & Greene, M. (2013). Scene categorization at large visual eccentricities. *Vision Research, 86*, 35-42.

Cajar, A., Engbert, R., & Laubrock, J. (2016). Spatial frequency processing in the central and peripheral visual field during scene viewing. *Vision Research, 127*, 186-197.

Calvo, M. G., Avero, P., & Nummenmaa, L. (2011). Primacy of emotional vs. semantic scene recognition in peripheral vision. *Cognition & Emotion, 25*(8), 1358-1375.

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., et al. (2018). Package 'pwr'. *https://cran.r-project.org/web/packages/pwr/index.html*

Coburn, A., Kardan, O., Kotabe, H., Steinberg, J., Hout, M. C., Robbins, A., et al. (2019). Psychological responses to natural patterns in architecture. *Journal of Environmental Psychology, 62*, 133-145.

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*(4), 544-564.

Cowey, A., & Rolls, E. (1974). Human cortical magnification factor and its relation to visual acuity. *Experimental Brain Research, 21*(5), 447-454.

Curcio, C. A., & Sloan, K. R. (1992). Packing geometry of human cone photoreceptors: Variation with eccentricity and evidence for local anisotropy. *Visual Neuroscience, 9*(2), 169-180.

D'Hondt, F., Lassonde, M., Collignon, O., Lepore, F., Honore, J., & Sequeira, H. (2013). "Emotions guide us": Behavioral and MEG correlates. *Cortex, 49*(9), 2473-2483.

Duncan, R. O., & Boynton, G. M. (2003). Cortical magnification within human primary visual cortex correlates with acuity thresholds. *Neuron, 38*(4), 659-671.

Eisenberg, M. L., & Zacks, J. M. (2016). Ambient and focal visual processing of naturalistic activity. *Journal of Vision, 16*(2), 5-5.

Finlay, D. (1982). Motion perception in the peripheral visual field. *Perception, 11*(4), 457-462.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science, 20*(4), 464-472.

Guo, K., Liu, C. H., & Roebuck, H. (2011). I know you are beautiful even without looking at you: Discrimination of facial beauty in peripheral vision. *Perception, 40*(2), 191-195.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17*(17), R751-R753.

Hansen, T., Pracejus, L., & Gegenfurtner, K. R. (2009). Color perception in the intermediate periphery of the visual field. *Journal of Vision, 9*(4), 26-26.

Howard, I. P. (1986). The perception of posture, self motion, and the visual vertical. *Handbook of Perception and Human Performance,* 18.1-18.61. Retrieved from https://ci.nii.ac.jp/naid/10004203146/en/

Irwin, D. E. (1998). Lexical processing during saccadic eye movements. *Cognitive Psychology, 36*(1), 1-27.

Kaplan, R., Kaplan, S., & Brown, T. (1989). Environmental preference: A comparison of four domains of predictors. *Environment and Behavior, 21*(5), 509-530.

Kardan, O., Demiralp, E., Hout, M. C., Hunter, M. R., Karimi, H., Hanayik, T., et al. (2015). Is the preference of natural versus man-made scenes driven by bottom–up processing of the visual features of nature? *Frontiers in Psychology, 6*, 471.

Kuraguchi, K., & Ashida, H. (2015). Beauty and cuteness in peripheral vision. *Frontiers in Psychology, 6*, 566.

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision, 9*(10), 6-6.

Laubrock, J., Engbert, R., & Kliegl, R. (2005). Microsaccade dynamics during covert attention. *Vision Research, 45*(6), 721-730.

Loschky, L. C., Sethi, A., Simons, D. J., Pydimarri, T. N., Ochs, D., & Corbeille, J. L. (2007). The importance of information localization in scene gist recognition. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1431.

Loschky, L., Boucart, M., Szaffarczyk, S., Beugnet, C., Johnson, A., & Tang, J. L. (2019). The contributions of central and peripheral vision to scene gist recognition with a 180° visual field. *Journal of Vision, 19*(15), 570-570.

Lukavský, J. (2019). Scene categorization in the presence of a distractor. *Journal of Vision, 19*(2), 6-6.

Makris, S., Hadar, A. A., & Yarrow, K. (2013). Are object affordances fully automatic? A case of covert attention. *Behavioral Neuroscience, 127*(5), 797.

Matin, L. (1982). Visual localization and eye movements. *Tutorials on motion perception* (pp. 101-156) Springer.

McColgin, F. H. (1960). Movement thresholds in peripheral vision. *Josa, 50*(8), 774-779.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences, 6*, 414-417.

Mullen, K. T. (1991). Colour vision as a post-receptoral specialization of the central visual field. *Vision Research, 31*(1), 119-130.

Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning and Memory, 6*(4), 391.

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41*(2), 176-210.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42*(3), 145-175.

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions.

Pannasch, S., & Velichkovsky, B. M. (2009). Distractor effect and saccade amplitudes: Further evidence on different modes of processing in free exploration of visual images. *Visual Cognition, 17*(6-7), 1109-1131.

Peirce, J. W. (2007). PsychoPy—psychophysics software in python. *Journal of Neuroscience Methods, 162*(1-2), 8-13.

Provis, J. M., Dubis, A. M., Maddess, T., & Carroll, J. (2013). Adaptation of the central retina for high acuity vision: Cones, the fovea and the avascular zone. *Progress in Retinal and Eye Research, 35*, 63-81.

Rigoulot, S., D'Hondt, F., Defoort-Dhellemmes, S., Despretz, P., Honoré, J., & Sequeira, H. (2011). Fearful faces impact in peripheral vision: Behavioral and neural evidence. *Neuropsychologia, 49*(7), 2013-2021.

Rolls, E., & Cowey, A. (1970). Topography of the retina and striate cortex and its relationship to visual acuity in rhesus monkeys and squirrel monkeys. *Experimental Brain Research, 10*(3), 298-310.

Rooney, K. K., Condia, R. J., & Loschky, L. C. (2017). Focal and ambient processing of built environments: Intellectual and atmospheric experiences of architecture. *Frontiers in Psychology, 8*, 326.

Roorda, A., & Williams, D. R. (1999). The arrangement of the three cone classes in the living human eye. *Nature, 397*(6719), 520.

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annual Review of Vision Science, 2*, 437-457.

Rossion, B., & Pourtois, G. (2004). Revisiting snodgrass and vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception, 33*(2), 217-236.

Rousselet, G., Joubert, O., & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes? *Visual Cognition, 12*(6), 852-877.

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological Science, 5*(4), 195-200.

Shenhav, A., Barrett, L. F., & Bar, M. (2013). Affective value and associative processing share a cortical substrate. *Cognitive, Affective, & Behavioral Neuroscience, 13*(1), 46-59.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences, 1*(7), 261-267.

Thorpe, S. J., Gegenfurtner, K. R., Fabre-Thorpe, M., & BuÈlthoff, H. H. (2001). Detection of animals in natural images using far peripheral vision. *European Journal of Neuroscience, 14*(5), 869-876.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature, 381*(6582), 520.

Tran, T. H. C., Rambaud, C., Despretz, P., & Boucart, M. (2010). Scene perception in age-related macular degeneration. *Investigative Ophthalmology & Visual Science, 51*(12), 6868-6874.

Trevarthen, C. B. (1968). Two mechanisms of vision in primates. *Psychologische Forschung, 31*(4), 299-337.

Unema, P. J., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition, 12*(3), 473-494.

Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology, 43*, 184-195.

van Diepen, P., & d'Ydewalle, G. (2003). Early peripheral and foveal processing in fixations during scene perception. *Visual Cognition, 10*(1), 79-100.

Wästlund, E., Shams, P., & Otterbring, T. (2018). Unsold is unseen… or is it? examining the role of peripheral vision in the consumer choice process using eye-tracking methodology. *Appetite, 120*, 49-56.

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods, 42*(3), 671-684.

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). Sun database: Large-scale scene recognition from abbey to zoo. Paper presented at the *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 3485-3492.

Yamamoto, N., & Philbeck, J. W. (2013). Peripheral vision benefits spatial learning by guiding eye movements. *Memory & Cognition, 41*(1), 109-121.

Zele, A. J., & Cao, D. (2015). Vision under mesopic and scotopic illumination. *Frontiers in Psychology, 5*, 1594.