Foveal and Extrafoveal Processing of Facial Features for Emotion Recognition

DURAN, NAZIRE

Foveal and Extrafoveal Processing of Facial Features for Emotion Recognition

Nazire Duran, BSc., MSc.

This thesis is submitted for the degree of Doctor of Philosophy in the Department of Psychology at Durham University

2020

Foveal and Extrafoveal Processing of Facial Features for Emotion Recognition

Nazire Duran

Abstract

This thesis focuses on the differential contributions of foveal and extrafoveal processing of facial features to the identification of facially expressed emotions. Each facial expression of the six basic emotions – happy, sad, fearful, surprised, disgusted and angry – has distinctive features that are informative for their recognition. We investigated whether: (1) foveal processing of emotion-informative features enhances emotion recognition, relative to extrafoveal processing of those features; (2) this is due to high spatial frequency (HSF) processing by the fovea; (3) eye movements target emotion-informative features when not initially fixated; (4) gaze patterns while viewing facial expressions are specific to each expression; and (5) preferential visual sampling of emotion-informative features is linked to better emotion recognition.

Across four experiments, expressions of anger, surprise, fear, sadness and disgust were presented briefly, precluding eye movements, at positions which ensured that either eye, central brow, either cheek or the mouth fell on the fovea. Enforced fixation on the mouth improved recognition accuracy for fearful, surprised and disgusted expressions and reduced misclassification of fear as surprise and disgust as anger (Experiments 1 and 2). Enforced fixation on the brow and mouth led to higher anger recognition accuracy compared to the cheeks (Experiment 2). Intensity of expressions did not modulate the effect of initially fixating on emotion-informative features on emotion recognition. There were fewer neutral responses for anger at the brow region (Experiment 3). There were also fewer neutral responses for fear and surprise at the mouth (Experiment 3). While filtering out the HSF from the fixated emotion-informative features did not affect emotion recognition, occluding them impaired it (Experiment 4). Reflexive first saccades were more often directed upwards from lower features than downwards from upper features, yet more detailed analysis showed that they did not target emotion-informative features as suggested by previous research (Experiments 1, 2 and 3). Finally, when allowed to view faces for 5s, gaze patterns reflected the distribution of emotion-informative features and there was a positive correlation between time spent looking at the mouth and disgust recognition (Experiment 2).

Overall, we found that foveal processing of some emotion-informative features improved emotion recognition, however, this was not due to HSF processing at the fovea. Additionally, instead of being guided by emotion-informative facial features, initial saccades might be guided more strongly by the centre-of-gravity effect or a general top-down knowledge of face configuration. The saccade patterns observed also suggest that observers seek to sample as much of the face space as possible in order to classify its expression.

Table of Contents

List of Tables

List of Figures

Statement of Copyright

Acknowledgements

First, I would like to thank my supervisor, Dr. Anthony Atkinson, for sharing this fascinating area of research with me. Your support has been invaluable throughout. I would like to acknowledge Prof. Hannah Smithson (Oxford University) who has been involved in the coding of the experiments reported in this thesis. A huge thank you to Elaine Stanton whose bespoke programmes allowed the analysis of the results.

I want to thank all my friends in the PhD office, especially Emine, Jasmin, Liz and Linda, who were there for me through thick and thin. You have all been a shoulder to cry on and someone to vent to when it all got a bit too much, but also the primary source of laughter during my PhD years. I have made the best memories with you, you are an amazing and inspirational bunch!

Finally, Kris, I lost count of how many times I came to you saying 'I can't do this!' Thank you for believing in me more than I believed in myself. Your presence in my life has kept me sane and helped me refocus my mind – thanks for all the support, fun times and video games! You are a wonderful boyfriend.

Dedication

I would like to dedicate this thesis to the two women who are the pillars of my life:

To my grandmother, Nazire, who never had the chance to go to school but her dedication and desire to learn to read and write, even at the age of 90, has inspired me all my life. I am proud to carry your name and put the Dr. title in front of it.

To my mother, who always pushed me to do better, supported me at every single step and continues to support me every day. I couldn't have asked for a better mom.

Your words continue to shape my life to this day and always will. I love you both.

Chapter 1

General Introduction

The overall aim of this thesis is to investigate the distinct contributions of foveal and extrafoveal visual processing of facial features to the recognition of facially expressed emotions. More specifically, there are three key sub-aims: (1) We investigate whether foveal processing of features previously deemed informative for particular emotional expressions will lead to enhanced recognition performance for those emotions, relative to fixation of less informative features (and thus extrafoveal processing of the informative features). (2) We ask whether any enhanced emotion recognition performance as a consequence of foveal processing of the informative facial features is due to the ability of the fovea to process fine visual detail. (3) We investigate whether observers seek out informative facial features when they are not initially fixated but are visible in the extrafoveal visual field.

Faces convey valuable social information about our counterparts, which we must decode in order to successfully navigate our social environment. A range of information can be transmitted by a face including a person's identity, gender and emotional state. Several roles have been suggested for facial expressions in social communication. Susskind, Lee, Cusi, Feiman, Grabski and Anderson (2008) suggested a sensory-regulatory role for facial expressions where the facial morphology of the face displaying an expression either increases or decreases sensory exposure. On the other hand, it is also suggested that facial expressions play various more social-communicative roles (Bavelas, Black, Lemery, & Mullett, 1986; Frith, 2009; Jack & Schyns, 2015; Parkinson, 2005). The communication of a social signal, such as facial expression of an emotion, is achieved by an expresser sending the emotion signal and a receiver decoding this signal. The experiments in the following chapters will focus on the social-communicative role of facial expressions, especially on the process of decoding/recognising a facial expression.

Informative Features for Facial Expression Recognition

Ekman, Sorenson and Friesen (1969) conducted a large-scale cross-cultural study of emotion recognition across America, Brazil, and Japan (literate cultures), and New Guinea and Borneo (pre-literate cultures) and showed that expressions of happy, fear, sadness, surprise, disgust and anger were categorised similarly across these cultures. These literate and pre-literate cultures were found to associate certain facial muscle movements to certain categories of emotions in the same way, indicating that the recognition of these six expressions of emotion is universal. These six emotions came to be known as the six basic emotions. Ekman and Friesen (1976) compiled a comprehensive manual called the Facial Action Coding (FAC) manual listing facial muscles, their actions and the effect of their action on facial appearance. Individual actions of all identified facial muscles were numbered and called action units (AU). For example, AU12 is termed the lip corner puller and is tied to the zygomatic major muscle. Using this FAC manual, the AU composition of each of the expressions of the six basic emotions was compiled. Each of the expressions of the six basic emotions has a unique combination of AUs leading to the unique appearance of the facial displays of each of the basic emotions. For example, a facial expression of happiness results from the action of AU6 (cheek raiser) and AU12 (lip corner puller) combined and the facial expression of anger results from the combination of AUs 4 (brow lowerer), 5 (upper lid raiser), 7 (lid tightener) and 23 (lip tightener). Ekman, Friesen and Hagar (1978) described prototypical AU patterns for each of the six basic emotions; however, they also allowed for variants of these which included additional or substitutional AUs to the prototypical patterns that lead to the same emotion categorization when presented to observers. Despite the widespread use of these six basic emotions in emotion recognition research, the universality of these basic emotions has recently been challenged (Jack, Sun, Delis, Garrod & Schyns, 2016). The main finding that led to this argument was lower accuracy for expressions of fear and disgust displaying the same AU pattern in Eastern cultures compared to Western cultures (Jack, Blais, Scheepers, Schyns & Caldara, 2009). This led to the suggestion that there are fewer than six emotions that are universal across cultures and these exclude disgust and fear (Jack, Garrod & Schyns, 2014)[1].

In emotion recognition tasks where the facial expressions of the six basic emotions are presented to the observers followed by a forced choice between a small set of emotion words, happy expressions are usually recognised better and faster compared to other facial expressions, both when presented at the centre of the screen (Calder, Young, Keane & Dean, 2000; Calvo & Lundqvist, 2008; Palermo & Coltheart, 2004; Smith & Rossit, 2018) and when presented peripherally (Calvo, Fernández-Martín & Nummenmaa, 2014; Calvo, Nummenmaa & Avero, 2010; Smith & Rossit, 2018). Fearful facial expressions are recognised poorest compared to all other expressions (Calvo & Lundqvist, 2008; Guo, 2012; Palermo & Coltheart, 2004). However, there is evidence to suggest that fearful expressions are better detected (i.e., perceived as an expressive face) than they are recognised (i.e., recognised as a fearful expression) (Smith & Rossit, 2018). As facial expressions are displayed with increasing intensity, the recognition of the expression increases (Blairy, Kleck & Hess, 1997; Calder, Rowland, Young, Nimmo-Smith, Keane & Perrett, 2000; Guo, 2012). Systematic misclassifications have been recorded between

[1] While we fully acknowledge these arguments and will refer to them throughout the chapters, the experiments reported in this thesis will investigate emotion recognition using images of facial expressions posed according to the criteria set out by Ekman and colleagues (1969). This is to ensure that our results are comparable to the previous studies of emotion recognition.

expressions of fear and surprise (Chamberland, Roy-Charland, Perron & Dickinson, 2017; Du & Martinez, 2011; Ekman, Friesen, O'Sullivan, et al., 1987; Gagnon, Gosselin, Hudon-ven der Buhs, Larocque & Milliard, 2010; Guo, 2012; Roy-Charland, Perron, Beaudry & Eady, 2014; Roy-Charland, Perron, Young, Boulard & Chamberland, 2015a; Young et al., 1997) as well as between anger and disgust (Du & Martinez, 2011; Ekman, Friesen, O'Sullivan, et al., 1987; Guo, 2012; Widen, Russell & Brooks, 2004).

There is debate in the literature as to whether facial expressions are analysed featurally (Chen & Chen, 2010) or holistically (Calder, Young, Keane & Dean, 2000), or whether both featural and holistic processing are involved, with the relative weighting depending on the particular emotion presented (Beaudry, Roy-charland, Perron & Cormier, 2014; Calvo, Numenmaa & Avero, 2010; Tanaka, Kaiser, Butler & Grand, 2012). These concepts were originally applied to face processing in the context of identity recognition. Tanaka and Farah (1993) showed that parts of a face (i.e., eyes, nose, mouth) are identified better when they are presented in the context of a whole face compared to when they are presented alone. They further showed that this effect was not present for whole houses and house parts. In light of these results, Tanaka and Farah (1993) suggested that faces are processed more holistically than objects. This holistic face representation means that face representation cannot be broken down into its constituent features. Previously, it has been shown that faces seen upside-down are less well-recognised compared to other inverted objects (Yin, 1969). Tanaka and Farah (1993) suggested that the reason behind this effect is the disruption of holistic processing. They supported this suggestion by showing that, for inverted faces, there was no difference between recognition of a face part when this was presented in isolation and when presented within a whole face. Therefore, in contrast to upright faces, inverted faces were represented in a more featural way. Tanaka and Sengco (1997) showed that the holistic perception of the face involves the configural properties of the face as well – i.e. the spatial relations between facial features. Tanaka and Sengco (1997) showed that recognition of a face part was lower when it was presented within a face with a new configuration – i.e., when the spatial relationships between facial features were altered. The holistic processing of upright faces is further supported by Young, Hellawell and Hay (1987). Young et al. (1987) showed that when upper and lower parts of two different face identities were aligned to form a *composite face*, participants were less able to recognise the identity of the top half of the face compared to when the two parts were misaligned. Young et al.'s (1987) results indicate that, when aligned, the two parts of the composite face combines to form a new face. The holistic processing of this new face interferes with the identification of the identity of the individual parts (upper and lower). In line with the suggestion that inversion of a face disrupts holistic processing, the composite effect was not present when the composite faces were inverted. While we acknowledge the contribution of holistic and/or featural strategies to emotion recognition, the experiments described in this thesis do not seek to directly investigate which of these strategies are involved in emotion recognition. We are interested in the question of whether facial features contribute to emotion recognition from a fully visible face where neither holistic nor featural properties of the face are manipulated.

Due to the unique AU pattern of each of the six basic expressions, facial features provide a signal that can be decoded in order to make emotion recognition judgements. Gosselin and Schyns (2001) developed a technique called *Bubbles* to illustrate that the nature of the categorization task changes the subset of visual information used from a face. Using this technique, the face image is presented to the observer through a mask that is punched with Gaussian apertures only revealing a subspace of the whole face. The original image is sampled at a range of spatial frequencies and the size of the Gaussian apertures is adjusted to reveal the same cycles per visual angle at each spatial frequency. The spatial frequencies are recombined with the corresponding masks to produce the final image. The revealed information at each spatial frequency is randomized for every face image leading to a thorough exploration of a face after a series of trials. The technique is self-calibrated so that the observers maintain 75% emotion recognition accuracy. The masks that lead to correct categorization of the underlying stimulus are summed and divided by all the masks presented to produce *effective masks* which contain the information the observer found 'diagnostic' for completing the task. Within *the diagnostic framework theory* (Gosselin & Schyns, 2002), effective masks represent the diagnostic or *informative* information that should be processed to match the available information to the representation of that information in memory. Gosselin and Schyns (2001) showed that when viewing the same face, observers use information from the mouth region to decide if the face is expressive whereas they use information from eye and mouth regions when identifying its gender. Furthermore, observers differentially use information from a range of spatial scales that make up the face stimuli.

In light of these findings, the same group (Smith, Cottrell, Gosselin & Schyns, 2005) went on to investigate how the brain decodes facial expression signals transmitted by a face. Using the *Bubbles* technique, they found that observers used the mouth region for identifying happiness and surprise, the eyes for identifying fear, eyes and the brow for identifying anger, nose and mouth for disgust and brow and the sides of the mouth for identifying sadness. The authors further pointed out that there is little overlap between the diagnostic regions of these emotions as used by human observers even though the information available to resolve the task extracted by a model filtering function show overlapping areas across emotions (Figure 1). This is apparent in the case of fear and surprise: Both expressions are transmitted with available

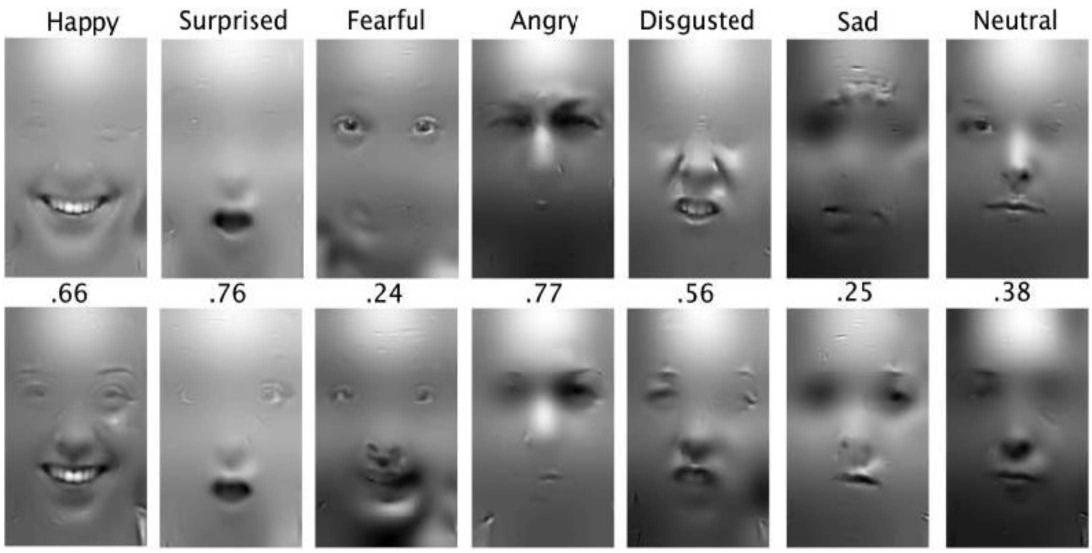

Figure 1: Top row represents the diagnostic filtering masks produced by the Bubbles technique by the human observers and the bottom row represents the effective masks produced by the model observer. The numbers represent the correlation between human and model filtering functions. Image from Smith et al. (2005).

information around the eyes and the mouth; however, human observers made use of the eye region to identify fear and the mouth region to identify surprise to decode the transmitted information. These results revealed that there exists a subset of diagnostic features on expressive faces that observers use to categorize the six basic emotions. Additionally, diagnostic information from these features are present at different spatial frequencies, for example high spatial frequency (HSF) information from the brow region for anger and HSF information from eyes for fear.

Smith et al. (2018) used the *Bubbles* technique to further investigate the differences in information use during emotion recognition between younger and older observers. While the older observers required more visual information to be revealed by the bubbles compared to younger observers, the patterns of information use corroborated the results from Smith et al. (2005) and showed similar information use across the age groups for happy, fearful, angry and sad facial expressions. Specifically, information from the smiling mouth was used across all spatial frequency bands for recognition of happy expressions while the HSFs from the eye region were no more helpful than the information from the hairline. High and mid-spatial frequency information from the eyes and lower spatial frequency information from the mouth region were found to be informative for fearful faces. The furrowed eyes and the taut mouth proved to be equally informative for the recognition of angry faces; however, the mouth region was used less at lower spatial frequencies compared to the eyes. Finally, for sad facial expressions, the eyebrows and the mouth were informative with the eyes being more useful at high spatial frequencies compared to the mouth.

A study by Wegrzyn, Vogt, Kireclioglu, Schneider and Kissler (2017) also investigated whether expression recognition performance was reliant on certain facial features for each expression. The method utilised by Wegrzyn et al. (2017) is reminiscent of the method used by Smith et al. (2005): The images of a male and a female face expressing the six basic emotions and neutral poses were obscured by a 6×8 grid of white tiles which were removed gradually (1 tile per second) in a random manner to reveal the underlying face. When the expression was recognised, the participants were asked to press a button to stop the removal of tiles and to categorise the face. Each tile was weighted according to how quickly and accurately it led to a response. Figure 2 shows the highest-weighted tiles superimposed on the respective facial expressions.

Figure 2: The red colour represents the tiles that were weighted highest for each of the expressions and face identities. The colours represent the min-max-scaled weights within each face. Image from Wegrzyn et al. (2017).

The recognition of happy and disgusted faces was shown to rely on the tiles in the bottom half of the faces, especially the tiles around the mouth. On the other hand, the recognition of anger, fear and sadness relied on the tiles in the upper half of the face, especially the tiles around the eyes. However, since the study only used 2 face images (a male and a female), some of these results were inconclusive. For example, while the recognition of surprise was reliant on the tiles in the bottom face half for the female face, this was not the case for the male face. Additionally, reliance on the eyes for anger for the female face did not have the same magnitude as it did for the male face. In addition to the comparison of upper and lower tiles, Wegrzyn et al. (2017) assigned hand-drawn action units to each of the 48 tiles and compared the weight of these tiles to the average weight of the tiles not belonging to an action unit. In line with the analysis of the upper and lower parts of the face, the action units around the mouth region were more informative for happy and disgust recognition than the non-action units. Fear, anger and sadness had the most weighted tiles around the eyes: The brows were rated most highly for sadness and various action units of the eyes were weighted highly for fear and anger. Both the eye and the

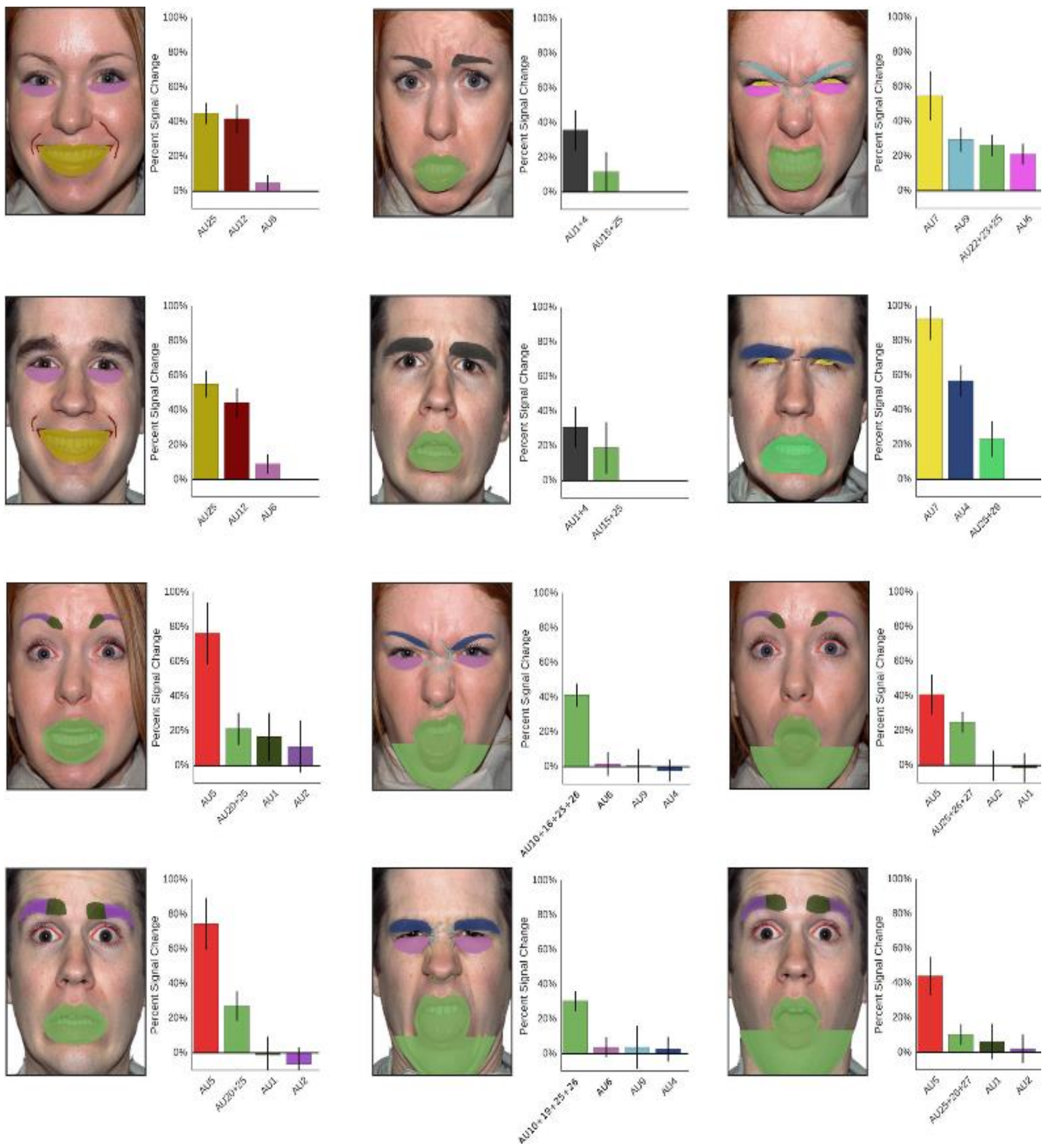

Figure 3: The drawings on the face images represent the AUs and the colours of the AUs correspond to the values on the bar graphs next to each face. The bar graphs represent the weights of the AUs relative to the average weight of the non-AU tiles. Image from Wegrzyn et al. (2017).

mouth action units were weighted highly for surprised expressions. However, as can be seen in Figure 3, the brows and the glabellar frown lines were also weighted quite highly for the recognition of anger as also suggested by Smith et al. (2005).

These studies (Smith et al., 2005, 2018; Wegrzyn et al., 2017) provide strong support for the presence of facial regions or facial features that are informative for the recognition of each of the six basic emotions. Nonetheless, the combination of facial expressions used in an experiment can affect what facial feature becomes informative for emotion recognition (Smith & Merlusca, 2014). Smith and Merlusca (2014) used the *Bubbles* technique to investigate what visual information would become informative for recognition of fearful, angry and disgusted expressions when the alternative categories were varied. They found that while high spatial frequency information from the eye region was consistently useful for the recognition of fearful faces, in keeping with Smith et al. (2015) and Smith et al. (2018), when they were compared to only neutral, anger and neutral, happy and neutral and all other expressions, low spatial frequency information from the mouth also became informative. Additionally, low spatial frequency information from different sub-regions of the mouth was informative with different comparisons: While the centre of the mouth was informative when fear was compared against neutral and angry faces, only the corners of the mouth were informative when compared to neutral and happy faces. High spatial frequency information from the brow region and the taut eyes was informative for anger when it was being compared to other expressions but when it was being compared to neutral faces alone, observers made use of a broader range of facial features. For disgusted expressions, high spatial frequency from the wrinkled nose and mouth region was consistently informative for recognition regardless of the comparison condition.

Face size and holistic processing

As discussed previously, when presented behind a mask, sampling of certain facial features is informative for the recognition of the underlying facial expression. However, under natural viewing conditions – i.e., when faces are seen in full view – most or all facial features are available for visual sampling. In visual tasks such as face perception, eye movements are linked to allocation of overt attention and information sampling (Findlay & Gilchrist, 2003), hence gaze patterns confer a reliable index of what facial regions/features conveyed useful information for the resolution of any face-related task during face viewing. Our eyes process the image we are looking at with a series of fixations and saccades. Several fixations can be made onto any given visual object depending on the size of the object in the visual field. One of the determinants of the retinal size of the observed face is viewing distance. Face size decreases with increasing viewing distance. Furthermore, as face size decreases, the high spatial frequency information from the image is stripped off, leaving mostly low spatial frequencies accessible (Loftus & Harley, 2005; Sowden & Schyns, 2006).

Goffaux and Rossion (2006) suggest that low spatial frequencies support the holistic processing of faces, implying that at large viewing distances faces are processed holistically. In line with this suggestion, McKone (2009) showed that holistic processing takes place at viewing distances starting from 2 meters and operates up to 10m. The results from the McKone (2009) study suggest that while holistic processing operates at larger distances, another face processing strategy must be at play at distances that are relevant to social communication such as facial expression recognition and speech recognition. However, several researchers found that while the critical spatial frequencies for face recognition increase for small faces that simulate faces seen at a distance, this increase is halted at face sizes of around 4.7-6 degree of visual angle

(Oruç & Barton, 2010; Yang, Shafai & Oruç, 2014). These researchers suggest that the reason for this shift is due to a shift from a part-based, analytical processing at large distances to holistic processing strategy for faces seen at social communicative distances.

During face-to-face interactions, faces will subtend visual angles of approximately 5 to 13 degrees on the retina at a range of distances between 45cm and 213 cm (Atkinson & Smithson, 2013). Therefore, according to the results of Oruç and Barton (2010) and Yang et al. (2014), holistic processing contributes more to the processing of faces seen at distances relevant to face-to-face interactions since human observers seem to rely on low spatial frequencies from these faces. These studies used upright faces with neutral expressions, however, using the Bubbles technique, Smith and Schyns (2009) showed that facial expressions are recognised differentially across a range of viewing distances. While happiness, surprise, anger and disgust contained diagnostic information at LSFs, making distal recognition of these emotions possible, the recognition of fear and sadness depended more on HSFs making them more suitable for proximal recognition. While Oruç and Barton (2010) and Yang et al. (2014) suggest that observers preferentially rely on lower spatial frequency information (more specifically SFs peaking around 8 cycles/face-width) for face processing at socially relevant distances, this strategy might prove harmful for recognition of proximal expressions such as fear and sadness whose informative SFs lie above 30 cycles/face. Therefore, recognition of facial expressions might require processing of higher spatial frequencies compared to identification of faces at socially relevant distances. During active vision, fixations bring an area of the retina known as the fovea to the object/feature of interest for high resolution processing and saccades move the eye to the next target to be fixated. Given that the fovea corresponds to the central 1.7° of the visual field (Wandell, 1995), at such interpersonal distances the whole face cannot fall within the fovea with a single fixation on the face; only a part of the face will project to the fovea, with the rest of the face falling on the extrafoveal retinal regions[2]. Therefore, when high spatial frequency information is required from a face/facial feature, fixations are directed to that face/facial feature.

[2] At this junction, it is important to point out that holistic processing of faces can proceed with one or two fixations to the face (Hsiao & Cottrell, 2008), however the main interest of the studies in this thesis is the contribution of foveal and extrafoveal processing of facial features to facial expressions recognition. Therefore, while we fully acknowledge the different strategies that can be utilised to achieve face perception related tasks, we will not delve into the relative contribution of these processes to our results. Additionally, it is worth noting that even when face recognition performance saturated after approximately 2 fixations, most observers made approximately three fixations before responding (Hsiao & Cottrell, 2008).

Foveal and Extrafoveal Visual Processing

Visual processing across the retina is not homogenous. The retina is divided into specialised regions which can broadly be defined as foveal and extrafoveal regions. There are fundamental differences between foveal and extrafoveal vision, both qualitatively and quantitatively (Rosenholtz, 2016; Strasburger, Rentschler & Jüttner, 2011). The fovea has the highest photoreceptor cell density across the retina making foveal vision capable of the highest visual acuity processing. As well as the density, the foveal photoreceptors have relatively smaller receptive fields allowing the processing of fine detail. Additionally, owing to the greater magnification factor of the fovea compared to the extrafoveal retina, the foveated part of the visual field is processed by a proportionately larger area of the visual cortex than is the rest of the visual field (Azzopardi & Cowey, 1993). To achieve the same level of visual performance in the extrafoveal region as the foveal region, the retinal size of an item in the extrafoveal region needs to be enlarged, to produce a cortical representation of comparable size (Virsu & Rovamo, 1979). For example, Yu and Chung (2011) showed that people with central vision loss are still capable of both identifying faces and accurately judging facial expressions in the periphery as long as the faces are appropriately scaled. Beyond the fovea, there is a drop in visual acuity and contrast sensitivity (Robson & Graham, 1981). The density of photoreceptor cells decreases and the size of the receptive field of photoreceptor cells increases with eccentricity from the fovea (Jonas, Schneider & Naumann, 1992; Song, Chui, Zhong, Elsner & Burns, 2011). Song et al. (2011) showed that the packing density of cone cells decreases inhomogenously across the vertical and horizontal meridians with the decrease being steeper with increasing eccentricity from the centre across the vertical meridian. In the periphery, in contrast to foveal photoreceptor cells, several photoreceptor cells converge onto a single ganglion cell (Banks, Sekuler & Anderson, 1991). Hence, the visual image processed by each ganglion cell receiving signals from extrafoveal regions is larger, leading to a loss in visual resolution capabilities.

Extrafoveal vision is also highly susceptible to a phenomenon termed *visual crowding* (Rosenholtz, 2016) which leads to a decline in the recognition of a target in the periphery due to the presence of adjacent objects that are termed flankers. Visual crowding leads to a steeper decline in the recognition of the crowded target than can be expected solely due to an increase in eccentricity (Bouma, 1970). The crowding of a target object is resolved when the centre-to-centre distance between the target and the flankers is increased. This space between target and flankers is called critical spacing and it increases with increasing eccentricity (Pelli & Tillman, 2008). Crowding is also shown to affect the recognition of a face part due to the presence of other facial features within the critical spacing (Martelli, Majaj & Pelli, 2005) .

Due to these differences between foveal and extrafoveal visual processing, on a face which covers 5 to 13 degrees of visual angle on the retina, detailed visual processing will be restricted to the fixated feature and visual information from the rest of the face will be relatively blurred. Therefore, fixating on a facial feature that is suggested to contain useful visual information, especially at higher spatial frequencies, might confer a recognition advantage through detailed processing that cannot be achieved by the extrafoveal retina.

Eye movements during face-related tasks

While viewing faces for face-related tasks, people generally follow a T-pattern where they fixate the eyes first followed by the nose and the mouth (Henderson, Williams & Falk, 2005; Yarbus, 1967) (but note cultural differences between the eye movement patterns of Eastern Asian (EA) and Western Caucasian (WC) observers for face recognition (Blais, Jack, Scheepers, Fiset & Caldara, 2008; Kelly, Miellet & Caldara, 2010) as well as idiosyncratic biases (Leonards & Scott-Samuel, 2005) and idiosyncratic and stable scanning patterns that diverge from the T-pattern (Mehoudar, Arizpe, Baker & Yovel, 2014). Hsiao and Cottrell (2008) suggest that two fixations landing at the centre of the nose (which corresponds roughly to the centre of the face) is adequate for face identification. The study consisted of a study phase where participants were shown faces to be remembered later and a test phase where they needed to indicate whether the face was old or new. In both phases, participants were asked to fixate centrally. Following fixation, the faces were presented either on top or the bottom of the screen. In the test phase, the participants were either restricted to making one, two or three fixations or their eye movements were unrestricted. Results show that discrimination performance was already above chance level after one fixation to the face but improved when two fixations were allowed. After two fixations, no further improvement was recorded. Since the first fixation into the face image was from a peripheral location, Hsiao and Cottrell (2008) suggested that the location of the first two fixations (i.e., the centre of the nose) reflects the *preferred landing location* of the initial fixations on a face. An implication of this study is that the *preferred landing location* allows access to holistic/configural information about the face which then guides face-related judgements. This means that since two fixations to the centre of the face suffice for recognition, eye movements to examine individual features might not be necessary. Instead, given a large perceptual span and taking into account the drop in visual acuity from fovea to the periphery, a more effective strategy might be to choose a *centre of information* to obtain enough face information to accomplish the task at hand at single fixation. Hsiao and Cottrell (2008) suggest this *centre of information* might be the centre of the face where observers preferred to fixate.

Similar to Hsiao and Cottrell (2008), Peterson and Eckstein (2012) investigated where the first fixation fell within the face in three different tasks: identification, gender and emotion recognition. Starting from either of eight peripheral starting fixation locations, participants were given 350ms to make a single fixation onto the target face. Similar to Hsiao and Cottrell (2008), Peterson and Eckstein (2012) found that observers fixated just below the eyes for all the tasks. Peterson and Eckstein's (2012) results strengthen the suggestion that observers select their first fixation location in order to optimize the information obtained from the face at a single fixation instead of fixating individual facial features. While Peterson and Eckstein's results might suggest that fixating on individual features might not be necessary for emotion recognition performance, they show that the preferred fixation position shifted downwards for the emotion recognition task. Results from the *Bubbles* study by Gosselin and Schyns (2001) showed that human observers relied on the mouth region when deciding whether a face was expressive or not while they needed additional information from the eye region to determine the gender of the same face. Therefore, a downward shift in the preferred fixation position for the emotion recognition task in Peterson and Eckstein's (2012) study indicates that, for emotion recognition, observers sought information from the mouth that could not be obtained by fixating the preferred fixation location for gender and identity recognition. Additionally, despite the ability of the observers in Peterson and Eckstein's (2012) study to categorise the given facial expressions at above chance levels, the proportion correct for the emotion task was low ($M=.542 \pm .015$) when compared to other studies of emotion recognition. Therefore, it is possible that while the processing of informative facial features in the peripheral/parafoveal visual field leads to adequate recognition performance, foveating these features might further improve emotion recognition. Furthermore, Peterson and Eckstein (2012) only presented face images for 350ms which only allowed the execution of a single saccade into the face image from a peripheral fixation location. This early fixation in face viewing might have been affected by initial fixation bias (Arizpe, Kravitz, Yovel & Baker, 2012) or the centre-of-gravity bias (Bindemann, Scheepers & Burton, 2009) instead of reflecting a genuine preference to sample information from the recorded facial region.

Preferential sampling of informative facial features during expression-related tasks

During free-viewing paradigms of emotion recognition where the observers are allowed to make as many fixations as they want during a fixed period, eye movements are specific to the presented facial expression. Eisenbarth and Alpers (2011) showed participants images of angry, fearful, happy, sad and neutral facial expressions for 2.5 seconds and asked them to rate these expressions on valance and arousal. Overall, the observers looked more at the eyes and the mouth of the faces; however, this gaze pattern was modulated by the expression of the face image. Observers spent less time looking at the eyes and more time looking at the mouth of happy expressions consistent with the suggestion that the mouth is the informative facial feature for happy faces. The eye and mouth regions were fixated equally for happy, fearful and neutral faces while the eye region was fixated more for the angry and sad expressions. Guo (2012) recorded the eye movements of the observers during an emotion categorisation task. The observers saw expressions of the six basic emotions at 5 different intensities (20, 40, 60, 80 and 100% creating using a morphing technique) and were asked to categorise the facial expression. They found that overall, regardless of facial expressions, there was a higher proportion of fixations and longer viewing time within the eye region followed by the nose and the mouth regions. This overall finding agreed with Eisenbarth and Alberts (2011). Again, the eye movement patterns were specific to each expression: For happy and surprised expressions, the mouth region received the highest proportion of fixations and viewing time; while for disgusted and sad expressions the nose received the highest proportion of fixations and viewing time. For angry, fearful and surprised expressions the eye region received the highest proportion of fixations and viewing time.

Schurgin et al. (2014) conducted a thorough experiment investigating eye movement patterns while observers completed an expressive/non-expressive task for facial expressions of anger, fear, happiness, sadness, disgust and shame. Each emotion was presented at four different levels of intensity and was to be distinguished from a neutral face in separate blocks. From an initial number of 21 regions of interest (ROIs), the authors isolated 5 main ROIs that received 88.03% of total fixation time: eyes, upper nose, lower nose, upper lip and the naison. Like the studies mentioned above, time spent fixating on these regions was affected by the presented emotion. The percentage of fixation time on each ROI was compared to the mean percentage of fixation time on that ROI collapsed across all expressions: For happy and disgusted faces, observers spent longer looking at the upper lip and shorter on the eyes. For fearful, angry, sad and shameful faces, observers fixated longer on the eyes. Additionally, observers spent less time fixating on the upper lip for sad and angry faces. Furthermore, as the intensity of the facial expressions increased, more fixations were directed towards the diagnostic features, especially for angry, happy and disgusted faces. The facial regions observers preferred to fixate when viewing facial expressions largely corresponded to the facial regions suggested to be informative of each facial expression. This provides strong support for the preferential processing of informative features when the whole face is visible with all the facial features available for visual sampling. However, they do not provide a direct link between fixation behaviour and emotion recognition ability. Therefore, we do not know whether fixations on informative facial features are functionally important for emotion recognition.

Foveal and extrafoveal processing of informative facial features

The free-viewing methodology used in studies such as Schurgin et al. (2014), Guo (2012) and Eisenbarth and Alpers (2011) allows observers to shift overt attention to different facial features. This makes it difficult to pinpoint the contribution of fixating an individual informative feature over non-informative features. In order to investigate the individual contributions of foveal and extrafoveal processing of facial features to emotion recognition, we require a method which allows us to control what information is being presented at the fovea and what is being presented extrafoveally. One such well-established method is gaze-contingent viewing. This method has been utilised for reading (Rayner, 1994), scene viewing (Cajar, Engbert & Laubrock, 2016) and more relevant to our purposes, face recognition (Caldara, Zhou & Miellet, 2010; Miellet, He, Zhou, Lao & Caldara, 2012). Previous studies have shown that there are differences between the eye movement strategies of Western Caucasian (WC) and Eastern Asian (EA) observers: During face recognition, WC observers prefer to fixate the eyes and the mouth whereas the EA observers fixate a more central location on the face. This led to the suggestion that while WC observers prefer to fixate and foveally process facial features such as eyes and mouth, the EA observers fixate a more central location on the face to obtain information from facial features extrafoveally. Caldara et al. (2010) investigated the eye movement patterns of EA and WC observers in a face recognition task under four conditions: natural vision, 2-degree, 5-degree and 8-degree gaze-contingent *spotlights.* The spotlight moves with the fixation of the observer and allows vision through an aperture of the chosen size while blocking extrafoveal vision outside this aperture. The same group of researchers (Miellet et al., 2012) investigated the eye movement patterns of WC and EA observers during face recognition using a gaze-contingent *blindspot* method as well (same size as the spotlights). In contrast to the *spotlight,* the *blindspot* blocks central vision with a blindspot while leaving extrafoveal vision un-occluded. Both studies showed that in the natural viewing condition, WCs looked more at the eyes and partially at the mouth and EAs looked more at the centre of the face. Caldara et al. (2010) found that with the gaze-contingent *spotlight*, both WC and EA observers fixated on the eyes and mouth in the 2- and 5-degree spotlight conditions. With this gaze-contingent spotlight where extrafoveal vision is blocked, the EA observers started relying on the foveal processing of facial features they normally process extrafoveally. On the other hand, in the 8-degree spotlight condition where extrafoveal processing of the eyes and mouth is made possible EA observers fixated the nose more compared to the WC observers. The gaze-contingent spotlight condition showed the flexible information use in EA observers in the absence of extrafoveal information. Miellet et al. (2012) extended these results by showing flexible use of visual information in WC observers with the gaze-contingent *blindspot*: When no foveal information was available from the fixated facial features (especially in the 5- and 8-degree blindspots), WC observers started fixating the nose in order to obtain extrafoveal information from the eyes and the mouth. While the gaze-contingent manipulation is an established way of manipulating what visual information is presented foveally and extrafoveally, it does not simulate the natural way we look at faces.

Another way of selectively presenting visual information foveally is the *brief-fixation paradigm* used by Gamer and Büchel (2009). Using eye-tracking and functional magnetic resonance imaging (fMRI), Gamer and Büchel (2009) investigated whether the amygdala was involved in detecting and orienting towards salient features of facial expressions, especially the eye region of fearful faces. For this purpose, they used a novel paradigm: Each trial started with a central fixation cross and the following face image was randomly shifted upwards or downwards such that the location of the fixation cross corresponds to either the left/right eye or the mouth. Following fixation of a random duration between 2 and 12 seconds, a face image displaying neutral, fear, happy or angry expression was presented for 150ms. The faces were presented for this brief duration to prevent participants moving their eyes to and thus fixating another part of the face. Two aspects of this paradigm allow the discrimination of the contribution of foveal and extrafoveal processing: Since there are no eye movements during face presentation, foveal information can only be obtained from the fixated feature (i.e., the eyes or the mouth). Secondly, the role of extrafoveal processing can be investigated through the examination of reflexive saccades. Gamer and Büchel (2009) defined reflexive saccades as the saccades triggered by the presentation of the face but occurring within 1000ms after the face disappeared. The direction of these reflexive saccades gives an indication of what extrafoveally presented facial feature the observers are seeking out. Using this paradigm, Gamer and Büchel (2009) found that initially fixating on the eyes or the mouth did not improve emotion recognition for any of the presented expressions. The only effect of initial fixation was higher intensity ratings when initial fixation location was the eyes compared to the mouth. However, they found that the proportion of saccades going upwards from the mouth region was higher compared to the proportion of saccades going downwards from the eyes. Furthermore, this effect was modulated by expression where the proportion of upward fixation changes from the mouth was greater for neutral and fearful expressions compared to happy expressions. Gamer and Büchel (2009) also found higher amygdala activation for fearful faces when the mouth region was initially fixated compared to when the eye region was fixated. This difference was not present for happy faces. Furthermore, Gamer and Büchel (2009) also showed that amygdala activity was positively correlated with upwards shifts from the mouth region: Participants who made more fixation change upwards from the mouth region showed higher (right) amygdala activation. These results suggest that amygdala might be involved in orientation of gaze towards the eyes of fearful faces, which are the informative facial features for fear recognition. The role of amygdala in fear recognition was previously established by studies of patients with amygdala damage who showed selective impairment of fear recognition (Adolphs, Tranel, Damasio & Damasio, 1994; Adolphs et al., 1999) and an inability to fixate the eye region spontaneously during an emotion recognition task (Adolphs, Gosselin & Buchanan, 2005; Gamer, Schmitz Tittgemeyer, & Schilbach, 2013). While these results suggest that there is no contribution to emotion recognition from foveal processing of the eyes or the mouth, the extrafoveal informative facial features, such as the eyes for fear, are being sought out by observers during their first saccades.

Using the same paradigm as Gamer and Büchel (2009), Scheller, Büchel and Gamer (2012) also failed to find an effect of initially fixating either the eye region or the mouth region on emotion recognition accuracy for happy, fearful and neutral facial expressions. Additionally, in line with Gamer and Büchel (2009)'s findings, Scheller et al. (2012) also found fewer saccades leaving the eye region compared to saccades targeting this region. This effect was again modulated by expression: participants made more saccades upwards from the mouth for fearful and neutral expressions, but the pattern was reversed for happy expressions. Additionally, this preference to shift gaze upwards from the mouth region was present for longer presentation time and for gender discrimination and passive viewing tasks. The preference to saccade upwards from the mouth region of briefly presented faces was replicated again by Gamer et al. (2013). Gamer et al. (2013) investigated emotion recognition for angry, happy, fearful and neutral expressions for patient MW who has right amygdala damage and control patients. They found that control patients showed a preference to saccade upwards from the mouth region for briefly presented expressions but less so for happy expressions. They also showed that patient MW showed less reflexive fixation changes overall, regardless of initial fixation location. However, when the presentation duration was longer, MW made a similar number of fixations to the eye and mouth regions as the controls. Kliemann, Dziobek, Hatri, Steimke and Heekeren (2010) and Kliemann, Dziobek, Hatri and Heekeren (2012) investigated emotion recognition and gaze patterns for neurotypical (NT) adults and adults with ASD. Both Kliemann et al. (2010) and Kliemann et al. (2012) used a similar paradigm to Gamer and Büchel (2009) with fearful, happy and neutral facial expressions. Both studies showed a greater preference to gaze upwards towards the eyes in NT adults with this effect being less prominent for happy expressions. Kliemann et al. (2010) also showed an effect of initial fixation location on emotion recognition: They showed that both NT and ASD participants were better at recognising fear when fixating the eye region and better at recognising neutral expressions when fixating the mouth region. Kliemann et al. (2012) also showed an effect of initial fixation location on emotion recognition accuracy: Initially fixating the eye region led to improved recognition for happy expressions compared to fixating the mouth in both NT and ASD participants.

While Gamer and Büchel (2009) controlled the initial fixation location so that when the faces were shifted upwards they were also shifted horizontally in order to coincide the initial fixation location with either the left or the right eye, this control manipulation was not carried out for any of the subsequent studies that followed Gamer and Büchel (2009)'s paradigm, including Scheller et al. (2012), Gamer et al. (2013), and Kliemann et al. (2010, 2012). Therefore, when initial fixation is said to be at the eye region, in practice (with the exception of Gamer and Büchel, 2009), this position falls between the two eyes, i.e., on the bridge of the nose. It has to be noted that there is need to better classify the trajectories of the gaze shifts defined as 'upwards from the mouth' and 'downwards from the eyes' since the potential endpoint of the reflexive saccade is vague with the broad visual angle range chosen to define gaze shifts. Moreover, there is need to go beyond the broadly defined eye and mouth regions and to investigate more diagnostic initial fixation points as suggested by the *Bubbles* technique, such as the brow region for angry faces.

Neath and Itier (2014) also investigated the effect of initial fixation location on emotion identification and, similar to Gamer and Büchel (2009), failed to find an advantage of initially fixating diagnostic features on emotion identification. Faces displaying neutral, disgusted, fearful, happy and surprised expressions were presented for brief durations (16.67ms, 50ms and 100ms over three experiments) and the initial fixation was aligned to the forehead, either eye, either cheek, mouth or chin. The only reliable difference found was that initially fixating the forehead region led to worse performance than any other initial fixation location. However, since the location of the faces on the screen always remained the same, this raises the possibility that observers might have covertly attended to other locations on the face contributing to performance accuracy.

Neath and Itier's (2014) study addressed the issue of the limited number of initial fixation locations in the previous studies, however; the informativeness of ROIs chosen for the study should also be considered: The mouth ROI covers the central part of the mouth even though the whole smile is suggested to be the diagnostic feature of happy faces (Smith et al., 2005). Furthermore, the fact that there was no difference between emotion recognition performance at any initial fixation, regardless of diagnosticity, might indicate that extrafoveal detection of the diagnostic feature contributes to emotion recognition performance at non-diagnostic initial fixation points. This possibility can be supported by the observation that initially fixating the forehead region led to poorest performance. The forehead region is furthest away from all the internal facial features; therefore, fixating on the forehead diminishes the quality of the extrafoveal information that can be obtained from other facial features. And, since eye movements following initial fixation were not examined, the contribution of extrafoveal information to eye movement guidance cannot be identified or compared to Gamer and Büchel's (2009) findings.

Functional role of informative facial features in emotion recognition

Studies where initial fixation location is controlled do not show an unequivocal contribution to emotion recognition accuracy of foveally processing informative facial features. While some studies show no effect of initially fixating on the informative features on emotion recognition (Gamer & Büchel, 2009; Scheller et al., 2012; Neath & Itier, 2014), other studies suggest fixating the eye region is informative for fear recognition (Kliemann et al., 2010) or for happiness recognition (Kliemann et al., 2012). Beyond paradigms where initial fixation location was controlled, Vaidya, Jin and Fellows (2014) investigated whether the sum of fixations to the eye, nose and mouth regions could predict participants' performance in an emotion rating task. They modelled a linear relationship between the sum of fixations to the eye, nose and mouth regions of extreme and subtle happy, fearful, disgusted and surprised expressions in order to predict emotion detection performance. They showed that the sum of fixations within the eye, nose and mouth regions was not a good predictor of recognition of most full-blown facial expressions. For full intensity fearful expressions, the sum of fixations to the nose and the mouth region was found to contribute more to fear recognition compared to the sum of fixations to the eyes. The sum of fixations within the eye region was found to predict emotion recognition performance for subtle fear, disgust and surprised expressions. For subtle fear and disgusted expressions, mouth fixations were also found to contribute to explaining emotion recognition performance.

Overall, the evidence regarding the functional role of informative facial features to emotion recognition is far from conclusive. However, research on individuals with bilateral amygdala damage and individuals with neurodevelopment disorders, such as ASD show that atypical gaze strategies lead to emotion recognition impairments. A bilateral amygdala lesioned patient, SM, was shown to be impaired in recognising fearful expressions compared to controls (Adolphs, Tranel, Damasio & Damasio, 1994). Kennedy and Adolphs (2010) showed that SM fixated less on the eye region when looking at faces and this effect was most pronounced for first fixations. They also showed that when SM viewed faces through a gaze-contingent spotlight, she made more fixations to the eyes. Therefore, it is possible that when viewing faces naturally, SM might be getting visual information from eyes extrafoveally and once the gaze-contingent spotlight blocks extrafoveal face information, SM uses her top-down knowledge of the location of the eyes to look there to obtain visual information. Adolphs, Gosselin, Buchanan, Tranel, Schyns and Damasio (2005) investigated the fear recognition impairment of SM using the *Bubbles* technique and found that in contrast to control subjects, SM did not make use of high spatial frequency information from the eye region when recognising fear. Additionally, when viewing images of all six facial expressions, SM made fewer fixations to the eyes for all expressions compared to control subjects. Finally, Adolphs et al. (2005) found that SM's impairment in fear recognition was reversed when explicitly instructed to fixate the eye region. So, while SM might be making use of information from the eyes extrafoveally, direct fixation on the eye region improves her emotion recognition.

A large number of studies investigating emotion recognition abilities in individuals with Autism Spectrum Disorder (ASD) have shown that they are impaired in recognition of emotion from facial expressions, especially negative emotions such as fear (Griffiths, Jarrold, Penton-Voak, Woods, Skinner & Munafò, 2019; Philip et al., 2010; Shanok, Jones & Lucas, 2019; Uljarevic & Hamilton, 2013). Studies suggest that this emotion recognition impairment is due to an atypical gaze strategy; for example, Spezio, Adolphs, Hurley and Piven (2007) used the *Bubbles* technique to show that individuals with high functioning autism relied more on the mouth region compared to the eyes and fixated the mouth more compared to controls. More importantly, combining the *Bubbles* technique with eye-tracking, Spezio et al. (2007) found that when ASD observers looked at the mouth region, there was more information at the left eye and they saccades away from the left eye and the right eye when there was expression information within the eyes. However, despite these behavioural differences, Spezio et al. (2007) did not find any difference in the recognition of facial expressions of happiness and fear between ASD and control subjects. On the other hand, it has to be noted that the *Bubbles* technique is a self-calibrating technique which aims to maintain recognition around a certain level (~80% in Spezio et al. 2007). While Spezio et al. (2007) gives us an insight into how individuals with ASD sample information from facial expression images, there is need to link this behaviour to emotion recognition in the more usual case – when whole face is visible. Corden, Chilvers and Skuse (2008) showed that individuals with Asperger's syndrome spent less time fixating the eye region compared to the control group and there was a correlation between fixations made to the eyes and fear recognition accuracy in participants with Asperger's syndrome. Furthermore, using the brief-fixation paradigm, Kliemann et al. (2010) and Kliemann et al. (2012) showed individuals with ASD had a higher preference to gaze towards the mouth region when initially fixating the eye region and this pattern was consistent across emotions. Crucially, higher preference to look towards the eyes in individuals with ASD was correlated with better emotion recognition performance. Even though this correlation was absent in the neurotypical group, findings from adults with ASD indicate that gazing at or detecting diagnostic facial features contributes to accurate emotion recognition performance.

Outstanding questions and aims of this thesis

In summary, the human face transmits emotional information through the channel of facial features that are informative for specific emotions (Smith et al., 2005; Smith & Merlusca, 2015; Smith et al., 2018; Wegrzyn et al., 2017). There is some evidence to show that while observers have a tendency to sample information mainly from the eye region (Eisenbarth & Alpers, 2011; Guo, 2012), their eye movements are still influenced by informative facial features of facial expressions (Schurgin et al., 2014; Eisenbarth & Alpers, 2011). Despite their preferential sampling, evidence regarding the functional role of fixating informative facial features is equivocal. While studies of emotion recognition in patients with amygdala damage and individuals with ASDs suggest that fixation to the eye region plays a critical role in recognition of negative expressions, evidence from typical observers is less conclusive. The overarching aim of the experiments reported in this thesis is to examine the role of foveal and extrafoveal processing of informative facial features in facial expression recognition using the brief-fixation paradigm used by Gamer and Büchel (2009) with several modifications to address some of the issues raised previously: One of the novel contributions of this thesis will be to investigate the contribution of initially fixating informative features additional to the eyes and mouth, such as the brow region for angry faces. Another novel contribution will be a more detailed investigation of the seeking out of informative features in order to better classify what facial feature/information is being sought.

Chapter 2 will describe two experiments that investigate the contribution of foveal and extrafoveal processing of informative features to emotion recognition. Both experiments comprise a brief-fixation paradigm (Gamer & Büchel, 2009) and a free-viewing paradigm and each of the two experiments uses a different set of four expressions. As well as the contribution of foveated informative regions, these studies investigate whether observers seek out informative facial features when they are not initially fixated. We use a novel and more precise measure of saccade path to determine not only if reflexive saccades are going upwards or downwards (cf. Gamer & Büchel, 2009) but also if reflexive saccades are targeting a specific facial feature. Finally, the data from the long presentation paradigm is used to investigate gaze preference for informative features under free-viewing conditions and to establish whether there is a correlation between time spent fixating on an informative feature and emotion recognition accuracy.

In *Chapter 3,* we investigate the contribution of foveal and extrafoveal processing of informative features for lower intensity expressions using the brief-fixation paradigm alone. In this experiment, we wanted to know whether informative features continued to convey expression-informative signals at lower intensities - subtler versions of the prototypical expressions which are more likely to be encountered in everyday scenarios. Since a single facial feature in a lower intensity expression would convey more ambiguous visual information regarding the identity of the expression, we reasoned that foveating an informative feature would be more critical for lower intensity expressions (cf. Vaidya et al., 2014). In the same vein as the first two experiments, we also investigated seeking out of informative features in lower intensity expressions.

Finally, *Chapter 4* describes the fourth and final experiment where we investigated the underlying mechanism for the contribution of initially fixating on an informative feature that we found for some expressions in the previous experiments. Due to the differences in visual processing across the retina, we suggested that foveating an informative facial feature leads to improved recognition through processing of expression-related visual information at high spatial frequencies that would not be accessible when this feature is presented extrafoveally. We tested this theory using the brief-fixation paradigm with images of facial expressions where the initially fixated feature was either visible, occluded or low-pass filtered. The low-pass filtering removed the HSF from the fixated feature. The filtering simulated the loss of HSF that would happen at the informative feature when it was visible at the extrafoveal eccentricity when a non-informative facial feature was fixated. We measured emotion recognition and we hypothesised that, if our suggestion was accurate, emotion recognition in the low-pass filtered condition would be lower compared to the visible condition.

Chapter 2

Contribution of foveal and extrafoveal facial features to facial expression recognition: An investigation of emotion recognition performance and gaze behaviour using forced fixation and free viewing paradigms

Introduction

The two experiments reported in this chapter investigated the question of whether initially fixating on an informative facial feature improved emotion recognition over initially fixating less informative facial features. Furthermore, the effect of informative features when they were not fixated – i.e., are visible in the extrafoveal regions – was examined by investigating the saccade direction and saccade paths of reflexive saccades triggered by the onset of a briefly presented face. While previous studies did not find an effect on emotion recognition of initial fixation on informative facial features (Gamer & Buchel, 2009; Scheller, Büchel & Gamer, 2012; Neath & Itier, 2014), they either investigated a limited number of initial fixation locations (i.e. only eyes and mouth, Gamer & Buchel, 2009; Scheller et al., 2012) or initial fixation locations that were not shown to be informative for any of the studied expressions (i.e. forehead, chin, Neath & Itier, 2014). Saccade directions from initially fixated locations provide an index of what extrafoveal facial features are being sought out; however, previous studies (Gamer & Buchel, 2009; Scheller et al., 2012) used a vague measure of saccade direction, namely, the relative proportions of fixation changes upwards from the mouth and downwards from the eyes. Therefore, using a novel measure, we investigated whether these reflexive saccades are targeting specific facial features. Finally, we investigated whether different facial expressions elicit specific fixation patterns and whether these fixation patterns are related to emotion recognition performance.

Facial expressions of each basic emotion have unique facial appearances due to the distinctive activation of individual facial muscles or groups of muscles, called AUs (Ekman & Friesen, 1978). Certain facial features are shown to be more informative than others for recognition and discrimination of each expression from the others (see Figure 1; *Bubbles Technique* – (Smith, Cottrell, Gosselin & Schyns, 2005; Smith et al., 2018; Smith & Merlusca, 2014a) and AUs defining a facial expression are weighted more highly compared to non-action units when the task is to recognize an expression (Wegrzyn et al., 2017). For example, as can be seen in Figure 1 from Smith et al. (2005), the *Bubbles* technique (discussed in detail in the Introduction Chapter) showed that, for human observers, the brow region is informative for an angry face, especially at higher spatial frequencies. For fearful faces, the eye region was informative especially at high spatial frequencies. Furthermore, Smith et al. (2005) pointed out that the six basic emotions are transmitted with minimal overlap in signal except for fear and surprise.

Whereas the ideal observer analysis of fearful and surprised facial expressions showed that there were expression-informative signals in both the eye and the mouth regions of these expressions (Figure 1, bottom row), human observers were shown to make use of the eye region to recognise a fearful face and the mouth region to recognise a surprised face (Figure 1, top row). In line with the findings from the ideal observer analysis, fear and surprised expressions are often shown to be confused (Calder, Rowland, et al., 2000; Chamberland, Roy-Charland, Perron & Dickinson, 2017; Gagnon, Gosselin, Hudon-ven der Buhs, Larocque & Milliard, 2010; Jack et al., 2009, 2014; Roy-Charland, Perron, Beaudry & Eady, 2014; Roy-Charland, Perron, Young, Boulard & Chamberland, 2015a). The perceptual-attentional hypothesis suggests that this confusion is due to the lack of attention to distinctive features of the target expression due to the overall visual similarity of the two expressions. This line of research also showed that the presence of the lip stretcher action unit in the fearful expression reduces the misclassification of fear as surprise (Roy-Charland, Perron, Young, Boulard & Chamberland, 2015b).

Eye-tracking studies show that informative facial features modulate eye movements when observers view images of facial expressions (Eisenbarth & Alpers, 2011; Guo, 2012; Schurgin et al., 2014). Eisenbarth and Alpers (2011) showed participants images of angry, fearful, happy, sad and neutral facial expressions for 2.5 seconds and asked them to rate these faces on their valance and arousal. Observers spent less time looking at the eyes and more time looking at the mouth of happy expressions (Eisenbarth & Alpers, 2011), consistent with the suggestion that the mouth is the informative facial feature for happy faces. In Schurgin et al.'s (2014) study, participants saw each of the six basic expressions at various intensities in separate blocks along with neutral expressions. After the three-second presentation of each face, participants were asked to decide if the face showed any expression or not. They also showed that the total fixation duration within the eye region was longer for angry, fearful and sad faces and longer for the mouth of joyful and disgusted faces. Interestingly, they also showed that observers looked at the nasion (bridge of the nose, between two eyes) of neutral faces when these were presented in a block of angry and neutral faces. Despite not having any distinctive properties in a neutral face, the nasion and the region just above (glabella) contain useful visual information for the recognition of angry faces as suggested by Smith et al. (2005). This result suggests that observers rely on the nasion when deciding if a face is angry or neutral, despite not fixating within this region for angry faces more compared to other expressions. Fixation patterns indicate which facial features were foveated and therefore were processed in finest visual detail. Thus, eye-tracking studies suggest that certain facial features, seemingly the suggested emotion-informative facial features, play a role in guiding eye movements when people view facial expressions. However, since the task in the aforementioned studies was not expression recognition, it is difficult to argue whether fixating on emotion-informative facial features contribute to emotion recognition performance.

Evidence that fixation of these emotion-informative facial features aids emotion recognition comes from studies that have examined the eye movement patterns of patients with various neurological conditions. Patients with unilateral or bilateral amygdala damage are shown to have impaired emotion recognition, particularly fear recognition, and this impairment was shown to be related to a failure to spontaneously shift gaze towards the eyes (Adolphs, Tranel, Damasio, & Damasio, 1994; Adolphs et al., 2005; Gamer & Buchel, 2009; Gamer, Schmitz, Tittgemeyer & Schilbach, 2013) Adolphs et al. (2005) further showed that the patient SM (with bilateral amygdala damage) was capable of fear recognition upon instruction to fixate the eye region. Recognition of negatively-valanced emotions (i.e. fear, anger, sadness), especially fear, was shown to be impaired in individuals with Autism Spectrum Disorders (ASDs) and the impairment is linked to a decreased preference to fixate the eye region of the facial expressions (Corden, Chilvers & Skuse, 2008; Kliemann, Dziobek, Hatri, Steimke & Heekeren, 2010). These studies suggest that fixating on the informative facial features is key to successful emotion recognition and failure to do so leads to impairments in emotion recognition.

At normal interpersonal distances, when one facial feature is fixated, that feature falls on the fovea and can be processed with high visual acuity while the rest of the face remains in the extrafoveal regions. Since there is a decline in the visual resolution capabilities of the retina with eccentricity from the fovea, the ability of retinal cells to process high spatial frequencies drops and increasingly lower spatial frequencies are processed from non-fixated facial regions. Several studies investigated the recognition of emotions from briefly presented faces where the initial fixation location was manipulated (Gamer & Buchel, 2009; Scheller et al., 2012; Neath & Itier, 2014; Klienmann et al. 2010). Use of these brief-fixation paradigms allows the investigation of the separate contributions of foveal and extrafoveal processing of facial features to emotion recognition. For example, Gamer and Buchel (2009) investigated whether initially fixating on either the eyes or the mouth of neutral, angry, fearful and happy expressions affected emotion recognition and amygdala activation. The faces were presented for 150ms at a screen location that ensured participants fixated either the mouth or one of the eyes. The brief presentation duration was insufficient for participants to make an eye movement to other parts of the face and therefore foveal processing was only possible at the fixated region. Several studies using this or similar paradigms showed no effect of initial fixation location on recognition for any of the studied expressions (Gamer & Buchel, 2009; Scheller et al., 2012; Neath & Itier, 2014) except Kliemann et al. (2010) who showed that fixating on the eyes led to better fear recognition and fixating on the mouth led to better neutral face recognition in both neurotypical and autistic groups. Majority of the aforementioned brief-fixation paradigms used a limited number of facial expressions, namely happy, fear and neutral and in the case of Gamer and Buchel (2009), angry faces. Only Neath and Itier (2014) investigated the recognition of all six of the basic emotions. Additionally, initial fixation locations used in these studies were limited to either the eyes, the area between the eyes or the mouth (Gamer & Buchel, 2009; Scheller et al., 2012; Klienmann et al., 2010, 2012) or relatively less informative for the studies expressions (Neath & Itier, 2014). In the current study, we aim to examine a wider range of expressions and initial fixation locations to understand how emotion recognition accuracy varies as a function of fixation location and expressed emotion.

While the effect of initially fixating on a facial feature informs us about whether foveal processing of that feature is beneficial for recognition of target expressions, the subsequent saccades inform us about what extrafoveal information is being sought out following initial fixation on a facial feature. Gamer and Buchel (2009) and Scheller et al. (2012) revealed that there was a higher proportion of fixation changes upwards from the mouth compared to fixation changes downwards from the eyes. This effect was more pronounced for fearful faces (Gamer & Buchel, 2009; Scheller et al., 2012) and observers directed gaze towards the mouth region for happy faces more strongly (Scheller et al., 2012). Kliemann et al. (2010) showed a similar pattern with the group of neurotypical participants but not with the group of participants with ASD. Furthermore, they showed that the propensity to direct fixation towards the eyes was related to better emotion recognition accuracy. These fixation changes after the disappearance of a face suggests that these saccades might have been triggered in response to the face and seek out the informative facial features of the expression (i.e., the eyes of the fearful expressions) when not initially fixated. In a similar vein, Bodenschatz, Kersting and Suslow (2018) showed that when primed with fearful faces, observers fixated the eye region of a subsequently presented neutral face quicker and dwelled on this region for longer whereas they fixated the mouth region quicker and dwelled on this region for longer when primed with a happy face compared to a sad prime. Taken together, these results suggest that the appearance of informative facial features outside foveal vision can trigger reflexive eye-movements towards these features.

On the other hand, the *centre-of-gravity effect* has been suggested to modulate the targets of early saccades so that they land on the geometric centre of a visual stimulus (Coren & Hoenig, 1972; Findlay, 1981, 1982; Findlay & Gilchrist, 1997; He & Kowler, 1989). Bindemann, Scheepers and Burton (2009) investigated the centre-of-gravity effect for faces presented in different viewpoints. In a series of two experiments, participants were asked to either identify the gender of a face or to freely view a face without any task. The initial fixation location was at a peripheral location, outside the face boundaries and the faces were presented at different screen locations. They showed that within 250ms of face onset, fixations targeted the centre of the face in different viewpoints. This coincided with the area between the two eyes in the frontal view, the innermost eye in the mid-profile view and the area between the ear and the visible eye in the profile view. Therefore, Bindemann et al. (2009) argue that the early fixations to a face are mostly driven by the centre-of-gravity effect rather than by a facial feature. Since the reflexive saccades/fixation changes measured by the previous studies using the brief-fixation paradigm (e.g. Gamer & Buchel, 2009) occur within this time window, it is possible that the centre-of-gravity effect might be partially influencing the reflexive saccades measured in these studies. It must be noted that Bindemann et al. (2009) showed the centre-of-gravity effect for gender discrimination and general face perception with no task constraints. The application of this centre-of-gravity effect to images of facial expressions is not directly researched; however, Guo and Shaw (2015) investigated the fixation patterns for the six basic emotions in three different viewpoints which are the same as Bindemann et al. (2009) – frontal, mid-profile and profile views. They observed that central features such as the cheeks and ears attracted more fixations in the profile view compared to the other views. Therefore, it is possible that the centre-of-gravity effect influences fixations even when expression-informative features are present, which might have influenced/polluted the direction of the fixation changes reported by previous researchers. In the studies reported in this chapter, we aim to address this confound by using a saccade path measure that quantifies which facial feature reflexive saccades are targeting.

According to the model of saccade generation by Findlay and Walker (1999), the metrics of a saccadic movement are determined by processes that control when and where a saccade will be made. The aforementioned research by Gamer and Buchel (2009), Scheller et al. (2012), Kliemann et al. (2010) and Bodenschattz et al. (2018) suggests that the 'where' component of a saccadic movement might be determined by the location of an expression-informative facial feature within a face, however; this previous research did not investigate whether the 'when' aspect of a saccade can also be affected by emotion-informative facial features. Findlay and Walker's (1999) model suggests that one factor that can influence when a saccade is made is ongoing cognitive and perceptual processing. Therefore, if the current fixation is on a stimulus that requires complex cognitive and perceptual processing, the eye will remain fixated at that region for longer before moving on to the next target location. Arizpe, Kravitz, Yovel and Baker (2012) conducted a facial recognition task where they manipulated initial fixation to either fall on the centre of the face or on a peripheral location which did not correspond to an internal facial feature. They found that the saccade latencies from the central location on the face were longer compared to peripheral fixation locations. The central location on a face is suggested to contain most visual information (Hsiao & Cottrell, 2008; Or, Peterson & Eckstein, 2015; Peterson & Eckstein, 2012); therefore, considered within the saccade generation model by

Findlay and Walker (1999), more perceptual processing might be required at this location leading to longer saccade latencies. However, since Arizpe et al. (2012) focused on face recognition, whether emotion-informative facial features will create a similar effect on saccade generation remains to be investigated.

Experiment 1

Results from the brief-fixation paradigms using typically developing participants suggest no contribution to emotion recognition of the facial feature initially fixated; however, these studies either investigated a limited number of facial expressions (Gamer & Buchel, 2009; Scheller et al., 2012) or they investigated initial fixation locations that did not align with what was suggested by Smith et al. (2005) and Wegrzyn et al. (2017) (e.g., Neath & Itier, 2014). Furthermore, the seeking out of informative facial features indexed by the relative proportions of fixation changes upwards or downwards from different facial features constitutes a vague measure in the sense that upwards from the mouth or downwards from the eyes might target any number of facial features or non-features. In this experiment, we aimed to expand on the results of studies investigating the effects on emotion recognition accuracy of a single fixation on emotion-informative facial features versus less informative features. We used additional initial fixation locations and expressions. Furthermore, we investigated the target of reflexive saccades using a more refined measure of saccade direction. We used a modified version of Gamer and Buchel's (2009) brief-fixation paradigm with angry, fearful surprised and sad expressions presented briefly at initial fixation locations that corresponded to either eye, the central lower brow, the central mouth or either cheek. We chose the brow region as an initial fixation location since the Smith et al. (2005) study suggested that this region contains informative visual information at high spatial frequencies for anger recognition. Most previous research included the brow region as part of the region of interest circumscribed as eyes (e.g., Beaudry, Roy-charland, Perron & Cormier, 2014; Eisenbarth & Alpers, 2011), therefore it is possible that this might have undermined the informativeness of this region for angry expressions. We hypothesized that initially fixating on the brow region will improve emotion recognition for anger compared to other initial fixation locations. We chose fearful and surprised expressions due to their established confusability. Since previous research suggests that processing of information from the mouth region might lead to the resolution of this confusion, we hypothesised that initially fixating on the mouth region will improve emotion recognition accuracy for these expressions. Finally, we chose expressions of sadness since the results from Smith et al. (2005) suggested that there is high spatial frequency information at both the brow and mouth regions that is informative for sadness. We chose the cheeks to be relatively uninformative regions for the chosen expressions.

In this experiment, we further employed a free-viewing paradigm that replicated the brief-fixation paradigm in every aspect except that the faces were presented for longer. When a face is presented for a few seconds, several eye movements are made allowing the observer to fixate on multiple facial features. In general, when viewing faces, people tend to make a triangular pattern of fixations whereby they look at the eye region the most frequently and for longer, followed by the nose and the mouth (Bindemann et al., 2009; Henderson et al., 2005; Sekiguchi, 2011). As discussed previously, this pattern of eye movements can be modulated when the viewed face is showing an expression (Calvo, Fernández-Martín, Gutiérrez-García & Lundqvist, 2018 - dynamic facial stimuli; Eisenbarth & Alpers, 2011; Schurgin et al., 2014 - static facial stimuli).

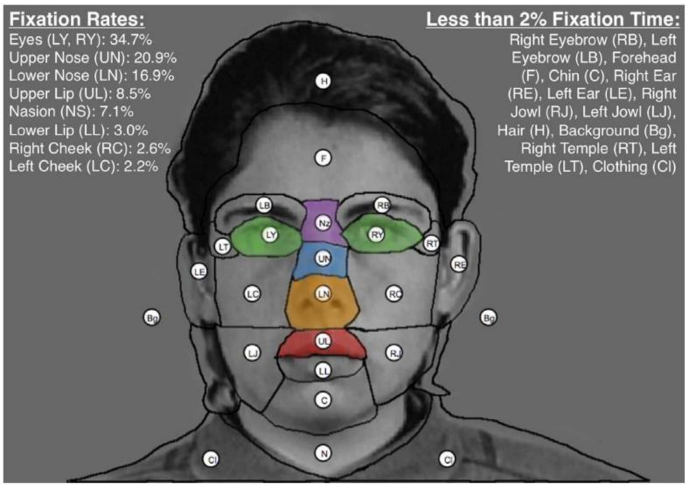

Figure 4: Figure taken from Shurgin et al. (2014) shows the regions of interest investigated in color.

Research has found that fixations tend to accumulate around the eye region for the recognition of angry and sad expressions, the mouth for happy expressions, the mouth and nose for disgust expressions and the eyes and mouth for fear and surprised expressions (Beaudry et al., 2014; Schurgin et al., 2014). The eye region of interest in these studies almost always consists of the two eyes, eyebrows and the region between the two eyes, which makes it difficult to pinpoint whether there is a subregion that people preferentially fixate within the large area defined as the 'eyes'. For example, Shurgin et al. (2014) used more circumscribed regions of interest (as can be seen in Figure 4) and showed that the nasion attracted more fixation time when discriminating angry expressions from neutral faces. Additionally, very few studies investigated whether there is a relationship between fixating these informative facial regions and emotion recognition accuracy.

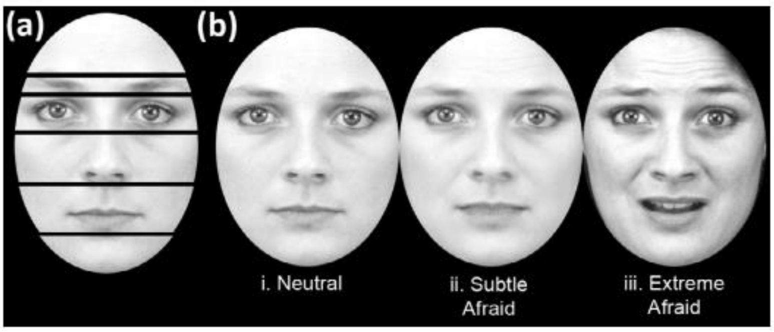

Figure 5: Figure taken from Vaidya et al. (2014). Image (a) shows the region of interest chosen for the investigation of the number of fixations.

Wong, Cronin-Golomb, and Neargarder (2005) found that fixating the top halves of facial expressions more than the bottom halves was correlated with better accuracy for anger, fear and sadness recognition and fixating the lower half of the face more frequently was correlated with higher disgust accuracy. On the other hand, Vaidya, Jin and Fellows (2014) showed that sum of fixations to the eyes (including the space between the eyes but not the eyebrows), nose (including the cheeks) and mouth regions failed to predict the accuracy score when expressions are presented at full intensity but when the expressions are presented at lower intensities, the number of fixations to the eye region was the most predictive of emotion recognition accuracy, specifically for fearful and disgusted faces. However, the regions of interest in this study were strips of facial regions (seen in Figure 5(a)) – for example the nose region included the nose and the rest of the face in the same horizontal line. This might have reduced the predictiveness of the fixations for emotion recognition since fixations that landed on the cheeks as well as the nose will be entered into the model rather than fixations that landed on the facial feature of interest alone. The eye strip was the only region of interest where most of the interest area included the facial feature of interest alone, therefore fixations are more likely to fall on either eye compared to a featureless space on the face. In the present study, we wanted to further investigate the eye movement patterns specific to the expressions studied and investigate whether there is a correlation between fixating on certain, more circumscribed facial features that were found to be informative in the brief fixation paradigm (eyes, brow, nose, and mouth) and emotion recognition accuracy. We also wanted to investigate whether any effect of initial fixation found in the brief fixation paradigm would persist in the long presentation paradigm.

For the brief-fixation paradigm, we hypothesized that initially fixating on the brow region will improve emotion recognition accuracy for angry facial expressions and initially fixating on the mouth will reduce the misclassification of fearful and surprised faces. We expected to find a higher proportion of saccades going upwards from the mouth compared to downwards from the eyes, in line with previous findings. We also expected that this will be modulated by the location of informative facial features for the target expressions. Using a more precise measure of saccade direction, we aimed to map the direction of reflexive saccades onto six possible saccade paths leading to six possible facial features. With this measure of saccade path, we expected that the paths of reflexive saccades will be most similar to paths leading to informative features. When these images are presented for a longer period, we expected the effect of initial fixation to be negligible since several eye-movements will allow extraction of expression information from several facial features. We also expected the total fixation duration within a facial feature to be modulated by the expression: We expected observers to look at the brow region longer when viewing angry expressions compared to any other expression. We also expected longer total fixation duration for the mouth region for fearful and surprised expressions. Furthermore, we expected to find a positive correlation between the amounts of time spent fixating an informative region and emotion recognition performance.

Methods

Participants

A total of 33 participants took part in the brief-fixation paradigm (female= 31, male=2; mean age= 20.73 years) and 36 participants (female= 34, male=2; mean age= 20.47 years) took part in the long-presentation paradigm. The 33 participants who completed the brief-fixation paradigm also completed the long presentation paradigm. Sample size was selected to be similar to the sample size of studies with similar design. All participants had normal or corrected-to-normal vision. All participants gave written consent to take part in the experiment and were rewarded participant pool credit for their participation. The study was approved by the Durham University Psychology Department Ethics Sub-committee.

Stimuli

Brief-fixation paradigm

24 facial identities (12 males, 12 females) were chosen from the Radbound Face Database (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & van Knippenberg, 2010). All faces were of Caucasian adults with full frontal pose and gaze. Angry, fearful, surprised and sad expressions from each identity were utilised leading to a total of 96 images being used in each experiment. In Experiment 1 and all following experiments, happy expressions were not investigated. The reasons for this were two-fold: Firstly, majority of previous research using a similar paradigm (e.g. Gamer & Buchel, 2009; Scheller et al., 2010; Kliemann et al., 2010; Kliemann et al., 2012) have used fearful, happy and neutral expressions (and in the case of Gamer and Buchel (2009), angry expressions). Since the informative regions for happy and fearful expressions are on vertically opposite sides of the face (i.e. the smile and the wide open eyes), this might have accentuated the difference in the saccade direction between these expressions and limited the generalizability of the findings regarding seeking out of informative facial features (i.e the eyes for fearful expressions). Therefore, we wanted to use expressions that have informative facial features in similar face regions (for example, eyes for both fear and surprise). Secondly, Atkinson and Smithson (2020) have used the brief fixation paradigm used here in order to investigate the contribution of initially fixating informative features to the recognition of happy expressions in a subset of happy, fearful, surprised, neutral and angry expressions. Therefore, we wanted to further investigate the replicability of the results pertaining to fear, surprise and anger when these need to be recognized among different sets of expressions given that informative features for facial expressions are flexible depending on the subset of expressions used (Merlusca & Smith, 2014). The images used were spatially aligned by Lagner et al. (2010) and this procedure was detailed in Langner et al. (2010). The images were cropped from their original size to 384 (width) × 576 (height) pixels, so that the face took up more of the image

than in the original image set. Each image therefore subtended 14.9 (w) × 22.3 (h) degrees of visual angle at a viewing distance of 57cm on a 1024 × 768 resolution screen.

Long-presentation paradigm

A subset of the face image set used in the brief fixation condition was used for the long presentation condition of this experiment, such that there were 12 facial identities (6 males, 6 females). Each individual was presented in each of 4 expressions (angry, fearful, surprised and sad) leading to a total of 48 images.

Design

Participants first completed the long-presentation paradigm and were invited to take part in the brief-fixation paradigm on a later date. A within-subjects design was used both for the long-presentation and the brief-fixation paradigms. The independent variables were Expression (anger, fear, surprise and sadness) and Initial Fixation Location (eyes, brow, cheeks, and mouth). For the long presentation condition, the analysis of fixation duration included an independent variable called Region of Interest (left + right eyes, central brow, nose and mouth).

Brief-fixation paradigm

There were 4 blocks of 96 trials. Over the course of 4 blocks, participants were presented with each expression at each initial fixation location 24 times (once for each identity), with the eye and the cheek fixation locations selected equally (12 times) on the left and right. The stimuli were pseudo-randomly ordered across the 4 blocks, with a new pseudo-random order for each participant.

Long-presentation paradigm

There were 4 blocks of 48 trials. Within each block, each expression was presented at each initial fixation location 12 times. For the eye and the cheek fixation locations, the left and right sides were presented 6 times each within each block. The order of the presentation was randomised for each participant. Since the presentation duration for the expressions in the long presentation paradigm is longer compared to the brief fixation paradigm, the number of trials is halved to keep the overall duration of both experiments similar. Due to a counterbalancing error, half the expressions were presented only on one side of the initial fixation locations that had left and right sides. More specifically, due to this counterbalancing error, angry and surprised expressions were only presented at initial fixation locations on the left (left eye and the left cheek) and fearful and sad expressions were only presented on the right initial fixation locations (right eye and right cheek). This error was rectified in Experiment 2.

Procedure

The experiment was executed and controlled using the Matlab® programming language with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al, 2007). To control stimulus presentation and to measure gaze behaviour, we used an EyeLink 1000 desktop-mounted eye-tracker (SR Research, Mississauga, ON, Canada), with data sampled at 500Hz. Each stimulus block was started with the default nine-point calibration and validation sequences. Recording was binocular but only data from the left eye was analysed since data from the left eye started the trial in a gaze-contingent manner. Default criteria for fixations, blinks, and saccades implemented in the Eyelink system were used.

Brief-fixation paradigm

Figure 6: The location of the six initial fixation locations. The coordinates of these fixation locations were constant across the images.

Each trial started with a fixation cross located at one of 25 possible locations on the screen. This was to make both the exact screen location of the fixation cross and the to-be-fixated facial feature unpredictable. These 25 possible locations for the fixation cross were at 0, 25, or 50 pixels left or right and up or down from the center of the screen. These fixation-cross positions were pseudo-randomly ordered across trials. Faces were presented in a gaze-contingent manner on each trial: The participants needed to fixate within 30 pixels (1.16 degrees of visual angle) of the fixation cross for 6 consecutive eye-tracking samples following which a face showing one of the four target expressions was presented. The location of the fixation cross corresponded to one of the informative (eyes, brow or mouth) or non-informative (cheeks) initial fixation locations for each of the

Figure 7. Examples of angry, fearful, surprised and sad expressions (from left to right) from the Radbound Face Database used for Experiment 1.

presented faces. The initial fixation locations can be seen in Figure 6 and an example of each expression used in this experiment can be seen in Figure 7. The initial fixation locations of the eyes were chosen due to their informativeness for expression of fear and sadness. The cheek locations were chosen to be relatively non-informative facial features for the chosen expressions. The positions of the cheeks were equidistant from the center of the face as the eyes. The initial location of the brow was chosen to be informative for anger and sadness. The central mouth location was chosen to be informative for surprise (and for disgust in Experiment 2), this location was at equal distance from the center of the face as the brow location. At the specified initial location positions, the visual angle of the fovea would cover most of the initially fixated feature. The face was presented for 82.4ms (7 monitor refreshes) on a monitor with an 85Hz refresh rate. Following the face presentation, the participant pressed a key on a QWERTY keyboard to indicate their answer. The row of number keys near the top of the keyboard were used, with 4 for anger, 5 for fear, 8 for surprise and 9. The keys were labelled A, F and Su and Sa from left to right and the order of these keys remained the same for each participant. Participants pressed the A and F keys with the left and the Su and Sa keys with the right hand. This configuration was chosen to optimize the reach of the participants to the keyboard from either side of the chinrest. Participants were asked to memorize the keys so as not to look down towards the keyboard during the experiment; all participants could memorize the keys. A valid response needed to be registered for the next trial to begin.

Long-presentation paradigm

For the long-presentation paradigm, everything was the same as for the brief-fixation paradigm except the face images were presented for 5 seconds. The participants were asked to give a response as soon as they decided what expression the face was showing using the A, S, Su and Sa key. Regardless of when the button press was made, the face image remained on the screen for 5 seconds.

Data analysis

Behavioural data analysis

For the analysis of behavioural data (i.e., accuracy), trials with reaction times shorter than 200ms were disregarded as automatic responses not reflecting a genuine perceptual response for both the brief fixation and the long-presentation paradigms. Following the removal of these outliers, unbiased hit rates were calculated using the formula supplied by Wagner (1993).[3] The

[3] The "unbiased hit rate" (H_u) accounts for response biases in classification experiments with multiple response options (Wagner, 1993). H_u for each participant is calculated as the squared frequency of correct responses for a target emotion in a particular condition divided by the product of the number of stimuli in

unbiased hit rates for the brief-fixation paradigm and the long-presentation paradigm were then arc sine square-root transformed in order to better approximate a normal distribution of the unbiased hit rate data. Emotion recognition accuracy was analysed by a 4 × 4 repeated measures ANOVA with emotion and initial fixation location as factors for each of the paradigms. Unless otherwise stated, the data from the left and right sides of the face (when the fixation location was an eye or cheek) were collapsed for the purposes of these analyses. Where the assumption of sphericity was violated, the Greenhouse-Geisser corrected F value is reported. Significant interactions were followed up by one-way repeated measures ANOVAs. Significant main effects were followed up with Bonferroni-corrected pairwise comparisons. Planned comparisons (one-tailed paired samples t-tests) were conducted to test the hypotheses relating to fixation location specific to each expression separately: For angry expressions, since we expected the brow region to be informative, we compared the accuracy when initially fixating the brow to initial fixation at the eyes, cheek and the mouth. For fearful and surprised expressions, we expected the mouth region to be informative, therefore we compared accuracy when initially fixating the mouth to initially fixating eyes, brow and cheeks. Finally, since the informative feature for sad expressions can be either eyes or the mouth, we did not have any a priori planned comparisons. Where the data failed to meet the normality assumption required for t-tests, the non-parametric alternative Wilcoxon signed ranks tests were used. A Bonferroni-corrected α was used to correct for multiple comparisons and a p-value of 0.017 (.05/3 comparisons) was used as significance criterion for the planned comparisons. No reaction time data analyses were included since reaction time data obtained in both studies were too noisy.

Eye-movement analysis

All eye movement data analysis was conducted on all trials regardless of accuracy.

Reflexive saccade selection criteria

For the brief-fixation paradigm, reflexive saccades were defined as saccades that were triggered by onset of a face image and executed following face offset. Accordingly, reflexive saccades used for the reported analyses were chosen as those that happened within the 82.4ms to 1082.4ms window after face onset (similar to e.g., Gamer & Buchel, 2009; Scheller et al., 2012). In other words, the reflexive saccades used were the ones that were executed within 1000ms after face offset. All reflexive saccades that had amplitudes smaller than 0.5 degrees were disregarded. Finally, of all the saccades that complied with these criteria, only the first saccade was used. After this data reduction, participants who had reflexive first saccades on

---

that condition representing this emotion and the overall frequency that that emotion category is chosen for that condition. H_u ranges from 0 to 1, with 1 indicating that all stimuli in a given condition representing a particular emotion have been correctly identified and that that emotion label has never been falsely selected for a different emotion.

fewer than 20% of the total trials per block (i.e., 20% of 96) in any one of the 4 blocks were removed from further analysis. All saccade direction related analyses were carried out on this set of data.

The latency of the reflexive saccades was calculated as the time elapsed between face offset and the start of the saccade.

Saccade Path Analysis

As discussed in the introduction section, in this study, we aimed to further examine the path reflexive saccades take to investigate what facial features they are targeting. For the purpose of this analysis, we only used the reflexive saccades from the brief-fixation paradigm, validated according to the saccade selection criteria outlined above. To estimate the paths of the reflexive saccades, a method created and previously used by Atkinson and Smithson (2020) was used. Six vectors were plotted from the starting coordinates of each saccade to the coordinates of the six possible initial fixation locations. These make up the saccade path vectors. Then, the dot products of the reflexive saccade vector and the six possible saccade path vectors were calculated and normalised to the magnitude of the saccade path vectors. This measure represents the similarity between the reflexive saccade path and the possible saccade path vectors. Due to the normalisation, a saccade path value of 1 indicates complete similarity of length and direction of the reflexive saccade and saccade path vector while a value of -1 indicates similar length but opposite direction. Saccade path values larger than 1 indicates that the reflexive saccade is longer than the saccade path vector. Therefore, taking the size of the face into account, saccade values larger than an absolute value of 1.5 were removed as outliers.

ROI selection

For the long-presentation paradigm, the regions of interest (ROI) were drawn free-hand using a bespoke C++ programme. The ROIs were both eyes, the central brow region, the nose and the mouth. Since the ROIs were drawn free-hand, the area of these regions was not stable across all the faces; however, free-hand drawing of the ROIs allowed for the individuality of facial features across the faces to be captured. The ROIs examined in the experiment can be seen in Figure 8.

Mean Total fixation duration

Total fixation duration was calculated as a function of ROI. The durations of all the fixations falling within an ROI were summed per trial per participant. These fixation durations were then averaged per participant for each condition of the experiment. Following the calculation of the mean total fixation durations, these values were normalised to the area of the ROI in question.

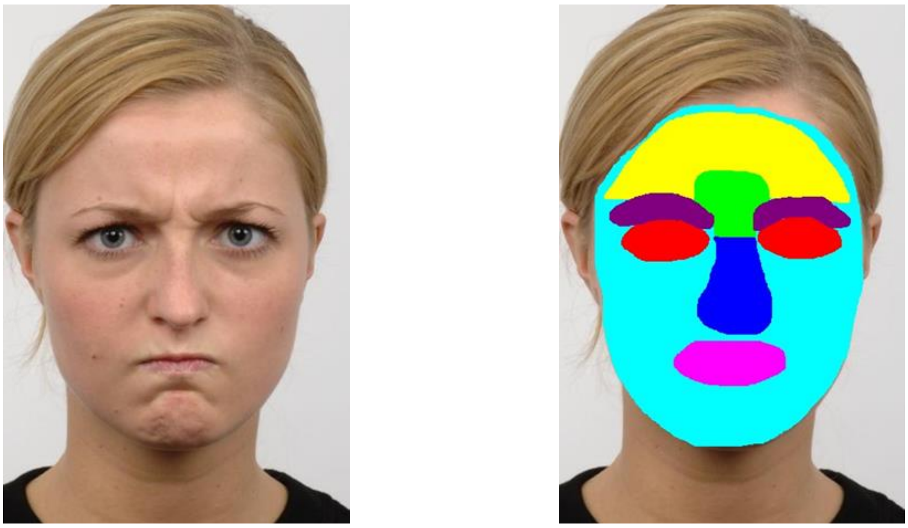

Figure 8. The regions of interest (ROI) drawn by the bespoke programme. The red region denotes the eye, the green region denotes the brow, blue region denotes the nose and the purple region denotes the mouth. The sizes of these regions are dependent on the underlying expressions and the shape of the facial feature. The forehead (yellow), eyebrows (deep purple) and the face (cyan) regions are not included in the analysis of total fixation duration.

Previous research looking at fixations (durations and numbers) have either used non-normalised data (e.g. Schurgin et al., 2014; Eisenbarth et al., 2011) or normalized the number of fixations at each ROI to the total number of fixations in each trial (eg. Guo 2012; Guo & Shaw, 2015). While Guo (2012) and Guo and Shaw (2015) kept the sizes of the ROIs across all faces and expressions consistent, the ROIs used in Experiments 1 and 2 of this chapter were different across faces and expressions. For this reason, we wanted to take into account the difference in sizes between ROIs within each face and across all faces therefore we chose to normalize the fixation duration according to ROI size. The area for each ROI of each expression was calculated as the average of all the faces presented to the participants. For example, the area for the mouth region of angry faces is the mean area of the mouth regions of the 12 faces displaying the angry expression. The mean total fixation duration value was then divided by the mean ROI area for the respective condition (i.e., mean total fixation duration for the mouth area of the angry faces was divided by the mean area of the mouth region of the 12 faces). All the analyses, therefore, were conducted on the area-normalised mean total fixation duration using a 4 × 4 (emotion × ROI) repeated measures ANOVA. Where the assumption of sphericity was violated, the Greenhouse-Geisser corrected F value is reported.

Results

Emotion Recognition Accuracy

Brief-fixation paradigm

Overall, the average reaction time for this task was 1.08sec (SD=0.78). A visualization of the reaction time distribution and average reaction time for each condition can be seen in the Appendix (Table A1, Figure A1). An examination of the reaction times revealed no outliers; therefore, all trials were used for the analysis of emotion recognition accuracy. General accuracy was high at an average of 83% (SD = 0.06) across all participants. The unbiased hit rates are summarized in Figure 9. There were main effects of emotion ($F_{(1.66,\ 53.06)}$=17.02, p<.001, η_p^2=.35) and initial fixation ($F_{(3,\ 96)}$=3.43, p=.02, η_p^2=.10) and a significant interaction

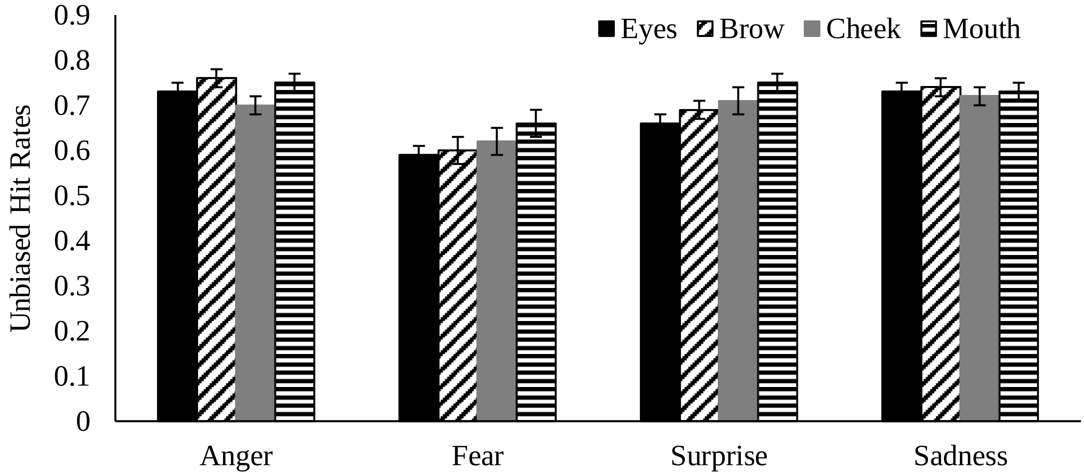

Figure 9. The mean unbiased hit rates for each expression at different initial fixation locations. The error bars represent the standard error of the mean.

between these factors ($F_{(5.49,\ 175.63)}$=3.68, p=.003, η_p^2=.10). Fear was the least well recognised facial expression compared to the others (all *p*s< .001). Initially fixating on the mouth region led to higher recognition accuracy compared to the cheeks (p = .04).

There was a main effect of initial fixation for angry expressions ($F_{(3,96)}$=3.14, p=.029, η_p^2=.09). Pairwise comparisons showed that initially fixating on the mouth led to significantly higher anger accuracy compared to initially fixating on the cheeks (p=.04). To further investigate our hypothesis regarding initially fixating on the brow region for angry faces, 3 one-tailed paired-samples t-tests were carried out comparing anger accuracy at the brow to accuracy at the other initial fixation locations. Anger recognition accuracy was significantly higher with initial fixation at the brow compared to initial fixation at the cheeks ($t_{(32)}$=2.45, p=.01). There were no further significant comparisons.

The main effect of initial fixation for fearful faces ($F_{(3,\ 96)}$=3.83, p<.05, η_p^2=.11) indicated that initially fixating on the mouth led to higher fear recognition compared to initially fixating on the eyes (p= .033). Planned comparisons showed that initially fixating on the mouth led to improved fear recognition compared to fixating the eyes ($t_{(32)}$=2.98, p=.003) and the brow ($t_{(32)}$=2.55, p=.008). There was also a trend towards a difference between the mouth and the cheeks ($t_{(32)}$=2.04, p=.025).

The significant main effect of initial fixation for surprise ($F_{(3, 96)}$=5.78, p=.001, η_p^2=.15) showed that initially fixating on the mouth led to higher emotion recognition compared to fixating the eyes (p= .001) and the brow (p= .02). Planned comparisons corroborated the results of pairwise comparisons with significantly higher surprise accuracy when initially fixating the mouth compared to eyes ($t_{(32)}$=4.28, p<.001) and the brow ($t_{(32)}$=3.19, p=.002). No difference in accuracy was found between fixation at the mouth and fixation at the cheeks ($t_{(32)}$=1.51, p=.071).

Finally, there was no effect of initial fixation location for sad expressions ($F_{(3, 96)}$=.89, p=.45, η_p^2=.03).

To further investigate the effect of initial fixation location, we also examined the confusion matrices to identify whether there were systematic misclassifications of the target expressions as another emotion. Table 1 shows the confusion matrices for each expression at all initial fixation locations. First, given that previous research suggest that fear and surprise are confusable

Table 1. Confusion matrices for anger, fear, surprise and sadness at each initial fixation location, representing the % of times each emotion category was given as an answer to the displayed expression.

Displayed Expression		**Response Given/%**			
	Initial Fixation	**Anger**	**Fear**	**Surprise**	**Sadness**
Anger	Eyes	**84.22**	2.90	1.64	11.24
	Brow	**86.74**	2.15	1.77	**9.34**
	Cheek	**82.32**	2.53	1.01	14.14
	Mouth	**85.98**	1.52	0.25	12.25
Fear	Eyes	3.66	**70.45**	22.98	2.90
	Brow	3.41	**69.07**	23.23	4.29
	Cheek	4.17	**71.46**	20.58	3.79
	Mouth	2.78	**74.87**	16.79	5.56
Surprise	Eyes	1.14	10.48	**86.24**	2.15
	Brow	1.26	7.95	**88.89**	1.89
	Cheek	1.64	9.22	**87.88**	1.26
	Mouth	1.39	8.59	**89.14**	0.88
Sadness	Eyes	9.22	1.64	2.90	**86.24**
	Brow	9.34	1.39	2.53	**86.74**
	Cheek	10.48	1.64	1.64	**86.24**
	Mouth	9.47	1.39	1.89	**87.25**

expressions, we conducted a one-way ANOVA to investigate whether there was a difference between the percentages of each emotion category used as a response for fearful and surprised expressions. There is a main effect of response for fearful expressions ($F_{(1.29, 41.16)}$=57.05, p<.001, η_p^2= .64) which confirmed our expectation that there were more surprise responses to fearful expressions compared to anger (p< .001) and sad (p< .001) responses. There is also a main effect of response for surprised expressions ($F_{(1.26, 40.45)}$=21.83, p<.001, η_p^2= .41) which reveals a symmetrical pattern of misclassifications: Fear responses to surprise expressions were higher compared to anger (p< .001) and sad (p< .001) responses. Since these results suggest that fear is misclassified as surprise and surprise as fear, we go on to investigate whether the percentage of surprise responses to fear and percentage of fear responses to surprise are affected by initial fixation location, using a one-way repeated measures ANOVA comparing across initial fixation locations. The main effect of initial fixation on surprised responses given to fearful expressions was significant ($F_{(3, 96)}$=3.96, p=.01, η_p^2= .11) confirming our expectation that initially fixating the mouth reduced the percentage of surprise responses given to fearful expressions compared to the eyes (p=.037) and the brow (p=.015) but not the cheeks (p=.253). On the other hand, there was no effect of initial fixation locations for fear responses given to surprise expressions ($F_{(2.45, 78.26)}$=1.19, p=.32, η_p^2= .04).

Long-presentation paradigm

Overall, the average reaction times for this task was 1.82s (SD=0.92). A visualization of the reaction time distribution and the average reaction times for each condition can be seen in the Appendix (Table A2 & Figure A2). An examination of the reaction time data revealed no responses under 200ms, therefore no trials needed to be removed from the data. General

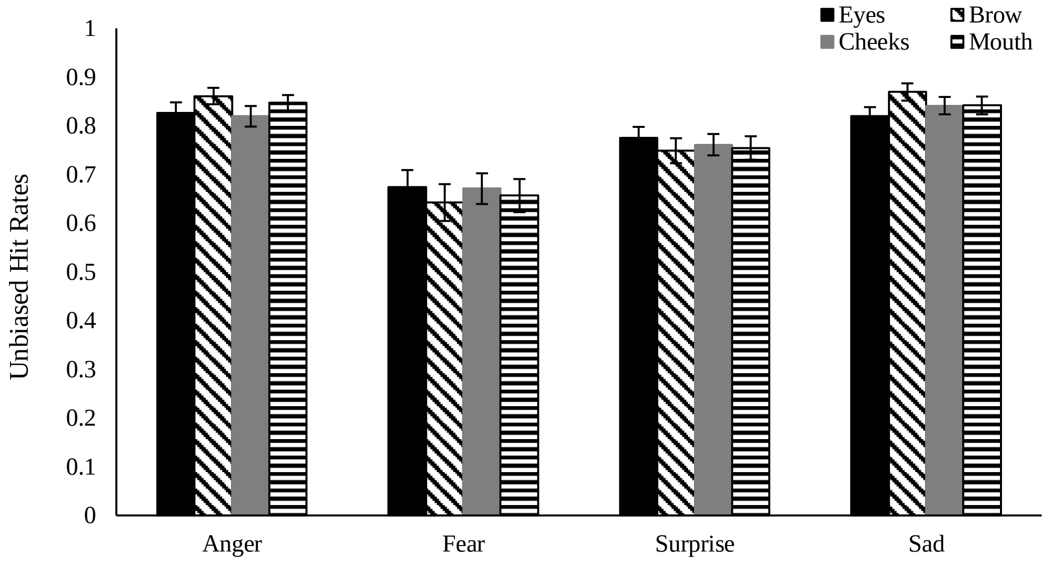

Figure 10. The mean unbiased hit rates for each expression at different initial fixation locations for the long-presentation paradigm. Error bars indicate standard error of the mean.

accuracy was high at an average of 87% (SD = 0.06) across all participants. Descriptive statistics for the unbiased hit rate data are shown in Figure 10. There was a main effect of emotion ($F_{(2.15, 75.24)}$=35.32, p< .001, $\eta_p{}^2$=.50) indicating that anger was better recognised compared to fear (p< .001) and surprise (p= .002). Sadness was also better recognised compared to fear (p< .001) and surprise (p< .001). Surprised expressions were also better recognised compared to fearful expressions (p<.001). No effect of initial fixation was found ($F_{(3, 105)}$=.31, p=.82, $\eta_p{}^2$=.009), nor an interaction between emotion and initial fixation location ($F_{(5.64, 197.33)}$=1.98, p=.07, $\eta_p{}^2$=.05) .

Confusion Matrices

Similar to the brief-fixation paradigm, we investigated the occurrence of any systematic misclassifications between emotions. Table 2 shows the confusion matrices. One-way repeated measure ANOVAs were used to compare the percentage of surprise responses to fear and percentage of fear responses to surprise expressions. There was a main effect of response for fearful expressions ($F_{(1.22, 42.68)}$=40.14, p<.001, $\eta_p{}^2$= .53). As expected, the surprise responses to fearful expressions was higher compared to angry (p<.001) and sad responses (p<.001). Similarly for surprised expressions, the main effect of response option ($F_{(1.07, 37.32)}$=23.29,

Table 2. Confusion matrices for anger, fear, surprise and sadness at each initial fixation location, representing the % of times each emotion category was given as an answer to the displayed expression.

Displayed Expression		**Response Given/%**			
	Initial Fixation	**Anger**	**Fear**	**Surprise**	**Sadness**
Anger	Eyes	**89.41**	2.03	0.68	7.66
	Brow	**91.22**	3.60	0.45	4.50
	Cheek	**88.96**	3.15	0.45	7.43
	Mouth	**92.34**	2.03	0.45	5.18
Fear	Eyes	2.03	**74.55**	18.69	4.05
	Brow	1.58	**72.07**	22.07	2.70
	Cheek	2.48	**75.00**	18.47	3.38
	Mouth	2.93	**73.87**	19.82	2.94
Surprise	Eyes	0.23	6.53	**92.34**	0.68
	Brow	0.45	7.66	**90.54**	1.13
	Cheek	1.35	7.21	**90.09**	0.68
	Mouth	0.90	7.21	**91.22**	0.45
Sadness	Eyes	5.41	1.58	0.68	**91.44**
	Brow	3.83	1.58	0.90	**93.24**
	Cheek	4.05	0.90	1.13	**93.92**
	Mouth	5.41	1.80	1.35	**91.22**

p<.001, η_p^2= .40) showed that the fearful responses to surprised expressions were higher compared to angry (p< .001) and sad responses (p< .001). Then, the effect of initial fixation location on the percentage of surprise responses to fear and fear responses to surprise were compared using a one-way repeated measures ANOVA using initial fixation location as a factor. There was no main effect of initial fixation location on surprised responses for fearful expressions ($F_{(3, 105)}$=.81, p=.49, η_p^2= .023) or for fear responses for surprised expressions ($F_{(3, 105)}$=.17, p=.91, η_p^2= .005). The numerical magnitude of these misclassifications is smaller compared to those for the brief-fixation paradigm.

Eye Movement Analysis

Brief Fixation Paradigm

Reflexive saccade data from three participants were removed from the analysis of eye-movement data since they did not meet the criteria for inclusion. The following analyses were conducted on the data from 30 participants. On average, for the 30 participants, 70% (range: 32%-98%) of all trials included a reflexive saccade.

Saccades Going Upwards from Mouth vs Downwards from Eyes

Another aim of the current study was to investigate whether the reflexive saccades were seeking out expression-informative facial features. The percentages of saccades going upwards from initial fixation on the mouth and downwards from initial fixation on the eyes were calculated relative to the total number of saccades from initial fixation on the eyes and mouth combined. A 4 × 2 repeated measures ANOVA was conducted using emotion and saccade direction as factors to compare the percentage of saccades going upwards from the mouth and downwards from the eyes for each emotion. Descriptive statistics can be seen in Figure 11. There was a main effect

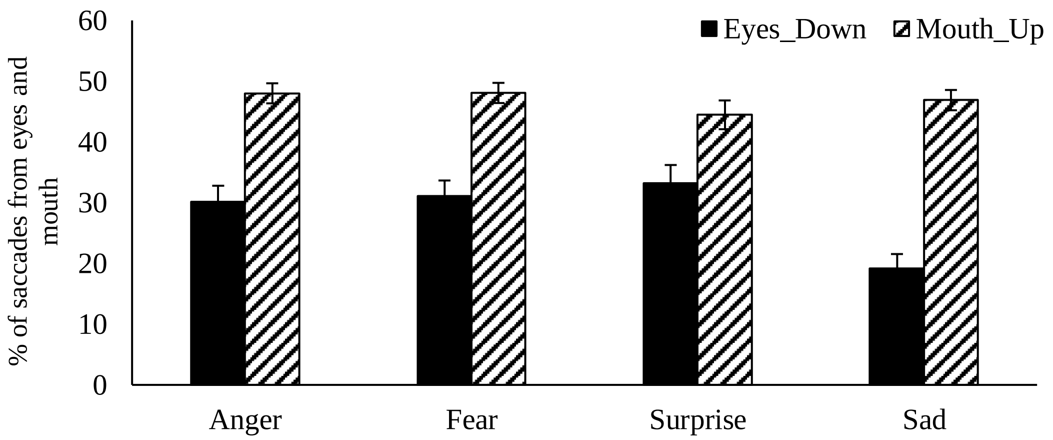

Figure 11: Percentage of reflexive saccades downwards from initial fixation locations on the eyes and upwards from initial fixation on the mouth. Error bars indicate standard error of the mean.

of emotion ($F_{(1.47, 42.69)}$=6.24, p= 0.008, $\eta_p^{\ 2}$=.18). There was also a main effect of saccade direction ($F_{(1, 29)}$=39.28, p< 0.001, $\eta_p^{\ 2}$=.58). As expected, the percentage of saccades going upwards from the mouth was significantly higher than the percentage of saccades going downwards from the eyes. The Expression × Saccade Direction interaction was also significant ($F_{(2.16, 62.52)}$=5.78, p= .004). To further investigate this interaction, two separate one-way ANOVAs were carried out to investigate the effect of expression on the percentage of saccades going upwards from the mouth and downwards from the eyes separately.

There was a main effect of expression for saccades going downwards from the eyes ($F_{(1.63, 47.30)}$=6.60, p= 0.005, $\eta_p^{\ 2}$= .19). The percentage of saccades going downwards from the eyes was lower for sad faces compared to surprised faces (p< .05). The main effect of expression did not reach significance for percentage of saccades going upwards from the mouth ($F_{(2.05, 51.59)}$=2.57, p= .08).

Four Wilcoxon signed-rank tests were carried separately comparing percentage of saccades going downwards from the eyes and saccades going upwards from the mouth for each emotion. For all expressions, the Wilcoxon signed-rank tests showed that there was a significantly higher percentage of saccades going upwards from the mouth compared to saccades going downwards from the eyes (anger: Z= -3.63, p<.001; fear: Z= -3.82, p<.001; surprise: Z= -2.53, p=.011; sad: Z= -4.68, p<.001).

Saccade Direction from Upper vs Lower Features

Since we have included additional initial fixation locations compared to the original paradigm by Gamer and Buchel (2009), we also calculated the percentage of saccades going downwards from upper initial fixation locations (i.e., the eyes and the brow) and upwards from lower initial

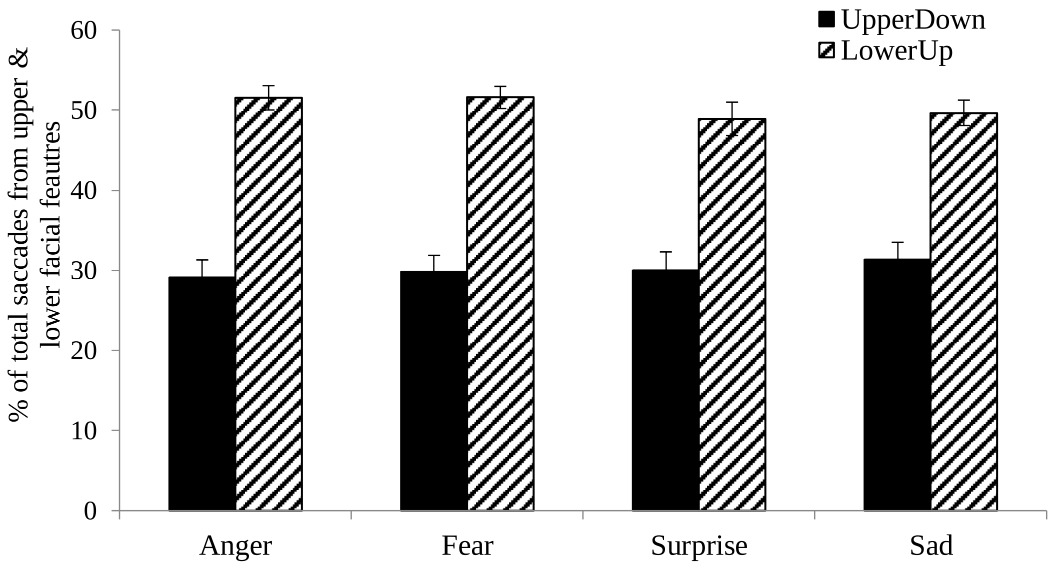

Figure 12: Percentage of reflexive saccades downwards from fixation on upper-face features or upwards from lower-face features. Error bars indicate standard error of the mean.

fixation locations (i.e., the mouth and the cheeks) as a percentage of the total number of initial saccades for each emotion per participant. A 4 × 2 repeated measures ANOVA was conducted using emotion and fixation location as factors to compare these values. Descriptive statistics can be seen in Figure 12. There was no main effect of emotion ($F_{(3, 87)}$=1.11, p=.35). There was a main effect of fixation location ($F_{(1, 29)}$=31.15, p< .001, $\eta_p{}^2$=.52). As expected, the percentage of saccades going upwards from lower features was significantly higher than the percentage of saccades going downwards from upper features. There was also a significant interaction ($F_{(3, 87)}$=2.81, p= .04, $\eta_p{}^2$= .11).

There was no main effect of expression on the percentage of saccades going downwards from upper features ($F_{(3, 87)}$=1.29, p= 0.28) but there was a significant effect of expression for saccades going upwards from lower facial features ($F_{(2.33, 67.43)}$=3.72, p= 0.02, η_p^2=.52). Despite the effect of expression; however, there were no significant pairwise comparisons. Only the comparison between anger and surprise approached significance (p=.09) suggesting a lower percentage of saccades going upwards from lower facial features for surprised compared to angry faces.

Wilcoxon signed-rank tests for angry (Z= -4.27, p<.001), fearful (Z= -4.55, p<.001) and sad (Z= -4.10, p<.001) expressions revealed significantly higher percentage of saccades going upwards from the lower features compared to saccades going downwards from the upper features. A paired samples t-test for surprised expressions showed that there was a higher percentage of saccades going upwards from the lower features compared to saccades going downwards from the upper features ($t_{(29)}$=4.05, p<.001).

Saccade Path Analysis

Collapsed across initial fixation location

Even though the analysis of saccade direction upwards and downwards from a facial feature is informative about what reflexive saccades might seek following initial fixation, there is need to analyse the direction of these saccades more precisely considering that the centre-of-gravity as well as the location of an informative facial feature might affect saccade target. We conducted an analysis of saccade direction using the saccade path measure described in Data Analysis to compare the paths of reflexive saccades to possible saccade paths targeting one of the other initial fixation locations. The mean saccade paths of the first saccades were calculated for each emotion and collapsed across all the initial fixation locations. A 4 × 6 repeated measures ANOVA was used to compare the mean saccade paths towards each of the six possible saccade targets for each expression. The descriptive statistics can be seen in Figure 13. There was no main effect of expression ($F_{(3, 87)}$= 1.58, p= .20, η_p^2= .05). The main effect of saccade target was significant ($F_{(1.82, 52.82)}$= 8.71, p= .001, η_p^2= .23). The pairwise comparisons indicated that the reflexive saccades towards the brow were stronger compared to all other locations (all ps< .001,

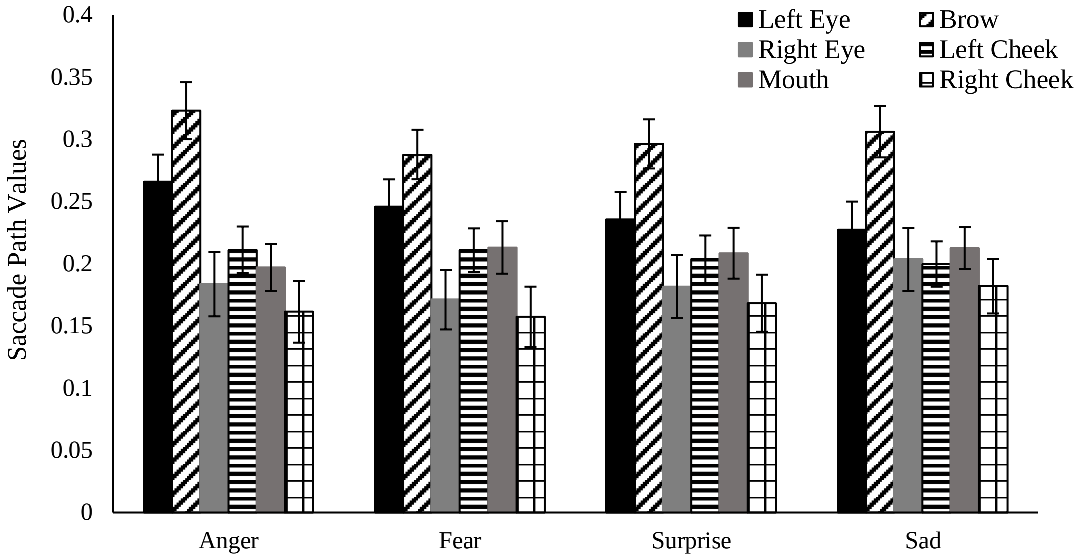

Figure 13. The mean saccade paths of reflexive saccades relative to the 6 target locations of interest, collapsed across initial fixation location. Error bars indicate standard error of the mean.

except brow vs. mouth where *p*= .004 and brow vs left eye where *p*= .008) and stronger towards the left eye compared to the left cheek (*p*= .013). There was a significant interaction between expression and saccade target ($F_{(7.39,\ 214.44)}$= 4.76, *p*< .001, η_p^2= .14), as would be expected given the hypothesis that reflexive first saccades target expression-informative facial features.

To follow up this significant interaction, 4 one-way repeated measures ANOVAs were conducted for each expression separately comparing saccade targets. There was a main effect of saccade target for all the expressions (anger: $F_{(2.00,\ 58.11)}$= 11.02, *p*< .001, η_p^2= .28; fear: $F_{(1.93,\ 56.10)}$= 7.93, *p*= .001, η_p^2= .22; surprise: $F_{(1.86,\ 53.95)}$= 7.35, *p*= .002, η_p^2= .20; sad: $F_{(1.87,\ 54.10)}$= 6.87, *p*= .003, η_p^2= .19). For all expressions except for fear, reflexive saccades were more strongly directed towards the brow compared to all other saccade targets, in keeping with the main effect of saccade target. For anger, reflexive saccades were also directed more strongly towards the left eye compared to the left cheek (*p*= .001). Even though the reflexive saccades were more strongly directed towards the brow for anger, since this was the case for all expressions, we cannot conclusively suggest that reflexive first saccades target the brow which is the expression-informative feature for anger. For fearful expressions, the reflexive saccades were directed more strongly towards the brow compared to the right eye (*p*< .001), left cheek (*p*= .001) and right cheek (*p*< .001). We expected that the mouth region would be the informative facial feature for fear recognition and found that initially fixating at the central mouth led to higher fear recognition accuracy compared to initially fixating at the eyes. However, inconsistent with the hypothesis that reflexive saccades will seek out emotion-informative features, we did not find that reflexive saccades targeted the mouth of fearful expressions.

Separately from each initial fixation location

For reflexive saccades initiating from the left eye, the interaction between expression and saccade target failed to reach significance ($F_{(2.88, 74.83)}$= 1.21, p= .31, η_p^2= .05) which goes against our prediction that reflexive saccades will target emotion-informative facial features. There was no main effect of expression ($F_{(1.80, 46.91)}$= 1.64, p= .21, η_p^2= .06). The main effect of saccade target ($F_{(1.07, 27.84)}$= 54.47, p< .001, η_p^2= .68) was significant. Pairwise comparisons indicated that the reflexive saccades from the left eye were directed towards the brow more strongly compared to all other saccade targets (all ps< .001). The saccades were also directed towards the right eye more strongly compared to the left cheek, mouth and right cheek (all ps< .001). Additionally, the reflexive saccades from the left eye were more strongly directed to the right cheek compared to the left cheek and the mouth (both ps< .001).

For reflexive saccades initiating from the brow, in line with our expectations regarding reflexive saccades targeting emotion-informative facial features, there was a significant interaction ($F_{(3.30, 85.88)}$= 5.25, p= .002, η_p^2= .17) between expression and saccade target. To investigate this interaction, four separate repeated measure one-way ANOVAs were conducted to compare saccade path values towards each saccade target from the brow for each expression separately. For angry expressions, there was a main effect of saccade target ($F_{(1.10, 28.56)}$= 6.96, p= .01, η_p^2= .21). Pairwise comparisons indicated that reflexive saccades from the brow region of angry expressions were more strongly directed towards the mouth region compared to the right cheek (p= .043). It is difficult to reconcile this result with our hypothesis since we expected the brow region to be emotion-informative facial feature for anger. There was a main effect of saccade target for fearful expressions as well ($F_{(1.13, 29.43)}$= 6.68, p= .01, η_p^2= .20). Reflexive saccades from the brow region of fearful faces were more strongly directed towards the left cheek (p= .039), mouth (p=.033) and right cheek (p=.030) compared to the right eye. We can tentatively suggest that this result is in line with our prediction since we expected the mouth region to be the emotion-informative region for fearful expressions. There was no main effect of saccade target for surprised expressions ($F_{(1.05, 27.20)}$= 2.39, p= .13, η_p^2= .08) or for sad expressions ($F_{(1.04, 27.04)}$= .83, p= .37, η_p^2= .03). Regarding the main effects, the effect of expression ($F_{(3, 78)}$= 1.80, p= .15, η_p^2= .07) was not significant. While the main effect of saccade target was almost significant following Greenhouse-Geisser correction ($F_{(1.05, 27.16)}$= 4.09, p= .052, η_p^2= .14), pairwise comparisons did not reveal any significant differences across saccade targets.

For reflexive saccades initiating from the right eye, the interaction between expression and saccade target was significant ($F_{(3.48, 90.56)}$= 3.75, p= .01, η_p^2= .13). To investigate this interaction, four separate repeated measure one-way ANOVAs were conducted to compare saccade path values towards each saccade target for each expression separately. For all four expressions, the main effect of saccade target was significant (anger: $F_{(1.53, 39.67)}$= 99.12, p< .001, η_p^2= .79 fear: $F_{(1.12, 29.09)}$= 91.49, p< .001, η_p^2= .78 surprise: $F_{(1.25, 32.50)}$= 84.68, p< .001, η_p^2= .77 sadness: $F_{(1.07, 27.90)}$= 48.58, p< .001, η_p^2= .65). However, despite this interaction and contrary to our expectations, pairwise comparisons for all the expressions reveal the same pattern of differences as the pairwise comparisons for the main effect of saccade target. There was no main effect of expression ($F_{(3, 78)}$= 0.95, p= .42, η_p^2= .04). There was a main effect of saccade target ($F_{(1.15, 29.79)}$= 108.78, p< .001, η_p^2= .81). Pairwise comparisons indicated that reflexive saccades from the right eye were more strongly directed towards the brow compared to all other saccade targets (all ps< .001) and towards the left eye compared to both cheeks and the mouth (all ps< .001). The saccades also target the left cheek more strongly compared to the mouth and the right cheek (ps< .001).

For reflexive saccades initiating from the left cheek, the interaction between expression and saccade target failed to reach significance following Greenhouse-Geisser correction ($F_{(2.71, 70.56)}$= 2.29, p= .09, η_p^2= .08). There was no main effect of expression ($F_{(3, 78)}$= 1.17, p= .33, η_p^2= .04) nor a main effect of saccade target ($F_{(1.04, 27.02)}$= 1.04, p= .32, η_p^2= .04).

For reflexive saccades initiating from the mouth, similar to reflexive saccades from the left cheek, there was no main effect of expression ($F_{(2.12, 55.09)}$=1.55, p= .22, η_p^2= .06) or saccade target ($F_{(1.08, 28.10)}$=0.99, p= .34, η_p^2= .04) nor an interaction ($F_{(3.81, 99.07)}$=1.19, p= .32, η_p^2= .04).

Finally, for reflexive saccades starting from the right cheek, the interaction between expression and saccade was almost significant following Greenhouse-Geisser correction ($F_{(2.90, 75.40)}$= 2.69, p= .054, η_p^2= .09). To investigate this interaction, four separate repeated measures one-way ANOVAs were conducted to compare saccade path values towards each saccade target for each expression separately. For angry expressions, there was a main effect of saccade target ($F_{(1.03, 26.84)}$= 6.99, p= .013, η_p^2= .21). The reflexive saccades from the right cheek of the angry expressions were directed more strongly towards the mouth region compared to the left cheek only (p= .037). This result is not in line with our expectations since we expected that the reflexive saccades from a less emotion-informative feature would be directed towards the more emotion-informative feature for anger, the brow. For fearful expressions, the main effect of saccade target ($F_{(1.06, 27.56)}$= 13.99, p= .001, η_p^2= .35) indicated that reflexive saccades were directed more strongly towards the mouth compared to all other saccade targets, in line with the suggestion that the mouth is the most emotion-informative region for fearful expressions (left eye: p=.003; brow: p= .005; right eye: p= .010; left cheek: p= .003) and towards the left cheek compared to left eye (p= .006), brow (p= .010) and right eye (p= .035). Finally, the saccades were more strongly directed towards the left eye compared to the brow (p= .033). There was also a main effect of saccade target for surprised expressions ($F_{(1.14, 29.68)}$= 8.95, p= .004,

$\eta_p{}^2$= .26). Saccades were more strongly directed towards the mouth compared to the left eye (*p*= .015), brow (*p*= .023) and left cheek (*p*= .015) and towards the left cheek compared to the left eye (*p*= .032). Finally, for sad expressions, there was a significant main effect of saccade target ($F_{(1.18, 30.58)}$= 5.55, *p*= .021, $\eta_p{}^2$= .18); however, pairwise comparisons revealed no significant comparisons between the saccade targets from the right cheek. There was no main effect of expression ($F_{(3,78)}$=0.49, *p*= .69, $\eta_p{}^2$= .02). There was a main effect of saccade target ($F_{(1.06, 27.43)}$= 10.10, *p*= .003, $\eta_p{}^2$= .28). Pairwise comparisons indicated that reflexive saccades from the right cheek were directed more strongly towards the mouth compared to all other saccade targets (left eye: *p*=.015; brow: *p*= .019; right eye: *p*= .044; left cheek: *p*= .011). These saccades were also directed more strongly towards the left cheek compared to the left eye (*p*= .031) and the brow (*p*= .044).

A few general patterns emerge from the analysis of reflexive saccades from each of the six initial fixation locations: Firstly, it appears that saccades from upper facial features such as the eyes and the brow target another upper facial feature more strongly compared to any of the lower facial features. Secondly, saccades starting from the left or right eye tend to target the cheek on the opposite side of the face. Finally, in contrast to our hypothesis that reflexive saccades will target emotion-informative features, targets of the reflexive saccades from none of the initial fixation locations were modulated by expression.

Long Presentation

Mean Total Fixation Duration

The mean total fixation duration was analysed using a 4 (Emotion) × 4 (ROI) repeated measures ANOVA. Descriptive statistics can be seen in Figure 14. The main effect of emotion was significant ($F_{(2.07, 72.33)}$=109.96, *p*<.001, $\eta_p{}^2$=.76). This indicated that the mean total fixation duration for the angry faces was longer than for all other expressions (*p*s<.01) and longer for sad faces compared to fearful and surprised faces (both ps<.01). The significant main effect of region of interest ($F_{(1.46, 51.10)}$=36.15, *p*< 0.001, $\eta_p{}^2$=.51) indicated that the eyes and the nose received longer mean total fixation duration compared to the brow and the mouth. The Emotion × ROI interaction was also significant ($F_{(4.25, 148.60)}$=6.12, *p*< 0.001, $\eta_p{}^2$=.15). To investigate this interaction further, separate one-way ANOVAs were carried out to investigate the effect of expression on each ROI separately.

For the eye region, the significant main effect of expression ($F_{(1.65, 57.80)}$=13.86, *p*< 0.001, $\eta_p{}^2$=.28) indicated that the mean total fixation duration was higher for angry and sad expressions compared to fearful (anger vs fear: p=.001; sad vs fear: p<.001) and surprised (anger vs surprise: p=.007; sad vs surprise: p=.005) expressions.

For the brow region, there was a main effect of expression ($F_{(2.15, 75.08)}$=24.68, p< 0.001, η_p^2=.41) indicating that the mean total fixation duration was higher for angry expressions compared to all others (all ps<.05). This is consistent with our initial hypothesis that observers will spend longer looking at the emotion-informative brow region of angry expressions. There were no more significant comparisons.

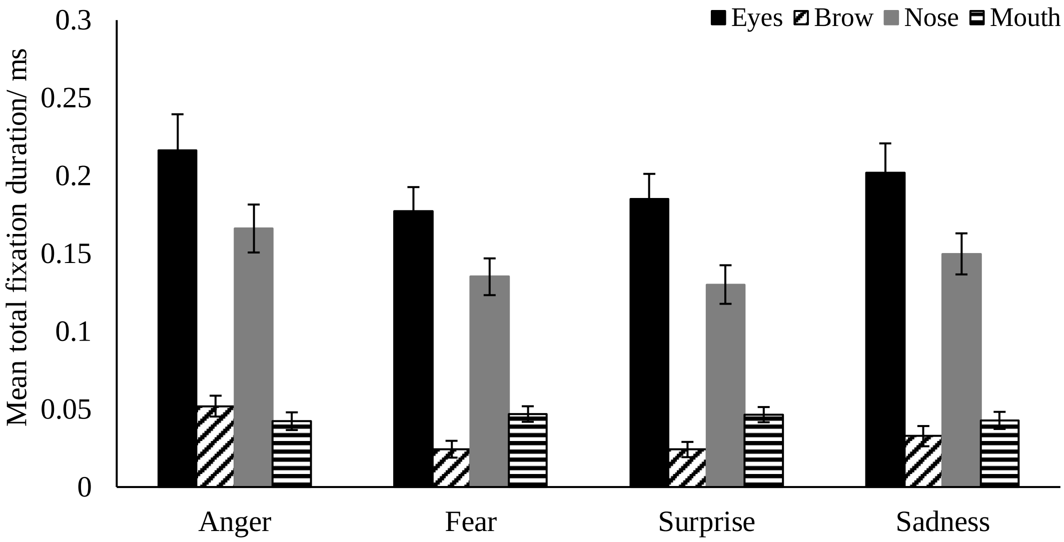

Figure 14. The average total fixation duration within each of the four regions of interest (eyes, brow, nose and mouth) for each expression. The values are area normalized. Error bars indicate standard error of the mean.

For the nose region, the significant main effect of expression ($F_{(3, 105)}$=22.59, p< 0.001, η_p^2=.39) showed that mean total fixation duration for the angry faces was higher compared to all other expressions (p< .001 for both fearful and surprised; p=.008 for sadness) and the mean total fixation duration for the sad faces was higher compared to fear (p=.017) and surprised expressions (p<.001).

Finally, there was no significant effect of expression for the mouth region ($F_{(2.55, 89.32)}$=0.85, p= .47, η_p^2=.03). Since we suggested that the mouth region would be informative for fearful and surprised expressions, the lack of an effect of emotion on the mean total fixation duration for the mouth region is inconsistent with our initial expectations.

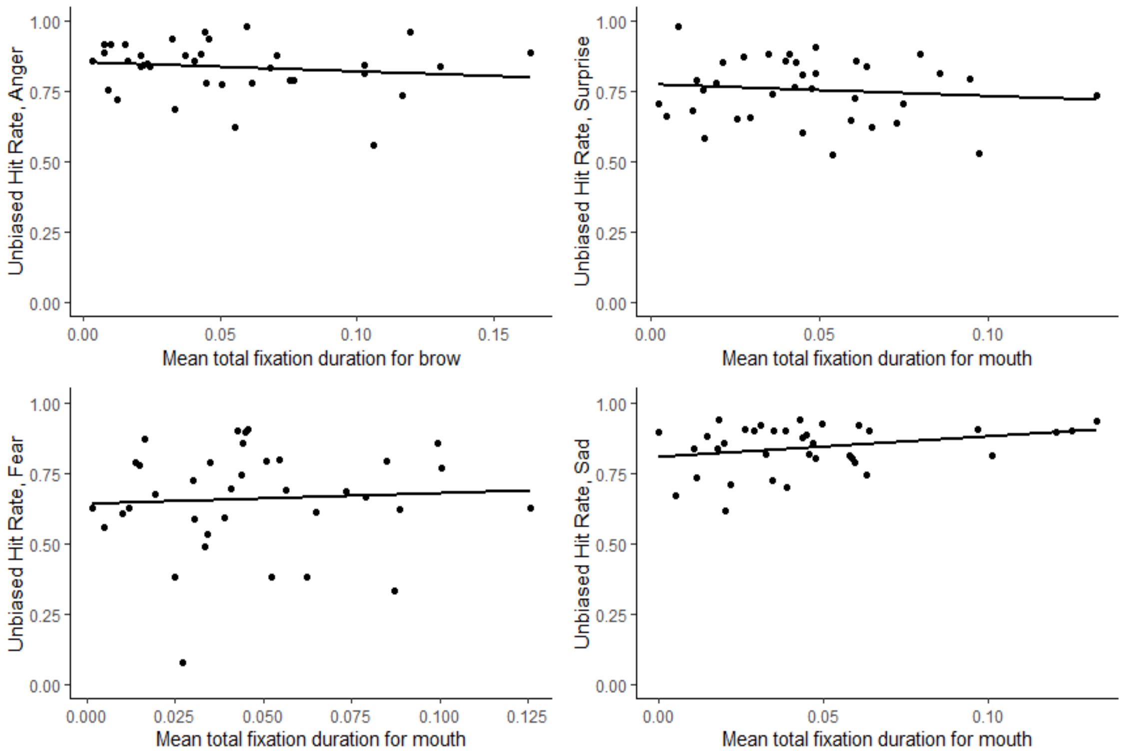

Figure 15. The correlation between unbiased hit rates for anger, fear, surprise and sadness and mean normalized total fixation duration in the long-presentation paradigm.

Total Fixation Duration and Emotion Recognition Accuracy

To address the question of whether the amount of time spent fixating an ROI contributes to emotion recognition performance, the relationship between total fixation duration in ROIs that are deemed informative for the target emotion (brow for anger, mouth for fear, surprise and sadness) and unbiased hit rates for that emotion was investigated by Pearson's correlations or, where the data is not normally distributed, the non-parametric equivalent, Spearman's rank order correlations. No significant correlations were found between anger accuracy and fixation duration within the brow region (rs = -.19, p = .26), or fear accuracy and mouth fixations (rs= .13, p = .46), or surprised accuracy and fixation duration on the mouth (r = -.11, p = .53) or for sadness and fixation duration on the mouth (rs = .21, p = .21). All the correlations can be seen in Figure 15.

Summary

Using a combination of briefly presented angry, fearful, surprised and sad expressions, we aimed to (1) investigate the contribution to emotion recognition of initially fixating an informative facial feature and (2) whether reflexive first saccades seek out informative facial features when these are not initially fixated. Additionally, by presenting the same expressions for longer in a separate experimental session, (3) we investigated whether the eye movements of observers were modulated by emotion-informative facial features and (4) whether the fixation duration within the emotion-informative features were correlated with emotion recognition accuracy.

With respect to the first aim, we found an interaction between expression and initial fixation location on recognition accuracy. Initially fixating the brow region of angry expressions led to numerically the highest recognition accuracy and improved emotion recognition compared to initially fixating a cheek. However, there were no other differences between the brow and other initial fixation locations for anger recognition. This partially supports our hypothesis that the brow region is informative for anger. For fearful and surprised facial expressions, initially fixating on the mouth region both led to better recognition compared to the eyes and the brow and a reduction in the misclassification of fearful expressions as surprised, as expected. No effect of initial fixation location was found for sad faces.

In both the brief fixation and the long-presentation paradigms, angry expressions were misclassified as sad expressions and fearful expressions were misclassified as surprised expressions. The misclassification between anger and sadness was reduced when initially fixated at the brow region for both the brief fixation and the long presentation paradigms. Additionally, only for the long presentation paradigm, initial fixation on the mouth also led to reduced misclassification of anger as sad. Initially fixating on the mouth region of the fearful faces reduced its misclassification as surprised only when the expressions were briefly fixated. The misclassification persisted regardless of initial fixation location in the long-presentation paradigm.

Regarding the second aim, in the brief-fixation paradigm, we found that observers preferred to shift their gaze upwards from lower facial features compared to downwards from upper face features. This was in line with findings in the literature (Gamer & Buchel, 2009; Kliemann et al., 2010; Kliemann et al., 2012; Scheller et al., 2012). As expected, this effect was modulated by expression. There was a reduced tendency to shift gaze downwards from the eyes for briefly presented sad faces compared to surprised faces. Using our more refined saccade path measure we found that reflexive saccades, collapsed across initial fixation location, more strongly targeted the brow region. Furthermore, reflexive saccades initiating from the left and right eyes were directed towards the brow and the opposite eye, which might explain why there was a lower preference to shift gaze downwards from the eyes compared to upwards from the mouth. It is possible that fixations initiated at the upper facial features target other upper facial features. Reflexive saccades initiated from the right cheek targeted the mouth region compared to all other possible targets and targeted the left cheek compared to the left eye and the brow. Combined, these results might suggest that saccades tend to target the closest facial feature to the previously fixated location. Further exploratory analyses on reflexive saccades were carried out and are reported in the Appendix. We found that the first saccades made by observers tended to end in the nose region regardless of the expression, which might be an indication that these early, reflexive saccades were influenced by the centre-of-gravity effect (Bindemann et al., 2009). Saccade latencies were found to be shortest from the cheeks, which we expected, given that these locations are less informative compared to the other initial fixation locations investigated.

With respect to the third aim, in the long-presentation experiment, mean total fixation duration results showed that observers preferred to look at the brow region of angry faces longer compared to the brow region of the other expressions. This is in line with the brow being more informative for anger recognition. While initially fixating on the brow did not improve anger recognition compared to all other initial fixation locations (only the cheeks), it seems that when presented for longer, observers sample information from the brow region of angry faces more so than they do for other expressions. Interestingly, there was no difference in the fixation patterns of fear and surprise which might explain why the misclassifications between these two expressions were not reduced in the long-presentation paradigm. It is possible that the observers were fixating on the facial features that are shared by both expressions at the expense of features that would discriminate them. Finally, regarding the fourth aim, we found no evidence to suggest a relationship between fixation duration within an informative facial feature and emotion recognition accuracy.

Experiment 2

In Experiment 2, we aimed to address the same research questions as in Experiment 1 using a different combination of facial expressions. We replicated Experiment 1 using angry, fearful, surprised and disgusted (replacing sad) facial expressions. The reasons for this were two-fold: Firstly, we wanted to replicate our results relating to anger, fear and surprise in the context of a different combination of expressions. The second motivation relates to the new expression of disgust. Smith et al. (2005) and Smith and Merlusca (2014) suggest that the mouth region and the wrinkled nose region is informative for the recognition of disgusted facial expressions. Furthermore, previous research also suggests that observers spend more time looking at the mouth regions of disgusted facial expressions (more specifically, the upper lip in Schurgin et al., 2014). Previous research also suggest high confusion rates between disgust and anger, especially disgusted expressions being misclassified as angry (Du & Martinez, 2011; Gagnon et al., 2010; Jack et al., 2009, 2014; Widen et al., 2004). Jack et al. (2014) found that the resolution of this misclassification occurs when the upper lip raiser AU is activated in dynamic unfolding of the expression. Therefore, we hypothesized higher emotion recognition accuracy for disgusted expressions and reduced misclassifications when initial fixation location falls at the mouth. We also expected the total fixation duration to be longer in the mouth region for disgusted expressions in the long presentation paradigm. Regarding reflexive saccades, we expected to find saccades more strongly directed to the mouth region for this expression.

Methods

Participants

A total of 40 participants took part in the brief-fixation paradigm (female= 31, male=9; mean age= 21.92 years) and of these 40, 39 participants also completed the long-presentation paradigm (female=30, males=9, mean age=22 years). Sample size was selected to be similar to the sample size of studies with similar design. All participants had normal or corrected-to-normal vision. All participants gave written consent to take part in the experiment and were rewarded participant pool credit for their participation. The study was approved by the Durham University Psychology Department Ethics Sub-committee.

Stimuli, Design and Procedure

The design and procedure of Experiment 2 was identical to Experiment 1 except the replacement of sad faces with disgusted faces from the same face database. With this change, the set of expressions used in this experiment are the most confused expressions. As discussed in detail in the Introduction section, fear is most often confused as surprise and disgust is often confused as anger. The facial appearance of these expressions is similar: therefore, this subset of expressions provides a good opportunity to investigate to what extent the informative facial features would contribute to emotion recognition when similar visual information is transmitted for multiple expressions and which facial feature aids in resolving the confusion among these expressions. An example of expressions presented to participants in Experiment 2 can be seen in Figure 16. In addition to the planned comparisons of Experiment 1, since we expected the mouth region to be informative for disgusted expressions, we conducted three one-tailed paired samples t-tests to compare disgust accuracy when initially fixating the mouth to initially fixating

Figure 16. Examples of angry, fearful, surprised and disgusted expressions (from left to right) from the Radbound Face Database used for Experiment 2.

the eyes, brow and cheeks. Additionally, for Experiment 2, the order of the brief fixation and long presentation sessions was counterbalanced for each participant so that half of the participants completed the brief-fixation paradigm first and long-presentation paradigm later and the other half did the opposite. This factor was not further analysed since we did not expect any order effects.

Results

Emotion Recognition Accuracy

Brief-fixation paradigm

Overall, the average reaction time for this task was 1.28s (SD= 1.18). The average reaction time for each condition and a distribution of the reaction time data can be seen in the Appendix, Table A5 and Figure A5. Removal of reaction times less than 200msec led to the removal of 0.02% of all trials. General accuracy for the brief-fixation paradigm was high at an average of 78% (SD = .11) across all participants. The unbiased hit rates are summarized in Figure 17. There was a main effect of expression ($F_{(1.63,\ 63.54)}$=6.96, p=.003, η_p^2=.15). Emotion recognition accuracy for surprised expressions was significantly higher than the accuracy for fearful expressions (p<.001). There was also a main effect of initial fixation location ($F_{(3,\ 117)}$=10.34, p< .001, η_p^2=.21). Initially fixating on the mouth region led to higher emotion recognition accuracy compared to initially fixating on the eyes and the brow (both ps<.001) but there was

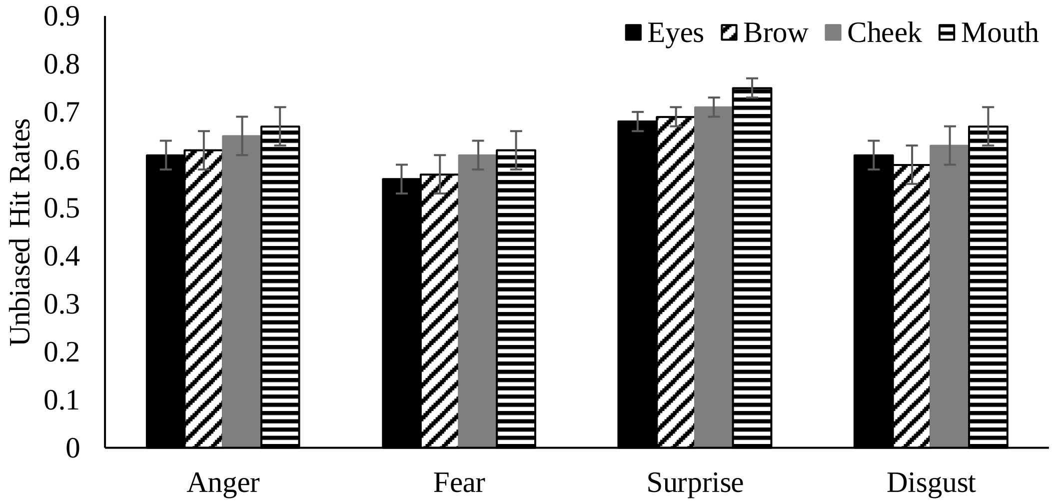

Figure 17. Mean unbiased hit rates at each initial fixation location for each expression in the brief-fixation paradigm. Error bars indicate standard error of the mean.

no difference between the mouth and the cheeks (p=.52). The interaction between emotion and initial fixation failed to reach significance ($F_{(4.83,\ 188.31)}$=0.42, p=.83, η_p^2=.01).

To further investigate our hypotheses relating to each expression separately, planned comparisons were carried out. This resulted in 3 one-tailed paired samples t-tests for each expression. Where the assumption of normality was violated, the non-parametric Wilcoxon signed ranks test was used. The p-values were accepted as significant at the Bonferroni-corrected level of $p<.017$.

Comparison of recognition accuracy for anger at the brow and eye initial fixation locations revealed no significant difference ($Z=-.25$, $p=.40$). There was a significant difference in accuracy between the brow and the cheeks ($Z=-2.49$, $p=.007$) and brow and mouth ($Z=-2.79$, $p=.003$). However, these differences were contrary to our expectations: anger accuracy was significantly *lower* with initial fixation at the brow compared to both the cheeks and the mouth.

We expected that initially fixating on the mouth would improve emotion recognition for fear and surprise. Emotion recognition accuracy for fear with initial fixation at the mouth was significantly higher compared to fixation at the eyes ($Z=-2.47$, $p=.007$) and the comparison with the brow approached significance ($Z=-1.87$, $p=0.03$). There was no difference between the mouth and cheeks ($Z=-.59$, $p=.28$). Similar results were found for surprised faces: Recognition accuracy was significantly higher with initial fixation at the mouth compared to the eyes ($Z=-3.37$, $p<.001$) and brow ($Z=-3.21$, $p< .001$) but there was no difference between the mouth and the cheeks ($Z=-1.74$, $p=.04$). These results support our hypothesis given that we suggested the mouth region would be informative for fear and surprise.

Finally, initially fixating on the mouth led to higher accuracy for disgusted expressions compared to initially fixating the eyes ($Z=-2.75$, $p=.003$) and the brow ($Z=-3.47$, $p<.001$) and there was a trend towards a significant difference between the mouth and the cheeks ($Z=-1.87$, $p=.03$). This was consistent with our hypothesis since mouth is suggested to be informative for disgust recognition.

Confusion Matrices

To investigate whether there were any systematic misclassifications between the target expressions, we computed confusion matrices as seen in Table 3. As suggested by previous literature, fear is most often misclassified as surprise and disgust as fear. Therefore, one-way repeated measures ANOVAs as response option as a factor were conducted for each expression to investigate whether there was a difference between the percentages of each emotion category used as a response for each expression. While the confusions for fear and disgust are mainly reported as asymmetrical, we wanted to explore this further, so we looked at the effect of response option for all expressions rather than only for fear and disgust. There was a main effect of response option for angry expressions ($F_{(1.30, 50.67)}=9.84$, $p=.001$, $\eta_p^2= .20$). There were more fearful ($p=.001$) and disgusted ($p<.001$) responses compared to surprised responses. While

Table 3. Confusion matrices for anger, fear, surprise and sadness at each initial fixation location, representing the % of times each emotion category was given as an answer to the displayed expression.

Displayed Expression		**Response given/%**			
	Initial Fixation	**Anger**	**Fear**	**Surprise**	**Disgust**
Anger	Eyes	**79.58**	6.77	1.15	12.50
	Brow	**80.83**	5.00	2.40	11.77
	Cheek	**81.46**	5.31	1.77	11.46
	Mouth	**81.04**	6.46	0.83	11.67
Fear	Eyes	3.02	**69.69**	22.08	5.21
	Brow	3.23	**68.33**	23.96	4.48
	Cheek	2.40	**71.98**	19.58	6.04
	Mouth	2.71	**72.19**	18.23	6.88
Surprise	Eyes	1.25	11.25	**85.00**	2.40
	Brow	0.94	8.54	**88.23**	2.29
	Cheek	0.63	9.48	**87.92**	1.88
	Mouth	0.94	7.60	**90.10**	1.35
Disgust	Eyes	23.85	0.83	1.67	**73.54**
	Brow	26.15	1.35	2.08	**70.42**
	Cheek	20.83	0.94	2.71	**75.52**
	Mouth	16.88	1.15	2.81	**79.17**

there was no significant difference between fearful and disgust responses, disgust responses (M=11.85%) were numerically almost twice as frequent as fear (M=5.89%) responses. There was a main effect of response option for disgust expressions as well ($F_{(1.05, 40.98)}=54.59$, $p<.001$, $\eta_p^2= .58$) which showed that there was a higher percentage of anger responses compared to fearful ($p<.001$) and surprised ($p<.001$) responses. There was a main effect of response option for fearful expressions ($F_{(1.30, 50.63)}=45.09$, $p<.001$, $\eta_p^2= .54$). Similar to Experiment 1, there were more surprise responses compared to angry ($p<.001$) and disgusted responses ($p<.001$). Additionally, there were more disgust responses compared to anger responses ($p=.034$) for fearful expressions. The significant effect of response for surprised expressions ($F_{(1.09, 42.37)}=36.06$, $p<.001$, $\eta_p^2= .48$) showed that there were more fearful responses compared to anger ($p<.001$) and disgust ($p<.001$). There were also more disgust responses compared to anger ($p=.006$).

Furthermore, one-way repeated measures ANOVAs using initial fixation location factor were conducted to investigate the effect of initial fixation location on the most common misclassifications. For angry expressions, there was no main effect of initial fixation location ($F_{(2.55, 99.42)}=.274$, $p=.812$, $\eta_p^2= .007$) on the percentage of disgust responses. On the other hand, the main effect of initial fixation location on the percentage of anger responses given to disgust expressions ($F_{(2.26, 87.97)}=8.76$, $p<.001$, $\eta_p^2= .183$) showed that initially fixating on the mouth region reduced anger responses compared to the eyes (*p*=.006) and the brow (*p*=.001) but not the cheeks (*p*=.056). Main effect of initial fixation on the percentage of surprised responses to fearful expressions ($F_{(2.49, 97.20)}$=2.80, *p*=.054, η_p^2= .067) show no differences among initial fixation locations. Main effect of initial fixation ($F_{(3, 117)}$=3.02, *p*=.033, η_p^2= .072) on the percentage of fear responses to surprised expressions show that there was significantly higher fearful responses to surprised expressions when fixating the eyes compared to the mouth (*p*=.033). This is somewhat contradictory to our findings for Experiment 1 since initial fixation on the mouth did not lead to reduced misclassification of fear as surprise.

Long-presentation paradigm

Overall, the average reaction time for this task was 1.85s (SD= 1.07). The average reaction time for each condition and a visualization of the reaction time data distribution can be seen in the Appendix, Table A6 and Figure A6. Removal of automatic responses from the data led to a removal of .01% of all trials. General accuracy was high with 83% (SD = 0.1). A 4 × 4 RM ANOVA was conducted using emotion and initial fixation location as factors. Illustration of the descriptive statistics for the unbiased hit rates can be seen in Figure 18. There was a main effect of emotion ($F_{(1.52, 57.89)}$ = 5.59, *p*= .01, η_p^2=.13). Surprise was better recognised compared to fear (*p*=.001) and disgust (*p*= .01). There was no effect of initial fixation location ($F_{(3,114)}$= 0.02, *p*=1.00). There was no significant interaction between emotion and initial fixation location ($F_{(5.12, 194.57)}$ = 0.84, *p*= .53).

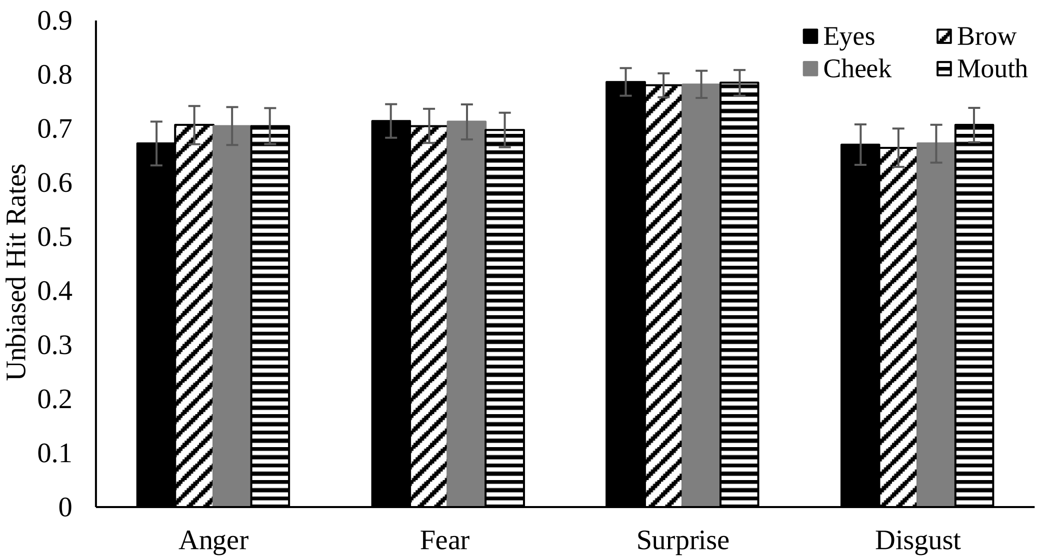

Figure 18: Unbiased hit rates at each initial fixation location for each expression in the long-presentation paradigm. Error bars indicate standard error of the mean.

Confusion Matrices

To investigate whether there were any systematic misclassifications between target expressions, we computed confusion matrices as seen in Table 4. Similar to the brief fixation paradigm, one-way repeated measures ANOVAs with response option as a factor were conducted on the percentage of emotion categories used as response options for each of the expressions. Then, one-way repeated measures ANOVAs using initial fixation location as a factor were conducted to investigate the effect of initial fixation location on the most common misclassifications for each expression. There was a main effect of response option for angry expressions ($F_{(1.51, 57.31)}$=7.61, *p*=.003, η_p^2= .17). There were more disgust responses to angry expressions compared to surprise responses (*p*=.002) but there was no difference between disgust and fear responses (*p*=.204). Symmetrical confusion for disgusted expressions was also present ($F_{(1.01, 38.40)}$=57.38, *p*<.001, η_p^2= .60) where anger responses to disgusted expressions were significantly higher compared to fear (*p*<.001) and surprised (*p*<.001) responses. There was a main effect of response option for fearful expressions ($F_{(1.37, 52.10)}$=34.70, *p*<.001, η_p^2= .48). There were more surprise responses to fearful expressions compared to anger (*p*<.001) and disgusted responses (*p*<.001). Additionally, there was a higher percentage of disgust responses to fearful expressions compared to angry responses (*p*=.001). There was also an effect of response option for surprised expressions ($F_{(1.26, 47.93)}$=13.82, *p*<.001, η_p^2= .27) and there was a higher percentage of fear (*p*<.001) and disgust (*p*<.001) responses to surprised expressions compared to anger responses.

Table 4. Confusion matrices for anger, fear, surprise and disgust at each initial fixation location, representing the % of times each emotion category was given as an answer to the displayed expression.

Displayed Expression		**Response given/%**			
	Initial Fixation	**Anger**	**Fear**	**Surprise**	**Disgust**
Anger	Eyes	**81.67**	5.63	1.67	10.63
	Brow	**85.83**	3.54	0.83	9.58
	Cheek	**85.42**	4.79	0.83	8.96
	Mouth	**83.75**	5	0.83	9.58
Fear	Eyes	1.25	**78.33**	15.83	4.38
	Brow	1.25	**77.71**	14.375	6.25
	Cheek	1.88	**77.71**	15.42	4.167
	Mouth	1.67	**77.08**	15.83	5
Surprise	Eyes	0.21	5	**91.46**	2.92
	Brow	0.42	7.29	**88.54**	2.92
	Cheek	0	5.83	**90.63**	3.33
	Mouth	0.83	6.04	**89.79**	2.5
Disgust	Eyes	20.21	0.21	0.21	**78.75**
	Brow	20.42	0.42	0.21	**78.13**
	Cheek	19.39	0.83	1.25	**77.92**
	Mouth	17.29	0	0.21	**81.88**

Initially fixating a facial features did not affect the percentage of disgust responses given to angry expressions ($F(_{2.44, 92.55})$=.27, p<.81, η_p^2= .007) or anger responses given to disgust expressions ($F_{(2.43, 92.33)}$=.92, p<.42, η_p^2= .024). Initially fixating a facial feature also did not affect the percentage of surprise responses given to fearful expressions ($F_{(3, 114)}$=.50, p=.68, η_p^2= .013) or the percentage of fear responses given to surprised expressions ($F_{(3, 114)}$=1.00, p=.40, η_p^2= .026).

Eye movement analysis

Brief-fixation paradigm

Reflexive saccade data from two participants were removed from the analysis of eye-movement data since they did not meet the criteria for inclusion. The following analyses were conducted on the data for 38 participants. On average, for the 38 participants, 82.74% of all trials included a reflexive saccade.

Saccades going upwards from the mouth vs downwards from the eyes

To investigate whether there were proportionately more saccades made upwards from the mouth than downwards from the eyes and whether this varied as a function of the expression, we calculated the number of initial saccades going downwards from the eyes and upwards from the mouth as a percentage of the total number of saccades starting from the eyes and the mouth combined. A 4 × 2 repeated measures ANOVA with emotion and initial fixation location as factors was conducted. The descriptive statistics can be seen in Figure 19. There was a main effect of emotion ($F_{(2.45, 90.64)}$=5.61, p<.001, η_p^2=.13) and a main effect of initial fixation ($F_{(1, 37)}$=5.67, p= .02, η_p^2=.13). The main effect of emotion indicated that there were fewer saccades for angry faces compared to fearful (p<.001), and disgusted faces (p<.001). The main effect of

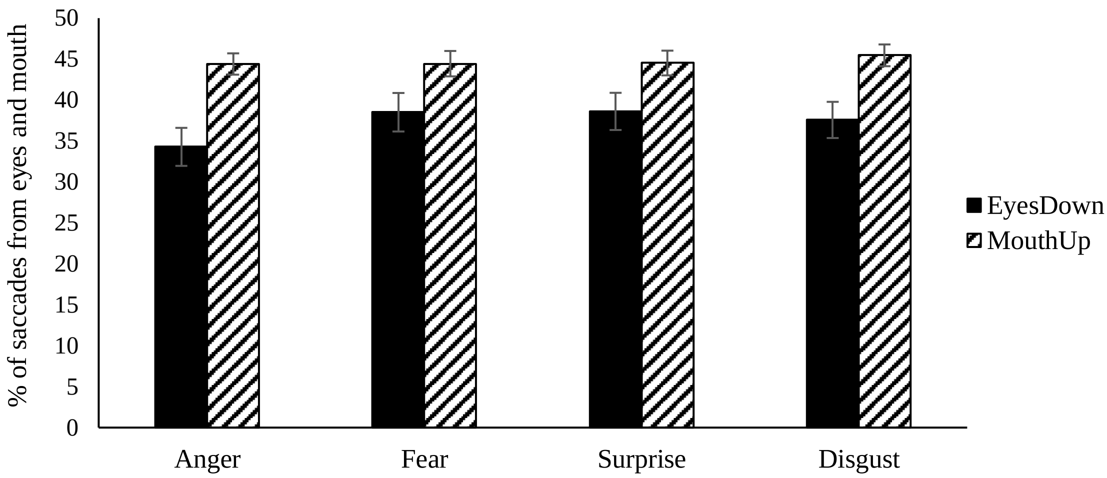

Figure 19: Percentage of reflexive saccades downwards from initial fixation locations on the eyes and upwards from initial fixation on the mouth. Error bars indicate standard error of the mean.

saccade direction indicated that there was a significantly higher percentage of saccades upwards from the mouth than downwards from the eyes (p=.01) in line with results from Experiment 1 and consistent with our expectations based on previous studies. There was no interaction between emotion and saccade direction ($F_{(3,111)}$=2.02, p= .12) which goes against our expectations that expression would modulate initial saccade direction. Since Gamer and Buchel (2009) found that the proportion of fixation changes from the mouth upwards was higher than the proportion of fixation going downward from the eyes and this effect was more pronounced for fearful faces, we wanted to further investigate whether the percentages of saccade direction were affected by facial expression used in this experiment as well. Two separate one-way ANOVAs were conducted for each saccade direction (eyes-down, mouth-up) comparing across emotions. There was a significant effect of emotion for saccades going downwards from the eyes ($F_{(3,111)}$= 4.19, p<.001, η_p^2=.10). This main effect indicated that there were fewer saccades going downwards from the eyes for angry faces compared to fearful and surprised faces. There was no main effect of emotion for saccades going upwards from the mouth ($F_{(3,111)}$= 0.60, p= .62).

A paired samples t-test for anger ($t_{(37)}$=3.21, p=.003) and a Wilcoxon signed-rank test for disgust (Z= -2.24, p=.03) showed that there was a higher percentage of saccades going upwards from the mouth compared to saccades going downwards from the eyes. Wilcoxon signed-rank tests for fearful (Z= -1.70, p=.088) and surprised (Z= -1.71, p=.088) expressions revealed no significant difference between saccade directions.

Saccade direction from upper vs lower features

Since we have used more initial fixation locations than the Gamer and Buchel (2009) study and other studies using the brief-fixation paradigm, we further analysed the percentage of the first reflexive saccades downwards from the upper features (i.e. eyes and the brow) versus upwards from lower features (i.e. cheeks and the mouth) for each emotion, with the percentages calculated relative to the total number of reflexive first saccades for each emotion. A 4 × 2 repeated measures ANOVA with emotion and fixation location (lower face, upper face) as factors was conducted. The descriptive statistics can be seen in Figure 20. There was no main effect of emotion ($F_{(2.54\ 94.05)}$=0.87, p=.44). The main effect of saccade direction ($F_{(1,\ 37)}$= 8.35, p< .001) indicated that there was a higher percentage of saccade going upwards from the lower features compared to downwards from upper features (p< .001). No interaction was found ($F_{(2.40,\ 88.96)}$= 2.23, p= .10).

Similar to the previous analysis, to further investigate whether expression had an effect on the direction of the first reflexive saccade, two separate one-way ANOVAs were conducted, one for each initial fixation location, comparing across emotions. The main effect of emotion

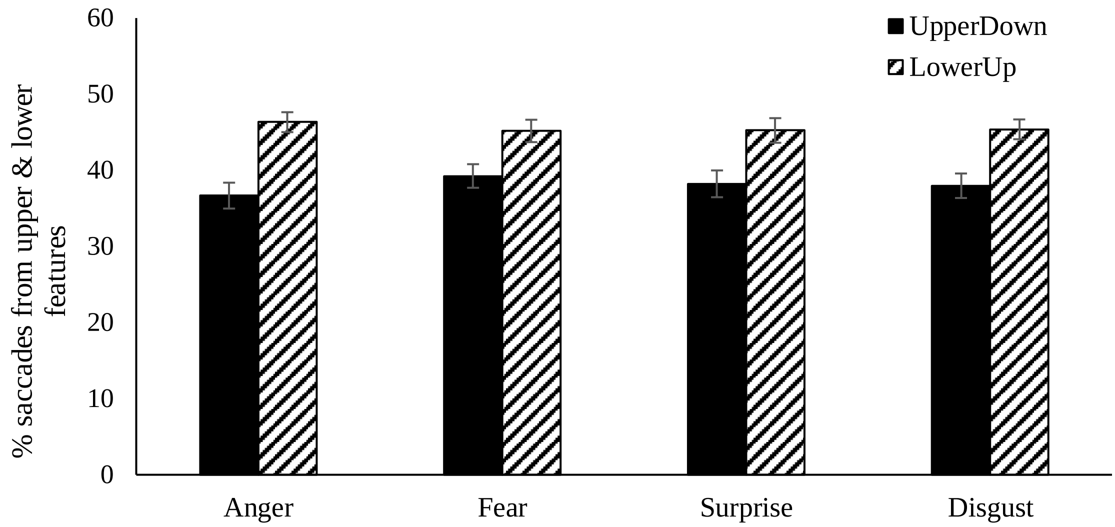

Figure 20: Percentage of reflexive saccades downwards from fixation on upper-face features or upwards from lower-face features. Error bars indicate standard error of the mean.

approached significance for saccades going downwards from the upper features ($F_{(1, 37)}$= 2.37, p= .07, η_p^2=.06) and pairwise comparisons showed that there were fewer saccades going downwards from upper features for angry faces compared to fearful faces (p=.02). No main effect of emotion was present for saccades going upwards from lower features ($F_{(2.43, 90.14)}$= 1.02, p= .39).

Wilcoxon signed-rank tests for anger (Z= -3.19, p=.001), fear (Z= -2.28, p=.023), surprise (Z= -2.02, p=.043) and disgust (Z= -2.24, p=.03) showed that there was a higher percentage of saccades going upwards from the mouth compared to saccades going downwards from the eyes.

Saccade Path Analysis

Collapsed across initial fixation location

Using a more precise saccade direction analysis, as described in the Data Analysis section, we wanted to investigate whether reflexive saccades targeted expression-informative facial features. The mean saccade paths of the reflexive saccades are shown in Figure 21 for each emotion, collapsed across initial fixation location; therefore, for each emotion there are 6 possible saccade targets (left and right eye, brow, left and right cheek, and the mouth). A 4 × 6 repeated measures ANOVA was run using emotion and saccade target as factors. The analysis was based on 38 participants. There was no main effect of expression ($F_{(3, 111)}$= 0.43, p= .73, η_p^2= .01). The main effect of saccade target failed to reach significance following Greenhouse-Geisser correction ($F_{(2.12, 78.26)}$= 2.77, p= .06, η_p^2= .07). The pairwise comparisons indicated that the reflexive saccades were more strongly in the direction of the brow than of the left (p= .002) and the right eyes (p= .017). There was also no significant interaction ($F_{(7.51, 277.83)}$= 1.51, p= .16, η_p^2= .04), contrary to the hypothesis that reflexive first saccades target expression-informative facial features.

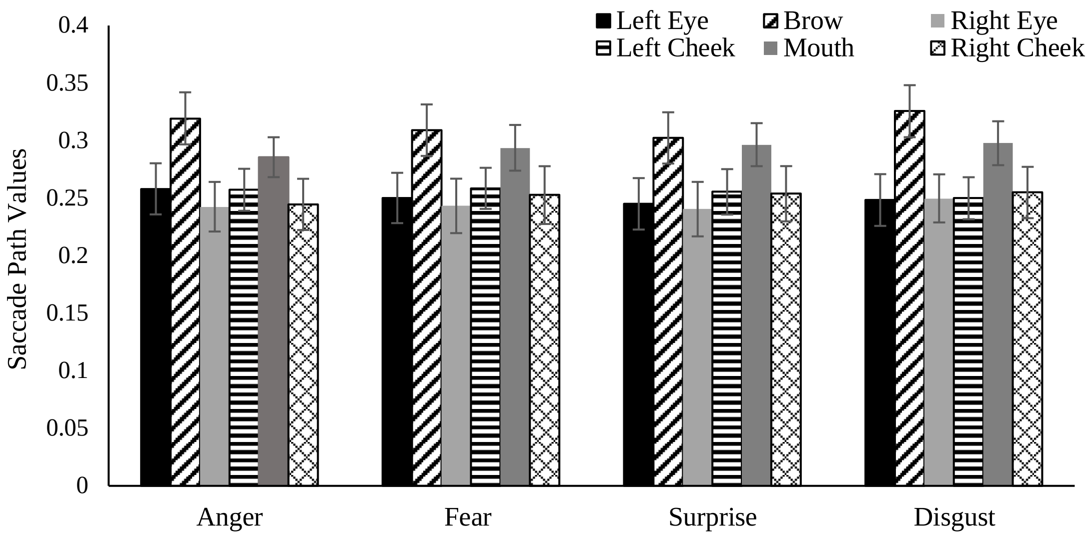

Figure 21; The mean saccade paths of reflexive saccades relative to the 6 target locations of interest, collapsed across initial fixation location. Error bars indicate standard error of the mean.

Separately from each initial fixation location

For reflexive saccades starting from the left eye, the interaction between expression and saccade target failed to reach significance after Greenhouse-Geisser correction ($F_{(3.85, 142.54)}$= 2.32, p= .06, η_p^2= .06), contrary to the hypothesis that reflexive first saccades target emotion-informative features. There was no main effect of expression ($F_{(2.39, 88.41)}$=0 .62, p= .57, η_p^2= .02), but there was a main effect of saccade target ($F_{(1.10, 41.06)}$= 24.08, p< .001, η_p^2= .39). Pairwise comparisons indicated that the reflexive saccades from the left eye were directed towards the right eye more strongly compared to the left cheek (p<.001), mouth (p<.001) and right cheek (p=.002) and more strongly towards the brow compared to the left cheek (p<.001) and mouth (p=.04). Additionally, the reflexive saccades from the left eye were more strongly directed to the right cheek compared to the left cheek (p< .001) and the mouth (p< .001).

For reflexive saccades starting from the brow, the interaction between expression and saccade target was not significant ($F_{(3.68, 136.25)}$= .1, p= .98, η_p^2= .003), contrary to the hypothesis that reflexive first saccades target emotion-informative features. There were no significant main effects of expression ($F_{(2.38, 88.13)}$= 1.23, p= .30, η_p^2= .03) or saccade target ($F_{(1.10, 40.86)}$= 2.63, p= .11, η_p^2= .07).

For reflexive saccades starting from the right eye, the interaction between expression and saccade target was not significant ($F_{(3.61, 133.53)}$= 1.72, p= .16, η_p^2= .04), contrary to the hypothesis that reflexive first saccades target emotion-informative features. There was no main effect of expression ($F_{(3, 111)}$= 1.77, p= .16, η_p^2= .05). There was a main effect of saccade target ($F_{(1.06, 39.07)}$= 26.22, p< .001, η_p^2= .42). Pairwise comparisons indicated that reflexive saccades from the right eye were more strongly directed towards the brow compared to all other saccade targets (all *p*s$< .001$) and towards the left eye compared to both cheeks and the mouth (all *p*s$< .001$). The saccades also targeted the left cheek more strongly compared to the mouth and the right cheek (*p*s$< .001$).

For reflexive saccades starting from the left cheek, the interaction between expression and saccade target was not significant ($F_{(3.34, 123.54)}= 1.02$, $p= .39$, $\eta_p^2= .03$), contrary to the hypothesis that reflexive first saccades target emotion-informative features. There was no main effect of expression ($F_{(2.52, 93.24)}= 1.07$, $p= .36$, $\eta_p^2= .03$). There was a main effect of saccade target ($F_{(1.05, 38.76)}= 10.04$, $p= .003$, $\eta_p^2= .21$). Pairwise comparisons indicated that reflexive saccades from the left cheek were more strongly directed towards the right cheek compared to the left eye ($p= .005$), brow ($p= .007$) and right eye ($p= .01$). Additionally, the reflexive saccades from the left cheek were more strongly directed towards the right eye compared to the brow and the left eye (both *p*s$= .004$). Finally, saccades were directed towards the mouth more strongly compared to the left eye ($p= .03$).

For reflexive saccades starting from the mouth, there was no interaction ($F_{(3.34, 123.57)}=1.02$, $p= .39$, $\eta_p^2= .03$), nor main effects of expression ($F_{(2.40, 88.94)}=0.31$, $p= .78$, $\eta_p^2= .008$) or saccade target ($F_{(1.09, 40.19)}=0.70$, $p= .42$, $\eta_p^2= .02$).

Finally, for reflexive saccades starting from the right cheek, there was no interaction between expression and saccade target ($F_{(3.22, 119.05)}= 1.05$, $p= .38$, $\eta_p^2= .03$), contrary to the hypothesis that reflexive first saccades target emotion-informative features. There was no main effect of expression ($F_{(3,111)}=0.85$, $p= .47$, $\eta_p^2= .02$). There was a main effect of saccade target ($F_{(1.04, 38.52)}= 10.99$, $p= .002$, $\eta_p^2= .23$). Pairwise comparisons indicated that reflexive saccades from the right cheek were directed more strongly towards the mouth compared to all other saccade targets (left eye: $p=.001$; brow: $p= .004$; right eye: $p= .037$; left cheek: $p< .001$).

Long presentation

Mean Total Fixation Duration

To investigate whether participants spent more time fixating the informative facial features for each of the expressions in the study, we calculated the mean total fixation duration for the eyes, brow, nose and mouth per participant. The mean total fixation duration was then normalized to the mean number of pixels within each ROI across all the faces presented in the study. Descriptive statistics can be seen in Figure 22. There were significant main effects of emotion ($F_{(2.00, 76.04)}=166.87$, $p< .001$, $\eta_p^2=.82$) and region of interest ($F_{(1.63, 62.21)}=40.13$, $p< .001$, $\eta_p^2=.51$). The main effect of emotion showed that the mean total fixation durations for the angry and disgusted faces were longer compared to surprised and fearful faces (all *p*s< 0.01) which did not

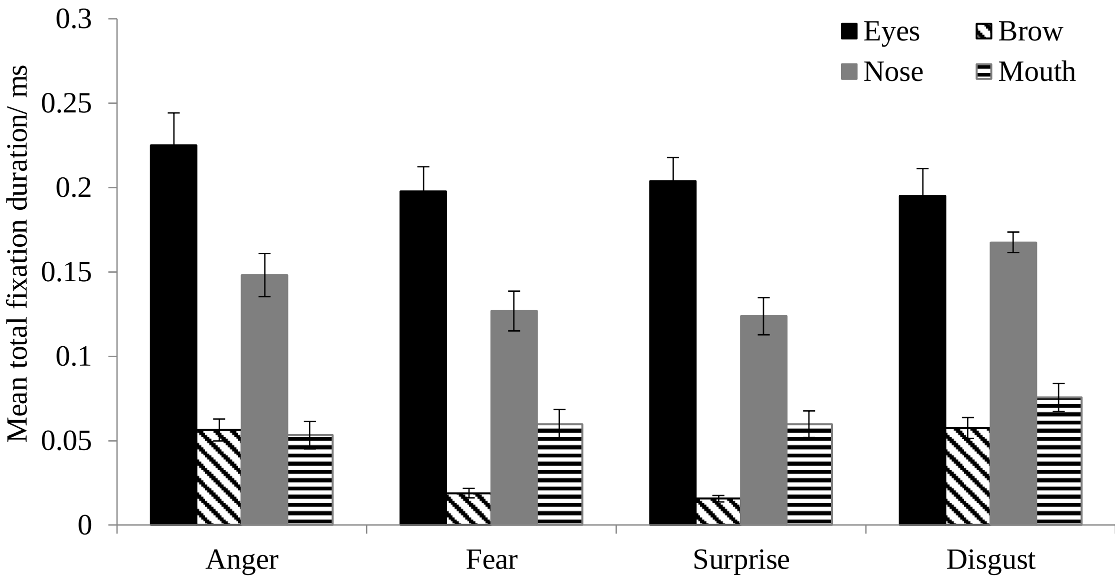

Figure 22; The average total fixation duration within each of the four regions of interest (eyes, brow, nose and mouth) for each expression. The values are area normalized. Error bars indicate standard error of the mean.

differ. The mean total fixation duration for the eye and nose regions (no difference, *p*=.10) was higher compared to the brow and the mouth regions (all *p*s<.01) which did not differ. Importantly, there was a significant interaction between emotion and ROI ($F_{(4.96,\ 188.59)}$=15.06, *p*< .001, η_p^2=.28). To further investigate the interaction effect, the mean total fixation duration within each ROI was compared across emotions. This led to 4 separate one-way ANOVAs looking at the effect of emotion for each ROI separately.

There was a significant effect of emotion for the eye region ($F_{(2.32,\ 87.98)}$=10.48, *p*< .001, η_p^2=.22) indicating that the mean total fixation duration for the eyes of the angry faces was significantly higher compared to all other expressions (fear: *p*=.001, surprise: *p*=.03, disgust: *p*<.001). No other comparison was significant.

There was also a significant effect of emotion for the brow region ($F_{(1.73,\ 65.59)}$=50.90, *p*< .001, η_p^2=.57). The mean total fixation duration for the brow region of angry and disgusted faces was longer compared to the fearful and surprised faces (all *p*s<.001) which did not differ. The longer mean total fixation duration for the brow region of angry expressions was in line with our expectation that observers would spend longer looking at expression-informative facial features. This result for disgusted expressions, on the other hand, was unexpected since we hypothesized that the mouth would be informative for recognition of disgusted expressions.

The significant effect of emotion on the mean total fixation duration for the nose region ($F_{(2.31,\ 87.56)}$=42.55, *p*< .001, η_p^2=.53) indicated that the mean total fixation duration for the nose region of disgusted faces was highest compared to all other emotions (all *p*s<.001). The nose region of angry faces also received higher mean total fixation durations than fearful (p<.001) and surprised expressions (p=.001) which themselves did not differ.

Finally, there was a significant effect of emotion on the mean total fixation duration for the mouth region ($F_{(2.52, 85.58)}$=19.14, p< .001, η_p^2=.34). Pairwise comparisons show that the mean total fixation duration for the mouth region of the disgusted faces was higher than all others (all ps<.001) which was in line with our expectation that observers would spend longer looking at expression-informative facial features. The mouth region of the fearful faces received longer mean total fixation durations than angry faces (p=.04) but there was no difference between fearful and surprised faces. The higher mean total fixation duration for the mouth region of fearful faces was also in line with our expectations, however while we expected a similar result for surprised expressions, we did not find this.

Total fixation duration and emotion recognition accuracy

To investigate whether there is a relationship between time spent fixating a region of informative value for the emotion and emotion recognition accuracy, Pearson's or Spearman's correlations (as appropriate) between the time spent fixating an informative region and recognition accuracy were computed. All computed correlations can be seen in Figure 23.

For angry expressions, we tested whether participants would be more accurate if they spent longer fixating the brow region, which is known to be informative for anger. There was no

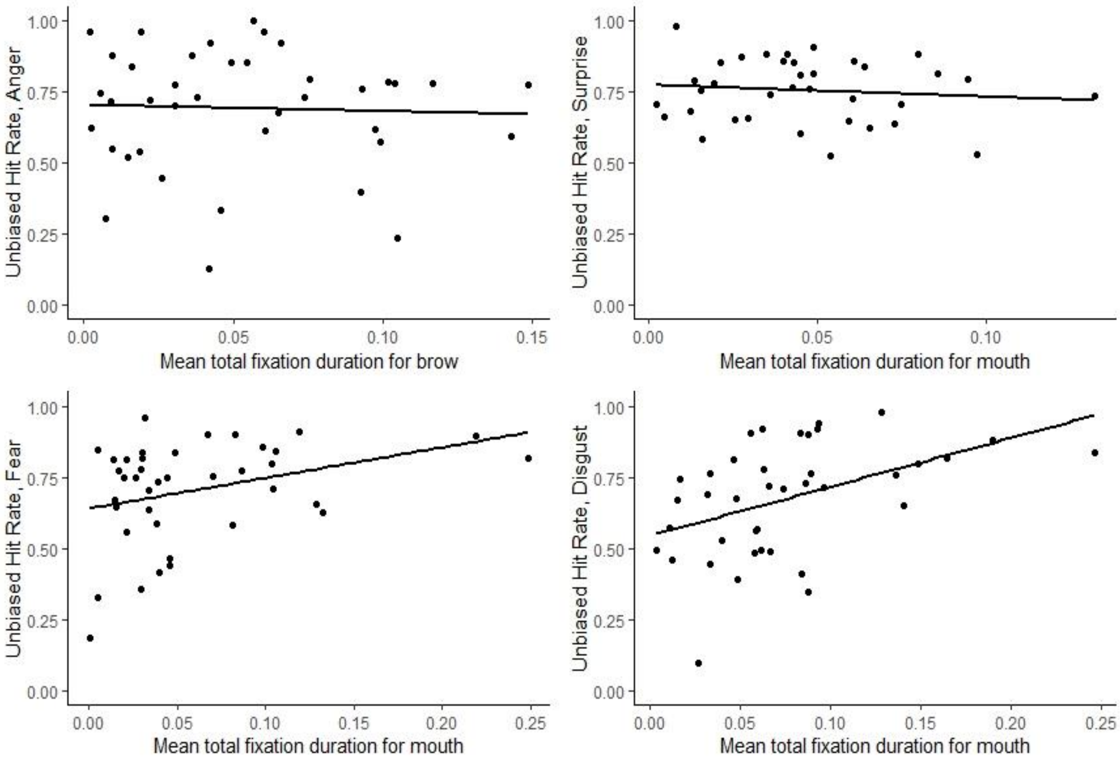

Figure 23: The correlation between unbiased hit rates for anger, fear, surprise and disgust and mean normalized total fixation duration in the long-presentation paradigm.

correlation between unbiased hit rates for anger and fixation duration on the brow (Spearman's *rho*$<$ .001, $p > .9$).

For fearful expressions, we tested whether participants would be more accurate if they spent longer fixating the mouth region, which is known to be informative for fear, especially when surprised expressions are also presented. The correlation between unbiased hit rates for fear and fixation time on the mouth approached significance (Spearman's *rho*= .29, p= .08).

For surprised expressions, we tested whether participants would be more accurate if they spent longer fixating the mouth region, which is known to be informative for surprised expressions, especially when fearful expressions are also presented. There was no correlation between accuracy and fixation duration on the mouth (Spearman's *rho*= .25, p= .12).

Finally, for disgusted expressions, we tested whether participants would be more accurate if they spent longer fixating the mouth region, which is known to be informative for disgusted expressions. There was a medium positive correlation between accuracy and fixation duration on the mouth (Spearman's *rho*= .49, p= .001) which is in line with our suggestion that the mouth region will be informative for disgust recognition.

Summary

In the brief-fixation paradigm in Experiment 2, we used a combination of briefly presented angry, fearful, surprised and disgusted expressions to investigate the contribution to emotion recognition of initially fixating an informative facial feature and whether reflexive first saccades seek out informative facial features when these are not initially fixated. Furthermore, in the long presentation paradigm, we investigated whether observers would spend longer in looking at emotion-informative facial features when allowed to view angry, fearful, surprised and disgusted expressions for longer. Finally, again in the long presentation paradigm, we investigated whether there was a relationship between how long observers looked at emotion-informative facial features and their emotion recognition performance.

In the brief-fixation paradigm, we found that initially fixating on the mouth region led to higher accuracy compared to initially fixating the eyes and brow. We also found that initially fixating the mouth region led to higher emotion recognition accuracy for fearful, surprised and disgusted expressions. In both the brief-fixation and the long-presentation paradigms, disgusted expressions were more often misclassified as angry than any other target expressions. Similar to Experiment 1, fearful expressions were more often misclassified as surprised than any other target expressions. Whereas the confusion between disgust and anger was alleviated when initial fixation was on the mouth in both paradigms, the confusion between fear and surprise was only alleviated when initial fixation was on the lower-face features (mouth and cheeks) in the brief-fixation paradigm but not the long-presentation paradigm.

Similar to Experiment 1, observers shifted their gaze upwards from initial fixation on the lower facial features (mouth and cheeks) than downwards from upper facial features (eyes and brow) in the brief-fixation paradigm. Additionally, observers preferred to shift their gaze downwards less for angry expressions compared to fearful and surprised faces. This is not in line with previous findings where researchers found that observers made higher proportions of fixation changes upwards from the mouth for fearful faces. It is possible that the mouth region in our study contained information useful for recognition of fear, which is supported by the reduction of confusion between fear and surprised expressions when initially fixating the lower facial features. The saccade path analysis showed that saccades initiated at the left or right eyes were directed towards the brow or the other eye, replicating and lending further support to our results from Experiment 1. Furthermore, saccades initiated at the cheeks, further sought out other lower facial features. Therefore, it is possible that instead of seeking out informative facial features, reflexive saccades might be targeting the closest facial feature. Further exploratory analyses on reflexive saccades were carried out and reported in the Appendix. We found some evidence to indicate that the first saccades were affected by the centre-of-gravity effect whereby most saccades ended in the nose region. However, there is also evidence that these first saccades were influenced by informative features of the expressions presented: more saccades ended in the mouth region for surprised and disgusted faces. Similar to our findings regarding saccade latencies in Experiment 1, the saccade latencies of the reflexive saccades from the cheeks were shorter compared to other initial fixation locations. Furthermore, we found that the saccade latencies were also shorter from the eyes compared to the brow. In Experiment 2, saccade latencies were also affected by expression: The saccade latencies from the brow region of angry faces were shorter compared to disgusted expressions. Additionally, saccade latencies from the cheeks of fearful expressions were longer than angry and disgusted expressions.

When allowed to view faces for 5 seconds, observers looked longer at the eye region of angry faces compared to all other expressions and brow region of angry faces longer compared to fearful and surprised expressions. Participants looked significantly longer at the mouth region of the disgusted faces compared to all other expressions and fearful expressions compared to angry faces. This is in line with what we expected regarding the distribution of informative facial features for these expressions. However, the fact that observers spent similar durations at the brow region of angry and disgusted faces might have led to the confusions between these two expressions, possibly due to their visual similarity (i.e., the furrows of the brow region). Similarly, and as pointed out in Experiment 1 as well, no differences in the total fixation durations spent in any of the ROIs were found between the fearful and surprised facial expressions. This lack of difference in the fixation patterns for fearful and surprised faces might have led to the confusion of these expressions since the observers might have neglected to fixate on the distinctive features of fear and surprise. Finally, we found that time spent fixating the mouth region was associated with higher emotion recognition accuracy for disgusted faces in the long-presentation paradigm.

Discussion

The main aims of the study were to investigate whether initially fixating an informative facial feature would improve the emotion recognition performance of observers and whether observers seek out informative facial features when these are not initially fixated. Furthermore, we were also interested in the distinctive gaze patterns elicited by the facial expressions and the relationship between the gaze patterns and emotion recognition. To this end, we employed two brief-fixation and long-presentation paradigms with different combinations of 4 out of the 6 basic emotions.

In the brief-fixation paradigm, we manipulated the initial fixation location to fall on the left or right eye, the central brow, the left or right cheek, or the centre of the mouth, in order to investigate the contribution of foveating an informative feature over a non-informative feature. The brief presentation time (~80ms) was used to ensure that participants did not have sufficient time to saccade to other locations on the face. Thus, we guaranteed that only one of the 6 facial features fell at the fovea on any given trial and thus that the rest of the face was projected extrafoveally. By examining the direction of the reflexive first saccades from face offset, we also investigated whether informative facial features influence reflexive eye movements. Finally, in the long presentation paradigm, we investigated the gaze patterns elicited by the target expressions when presented for a longer time (5s) and the link between amounts of time spent fixating informative facial features and emotion recognition performance in both the long-presentation and brief-fixation paradigms.

Contribution of foveal processing of informative features to emotion recognition

We found that overall, initially fixating on the mouth led to higher emotion recognition performance compared to the cheeks in Experiment 1 and to the eyes and the brow in Experiment 2. Previous brief-fixation paradigms failed to show an effect of initial fixation location on emotion recognition performance (Gamer & Buchel, 2009; Scheller et al., 2012) except for Kliemann et al. (2010) who showed that initially fixating on the eye region improved fear recognition. Neath and Itier (2014) reported that initially fixating on the nose region led to higher hit rates compared to initially fixating either the left or the right eye. In a further study, Neath-Tavares and Itier (2016) found that initial fixation on the nose and the mouth regions led to higher recognition accuracy for fearful, happy and neutral faces. Using the *Bubbles* technique, Blais, Roy, Fiset, Arguin and Gosselin (2012) found that the mouth was the most informative region for emotion recognition across the six basic expressions, plus pain and neutral expressions, both static and dynamic. They suggest that the preference to utilise the mouth region for their emotion recognition task might stem from the highly dynamic motion range of the mouth compared to the eyes or other facial features and the extension of this strategy to the visual processing of static expressions. Interestingly, research also suggests that the mouth region is more informative for the recognition of posed expressions whereas the eye region is the most informative for spontaneous expressions due to this dynamic motion range of the mouth (Duncan et al., 2017; Saumure, Plouffe-Demers, Estéphan, Fiset & Blais, 2018). Considering the nature of the stimuli used (posed expressions) and the combination of expressions included in our experiments, it is possible that the mouth region was more informative for the recognition of all the expressions included in both our experiments. This can further be supported by the reduction of misclassifications between most expressions (anger and disgust, fear and surprise) when the mouth region was initially fixated.

In the brief-fixation paradigm of both experiments, we found that initially fixating on the mouth led to better recognition for fearful expressions compared to fixating on the eyes and the brow. This is surprising, given that use of the Bubbles methodology has shown that high spatial frequency information from the eye region is informative for fearful expression recognition (Smith et al., 2005; Merlusca & Smith, 2015; Smith et al., 2018) and that the ability of a patient with bilateral amygdala damage to recognise fear in faces was restored when she was instructed to direct her gaze to the eye region (Adolphs et al., 2005). Using a similar paradigm, in which observers fixated one of the eye regions or the mouth, Gamer and Buchel (2009) and Scheller, Buchel and Gamer (2012) failed to find an effect of initial fixation location on fear recognition accuracy[4]. However, Smith and Merlusca (2015) state that low spatial frequency information at the mouth region becomes informative when fear is compared against neutral faces alone, angry and neutral faces and happy and neutral faces. The area of the mouth that becomes informative in different comparisons differs. Therefore, it is possible that in our study, when being compared to anger, surprise, sadness and disgust, the mouth region was more informative for fear recognition. However, the informative low spatial frequencies at the mouth cannot solely explain our findings since these would have been present extrafoveally even when the mouth region was not fixated. Smith et al. (2005) showed that expressive information transmitted for fear and surprise overlaps to a high extent with the eye region and mouth region transmitting informative information for both expressions. It is possible that in a subset of expressions

[4] It has to be noted that the 'eye region' used by Gamer and Buchel (2009) and Scheller et al., (2012) are different: While Gamer and Buchel (2009) shifted the faces so that the initial fixation will fall on either eye, Scheller et al. (2012) shifted the face images so that initial fixation will fall between the two eyes, which is more similar to the brow region used in this experiment.

containing fear and surprise, observers make use of high spatial frequency information from the mouth to recognise fear. This suggestion is supported by the finding that observers were less likely to misclassify fear as surprise when their initial fixation was on the mouth of fearful faces. The usefulness of the mouth region for fear recognition might help to reconcile the contradictory findings on the emotion recognition abilities of individuals with ASD. Some studies show that individuals with ASD prefer to look at the mouth region of faces at the expense of the eyes (Klin et al., 2002), a strategy that might allow them to extract relevant expression information from faces.

Initially fixating the mouth of the disgusted expressions also led to improved expression recognition, and this is in line with studies which suggest that the mouth region contains informative information for the recognition of disgusted faces (Smith et al., 2005; Wegrzyn et al., 2017). Furthermore, Jack et al. (2009) showed that Eastern Asian (EA) observers have poorer recognition of certain expressions compared to Western Caucasian (WC) observers, including disgust. This poor categorisation of disgust was also accompanied by its confusion with anger, which is also replicated in our study. EA observers preferred to look more at the eye region of faces at the expense of the mouth region in Jack et al.'s (2009) study. They showed that a model observer could closely replicate the recognition results of EA observers when sampling from the eye and eyebrow regions and the model observer's results were most dissimilar to the EA observers when sampling information from the mouth. Therefore, Jack et al. (2009) suggested that sampling information from visually similar facial features between anger and disgust, such as the eyes and eyebrows, at the expense of the distinctive facial features, such as the mouth, led to the confusion between certain expressions. Both obtaining expression-informative, high spatial frequency information from the mouth region at initial fixation and attending to the distinctive mouth region between anger and disgust early in visual processing could have resulted in improved recognition accuracy for disgust in the present study.

Initially fixating on the brow region led to higher emotion recognition accuracy for anger compared to the cheeks in Experiment 1, in line with our hypothesis. However, this result should be considered alongside the finding of higher anger recognition accuracy when initially fixating the mouth compared to the cheeks. It is possible that these results might allude to the non-informativeness of the cheeks rather than the informativeness of the brow or the mouth. On the contrary, in Experiment 2, we found that initially fixating on the brow region led to *lower* emotion recognition accuracy compared to initially fixation on the mouth or the cheeks. There might be several reasons for the lack of consistent recognition improvement when initially fixating the brow for anger: The overall recognition performance in both experiments was high, therefore, this ceiling effect might have prevented any contribution of initially fixating informative features of anger from being observed. Furthermore, Papinutto, Lao, Ramon, Caldara and Miellet (2017) suggested that an area of 7 degrees of visual angle around a single fixation location can lead to face recognition performance that is similar to performance during natural viewing. They termed this area the '*facespan*'. This study investigated the facespan for face recognition, therefore the application of this to emotion recognition should be approached with caution. The close proximity of the brow and the eyes might mean that the informative information from the brow region is still accessible via the extrafoveal visual field when fixating the eyes, diminishing the contribution of foveating the brow. It is also possible that the informativeness of the brow region for anger recognition is less pronounced for the set of expressions used in Experiments 1 and 2. In Experiment 1, it was found that initial fixation on the mouth, as well as the brow, led to higher anger accuracy compared to the cheek and in Experiment 2, initially fixating on the mouth and cheeks led to higher emotion recognition compared to the brow region. Additionally, contrary to what we hypothesized, the reflexive saccades did not differentially target the brow region of angry expressions, even from lower facial features. On the other hand, mouth was targeted from the brow and from the right cheek for angry expressions. Combined, it is possible that the mouth region, instead of the brow region, was informative for the recognition of angry expressions in the experiments reported in this chapter and this might be due to the subset of the basic emotions used in these experiments. Using the same brief-fixation paradigm with different combinations of emotions, Atkinson and Smithson (2020) found that initially fixating on the central brow of angry faces improved recognition and this was due to a reduction of misclassification of angry faces as neutral faces. This result was found in an experiment with angry, fearful, happy and neutral expressions, and in an experiment with angry, fearful, surprised and neutral expressions. Additionally, Shurgin et al. (2014) found that observers spent more time looking at the brow region of neutral faces when they were asked to identify angry faces of various intensities in a block. The main difference between these previous studies and the current study is the lack of neutral faces. According to Mersulca and Smith (2015), the informative features of expressions can be flexible and be determined by the combination of expressions they need to be discriminated from. For example, in Experiment 2, we used expressions of anger, fear, surprise and disgust and found that anger and disgust was confused with each other. The brow region of angry and disgusted expressions is similar, therefore, it is possible that initially fixating the brow region of angry faces might have contributed to the confusion of these expressions. This, in turn, might have hindered anger recognition among this combination of expressions rather than being informative. Examining the results of previous studies, it is possible that the brow region is particularly informative for discriminating angry faces from neutral faces and the lack of neutral faces along with the addition of disgust in Experiment 2 reduced the informativeness of this region for anger.

The confusion between fearful and surprised expressions found in both experiments is a prevalent finding in the literature (Calder, Rowland, Young, Nimmo-Smith, Keane & Perrett, 2000; Chamberland et al., 2017; Gagnon et al., 2010; Jack et al., 2009, 2014; Roy-Charland et al., 2014, 2015a). Jack et al. (2014) suggested that fearful and surprised expressions are confused due to the activation of the upper lid raiser and the jaw drop muscles early in the unfolding of these expressions. This suggests that the visual similarity of these expressions lead to their confusion. Furthermore, in the framework of attentional-perceptual limitation theory, since fear and surprise have more shared action units than unique ones, this leads to an inability to perceive the distinctive signals between the two expressions (Roy-Charland, Perron, Beaudry & Eady, 2014; Roy-Charland, Perron, Young, Boulard & Chamberland, 2015; Chamberland, Roy-Charland, Perron & Dickinson, 2016). Roy-Charland et al. (2014) and Roy-Charland et al. (2015) showed that when the lip stretcher action unit was present in the mouth region of fearful faces, observers were better able to recognize fearful expressions. Jack et al. (2009) found that EA observers confused fear and surprise due to an eye movement strategy that involved sampling information exclusively from the eye region. Using a model observer, Jack et al. (2009) also showed that sampling information from the mouth region of surprised and fearful faces decreases the confusion of these expressions. Our result agrees with these findings and extends them to show that initially fixating the mouth region decreases misclassification of briefly presented fearful faces as surprised, possibly by directing attention to the distinctive facial feature between these expressions. The confusion between fearful and surprised expressions seems to be somewhat asymmetrical (Du & Martinez, 2011): While we show that fear is confused as surprise and surprise as fear more frequently than other emotions, the magnitude of these confusions are not equal: Misclassification of fear as surprise is more prominent compared to misclassification of surprise as fear. We also find that in both Experiment 1 and Experiment 2, surprised expressions were better recognised compared to fearful expressions. The AUs involved in the prototypical expressions of surprise are a subset of the AUs making up fearful expressions. A prototypical fearful expression has two distinctive AUs not expressed in surprise: the brow lowerer and the lip stretcher. Roy-Charland et al. (2015b) showed that the presence of the lip stretcher AU, at the mouth region, contributes more to the discrimination of fearful expressions from surprise. This might help to explain the asymmetry of confusions between fear and surprise: Since there is additional information regarding identification of fearful expressions at the mouth region, initially fixating this region might aid fear recognition more than surprise recognition.

We also found that disgust was more often mistaken as anger than it was misclassified as another expression. Such confusions between these expressions also agree with previous findings in the literature (Jack, Garrod & Schyns, 2014; Jack, Blais, Scheepers, Schyns &

Caldara, 2009; Du & Martinez, 2011; Widen, Russel & Brooks, 2004; Gagnon, Gosselin, Buhs, Larocque & Milliard, 2009). Jack et al. (2014) suggest that anger and disgust are confused due to the activation of the nose wrinkler and the lip funneler early on in the unfolding of dynamic expressions, but this is then resolved by the activation of the upper lip raiser. In the same vein as the confusion between surprise and fear, Jack et al. (2009) suggested that sampling information from the eye region at the expense of the mouth region leads to the confusion patterns observed in EA participants. Our results show that when attention is directed to the mouth region of disgusted faces, disgusted expressions are less likely to be misclassified as anger.

Across the two experiments, while we established that initially fixating the mouth region benefitted the recognition of fear, surprise and disgust compared to initially fixating upper facial features (i.e. the eyes and the brow), we failed to find any difference in recognition accuracy for these expressions between initially fixating the mouth and the cheeks. Additionally, while initially fixating the mouth reduced the confusion of fear as surprise and disgust as anger, there was no difference in these misclassifications between initially fixating the mouth and the cheeks. While we have chosen the cheeks to be relatively non-informative facial features, it is possible that the chosen cheek location in this study is not as non-informative as expected. As can be seen in Figure 6, the location of the cheek positions is close to the nasolabial folds. Therefore, it is possible that this region confers useful visual information for certain expressions. The nasolabial folds were shown to be informative for disgust (Smith et al., 2005). Further, it is possible that the open mouth expressions of fear and surprise led to more mid- to high- SF information from the mouth to be accessible when initially fixating on the cheeks, making the cheeks as informative as the mouth.

Research showed that the perceptual sensitivity for each of the six basic emotions differ (Calvo & Lundqvist, 2008; Delicato, 2020; Maher, Ekstrom & Chen, 2014; Marneweck, Loftus & Hammond, 2013). Maher, Ekstrom and Chen (2014) showed that the perceptual sensitivity for happy expressions was lower compared to angry, fearful and sad expressions and higher for sad expressions compared to angry, fear and happy expressions. This finding was corroborated by Marneweck, Loftus and Hammond (2013) who also found lowest perceptual sensitivity for happy and highest for sad expressions. Both these studies measured perceptual sensitivity in a paradigm where participants were shown two faces of the same individual at two different intensity levels (neutral paired with a higher intensity and pair of expressions with two different intensities) and asked to identify which face showed an expression (or which face expressed the higher intensity expression). The perceptual sensitivity was defined as the lowest intensity for each expression to be discriminated from neutral. Defining perceptual sensitivity as the amount of neutral face in an expression which would lead to 75% correct recognition, Delicato (2020) showed high perceptual sensitivity for happy expressions. Delicato (2020) also showed lower sensitivity for fearful expressions compared to happy. Therefore, in line with Marneweck et al. (2013) and Maher et al. (2014), Delicato (2020) also showed that happy expressions can be recognized at lower intensities while more expressive information is needed for the recognition of fearful expressions. It should be noted that this experiment used only happy and fearful expressions, so it is not possible to say perceptual sensitivity for fear is the lowest since previous research showed that this was lower for sad expressions when fear was included in the subset of expressions examined (Maher et al., 2014). Regardless, Delicato (2020) found that the difference in perceptual sensitivity between happy and fearful expressions did not change with exposure duration (8, 83 and 200ms) but sensitivity for both expressions increased with longer exposure durations. Calvo and Lundqvist (2008) also showed the superiority of happy expressions in an experiment where participants were asked to identify expressions of all six basic emotions under different exposure conditions (25, 50, 100 and 250ms). The identification of happy expressions reached ceiling performance at 50ms presentation duration. In contrast, fearful expressions were generally recognized poorest and slowest. Additionally, there were further differences between the identification of fearful expressions and expressions of other emotions (anger, disgust, sad and surprise) in all exposure durations: Fearful faces were recognized less accurately compared to sad expressions at 25ms, compared to sad and angry expressions at 50ms, compared to all expressions at 100ms and compared to angry expressions in 500ms. Since our perceptual capacity to discriminate the expressions of each of the six basic emotions, and especially fearful and sad expressions, are fundamentally different, any research comparing emotion recognition performance between expressions under similar conditions should account for the confound this difference introduces into the results. While happy expressions were not studied in the experiments reported in Chapter 2, the difference in perceptual sensitivity between other expressions and fear might still indicate that our 82.4ms presentation duration did not allow enough time for all expressions to be processed in a comparable manner. This is supported by the overall worse recognition performance for fear recognition. However, lower sensitivity for fearful expressions compared to other expressions seems to be present regardless of exposure duration (Calvo & Lundqvist, 2008; Delicato, 2020). To account for the confound the difference in perceptual sensitivities between expressions would introduce, it would be best to limit comparisons to within-emotion in cases where fearful expressions are investigated. In the analysis of the contribution of initial fixation location to emotion recognition, we have done this. Since we compared the effect of initial fixation location on emotion recognition accuracy within expression (i.e. fear eyes compared to fear mouth), we can say with confidence that this effect was not affected by perceptual sensitivity. However, it still has to be noted that reaction times to fearful expressions in the brief fixation paradigm was longer which might suggest that with a longer presentation duration, it might have been possible to see a contributory effect of initial fixation locations other than the mouth. Nevertheless, future research, especially research using short presentation durations, could investigate expressions with similar perceptual thresholds or create expression-specific conditions which would equate the perceptual sensitivity for all expressions studied if comparison between expressions is necessary.

While there have been a considerable number of research conducted using a similar brief fixation paradigm (Gamer & Buchel, 2009; Scheller et al., 2012; Kliemann et al., 2010; 2012; Neath & Itier, 2014; Atkinson & Smithson, 2020), the results seem to be unequivocal and largely dependent on the set of expressions used. Studies using happy and fearful expressions found no effect of initial fixation location on emotion recognition accuracy (Gamer & Buchel, 2009; Scheller et al., 2012; Kliemann et al., 2010; 2012). Since this is a binary decision, both the eyes and the mouth are equally informative for recognition of both expressions. Since these two expressions do not share any action units, the informative region for fearful faces (the widened eyes) can also be used to decide that the expression is not happy. In the absence of another emotion response category, if the expression is not happy, it must be fear. This issue could be overcome by investigating the contribution of facial features to recognition of expressions when there are more than two emotion categories, especially emotion categories where sampling of a single feature would not eliminate all other response options in favor of only one option. This is what we tried to achieve in the experiments reported in this Chapter, especially with expressions that are commonly confused with each other. However, comparing our findings from Experiments 1 and 2 with each other and previous research, we see that the use of different subsets of facial expressions lead to different results regarding the informativeness of facial features. For example, while Atkinson and Smithson (2020) showed that initial fixation on the brow led to improved anger accuracy, here using the same paradigm with a different subset of expressions, we find only that initial fixation on the brow improves anger accuracy only compared to the cheeks when anger, fear, surprise and sad expressions are used. More strikingly, we find that initial fixation on the brow leads to a *decline* in anger recognition when a subset of anger, fear, surprise and disgust expressions are used. The dependence of findings on the subset of expressions used is a problem for the generalization of findings from this line of research and the formation of a unitary theory for the contribution of facial features to emotion recognition. Additionally, using different subset of expressions makes findings from different studies incomparable. Future research can commit to investigating expressions of all six basic emotions to make findings comparable. For example, here we find similar results to Atkinson and Smithson (2020) regarding the informativeness of the mouth region for the resolution of the misclassification of fearful expressions as surprise. And these findings also map onto the findings of Jack et al. (2009) who showed that the mouth region aids discrimination of fear from surprise.

Seeking out of informative features

We found that, in the brief-fixation paradigms for both experiments, initial saccades were directed upwards from initial fixation on the mouth and lower facial features (cheeks and mouth) more frequently than downwards from initial fixation on the eyes and upper facial features (eyes and brow). These findings replicate Gamer and Buchel (2009) and Scheller et al.'s (2012) findings of higher proportions of fixation changes upwards from the mouth than downwards from the eyes. Therefore, our results do indicate that the saccades of observers are reflexively directed upwards despite the lack of possibility to sample any further information from the upper face. Blais et al. (2012) suggested that since the eyes are smaller than the mouth, the eye region contains more high spatial frequency information that needs to be processed foveally, whereas the lower spatial frequency information from the mouth region can be processed extrafoveally. Therefore, it is possible that the propensity of observers to direct their gaze upwards from the mouth might stem from the need to foveate the eye region to obtain high spatial frequency information. We also found that the propensity to saccade upwards or downwards was modulated by the expression presented but this was not consistent across the two studies. Contrary to the findings of Gamer and colleagues, we did not find more saccades going upwards from the mouth for fearful faces compared to any of the other expressions. Combined with the finding that initially fixating on the mouth of fearful faces reduced their misclassification as surprise, it is possible that observers in our study found the mouth region more informative and therefore did not make as many saccades to the eye region as in the previous studies. By contrast, expression did have an effect on the percentage of saccades going downwards from the upper features. We found that there were fewer saccades going downwards from the eyes for sad faces for Experiment 1. In Experiment 2, there were more saccades going downwards from the eyes for fear and surprised faces compared to angry faces.

Altogether, our results suggest that observers make reflexive saccades upwards from the mouth region of expressive faces more often than they do downwards from the eye region; however, our results do not allow us to argue that these saccades are consistently targeting informative features. One possible reason for the difference between our studies and the previous ones is their use of a limited number of expressions, namely fear, happy and neutral (and in some cases, also anger). The informative regions for happy and fearful faces are spread out in different sections of the face (Calder et al., 2000; Beaudry et al., 2014). The most informative region for happy faces is shown to be the mouth (Smith et al., 2005; Calder et al., 2000; Beaudry et al., 2014). When the alternative emotion categories presented are happy and neutral expressions, the most informative region for fearful faces is the eye region (Smith & Merlusca, 2014). It is possible that this separation of informative features between these two expressions might have accentuated the difference in reflexive saccade patterns. Additionally, since the information in the eye region of fearful and surprised faces is similar, our participants might not have found the eye region more useful to identify fearful expressions than to identify any other expression leading to no clear difference in reflexive saccade patterns from the mouth region upwards.

To better classify the target of these reflexive saccades, we conducted a more fine-grained analysis of the paths of the reflexive first saccades. We measured the similarity of the first saccade vector to six possible saccade vectors which targeted one of the possible initial fixation locations (i.e., left or right eye, brow, left or right cheek, and mouth) on the face. This analysis showed that the saccades initiated at the left and right eyes were more strongly directed towards one of the other upper facial features (i.e., the opposite eye or the brow). This might be the reason behind the observation of lower proportion of saccades going downwards from the eyes. When we consider the upper and lower portions of faces, there are more facial features of interest in the upper part of the face (i.e., two eyes, eyebrows, wrinkles of the brow) compared to the lower part (i.e., mouth). Additionally, as stated above, the eye region contains more details in high spatial frequencies (Blais et al., 2012). Therefore, it is possible that observers examine the upper half of the face more extensively than the lower part simply due to the quantity of information within this region. This suggestion is also supported by the mean total fixation duration analysis which indicated that participants spent more time looking at the eye region of facial expressions indicating increased sampling of visual information from this region. Furthermore, reflexive saccades originating from the cheeks are directed more strongly towards one of the lower features (i.e., opposite cheek and mouth) than one of the upper features. Arizpe et al. (2012) showed that initial fixation location affected the target of subsequent saccades and showed that first fixations target the centre of the face, with a bias towards the location of the initial fixation so they end up close to the initial fixation location. When Scheller et al. (2012) shifted the centre of the face images to the upper and lower halves of the screen they found that reflexive saccades targeted the facial feature that was closest to the initial fixation location – in other words, when faces were presented in the upper half of the screen, initial fixations starting from the centre of the screen targeted the mouth region more compared to the eyes and vice versa. The initial saccades in both our experiments might be targeting the facial features that are spatially closest to the initially fixated location rather than being guided by expression-informative facial features.

Another interesting pattern of eye movement which emerged from the analysis of saccade paths is that reflexive saccades target facial features on the opposite side to their initial fixation. This was especially apparent for saccades from the eyes targeting the cheek on the opposite side of the face more strongly than the other lower saccade targets. This finding was in line with Arizpe et al.'s (2012) findings. Arizpe et al. (2012) showed that the fixations following the first one (fixations 2-5) tend to end at a location that is opposite to their peripheral initial fixation location. Additionally, we report additional analysis in the Appendix which provides some evidence for the suggestion that reflexive saccades target the centre of the face in line with the centre-of-gravity effect (Bindemann et al., 2009). This finding fits in with previous research that suggested that early fixations (within 0-250ms after face onset) target the geometric centre of faces (Bindemann et al., 2009). The reflexive saccades in our studies started less than 1000ms after the onset of the face stimuli, therefore it is possible that there is a mixture of early saccades that target the centre of the face and later saccades that might target expression-informative facial features among the reflexive saccades analysed. One important thing to note about the measurement of the centre-of-gravity effect in this study is that we used the nose region as the centre of the face. However, since the nose region covers a large area in the middle of the face, caution should be taken while interpreting the effect as the pure centre-of-gravity effect.

Previous research suggest that there is a perceptual bias to the left side of the face for many face-related tasks including expression recognition whereby performance of the observers is influenced more by the left side of the face (Burt & Perrett, 1997; Butler et al., 2005; Innes, Burt, Birch & Hausmann, 2015). This perceptual bias is also reflected in the eye movements of the observers. For example. Butler et al. (2005) showed that observers made more fixations to the left side of the face during a gender identification task. Additionally, for an expression recognition task similar to the one reported in this Chapter, Atkinson and Smithson (2020) found that saccades starting from the brow and the mouth targeted leftward and central features more (i.e. the mouth, brow, left eye and left cheek) compared to facial features on the right-side. In the saccade projection analyses reported in this Chapter, we report some similar findings: in Experiment 1, saccades were directed more towards the central brow and the left eye regardless of initial fixation location. Furthermore, from features on the right side (right eye and right cheek), central (mouth and brow) and left sided features were more strongly targeted compared to other features on the right side. This was especially apparent for fearful expressions. However, this potential leftwards bias was less pronounced in Experiment 2. While saccades from the right eye were still targeting leftward and central features more, saccades from the left cheek showed a bias towards the right side of the face. Despite some unequivocal finding, the strong leftwards bias reported in the literature should not be dismissed. If there is a functional difference between the left and right sides of the face, analysis reported here which collapses across initial fixation location, such as the analysis of initial fixation location on emotion recognition accuracy, might be masking the potential greater informativeness of the left side of the face (left eye and left cheek) over the right side of the face. Additionally, the analysis of the total fixation duration which treats the left eye and the right eye as a unitary ROI might be masking a leftward bias in eye movements.

We found that the latencies of the first reflexive saccades were affected by the initial fixation location on the face image for both experiments (as reported in the Appendix). Observers made quicker saccades from the initial fixation locations at the cheeks compared to all other initial fixation locations. It has been shown that latency to saccade to an emotional face (fearful and happy) presented briefly in the periphery is shorter compared to the latency to saccade to a neutral face, suggesting that saccade programming is affected by the emotional content of the face (Bannerman, Hibbard, Chalmers & Sahraie, 2012). It has also been suggested that saccade latency can be a reliable index of spatial attention orientation (Şentürk, Greenberg & Liu, 2016). These researchers showed that saccade latencies to a cued location are shorter compared to a non-cued location. It is possible, therefore, that in our studies the emotional content in the rest of the face acted as a cue for the selection of the next saccade target leading our participants to make a quicker saccade towards another, more expression-informative facial feature. Similarly, Arizpe, Kravitz, Yovel and Baker (2012) showed that the saccade latencies from a central location on the face were longer compared to peripheral fixation locations that did not correspond to any of the internal facial features. The central location on a face is suggested to contain most visual information making this the optimal fixation location for face-related tasks (Hsiao & Cottrell, 2008; Or et al., 2015; Peterson & Eckstein, 2012). Additionally, the saccade generation model by Findlay and Walker (1999) suggests that the trigger to move from a fixated position depends on the amount of cognitive and perceptual processing the currently fixated location requires. Since the initial fixation locations in both of our experiments consisted of internal features that are considered informative (i.e., the eyes, the brow and the mouth), it is possible that there was more information to be processed at initial fixation locations other than the cheeks, leading our observers to saccade away from these features much slower compared to the cheeks.

Gaze patterns specific to each expression and their relation to emotion recognition performance

In addition to the brief-fixation paradigm, our participants also carried out a free-viewing task for the same expressions where they were allowed to examine the facial images for five seconds. This way, the observers were able to sample visual information from any region of the face freely. Previous research suggests that the eye region is the most frequently visited and longest viewed facial feature, but the eye movements are still affected by the expressive content of the face (Eisenbarth & Alpers, 2011; Scheller et al., 2012; Schurgin et al., 2014). In both experiments, we found that observers spent more time fixating the eyes and the nose regions compared to the brow and mouth regions regardless of expression. We also found that total fixation duration was affected by the expression the face image displayed. In both studies, we found that observers spent more time looking at the brow region of angry faces compared to fearful and surprised faces. To the best of our knowledge, this is a novel finding, mainly because previous studies did not pick the brow region as a region of interest. Indeed, the bottom portion of our brow ROI is typically included in the eye region ROIs in those previous studies (e.g., Vaidya, Jin & Fellows, 2014; Eisenbarth & Alpers, 2011; Beaudry et al., 2014). One study that chose the nasion region as an ROI (Schurgin et al., 2014) did not find that observers spent more time fixating this region for angry faces, however; they did find that observers spent more time looking at the nasion of neutral faces when these were presented in the same block with angry faces. This might mean that the brow region is important for discriminating angry from neutral faces. In our experiments, we show that observers spend more time looking at the brow region of angry faces in the absence of neutral faces. Even though our brow ROI and Schurgin et al.'s (2014) nasion region might not be identical, they are overlapping. Our brow ROI encompasses the glabella which is above the nasion. There was also a task difference between ours and Schurgin et al.'s (2014) study. Where Schurgin et al. (2014) asked the participants to decide whether faces showed any expression, we asked our participants to explicitly choose an expression label for the faces they saw. Either one of these slight differences or a combination of them might have led to the differences between our two studies. Additionally, we found that there was no difference in the time spent looking at any of the ROIs between fearful and surprised faces in either study. Following on from the ideas of Jack et al. (2009), this finding can explain why the misclassification of fear and surprise is not alleviated in the long-presentation paradigm as opposed to the brief-fixation paradigm. In other words, using a similar visual sampling strategy for fear and surprise might have prevented the processing of the distinctive information of each expression therefore leading to a misclassification of the two regardless of initial fixation location. Observers also spent more time fixating the nose and mouth regions of disgusted faces compared to all other expressions. The wrinkles around the lower nose and the mouth region of disgusted faces are shown to be informative for the recognition of disgusted expressions (Smith et al., 2005; Smith & Merlusca, 2015). Our findings also agree with Shurgin et al. (2014) and further show that observers also sample information from the nose of the disgusted faces more compared to expressions of anger, fear and surprise. The mouth region chosen by Schurgin et al. (2014) is more specifically the upper lip and this coincides with Jack et al.'s (2014) finding that the activation of the upper lip raiser resolved the confusion of disgust as anger.

Thus, consistent with previous findings, we show that eye movements are modulated by the emotion-informative facial features. We further investigated whether there was a relationship between how long observers looked at the regions of interest and their recognition performance. Since research suggests that visual information can be extracted from a single fixation near the nose (Hsiao & Cottrell, 2008; Peterson & Eckstein, 2012; Or et al., 2015) and the location of fixations might not be perfectly linked to information extraction (Arizpe et al., 2012; Caldara et al., 2010), it is necessary to investigate whether these fixations are functionally related to emotion recognition. We only found a relationship between time spent looking at the mouth region of disgusted facial expressions and recognition accuracy. The general lack of a relationship between fixating facial features and emotion recognition performance agrees with the findings of Vaidya, Jin and Fellows (2014). Investigating the predictive value of fixating certain facial features on emotion recognition, Vaidya et al. (2014) failed to find a predictive relationship between fixations and emotion recognition performance for facial emotions expressed at their peak intensity. Vaidya et al. (2014) showed that visual analysis of all internal facial features equally benefits emotion recognition of extreme emotions better than fixating on individual features. While our finding of a lack of correlation between average total fixation duration at informative facial features and emotion recognition is in line with Vaidya et al. (2014)'s findings for extreme emotions, there is one major limitation that needs to be acknowledged. The average total fixation duration in both experiments was calculated across the whole viewing time of 5 seconds. Looking at the average reaction time for the long presentation paradigm for both Experiment 1 and 2, participants took around 1.8 seconds to respond and therefore to recognise the expression. It is possible that the eye movement patterns within this time will show a correlation with emotion recognition accuracy due to participants seeking visual information to recognise the target expression. Following response, in the remaining approximately 3 seconds, it is possible that the eye movement patterns reverted to the more general T-pattern observers adopt during face viewing and therefore might not reflect the eye movement strategy adopted to recognise an expression. On the other hand, Vaidya et al. (2014) showed that fixations can be predictive of emotion recognition performance when the expressions presented are subtle. While the presence of informative visual information in low spatial frequencies negate the necessity of foveal processing of certain facial features, it is possible that foveal processing is more instrumental for the recognition of subtle facial expressions.

Some parallels and inconsistencies between the brief fixation and long presentation paradigms need to be discussed. In the brief fixation paradigm of both experiments, initially fixating the mouth led to higher recognition accuracy for fear, surprise and disgust. However, except for Experiment 2, we did not find that mouth region was looked at for longer for fearful or surprised expressions. This goes against our expectation that observers preferentially sample informative facial features during emotion recognition. Additionally, since in the brief fixation paradigm we showed that initial fixation on the mouth reduces misclassification of fear as surprise, not preferentially sampling information from this feature in the long presentation paradigm led to this misclassification to remain consistent across all initial fixation conditions. Our findings for disgust expressions were consistent with our expectation: initial fixation on the mouth led to improved disgust accuracy and the mouth region was looked at for longer compared to other regions of interest. Nevertheless, while initial fixation on the mouth reduced misclassification of disgust as anger, this was not the case for the long presentation paradigm. Finally, for angry expressions, while we did not find that initial fixation at the brow improved emotion recognition accuracy, this region was looked at for longer for angry expressions compared to other expressions. These inconsistencies between the findings of the brief fixation and long presentation paradigms suggest that the informativeness of a facial feature might be modulated by presentation duration. While distinctive facial features such as the mouth might be informative when faces are seen at a glance, with no opportunity to integrate multiple cues from facial features with multiple fixations, when allowed to view expressions for longer, observers might prefer to sample from multiple facial features, possibly even obtaining information from distinctive features extrafoveally. However, it has to be noted that if preferential sampling of these distinctive features are abandoned, with more homogenous eye movement strategies adopted for all expressions, the misclassifications between expressions such as fear and surprise are more difficult to resolve, as indicated by our results.

Conclusions

Results from two experiments showed that initially fixating on the mouth region of facial expressions led to improved emotion recognition performance overall and led to significant contributions to the recognition of fearful, surprised and disgusted faces. While we found that initially fixating the brow improved anger accuracy compared to the cheeks (Experiment 1), we also found that initially fixating the brow impaired anger recognition (Experiment 2), possibly due to the similarity of this region in angry and disgusted expressions. Initially fixating the mouth region of fearful and disgusted expressions also reduced the confusion of these expressions as surprised and angry respectively, possibly by directing early attention to distinctive features of these pairs of expressions. We did not find a consistent contributory effect on emotion recognition of initially fixating expression-informative facial features for all the expressions included in our experiments. However, given that we have used expressions posed at their peak intensity, emotion recognition performance was almost at ceiling for most of the expressions. This might have masked the contribution of foveal processing of the respective informative features. In contrast to studies employing similar brief-fixation paradigms, we found that despite the propensity of fixation changes to upper facial features (i.e., the eyes), it is more likely that reflexive saccades target facial features that are spatially closest to the one previously fixated but there is also evidence that these saccades might be targeting the centre of the face image which is operationalised as the nose ROI in this study. Fixation patterns observed in the long-presentation paradigm agreed with findings in the literature which showed that while the eyes are the most frequently fixated facial features, fixations do accumulate on informative facial features such as brow for anger and the mouth for disgust compared to these regions in other expressions. Our results did not reveal a relationship between fixation patterns and emotion recognition performance, except for the finding that participants who spend more time fixating the mouth tended to be more accurate in recognising disgust. It is possible that the extreme nature of the presented expressions might have reduced or even eliminated the necessity of foveal processing. The next study will seek to address this limitation using subtler versions of these expressions which should eliminate the ceiling effects observed in the current experiments. We will seek to understand whether foveal processing of informative features contributes more to the recognition of subtle expressions.

Chapter 3

Contribution of foveal and extrafoveal facial features to recognition of facial expressions at varying intensities: An investigation of emotion recognition performance and gaze behaviour using forced fixation paradigm

Introduction

Facial expressions expressed at full intensities are generally recognized very well by observers. On the other hand, there is little research investigating the recognition of the more frequently encountered, lower intensity facial expressions.

There is consensus that increasing the intensity of an expression by means of 'morphing' gradually higher proportions of a prototypical expressive face into a neutral face (Hess et al., 1997) or exaggerating the metric differences of facial expressions relative to a norm (i.e., a neutral face or a composite facial expression) by caricaturing (Benson, Campbell, Harris, Frank & Tovée, 1999; Calder, Young, Rowland & Perrett, 1997; Calder et al., 2000) increases emotion recognition accuracy. However, in daily life social communications, most facial expressions humans need to decode are expressed in a subtle manner. How does emotion recognition performance decrease with decreasing emotional intensity?

Calvo, Avero, Fernández-Martín and Recio (2016) investigated the intensity at which each of the six basic emotions would be recognized above chance levels – i.e., the recognition threshold for each basic expression. The threshold was operationalised as the lowest intensity level at which the discrimination measure (non-parametric A') exceeded chance level which was 0.5. For this, they used a morphing technique to create facial expression morphs, that is, images that blend the full-intensity expression from a given individual with the emotionally neutral expression of that same individual at different relative strengths. Calvo et al. (2016) used morphed faces containing 20, 30, 40, 50, 60, 75 and 100% expressive information and asked observers to label each facial expression morph. The authors found that the recognition threshold for each of the six basic expressions was different: Happy faces were recognized significantly above chance at the lowest intensity levels (20 and 30%), anger, disgust, sadness and surprised expressions were recognized significantly above chance at 40% and fearful faces were recognized significantly above chance at 50% intensity. These authors also quantified the amount of change in the physical appearance of each level of intensity for each expression. This was achieved by software which uses the appearance of the facial features to determine the probability of an expression being present in a given face stimulus. The difference between these probability values for the neutral face and the expressive face at different intensity levels provides a measure of the physical change of a given expression from neutral at a given intensity level. They showed that the biggest change was for happy faces followed by disgusted faces, whereas the changes for angry, sad and fearful faces were similar to each other. Additionally, there was a significant difference in the amount of change from one intensity level to the next for all expressions. The authors suggest that the change in appearance of the facial expressions at different intensities can account for the recognition advantage for happy faces since this expression has both the greatest physical change *and* the lowest recognition threshold. However, since the similarity of the physical change among fear, anger and sadness does not translate into equal recognition threshold for these three expressions physical change cannot be the only explanation behind their differing threshold values. On the other hand, it must be noted that the linear increase in physical change is accompanied by a linear increase in emotion recognition accuracy, therefore, featural appearance of the expressions should not be disregarded as a contributory factor to emotion recognition.

Additional support for this statement comes from a study by Matsumoto and Hwang (2014) where they show that the presence and intensity of the prototypical AUs contributes to the recognition of the target expression. For this study, Matsumoto and Hwang (2014) defined subtle expressions as expressions that involve low-intensity/partial appearance changes in the face and they were interested in investigating whether these subtle expressions have signal values as emotions. In this study, in addition to the subtle versions of prototypical expressions, they also investigated the signal value of variants of prototypical expressions. The variants were created with a combination of critical AUs and acceptable AUs which do not add signal value to the expressions. An important contribution of the variant expressions in this study was showing what facial AUs contributed to above chance level recognition of facial expressions even when these were not in a full prototypical face format. According to the results, variants of anger containing the brow lowerer, upper lid raiser and lid tightener were judged as anger and the presence of the brow lowerer and the lid tightener was correlated with emotion recognition accuracy individually. Fearful expressions containing the lip stretcher or upper lid raiser were judged as fear indicating that the eyes or the mouth region can be informative for fear. The presence of the inner and outer brow raiser, brow lowerer and upper lid raiser were individually correlated with fear recognition, whereas, the presence of the lip stretcher was not. It should be noted, however, that these action units were shown at higher intensities compared to the other expression AUs. Expressions of sadness were judged as sadness higher than chance when only the inner brow raiser and brow lowerer or only lip corner depressor was present, and presence of the lip corner depressor was correlated with accuracy. Finally, surprised expressions containing the inner and outer brow raiser along with the upper lid raiser were judged as surprised and presence of the inner and outer brow raisers was correlated with recognition accuracy. These findings are more or less in line with previous findings about informative features in facial expressions and they extend these to show that they are still informative when they are present in a full face at lower intensities.

The aforementioned studies show that static facial expressions of lower intensities do still signal the six basic emotions through their differential informative facial features. They, however, cannot determine whether observers make use of these facial features when they are looking at faces expressing emotions at lower intensities. When expressions are at lower intensities, visual information from a single feature becomes more ambiguous (Guo, 2012); therefore, this might

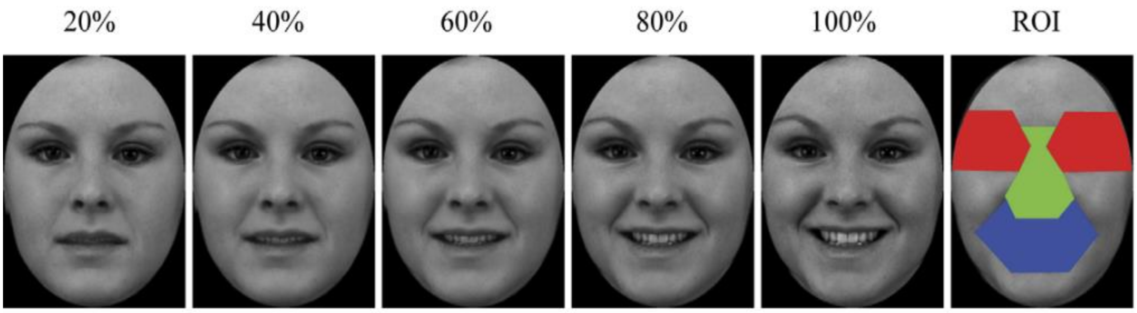

Figure 24: Image taken from Guo (2012). The facial expressions of different intensities used in Guo (2012) are illustrated along with the ROIs chosen where the red represents the eyes (including the eyebrows), green represents the nose and the blue represents the mouth.

lead to changes in information sampling strategies. Guo (2012) suggested that the ambiguity of subtle facial expressions might lead to observers making use of several features to make their emotion recognition judgement: therefore, leading them to use a holistic strategy. The participants in Guo's (2012) study were shown all six basic expressions at intensities varying from 20 to 100 in 20% steps and the viewing time and number of fixations to the eyes, nose and mouth ROIs (Figure 24) were recorded. It was found that emotion recognition accuracy increased, and reaction time decreased, with increasing expression intensity as expected and in line with previous research. Increasing intensity led to a decrease in viewing time and total number of fixations up to 60% intensity after which there was no difference in these measures. Observers directed the highest number of fixations and spent the longest viewing time at the eyes followed by nose and finally mouth. However, the distribution of viewing time and number of fixations was affected by the expression being viewed. Regardless of intensity, the mouth region of the happy and surprised faces attracted the highest proportion of fixations and longest viewing time. The nose region of sad and disgusted faces attracted the most fixations. The eye region attracted the highest viewing time and most frequent fixations for fearful, angry and surprised faces. However, since there was no analysis of any relationship between the amount of time spent fixating emotion-informative features and emotion recognition performance, it is difficult to conclude whether the increased time spent fixating a certain facial feature contributed to the accuracy of facial expression recognition. Additionally, the analyses of total viewing time and total number of fixations were conducted on all trials regardless of accuracy.

Combined with the finding of low recognition accuracy at lower intensities (especially for disgust and fear) and high misclassification between some expressions, it is possible that the holistic gaze strategy observed in this experiment is not optimal for low intensity facial expressions since it does not allow the extraction of distinctive visual information from the observed faces.

Vaidya, Jin and Fellows (2014) further investigated the importance of fixating informative facial features of low intensity facial expressions. They suggested that foveal processing of facial features might benefit recognition of lower intensity facial expressions more than full-blown prototypical expressions. According to the authors, it is possible for informative features in full-blown expressions to be accessible in the parafoveal region at lower frequencies making information sampling using multiple fixations redundant. On the other hand, when an emotion is expressed in a more subtle manner, there is less change in facial appearance from a neutral baseline compared to full-blown expressions. This subtle change in visual information in lower intensity expressions might not be accessible in the parafoveal regions and therefore require processing by the fovea to obtain high spatial frequency information. Vaidya et al. (2014) asked participants to perform two tasks: The first task was divided into 4 blocks of trials where participants saw neutral faces and subtle and extreme versions of happy, disgusted, fearful and surprised expressions. For each block, participants were given a target emotion out of the four and asked to judge the presented expressions for the target emotion. For example, on one block, participants rated happy, disgusted, fearful and surprised expressions for the target emotion of 'happiness' on a scale of 1 to 10. For the second task, they were asked to label these expressions as happy, surprised, disgusted or fearful. Faces were shown for 2 seconds in both tasks and the eye movements of the observers were recorded. The subtle expressions were chosen from 20 – 40% morphs that were of similar difficulty across expressions and actors. As expected, observers were better at recognizing extreme expressions compared to subtle expressions. However, there was a lack of difference between extreme and subtle intensities for fearful expressions. An ideal observer analysis was carried out on both the extreme and subtle expressions to determine the face region that helps distinguish expressive faces from neutral ones. For subtle expressions, the analysis showed that the eye region contained the most diagnostic information across high and low frequencies. The mouth region on the other hand only contained diagnostic information in higher frequencies. This pattern was similar across facial expressions. Extreme expressions showed a more diverse range of diagnostic regions. While fearful expressions showed similar information distribution to subtle expressions, there seemed to be more low frequency information at the nose and cheeks for extreme happy and disgusted faces. Finally, for the extreme surprised faces, eyes only had distinctive information at the highest frequencies and almost no information at the nose and cheeks.

Although informative, it should be noted that the ideal observer analysis in Vaidya et al.'s (2014) study searched for information that distinguished expressive faces from neutral faces.

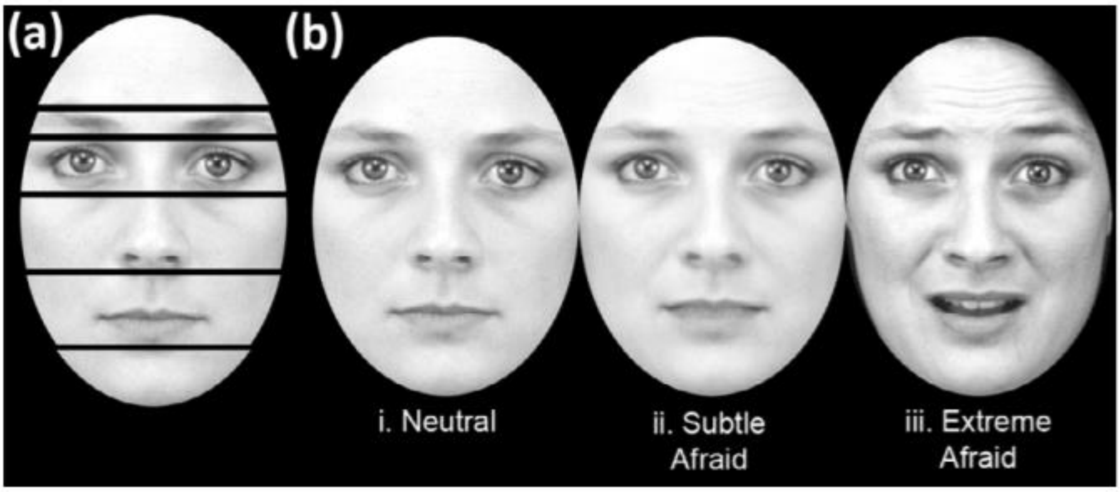

Figure 25; Image taken from Vaidya et al. (2014). (a) Denotes the regions of interest chosen by Vaidya et al. (2014) for the analysis of number of fixations and frequency of first and second fixations.

Since Smith and Merlusca (2014) showed that informative face regions change depending on which expressions the target is identified against, the informative regions of subtle expressions might differ when they are being distinguished from other expressions as opposed to neutral faces.

Vaidya et al.'s (2014) analysis of number of fixations to ROIs (which can be seen in Figure 25) showed that observers made more fixations to the mouth region of subtle happy and disgusted faces compared to neutral faces. The analysis of the frequency of first and second fixations to ROIs showed that first fixations were directed more to the mouth of extreme happy and surprised faces and the second fixations were directed to the nose of the extreme disgusted stimuli and to the mouth of subtle disgusted and happy stimuli. (All comparisons were made to the neutral faces.) The implication from this analysis is that the mouth region is a sought-out region for the subtle disgusted faces, which indicates the informativeness of this region for this expression. Additionally, in experiments reported in Chapter 2, we also found that initial fixation on the mouth in the brief-fixation paradigm improved disgust recognition accuracy and there was a positive correlation between total fixation duration within the mouth and disgust accuracy in the long-presentation paradigm. Finally, the researchers were interested in investigating whether number of fixations in an ROI had a cumulative effect in predicting participants' detection scores. They used the number of fixations to eyes, nose and mouth regions (see Figure 25(a)) in the emotion-rating task trials where the target emotion and the expression were congruent, to model a linear relationship between number of fixations to ROIs and detection scores for subtle expressions (it is worth noting here that Vaidya et al. (2014) also used the dwell time within the ROIs to model this relationship and found similar results.). The model, with positive weights given to fixations to ROIs, was able to predict the detection scores for subtle fear and disgust and showed a trend towards predicting surprise but failed to predict scores for subtle happy and extreme expressions except extreme fear. For subtle fear, the highest weight was given to the eyes followed by mouth and nose. For extreme fear on the other hand, less weight was given to the eyes compared to mouth and nose. The greater weight given to the mouth compared to the eyes of extreme fearful expressions also agrees with our finding of increased emotion recognition accuracy for fear when initially fixating the mouth region. The difference in the weights given to the eyes and the mouth region between subtle and extreme fear expression, on the other hand, suggests that there is a difference in informative features for extreme and subtle fear expressions. For subtle surprised and disgusted expressions, eye fixations were given the highest weight with less influence given to mouth and nose. The ability of the linear model to predict recognition scores for most of the subtle expressions and the failure of it to do so for the extreme expressions lends support to Vaidya et al.'s (2014) suggestion that the presence of lower spatial frequency information in parafoveal regions makes visual sampling by multiple fixations redundant and fixations are more important for detection of subtle expressions.

Altogether, these findings indicate that lower intensity expressions signal emotions through facial regions that are informative to the recognition of the expression and observers do sample information from these features through fixations. These fixations are more frequent, and they might aid recognition of subtle expressions to a higher extent than the recognition of full-blown expressions. These previous studies allowed the observers to freely view the faces, which allowed them to make multiple fixations on several facial features. Therefore, the question of which particular facial feature/region is leading to more accurate identification of the target low intensity expression remains to be answered. To answer this question, we employed the brief-fixation paradigm (Gamer & Büchel, 2009) to present angry, fearful, surprised and disgusted expressions displayed at lower intensities at informative or non-informative initial fixation locations and measure the recognition accuracy of our observers. We hypothesized that increasing expression intensity will increase emotion recognition. We also expected to see an interaction between the factors: Initially fixating on an informative facial feature will improve emotion recognition accuracy for lower intensity expressions while this effect might be non-existent in the higher intensity expressions. For example, while initially fixating on the brow region might not contribute to emotion recognition at 70% intensity, we expect to find this effect at lower intensity expressions. We expected that lowering the intensity of the expression will lower the extrafoveal expressive information; therefore, making the effect of foveating the informative facial feature more critical in resolving the recognition task. Regarding the systematic misclassifications of facial expressions reported in the previous chapter, we still expected to observe these misclassifications, but on a larger scale for lower intensity expressions since the distinguishing features will be less discernible. When the distinguishing facial feature, e.g., the mouth, is initially fixated, we expect these misclassifications to decline. Furthermore, in previous studies that have employed lower intensity faces (especially at 10-30% intensities), observers were not given an option to label these images as neutral even though those faces might not be perceived as transmitting any emotional signals. Therefore, in this study, we gave the observers the option to label an image as neutral if they believed it showed no expression, which prevents us from labelling a feature informative despite a lack of signal. Secondly, this allowed us to investigate whether fixating on an informative feature also aids the observers to detect a low intensity facial expression. Specifically, we expected a gradual decline in the use of the neutral response option with increasing expression intensity and fewer neutral responses when the initial fixation falls on the informative feature compared to the non-informative feature.

In addition to emotion recognition accuracy, we are also interested in investigating whether reflexive saccades would seek out informative features and whether this was dependent on expression intensity. As discussed in Chapter 2, Gamer and Buchel (2009) showed that when briefly presented with a facial expression with initial fixation at the eyes or the mouth, observers made a higher proportion of reflexive saccades upwards from the mouth region compared to downwards from the eyes. Furthermore, this tendency was stronger for fearful expressions compared to happy and neutral faces. In the previous chapter, we replicated the finding of higher percentage of saccades upwards from the mouth and lower features (mouth and cheeks). However, we failed to find an effect of expression on saccade direction for full-blown, prototypical facial expressions. Furthermore, we provided some evidence to suggest that the initial saccades targeted the centre of the face (i.e., the nose region) in line with the centre-of-gravity effect (Bindemann et al., 2009). Due to the nature of the expressions used in the previous experiment, saccading to an informative facial feature might not have been necessary since the required visual information could have already been accessed extrafoveally from the fixated location. As discussed previously, foveal processing of informative facial features might be more crucial for the recognition of lower intensity facial expressions (Vaidya et al., 2014). Therefore, it is possible that reflexive saccades would target expression-informative facial features in order to process them foveally. Lending support for this is Guo et al.'s (2012) finding of a higher number of fixations directed to lower intensity expressions. This finding shows that observers inspect lower intensity facial expressions more compared to prototypical expressions. Therefore, we expect that initially fixating on a non-informative feature will lead to reflexive saccades seeking out informative features in lower intensity expressions more compared to the higher intensity expressions.

Methods

Participants

A total of 41 participants took part in the experiment (female=35, male=6; mean age= 20.88 years). Sample size was selected to be similar to the sample size of studies with similar design. All participants had normal or corrected-to-normal vision. All participants gave written consent to take part in the experiment and were rewarded participant pool credit for their participation. The study was approved by the Durham University Psychology Department Ethics Sub-committee.

Stimuli

12 facial identities (6 males, 6 females) were chosen from the Radbound Face Database (Langner, Dotsch, Bijlstra, Wigboldus, Hawk & van Knippenberg, 2010). All faces were of Caucasian adults with full frontal pose and gaze. The images were cropped from their original size to 384 (width) × 576 (height) pixels, so that the face took up more of the image than in the original image set. We used the same subset of expressions used in Experiment 2 reported in Chapter 2 – anger, fear, surprise and disgust. The reason behind choosing these expressions is explained in more detail in Chapter 2. We use the same subset of expressions here in order to draw links between the results of this study and the results reported in Chapter 2. The same initial fixation locations as Experiments 1 and 2 were also used. An example of the locations of these initial fixations can be seen superimposed on different intensities of an angry expressions in Figure 26. The neutral pose and angry, fearful, surprised and disgusted poses from each facial identity were utilised to create the morphed stimuli. For each of the four expressions of each facial identity, we used the Psychomorph (Chen & Tiddeman, 2010)

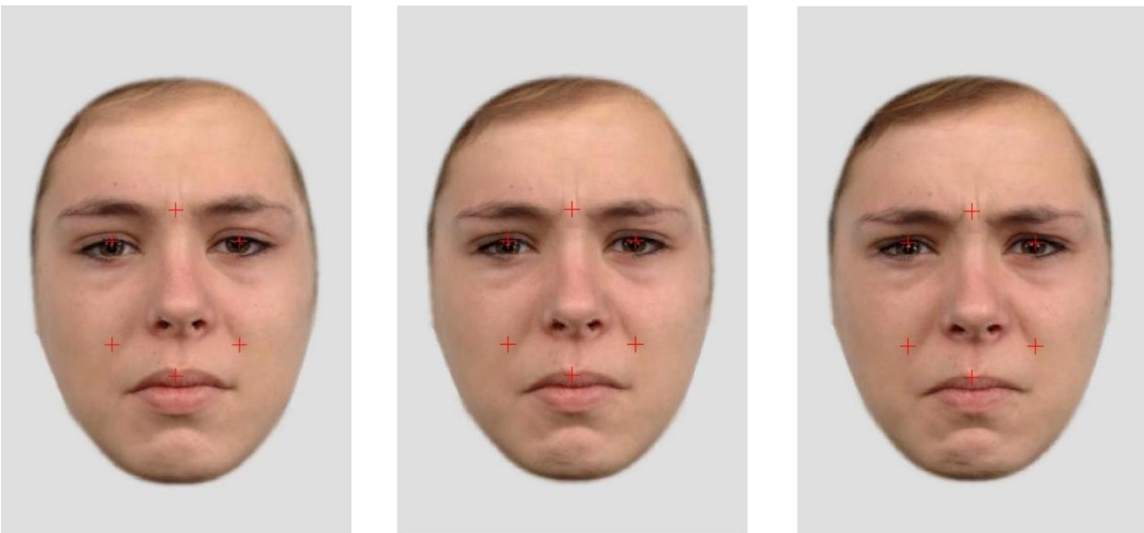

Figure 26. Example of the initial fixation locations superimposed onto an angry expression at 30%, 50% and 70% intensities (from left to right).

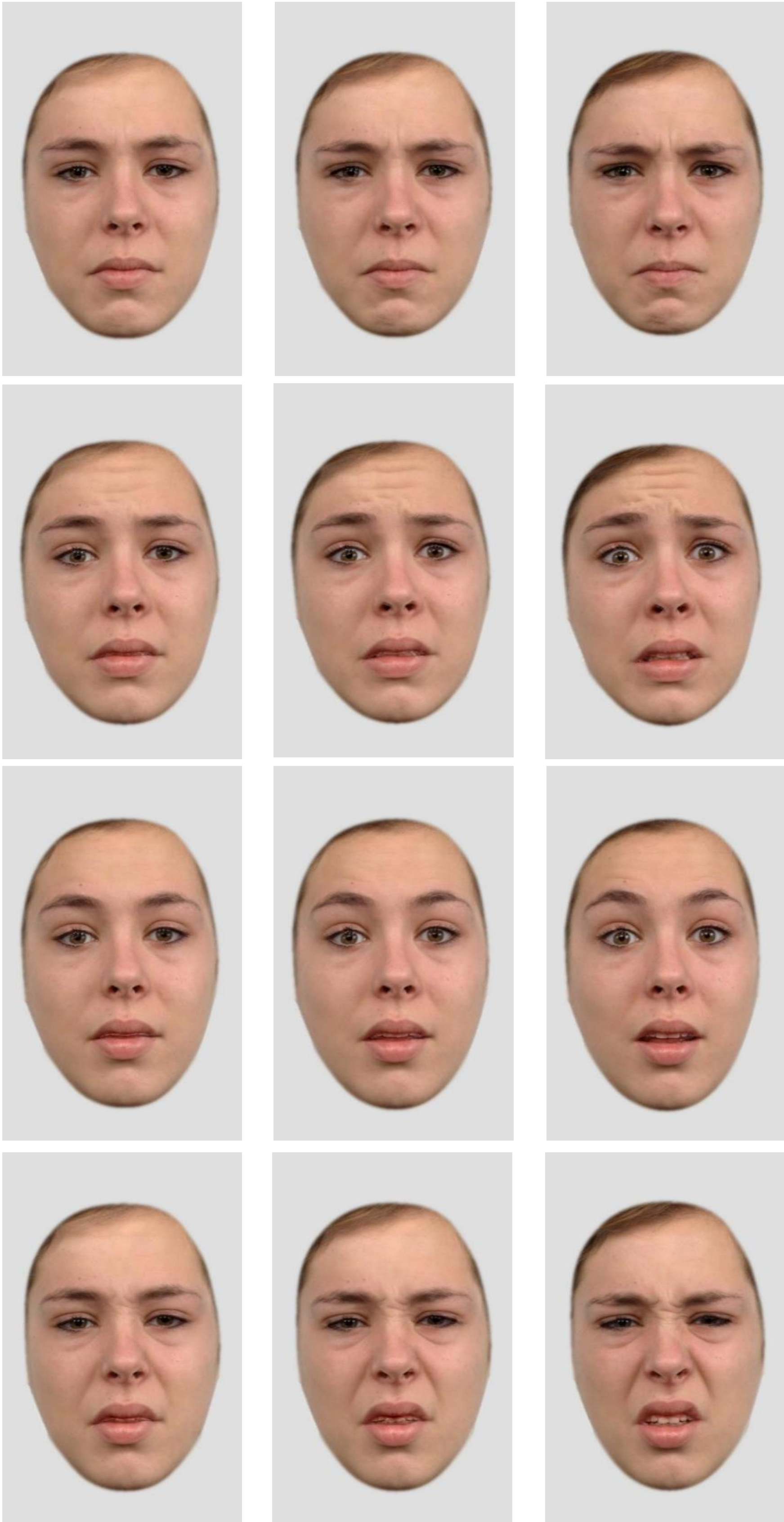

Figure 27. Examples of angry, fearful, surprised and disgusted expressions (from top to bottom) at 30%, 50% and 70% intensities (from left to right) from the Radbound Face Database used for Experiment 3.

software to create 3 levels of intensity by morphing the expressive face of each facial identity into the neutral pose by the same identity. Each expressive face was morphed from 0% (neutral) to 100% (full intensity) by 10% increments but only 30%, 50% and 70% morphs of each facial identity and each expression were used for the experiment. Figure 27 shows all expressions at the three different intensity levels posed by one of the face identities from the Radbound Face Database. The lowest intensity started from 30% since Wells et al. (2016) showed that expressions of 10% intensity was identified as neutral 80% of the time and 20% intensities were visually very similar to 10% intensities. We chose not to include the 100% (prototypical expressions) since Guo (2012) showed that there was no change in fixation patterns for expressions beyond 60% intensity. Furthermore, we already investigated the same research questions on the prototypical expressions in experiments reported in Chapter 2. Additionally, Schurgin et al., (2014) showed that fixation patterns change and reflect a sampling strategy where the observer is looking for informative features for expressions evolving from 20% to 40% to 60% intensity. Therefore, it is possible to see differential effects of initially fixating informative facial features on emotion recognition for our chosen intensity intervals. The external features of the faces were then masked by applying a uniform grey filter around the outside of the delineated face oval. Due to the morphing of expressive faces into neutral poses led to minor size differences between the face ovals. This did not affect the configuration of the inner facial features; therefore, the intended facial feature still fell on the fovea at the fixated location.

Design

There were three manipulations in this experiment: the Expression of the face (angry, fearful, surprised and disgusted), the Intensity of the expression (30%, 50% and 70%) and the Initial Fixation location (left and right eye, brow, left and right cheek, mouth). There were 4 blocks of 144 trials. In each block, each expression was shown at each initial fixation location and at each intensity three times. For the initial fixation locations that had a left and right side, each expression at each intensity was only presented at either the left or the right side in one block. In the following block, the same expressions at the same intensities were presented at the facial feature on the opposite side. This was counterbalanced for the expressions so that while 2 out of the 4 expressions were presented at the left initial fixation locations, the other two were presented at the right initial fixation location within each block. The order of face presentation within each block was randomised for each participant.

Procedure

The procedure was very similar to that for the brief-fixation experiments reported in the previous chapter and are reported here in full for completeness. The experiment was executed and controlled using the Matlab® programming language with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). To control stimulus presentation and to measure gaze behaviour, we used an EyeLink 1000 desktop-mounted eye-tracker (SR Research, Mississauga, ON, Canada), sampling at 500 Hz. Each stimulus block was started with the default nine-point calibration and validation sequences. Recording was binocular but only data from the left eye was analysed since data from the left eye started the trial in a gaze-contingent manner. Default criteria for fixations, blinks, and saccades implemented in the EyeLink system were used. Each trial started with a fixation cross located at one of 25 possible locations on the screen. This was to make both the exact screen location of the fixation cross and the to-be-fixated facial feature unpredictable. These 25 possible locations for the fixation cross were at 0, 25, or 50 pixels left or right and up or down from the center of the screen. These fixation-cross positions were randomly ordered across trials. Faces were presented in a gaze-contingent manner on each trial: The participants needed to fixate within 0.6 degrees of visual angle of the fixation cross for 6 consecutive eye-tracking samples following which a face showing one of the four target expressions was presented. The location of the fixation cross corresponded to one of the informative (eyes, brow or mouth) or non-informative (cheeks) initial fixation locations for each of the presented faces. The face was presented for 82.4 ms (7 monitor refreshes) on a monitor with an 85Hz refresh rate. Following the face presentation, the participant pressed a key on a QWERTY keyboard to indicate their answer. The row of number keys near the top of the keyboard was used, with 4 for anger, 5 for fear, 8 for surprise, 9 for disgust and 0 for neutral. The keys were labelled A, F and Su, D and N from left to right and the order of these keys remained the same for each participant. Participants pressed the A and F keys with the left and the Su, D and N keys with the right hand. This configuration was chosen to optimize the reach of the participants to the keyboard from either side of the chinrest. The 'Neutral' response option was added so as not to force participants to attach an expression label to a low intensity expression if they did not perceive any expression in the presented face image. Participants were told to press the N response key if they thought the face showed no expression. Participants were asked to memorize the keys so as not to look down towards the keyboard during the experiment; all participants could memorize the keys. A valid response needed to be registered for the next trial to begin.

Data Analysis

Behavioural Data Analysis

Trials with reaction times less than 200msec were removed from the dataset (0.03% of all trials). A 4×4×3 RM ANOVA using expression, initial fixation and intensity as factors was conducted on the remaining data. Significant two-way interactions were followed up with one-way repeated measures ANOVAs. Significant main effects were followed up with Bonferroni-corrected pairwise comparisons. Where the assumptions of equal variances were not met, Greenhouse-Geisser corrected test values are reported.

Eye-movement Data Analysis

Reflexive saccade selection criteria

As for the brief-fixation experiments reported in the previous chapter, reflexive saccades were defined as saccades that were triggered by onset of a face image and executed following face offset. Accordingly, reflexive saccades used for the reported analyses were chosen as those that happened within the 82.4ms to 1082.4ms window after face onset (similar to e.g., Gamer & Buchel, 2009; Scheller et al., 2012). In other words, the reflexive saccades used were the ones that were executed within 1000ms after face offset. All reflexive saccades that had amplitudes smaller than 0.5 degrees were disregarded. Finally, of all the saccades that complied with these criteria, only the first saccade was used. After this data reduction, participants who had reflexive first saccades on fewer than 20% of the total trials per block (i.e. 20% of 144) in any one of the 4 blocks were removed from further analysis. All saccade direction related analyses were carried out on this set of data.

The saccade latency of the reflexive saccades was calculated as the time elapsed between face offset and the start of the saccade.

Saccade Direction and Saccade Path Analysis

As with the brief-fixation experiments reported in the previous chapter, in this study, we also aimed to examine the direction (upwards vs downwards) and path of reflexive saccades to investigate what facial features they are targeting. For the purpose of this analyses, we only used the reflexive saccades validated according to the saccade selection criteria outlined above. Saccade direction was measured as the percentage of reflexive saccades going upwards or downwards from the fixated feature. This was calculated in the same way as the previously reported analysis: Saccades downwards from the eyes and upwards from the mouth were measured as a percentage of total saccades initiated from the eyes and mouth combined. Saccades going upwards from the lower features and downwards from the upper features are measured as a percentage of total saccades from the initial fixation locations at the respective face location (ie. upper: eyes and brow; lower: mouth and cheeks). To estimate the paths of the reflexive saccades, six vectors were plotted from the starting coordinates of each saccade to the coordinates of the six possible initial fixation locations. These make up the saccade path vectors. Then, the dot products of the reflexive saccade vector and the six possible saccade path vectors were calculated and normalised to the magnitude of the saccade path vectors. This measure represents the similarity between the reflexive saccade path and the possible saccade path vectors. Higher normalised dot product values represent greater similarity between the paths of the reflexive saccades and the each of the six possible saccade paths.

Results

Emotion Recognition Accuracy

We were interested in investigating whether initially fixating on an informative facial feature would influence the recognition accuracy for facial expressions and whether this was dependent on expression intensity. The average reaction time for this task was 1.48s (SD=1.22). The average reaction time for each condition and a visualization of reaction time distribution can be seen in the Appendix, Table A9 and Figure A9. We calculated the unbiased hit rates for each of the 48 conditions and conducted a 4x3x4 repeated measures ANOVA using expression, intensity and initial fixation as factors. We hypothesised that initially fixating on an emotion-informative feature would increase recognition accuracy for that emotion compared to a non-informative feature for a lower intensity expression where this effect might not be present for the same expression at higher intensities. Emotion recognition performance across experimental conditions is summarised in Figure 28.

There was a main effect of expression ($F_{(1.98, 79.53)}$= 9.61, p< .001, η_p^2=.19). Surprised expressions were better recognized compared to angry (p=.001) and fearful expressions (p<.001) and disgusted expressions were better recognised compared to angry expressions (p=.002). The main effect of intensity was significant ($F_{(1.69, 67.55)}$= 298.58, p< .001, η_p^2=.88). As expected, recognition accuracy at 30% intensity was the lowest, 50% intensity was higher compared to 30% (p<.001) and 70% intensity was higher compared to both 30 and 50% intensity levels (both ps< .001). There was also a main effect of initial fixation ($F_{(2.57, 102.80)}$= 9.29, p< .001, η_p^2=.19). Initially fixating on the mouth led to the highest accuracy compared to all other initial fixation locations (mouth vs eyes: p=.004, mouth vs brow: p<.001 and mouth vs cheek: p= .007).

The interaction between expression and intensity was significant ($F_{(3.57, 142.59)}$= 16.73, p< .001, η_p^2= .30). Four separate one-way ANOVAs were conducted on the unbiased hit rates across different intensity levels for each expression to investigate the interaction further. There was a main effect of intensity for all the expressions (anger: $F_{(1.58, 63.27)}$= 129.88, p<.001, η_p^2=.77; fear: $F_{(1.66, 66.54)}$= 117.48, p< .001, η_p^2=.75; surprise: $F_{(1.60, 64.09)}$= 173.19, p< .001, η_p^2= .81; disgust: $F_{(1.69, 67.81)}$= 39.31, p< .001, η_p^2=.50). Anger, fear and surprised expressions followed the same pattern as the main effect of intensity where 30% intensity expressions were recognised worst followed by 50% and 70% expressions were recognised the best. However, disgusted expressions showed a different pattern: recognition accuracy was worse for 30% intensity expressions compared to both 50% and 70% (both ps< .001) but there was no difference between 50% and 70% intensities (p=1.00).

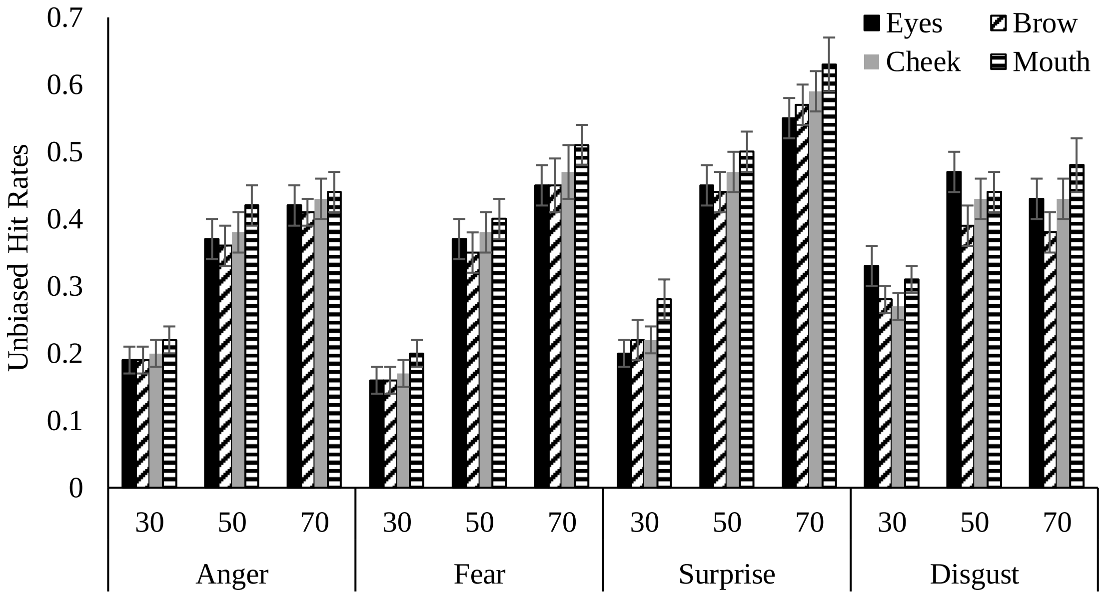

Figure 28; Mean unbiased hit rates at each intensity and initial fixation location for each expression. Error bars represent standard error of the mean.

Although the interaction between emotion and initial fixation only approached significance ($F_{(9, 360)}$= 1.76, p= .08, η_p^2=.04), given our hypotheses and in order to provide a more direct comparison with the findings of the experiments reported in the previous chapter, we followed-up this interaction with four separate one-way ANOVAs to compare the unbiased hit rates across initial fixation locations for each emotion separately. For angry faces, there was no effect of initial fixation ($F_{(2.54, 101.60)}$= 2.18, p= .11, η_p^2=.05). There was a main effect of initial fixation for fearful faces ($F_{(3, 120)}$= 3.49, p= .02, η_p^2=.08). Pairwise comparisons indicated that fear recognition accuracy was higher when initially fixating the mouth compared to when initially fixating the brow (p=.04). The significant main effect of initial fixation location for surprised faces ($F_{(3, 120)}$= 6.77, p< .001, η_p^2=.15) showed that initially fixating on the mouth led to higher recognition accuracy compared to the eyes (p<.001), brow (p=.01) and the cheeks (p=.04). Finally, for disgusted faces, the main effect of initial fixation ($F_{(3, 120)}$= 5.43, p= .002, η_p^2=12) showed that initially fixating the eye region led to higher accuracy compared to the brow (p=.01) and initially fixating on the mouth region led to higher accuracy compared to brow (p=.01).

There was no interaction between initial fixation and intensity ($F_{(6,240)}$= 0.59, p= .74, η_p^2= .015). Finally, contrary to our expectations, the three-way interaction also did not reach significance ($F_{(10.26, 410.42)}$= 0.60, p= .90, η_p^2= .015) indicating that the effect of fixation on emotion-informative features for facial expression recognition is not different across intensities.

Analysis of Neutral Responses

We were also interested whether initially fixating on an informative feature reduces the frequency with which observers responded with neutral to lower intensity facial expressions. We expected the frequency of neutral responses to decline with increasing intensity. Furthermore, we expected significantly fewer neutral responses when the initially fixated

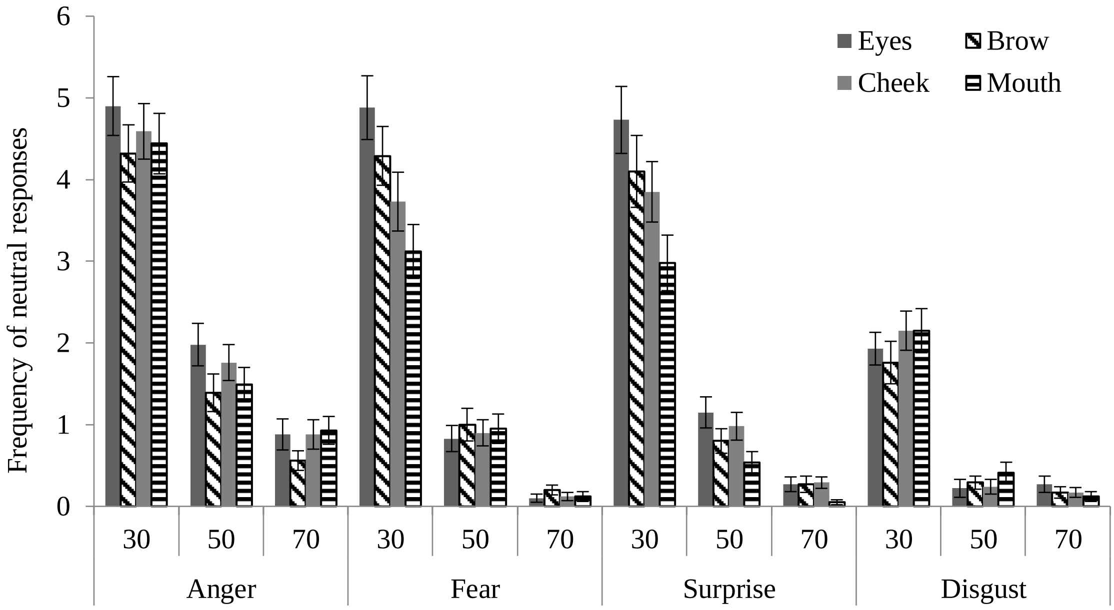

Figure 29; Average number of neutral responses given for each expression of different intensities at each initial fixation location. Error bars represent the standard error of the mean.

feature is informative compared to when it is non-informative. These data are summarised in Figure 29. We carried out a 4×4×3 RM ANOVA on the frequency of neutral responses using emotion, initial fixation and intensity as factors.

There was a main effect of emotion ($F_{(3, 120)}$= 56.30, p< .001, η_p^2=.59). The frequency of neutral responses was significantly higher for angry faces compared to all other expressions (all ps< .001) and lowest for disgusted faces (all ps <.001). The main effect of initial fixation ($F_{(2.28, 91.37)}$= 7.40, p= .001, η_p^2=.16) indicated that neutral responses were significantly more frequent when observers were initially fixating the eye region compared to the brow (p= .006) and mouth

(p= .003) regions. The frequency of neutral responses was lower when initially fixating the mouth region compared to the eyes (p=.003) and the cheek (p=.05). Finally, there was a main effect of intensity ($F_{(1.11, 44.31)}$= 214.93, p< .001, η_p^2=.84): Frequency of the neutral responses was higher for expressions presented at 30% intensity compared to 50% and 70% intensities (both ps< .001) and higher for 50% intensity expressions compared to 70% intensity (p< .001). Frequency of the neutral responses was lower for 70% intensity expressions compared to 30% and 50% intensity expressions (both ps< .001). There were significant interactions between emotion and initial fixation ($F_{(9, 360)}$= 4.25, p< .001, η_p^2=.10), emotion and intensity ($F_{(4.30, 171.81)}$= 30.49, p< .001, η_p^2=43) and initial fixation and intensity ($F_{(3.60, 143.92)}$= 5.96, p< .001, η_p^2= .13). These interactions will be discussed in the light of a significant three-way interaction ($F_{(9.03, 361.08)}$= 3.14, p< .001, η_p^2= .07).

To follow up the significant three-way interaction, two-way repeated measures ANOVAs with emotion and initial fixation were conducted separately for each intensity. When there was a two-way interaction between emotion and initial fixation, the interaction was followed up with one-way repeated measures ANOVAs using initial fixation as a factor for each expression separately. The analysis for 70% intensity is based on 39 of the 41 participants since 2 of the participants did not respond to any of the 70% expressions with the neutral option.

For 30% intensity, there was a main effect of expression ($F_{(3,120)}$= 55.10, p< .001, η_p^2=.58). Neutral responses were least frequent for disgusted expressions compared to all other expressions (all ps<.001). Neutral responses to angry faces were more frequent compared to fearful (p= .04) and disgusted expressions (p< .001). The main effect of initial fixation was also significant ($F_{(2.36, 94.56)}$= 8.59, p< .001, η_p^2= .18) which showed that fixating on the eyes led to the highest frequency of neutral responses (eyes vs brow: p=.02; eyes vs cheek: p= .01; eyes vs mouth: p< .001). The emotion and initial fixation interaction was also significant ($F_{(9, 360)}$= 4.18, p< .001, η_p^2=.10). There was no main effect of initial fixation for anger ($F_{(3,120)}$= 1.19, p=.32, η_p^2= .029) or disgust ($F_{(3,117)}$= 1.34, p=.27). There was a main effect of initial fixation for fear ($F_{(2.48, 99.03)}$= 10.06, p<.001, η_p^2=.20) and surprise ($F_{(2.39, 95.48)}$= 7.93, p<.001, η_p^2= .17). For fearful faces, initially fixating the mouth region led to fewer neutral responses compared to the eyes (p=.001) and brow (p=.007). For surprised expressions, fixating on the mouth region led to fewer neutral responses compared to the eyes (p< .001) and the cheeks (p= .04). The results for fearful and surprised expressions support our hypothesis that initially fixating an emotion-informative facial feature, the mouth in this case, would reduce the number of neutral responses for these expressions. However, initially fixating the emotion-informative facial feature for anger and disgust (the brow and the mouth, respectively) did not reduce the number of neutral responses for these expressions.

For 50% intensity, there was a main effect of emotion ($F_{(3,120)}$= 28.06, p<.001, η_p^2=.41) indicating that there were fewer neutral responses for disgusted expressions compared to all other expressions (all ps< .001) and more neutral responses for angry expressions compared to all others (all ps< .001 except anger vs surprise where p= .001). There was no main effect of initial fixation ($F_{(2.57, 102.74)}$= 1.65, p=.18, η_p^2= .040). The interaction reached significance ($F_{(6.55, 261.83)}$= 2.54, p= .02, η_p^2=.06). There was a near significant main effect of initial fixation for anger ($F_{(3,120)}$= 2.58, p=.06, η_p^2=.06). Pairwise comparisons revealed no difference between any fixation locations, however, numerical inspection of the data revealed that there were fewer neutral responses made to angry faces when they were presented with the brow at fixation. Planned comparisons comparing the neutral responses at the brow region to other initial fixation locations, again, revealed an almost significant decrease in neutral responses at the brow region for anger compared to the eyes (Z= -2.04, p=.02) and the cheeks (Z= -1.70, p= .05). While not significant, this trend is in the direction suggested by our hypothesis that initially fixating the emotion-informative brow will lead to less neutral responses for anger. The significant main effect of initial fixation for surprise ($F_{(3,120)}$= 6.03, p=.001, η_p^2=.13) showed that there were significantly fewer neutral responses when initially fixating the mouth compared to the eyes (p= .002) and the cheeks (p= .01). There was no effect of initial fixation for fear ($F_{(2.54, 101.77)}$= 0.27, p=.82, η_p^2= .007) or disgust ($F_{(3,120)}$= 0.97, p=.41, η_p^2= .024). Similar to 30% intensity expressions, initially fixating the emotion-informative mouth region for fearful expressions reduced neutral responses to fear. Contrary to out hypothesis and results from the 30% intensity expressions, there was no effect of initially fixating the emotion-informative facial feature for surprise.

For 70% intensity, there was a main effect of emotion ($F_{(1.58, 60.20)}$= 31.35, p<.001, η_p^2=.45) showing that neutral responses were most frequently made in response to angry expressions compared to all others (all ps< .001). There was no main effect of initial fixation ($F_{(3,120)}$= 1.03, p=.38, η_p^2= .026). Contrary to our hypothesis, there was no emotion and initial fixation interaction ($F_{(4.85, 184.21)}$= 1.90, p=.099, η_p^2= .048). Despite the lack of interaction, it is worth noting that there were considerably fewer neutral responses to angry expressions when initially fixating the brow region (M_{eyes}= .82, M_{brow}=.53, M_{cheek}=.92, M_{mouth}=.95) which is a trend in the hypothesised direction.

Confusion Matrices

We also investigated whether there was any systematic misclassification of the target expressions. Since we have used the same expressions as Experiment 2 reported in Chapter 2, we expected to see a similar pattern of confusions: fear confused as surprise and disgust confused as anger. We expected these misclassifications to decrease with increased intensity

Table 5. Percentage of each emotion response option chosen for each target expression presented at 30% intensity.

Displayed expression		**Response Given/%**				
	Initial Fixation	**Anger**	**Fear**	**Surprise**	**Disgust**	**Neutral**
Anger	Eyes	**35.23**	8.55	2.85	12.42	40.94
	Brow	**36.86**	7.94	4.07	15.07	36.05
	Cheeks	**36.25**	7.74	3.87	13.85	38.29
	Mouth	**34.01**	10.18	1.83	16.90	37.07
Fear	Eyes	6.53	**28.78**	15.31	8.57	40.82
	Brow	7.54	**29.53**	18.33	8.76	35.85
	Cheeks	8.15	**30.96**	15.27	14.46	31.16
	Mouth	6.53	**36.33**	18.16	12.86	26.12
Surprise	Eyes	5.71	20.82	**31.02**	2.86	39.59
	Brow	3.47	16.94	**36.53**	8.78	34.29
	Cheeks	4.29	18.57	**34.69**	10.20	32.24
	Mouth	3.47	21.84	**43.06**	6.73	24.90
Disgust	Eyes	24.69	7.35	6.94	**44.90**	16.12
	Brow	28.43	5.93	6.95	**43.97**	14.72
	Cheeks	25.25	6.72	4.68	**45.42**	17.92
	Mouth	16.33	7.76	7.76	**50.20**	17.96

and we also expected fewer misclassifications when the distinguishing feature (i.e., the mouth) was initially fixated regardless of intensity. Following the analysis conducted in the previous Chapter on the confusions, one-way repeated measures ANOVAs with response option as a factor were conducted on the percentage of emotion categories given as responses to each of the facial expressions. Then, further one-way repeated measures ANOVAs were conducted as initial fixation location as a factor to investigate whether informative facial features reduced the most common misclassifications. The following confusion matrices show the percentage of total responses given to each target expression when presented at each of the possible initial fixation locations. Different confusion matrices were derived for the three levels of intensity.

Confusion matrix for expressions presented at 30% intensity can be seen in Table 5. At 30% intensity, there was an effect of response option for all expressions (anger: $F_{(1.83, 73.04)}$=74.65, p<.001, η_p^2= .65; fear: $F_{(1.87, 74.78)}$=40.48, p<.001, η_p^2= .50; surprise: $F_{(1.68, 67.24)}$=46.81, p<.001, η_p^2= .54; disgust: $F_{(2.57, 102.90)}$=33.53, p<.001, η_p^2= .46). At this lowest intensity, all expressions were most frequently misclassified as neutral compared to all other emotions, except for disgust where there was no difference between misclassifications as neutral and anger. Angry expressions were more frequently misclassified as fear (p= .012) and disgust (p< .001) compared to surprise. However, contrary to our expectations, there was no effect of initial fixation location on the misclassification of anger as fear ($F_{(2.37, 94.72)}$=.99, p=.39, η_p^2= .024) or as disgust ($F_{(3, 120)}$=1.61, p=.19, η_p^2= .039). Fearful expressions were more often misclassified as surprise compared to anger (p< .001) and disgust (p= .022). Again, there was no effect of initial fixation on misclassification of fear as surprise ($F_{(3, 120)}$=.94, p=.43, η_p^2= .023). Surprised

Table 6. Percentage of each emotion response option chosen for each target expression presented at 50% intensity.

Displayed expression		**Response Given/%**				
	Initial Fixation	**Anger**	**Fear**	**Surprise**	**Disgust**	**Neutral**
Anger	Eyes	**56.62**	8.96	2.85	15.07	16.50
	Brow	**56.53**	9.80	2.04	20.00	11.63
	Cheeks	**57.14**	7.96	2.65	17.55	14.69
	Mouth	**58.45**	7.74	1.83	19.55	12.42
Fear	Eyes	3.67	**53.06**	27.55	8.78	6.94
	Brow	2.45	**51.22**	32.65	5.31	8.37
	Cheeks	4.29	**57.35**	20.82	10.00	7.55
	Mouth	2.65	**57.35**	21.84	10.20	7.96
Surprise	Eyes	4.28	19.14	**64.15**	2.85	9.57
	Brow	3.26	17.72	**67.41**	4.89	6.72
	Cheeks	1.64	22.49	**62.99**	4.70	8.18
	Mouth	3.07	19.22	**67.69**	5.52	4.50
Disgust	Eyes	28.37	2.24	3.88	**63.67**	1.84
	Brow	32.18	4.07	5.09	**56.21**	2.44
	Cheeks	29.12	3.46	4.07	**61.30**	2.04
	Mouth	22.15	3.46	4.67	**66.26**	3.46

expressions were most often misclassified as fear compared to anger and disgust (both *p*s < .001). There was no effect of initial fixation location for surprised expressions misclassified as fearful ($F_{(3, 120)}$=1.61, *p*=.190, η_p^2= .039). Finally, disgust expressions were most often misclassified as anger (vs. fear and surprise *p*s< .001; neutral, *p*=.038). Disgust was less frequently misclassified as anger when the mouth is initially fixated ($F_{(3, 120)}$=9.29, *p*<.001, η_p^2= .19) compared to the eyes (p= .015), brow (p< .001) and the cheeks (p= .006).

Confusion matrix for expressions presented at 50% intensity can be seen in Table 6. At 50 % intensity, there was an effect of response option for all expressions (anger: $F_{(2,25,\ 90.13)}$=24.16, *p*<.001, η_p^2= .38; fear: $F_{(1,82,\ 73.15)}$=46.26, *p*<.001, η_p^2= .54; surprise: $F_{(1,68,\ 66.99)}$=27.93, *p*<.001, η_p^2= .41; disgust: $F_{(1,37,\ 54.90)}$=93.64, *p*<.001, η_p^2= .70). Angry expressions were more often misclassified as disgust compared to fear (*p*=.001) and surprise (*p*< .001). Anger was also misclassified more as fear (*p*=.001) and neutral (*p*< .001) compared to surprise. Initial fixation had no effect on the misclassification of anger as fear ($F_{(2,33,\ 93.37)}$=.81, *p*=.47, η_p^2= .02) but initially fixating the eyes ($F_{(3, 120)}$=2.43, *p*=.07, η_p^2= .06) reduced the misclassification of anger as disgust compared to the brow (*p*= .081), albeit only approaching significance. Fear was most often misclassified as surprise (all *p*s< .001) and surprise was most often misclassified as fear (all *p*s< .001). While misclassification of fear as surprise was reduced (initial fixation main effect: $F_{(2,49,\ 99.51)}$=9.28, *p*=.001, η_p^2= .19) when mouth was initially fixated compared to the brow (*p*< .001) and when cheeks are fixated compared to the eyes (*p*=.023) and brow (*p*=. 004); initial fixation had no effect on surprise being misclassified as fear ($F_{(3, 120)}$=1.58, *p*=.20, η_p^2= .04). Disgust was most often misclassified as anger compared to other emotions (all

Table 7. Percentage of each emotion response option chosen for each target expression presented at 70% intensity.

Displayed expression		**Response Given/%**				
	Initial Fixation	**Anger**	**Fear**	**Surprise**	**Disgust**	**Neutral**
Anger	Eyes	**63.34**	7.33	3.67	18.33	7.33
	Brow	**65.10**	7.96	2.24	20.00	4.69
	Cheeks	**66.26**	7.32	2.64	16.46	7.32
	Mouth	**63.54**	8.76	2.85	17.11	7.74
Fear	Eyes	3.27	**60.20**	30.61	5.10	0.82
	Brow	2.04	**60.12**	30.47	5.73	1.64
	Cheeks	4.70	**61.15**	25.56	7.57	1.02
	Mouth	4.70	**67.28**	18.61	8.38	1.02
Surprise	Eyes	2.24	15.27	**77.80**	2.44	2.24
	Brow	1.23	15.54	**77.51**	3.48	2.25
	Cheeks	3.27	14.29	**77.55**	2.45	2.45
	Mouth	1.84	16.53	**78.78**	2.45	0.41
Disgust	Eyes	32.52	2.86	3.48	**58.90**	2.25
	Brow	39.18	1.84	2.86	**54.69**	1.43
	Cheeks	35.57	2.03	2.03	**58.94**	1.42
	Mouth	29.18	2.24	2.24	**65.31**	1.02

*p*s< .001). This misclassification was reduced when initially fixating the mouth ($F_{(1,97, 78.79)}$=4.80, *p*=.011, η_p^2= .11) compared to brow (*p*=. 015) and cheeks (*p*<.001) but the comparison between the mouth and the eyes shortly failed significance (*p*=.052).

Confusion matrix for expressions presented at 70% intensity can be seen in Table 7. At 70% intensity, there was still an effect of response option for all expressions (anger: $F_{(1,95, 78.13)}$=25.14, *p*<.001, η_p^2= .39; fear: $F_{(1,31, 52.40)}$=49.97, *p*<.001, η_p^2= .56; surprise: $F_{(1,16, 46.19)}$=27.22, *p*<.001, η_p^2= .41; disgust: $F_{(1,15, 45.98)}$=137.39, *p*<.001, η_p^2= .78). Anger was again most often misclassified as disgust (fear: *p*=.002; surprise: *p*< .001; neutral: *p*< .001). Initially fixating any feature did not affect the misclassification of anger as disgust ($F_{(3, 120)}$=1.21, *p*=.31, η_p^2= .029). Fear was most often misclassified as surprise (all *p*s< .001) and surprise was most often misclassified as fear (all *p*s< .001). Initially fixating the mouth ($F_{(3, 120)}$=8.02, *p*<.001, η_p^2= .17) reduced the misclassification of fear as surprise compared to initially fixating the eyes (*p*= .001) and brow (*p*=.004) but not the cheeks (*p*=.087). On the other hand, there was no effect of initial fixation location on the misclassification of surprise as fear ($F_{(2.61, 104.31)}$=.42, *p*=.71, η_p^2= .01). Disgust was most often misclassified as anger (all *p*s< .001). The main effect of initial fixation location on the misclassification of disgust as anger ($F_{(3, 120)}$=4.45, *p*=.005, η_p^2= .10) showed that initially fixating on the mouth led to reduced misclassification compared to fixating the brow (*p*=.004) only.

Eye-movement Analysis

The analysis of saccade data is based on 40 participants since one block of eye-movement data for one participant could not be recorded due to technical issues, therefore the data from that participant is excluded from the eye movement data analysis. On average, 83.71% of all trials across participants contained valid saccades given by the criteria in the Data Analysis section.

Percentage of Saccades Going Upwards from the Mouth and Downwards from the Eyes Gamer and Buchel (2009) investigated the percentage of reflexive saccades going upwards from the mouth and downwards from the eyes to investigate the seeking out of informative facial features for emotion recognition. Here, we replicate their analysis. We ran a 4×3×2 repeated

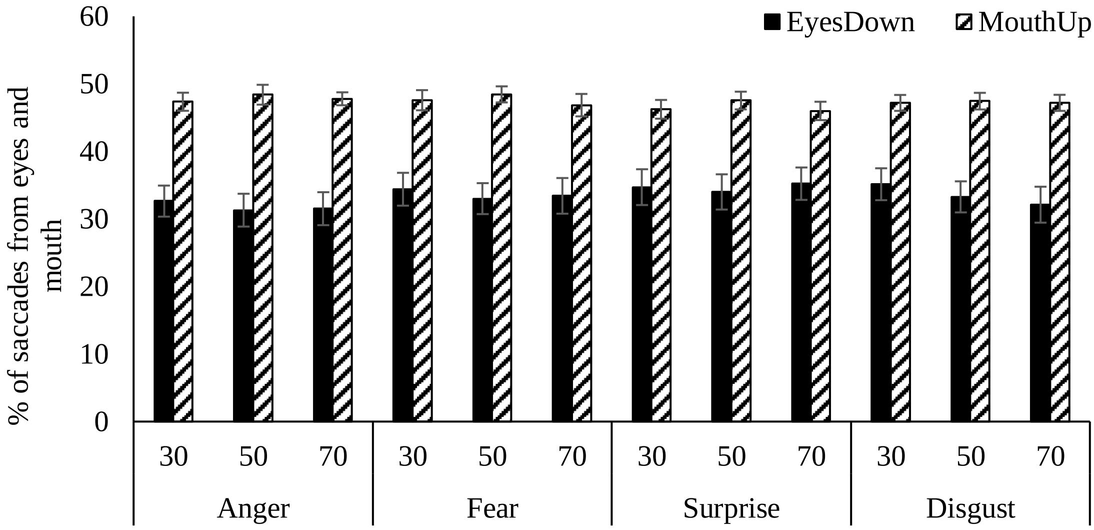

Figure 30. Percentage of reflexive saccades downwards from initial fixation locations on the eyes (EyesDown) and upwards from initial fixation on the mouth (MouthUp) for each expressions at all three intensities. Error bars indicate standard error of the mean.

measures ANOVA using expression, intensity and saccade direction as factors to investigate whether the percentages of saccades going upwards from the mouth and downwards from the eyes were affected by expression and intensity of expression. We expected to find a higher percentage of saccades going upwards from the mouth in line with both our previous findings and Gamer and Buchel (2009). We also expected to find a higher percentage of reflexive saccades towards the informative facial features for lower intensity facial expressions compared to the higher intensity facial expressions. Figure 30 summarises the descriptive statistics of this analysis. The values are calculated as a percentage of all saccades initiated from the eyes and the mouth initial fixation locations. There was no main effect of emotion ($F_{(3, 117)}$= 0.35, p=.79, η_p^2= .009) or intensity ($F_{(3, 117)}$= 0.46, p=.63, η_p^2= .012). The significant main effect of saccade direction ($F_{(3, 117)}$= 23.87, p<.001, η_p^2=.38) indicated that the percentage of saccades going upwards from the mouth was significantly higher than the percentage of saccades going downwards from the eyes which is in line with our hypothesis. Contrary to our hypothesis, there were no significant two-way interactions (emotion x intensity: $F_{(6, 234)}$= 1.03, p=.41, η_p^2= .026; emotion x saccade direction: $F_{(3, 117)}$= 1.13, p=.34, η_p^2= .028; intensity x saccade direction: $F_{(2, 78)}$= 0.47, p=.63, η_p^2= .012) and the three-way interaction also failed to reach significance ($F_{(4.74, 184.70)}$= 1.72, p=.14, η_p^2= .042) which suggests that the direction of the reflexive saccades were not modulated by expression or intensity.

Saccades Going Upwards from Lower Features and Downwards from the Upper Features

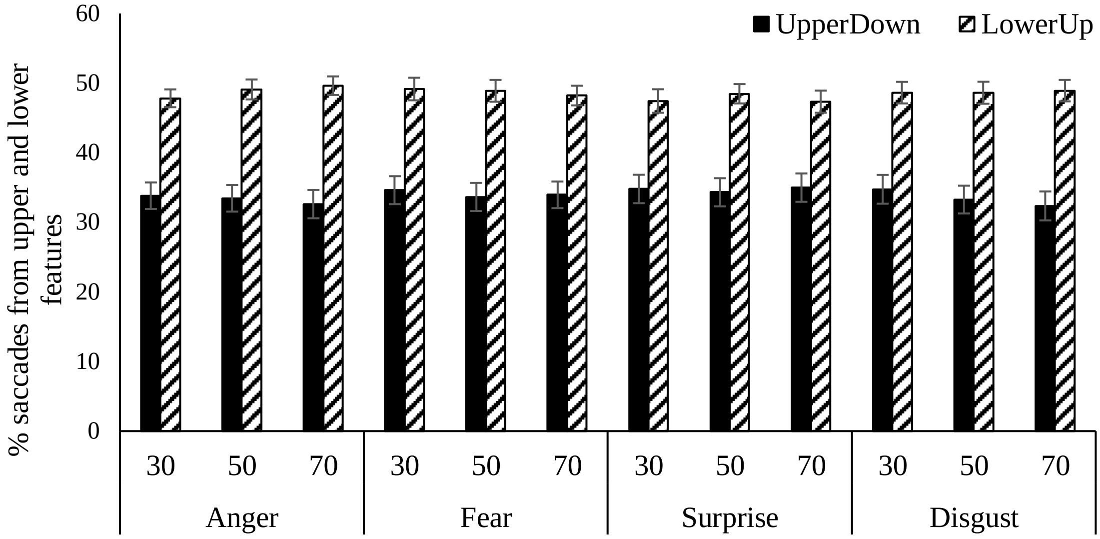

Figure 31. Percentage of reflexive saccades downwards from fixation on upper-face features (UpperDown) or upwards from lower-face (LowerUp) features for each expression and intensity. Error bars indicate standard error of the mean.

Since we used additional initial fixation locations to Gamer and Buchel (2009), we ran a further analysis to include these additional locations. We ran a 4×3×2 repeated measures ANOVA using expression, intensity and saccade direction as factors to investigate whether percentage of saccades going upwards from the lower facial features (i.e., the cheeks and the mouth) and downwards from the upper facial features (i.e., the mouth and the cheeks) were affected by expression and intensity of expression. Figure 31 summarises the descriptive statistics for this analysis. There was no main effect of emotion ($F_{(3, 117)}$= 0.44, p=.73, η_p^2= .011) or intensity ($F_{(3, 117)}$= 0.86, p=.43, η_p^2= .021). The significant main effect of saccade direction ($F_{(3, 117)}$= 26.31, p<.001, η_p^2= .40) indicated that the percentage of saccades going upwards from the lower features was significantly higher than the percentage of saccades going downwards from the upper features corroborating the results from the previous analysis and in line with our hypothesis. The interaction between expression and saccade direction tended towards significant ($F_{(3, 117)}$= 2.27, p=.08, η_p^2= .06). No other two-way interaction (expression x intensity: $F_{(4.51, 175.50)}$= 0.91, p=.47, η_p^2= .023; intensity x saccade direction: $F_{(2, 78)}$= 1.64, p=.20, η_p^2= .040) or the three-way interaction reached significance ($F_{(6, 234)}$= 0.72, p=.64, η_p^2= .018).

Two one-way repeated measures ANOVAs were conducted to follow-up the trend towards an interaction between expression and saccade direction, looking at the effect of expression for each saccade direction individually. There was no effect of expression on saccades going upwards from the lower facial features ($F_{(3, 123)}$= 1.81, p=.15, η_p^2= .042) or saccades going downwards from the upper facial features ($F_{(3, 123)}$= 1.86, p=.14, η_p^2= .043). The pairwise comparisons, on the other hand, showed that there were a higher percentage of saccades going downwards from the upper facial features for surprised compared to angry faces (p=.04) which is in line with our suggestion that reflexive saccades will target emotion-informative facial features (i.e. mouth) but only for surprised expressions.

Saccade Path Analysis

Collapsed across initial fixation

We investigated the paths of reflexive saccades collapsed across initial fixation location using a 4 (Expression) × 3 (Intensity) × 6 (Saccade Target) repeated-measures ANOVA (Figure 32). We expected saccades to be more strongly directed towards informative features of lower intensity expressions (i.e., 30% intensity) while this effect might not be present or less strong in higher intensity expressions (i.e., 70% intensity). There was no main effect of expression ($F_{(2.46, 96.02)}$= 0.35, p=.75, η_p^2= .009) or intensity ($F_{(1.96, 76.39)}$= 0.73, p=.48, η_p^2= .018). There was a main effect of saccade target ($F_{(1.97, 76.39)}$= 18.38, p< .001, η_p^2= .32) which suggests that the reflexive saccades were targeting the brow region more compared to all other saccade targets (all ps< .001 except brow vs left cheek: p= .008, brow vs mouth: p= .003). And, reflexive saccades were least strongly directed towards the right cheek compared to all other saccade targets (all ps< .001 except right cheek vs left eye: p=.003).

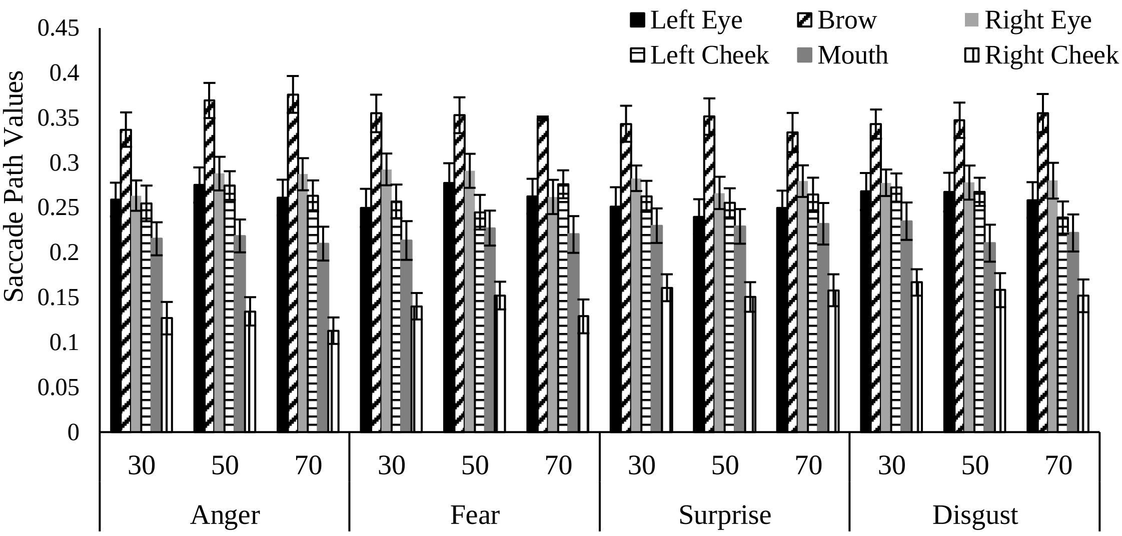

Figure 32. Mean saccade paths of reflexive saccades for each expression at each intensity towards each saccade target collapsed across initial fixation. Error bars represent standard error.

There was no interaction between expression and intensity ($F_{(4.64, 181.05)}$= 1.80, p=.12, η_p^2= .044) or intensity and saccade target ($F_{(5.81, 226.58)}$= 0.56, p=.76, η_p^2= .014). There was a significant interaction between expression and saccade target ($F_{(8.15, 317.85)}$= 5.65, p= .008, η_p^2= .06). Contrary to our expectations, there was no three way interaction among expression, intensity and saccade target ($F_{(14.43, 562.71)}$= 1.53, p=.09, η_p^2= .038) suggesting that the intensity did not modulate the effect of expression and initial fixation on the saccade path values.

In order to further investigate the expression, saccade target interaction, four two-way repeated measure ANOVAs were conducted comparing the saccade targets for each expression separately. There was a main effect of saccade target for all expressions (anger: $F_{(2.10, 81.72)}$= 22.09, p< .001, η_p^2= .36; fear: $F_{(2.07, 80.86)}$= 17.66, p< .001, η_p^2= .31; surprise: $F_{(2.02, 78.87)}$= 13.98, p< .001, η_p^2= .26; disgust: $F_{(2.09, 81.47)}$= 15.31, p< .001, η_p^2= .28). For fear, surprise and disgust, reflexive saccades towards the brow region were strongest compared to all other saccade targets and least strong towards the right cheek compared to all other saccade targets. For angry expressions, saccades were directed towards the brow more strongly and towards the right cheek least strongly. Additionally, only for angry expressions, reflexive saccades were directed towards the left cheek more strongly than towards the mouth.

Separately from each initial fixation location

For initial saccades starting from the left eye, there was no main effect of expression ($F_{(3, 102)}$= 1.44, p=.24, η_p^2= .04) or intensity ($F_{(2, 68)}$= 0.22, p=.80, η_p^2= .006). There was a main effect of saccade target ($F_{(1.17, 35.76)}$= 37.48, p< .001, η_p^2= .52) which indicated that saccades from the left eye are more strongly directed towards the brow compared to all other targets (all ps< .001 except brow vs right eye: p= .002) and towards the right eye compared to the cheeks and the mouth (all ps< .001). None of the two way interactions reached significance following Greenhouse-Geisser correction (expression × intensity: $F_{(4.28, 145.63)}$= 0.16, p=.96, η_p^2= .005; expression × saccade target: $F_{(3.97, 134.80)}$= 1.08, p=.37, η_p^2= .031; intensity × saccade target: $F_{(2.98, 101.38)}$= 1.91, p=.13, η_p^2= .053). The three way interaction was significant ($F_{(7.82, 265.70)}$= 2.00, p=.05, η_p^2= .055).

To investigate this interaction, four separate 3 × 5 repeated measures ANOVAs were conducted for each expression separately. There was no intensity × saccade target interaction for anger ($F_{(2.99, 101.56)}$= 0.66, p=.58, η_p^2= .019), fear ($F_{(2.60, 88.47)}$= 0.83, p=.46, η_p^2= .024) or surprised expressions ($F_{(3.30, 112.23)}$= 1.17, p=.33, η_p^2= .033) opposing our prediction that intensity will affect the way emotion-informative features are targeted. The interaction between intensity and saccade target for initial saccades starting from the left eye was significant for disgusted expressions ($F_{(3.01, 102.32)}$= 5.05, p= .003, η_p^2= .13). For disgusted expressions at 30% intensity, the initial saccades from the left eye more strongly targeted the brow (left cheek: p= .034;

mouth: p= .004; right cheek: p= .006) and the right eye (left cheek: p= .022; mouth: p< .001; right cheek: p< .001) compared to the cheeks and the mouth. They also targeted the right cheek more strongly compared to the mouth (p =.023). For disgusted expressions at 50% intensity, the initial saccades from the left eye more strongly targeted the brow compared to all other saccade targets (right eye: p= .019; left cheek: p= .035; mouth: p= .001; right cheek: p= .001) and targeted the right eye more strongly compared to the mouth (p= .002) and the right cheek (p< .001). Finally, for disgusted expressions at 70% intensity, the initial saccades from the left eye targeted the brow (all ps< .001) and the right eye (all ps< .001) more strongly compared to the cheeks and the mouth and targeted the right cheek more strongly compared to the left cheek (p< .001) and the mouth (p= .001). Overall, for disgusted expressions, the results provided no support for our prediction that reflexive saccades will target the emotion-informative mouth region when they initiated from a less informative facial feature such as the eyes. Even when the intensity increased, leading to a more distinctive appearance of the informative mouth region, the reflexive saccades from the left eye targeted upper features more strongly.

For initial saccades starting from the brow, there was no main effect of expression ($F_{(3, 99)}$= 2.17, p=.097, η_p^2= .06) or intensity ($F_{(2, 66)}$= 0.16, p=.85, η_p^2= .005). There was a main effect of saccade target ($F_{(1.45, 47.94)}$= 4.09, p= .034, η_p^2= .11) which indicated that saccades from the brow were directed towards the left cheek more strongly compared to the right eye (p= .016) and the right cheek (p= .018). Saccades from the brow were also more strongly directed towards the mouth compared to the right eye (p= .016) and the right cheek (p= .004). There were no two-way or three-way interactions for saccades starting from the brow (expression x intensity: $F_{(4.28, 141.29)}$= 1.10, p=.36, η_p^2= .032; expression x saccade target: $F_{(4.73, 156.14)}$= 1.30, p=.27, η_p^2= .038; intensity x saccade target: $F_{(4.02, 132.57)}$= 1.72, p=.15, η_p^2= .049; expression x intensity x saccade target: $F_{(8.51, 280.97)}$= 1.24, p=.27, η_p^2= .036) which indicated that neither expression nor intensity had an effect on what features reflexive saccades targeted.

From the right eye, there was no main effect of expression ($F_{(2.37, 80.52)}$= 2.65, p=.068, η_p^2= .072) or intensity ($F_{(2, 68)}$= 1.27, p=.29, η_p^2= .036). There was a main effect of saccade target ($F_{(1.31, 44.65)}$= 78.54, p< .001, η_p^2= .70) which indicated that saccades from the right eye was more strongly directed towards the left eye (left cheek: p= .007; mouth: p< .001; right cheek: p< .001) and the brow (left cheek: p= .021; mouth: p< .001; right cheek: p< .001) compared to the cheeks and the mouth. The saccades from the right eye were also directed towards the left cheek more strongly compared to the right cheek (p< .001) and the mouth (p< .001). None of the interactions reached significance following Greenhouse-Geisser correction (expression × intensity: $F_{(5.00, 170.12)}$= .53, p=.75, η_p^2= .015; expression × saccade target: $F_{(5.37, 182.52)}$= 1.89, p=.09, η_p^2= .053; intensity × saccade target: $F_{(4.11, 139.88)}$=.35, p=.85, η_p^2= .010; expression ×

intensity x saccade target: $F_{(9.04, 307.29)}$= 1.24, p=.27, η_p^2= .035) again indicating that neither expression nor intensity had an effect on what features reflexive saccades targeted.

From the left cheek, there was no main effect of expression ($F_{(3, 102)}$= .052, p=.98, η_p^2= .002). There was a main effect of intensity ($F_{(2, 68)}$= 3.13, p=.05, η_p^2= .084) however the pairwise comparisons revealed no significant difference between the intensities. There was also a main effect of saccade target ($F_{(1.24, 42.27)}$= 5.30, p=.02, η_p^2= .14) which indicated that saccades from the left cheek were more strongly directed towards the right eye compared to the left eye (p= .016) and the left cheek (p= .029) and towards the brow compared to left eye (p< .001). The saccades were also more strongly directed towards the right cheek compared to the mouth (p= .012).

For initial saccades from the mouth, there was no main effect of expression ($F_{(3, 102)}$= 0.14, p=.93, η_p^2= .004) or intensity ($F_{(2, 68)}$= 0.57, p=.57, η_p^2= .016). There was a main effect of saccade target ($F_{(1.58, 53.82)}$= 31.69, p< .001, η_p^2= .48) which indicated that saccades from the mouth were least strongly directed towards the right cheek compared to all other saccade targets (all ps< .001). The saccades from the mouth were also more strongly directed towards the left cheek compared to the right cheek (p< .001). None of the interactions reached significance following Greenhouse-Geisser correction (expression × intensity: $F_{(4.46, 151.64)}$= 1.041, p=.39, η_p^2= .030; expression × saccade target: $F_{(4.95, 168.21)}$= 2.21, p=.06, η_p^2= .061; intensity × saccade target: $F_{(3.64, 123.78)}$=0.50, p=.72, η_p^2= .014; expression × intensity × saccade target: $F_{(8.48, 288.40)}$= 0.89, p=.53, η_p^2= .025) contradicting our predictions that reflexive saccade paths will be modulated by intensity and the expression of the face.

Finally, for saccades from the right cheek, there was no main effect of expression ($F_{(3, 102)}$= 2.25, p=.087, η_p^2= .062), intensity ($F_{(2, 68)}$= 1.65, p=.20, η_p^2= .046) or saccade target ($F_{(1.08, 36.57)}$= 2.20, p= .146, η_p^2= .061). None of the interactions reached significance following Greenhouse-Geisser correction (expression × intensity: $F_{(4.11, 139.84)}$= 1.65, p=.16, η_p^2= .046; expression × saccade target: $F_{(4.66, 158.53)}$= 1.78, p=.125, η_p^2= .050; intensity × saccade target: $F_{(3.18, 108.17)}$=1.15, p=.34, η_p^2= .033; expression × intensity × saccade target: $F_{(7.08, 240.60)}$= 0.95, p=.57, η_p^2= .027) again contrary to our expectations regarding the effect of expression and intensity on reflexive saccade paths.

Discussion

The main aim of this experiment was to investigate whether initially fixating on the emotion-informative features of briefly presented expressions improved emotion recognition for emotions expressed at lower intensities. To this end, we used a brief-fixation paradigm (Gamer & Buchel, 2009, Gamer et al., 2012) to present angry, fearful, surprised and disgusted expressions at three different intensity levels (30%, 50% and 70% morphs between the full-intensity and neutral expressions) for approximately 80ms and asked participants to label the presented expressions either with the relevant emotion label or with the neutral label.

As expected, emotion recognition performance of participants showed a gradual increase with increasing intensity, in line with the results of most previous research (Guo, 2012; Vaidya et al., 2014; Hess et al., 1997; Wells et al., 2016). The only exception to this was the disgusted expression where there was no difference between the 50% and 70% intensities. Despite showing an overall cumulative effect of increasing expression intensity for all expressions, Hess et al. (1997) also failed to show a difference in decoding accuracy for disgust between 60% and 80% morphs of the expression. It is possible that after 50% expression intensity, disgust recognition reached its peak and additional expressive information brought about by the higher intensity morph did not have any further informative value. However, Hess et al. (1997) found a further increase in disgust recognition when the intensity increased from 80 to 100%, which could not be tested here since the highest intensity used was 70%. Vaidya et al. (2014) also found that the recognition accuracy for subtle disgust (range between 20-40% intensity) was higher compared to subtle fear and surprise. Therefore, it is possible that lower intensity expressions of disgust simply carry a high amount of expression-informative information compared to other expressions.

The increase in performance with increasing intensity shows that as the facial features that define facial expressions become less ambiguous, emotion recognition performance increases.

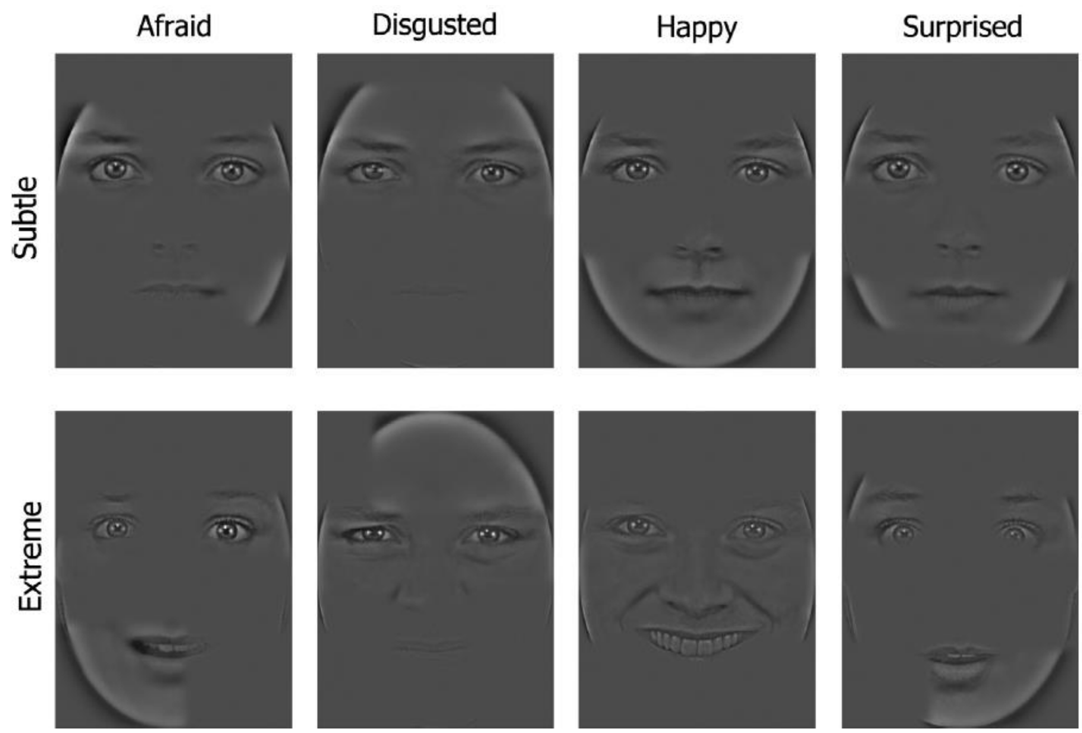

Figure 33. Results of the ideal observer analysis from Vaidya et al. (2014) showing where the informative features for discriminating a facial expression from a neutral pose lie.

Although this effect of expression intensity differed slightly across emotions, it did not differ as a function of initial fixation location, either overall or in an emotion-specific manner. Even though we did not find any effect of intensity as a function of initial fixation location on emotion recognition accuracy, our hypothesis that initially fixating on the informative facial features would improve emotion recognition accuracy is partially supported since initially fixating on the mouth improved recognition accuracy for disgust, surprise and fear regardless of intensity. Previous research has shown that, despite being more ambiguous, informative facial features at lower intensities can still attract relatively more fixations. For example, Guo (2012) found that mouth region of surprised faces received the highest number of fixations and longest viewing time and the eyes of angry faces attracted the highest number of fixations and longest viewing time for all the five levels of intensities they investigated (20, 40, 60, 80 and 100%). Vaidya et al. (2014) also found that more fixations were directed to the mouth region of subtle disgusted and fearful expressions compared to neutral poses (though not significantly). Schurgin et al. (2014) found that with increasing intensity from 20% to 60% intensities, the fixations to disgusted faces shifted from the eye region to the upper lip and a similar shift was observed for fearful faces. Furthermore, Matsumoto and Hwang (2014) showed that despite the presence of an AU and the intensity of the activation of the AU being separately correlated with emotion recognition performance, when the effects of both are considered together in a multiple regression model, the intensity of the AU does not have an effect over and above the presence of the AU. Therefore, it is fair to conclude that facial features still transmit expression-informative signals even at low intensities and initially fixating on these regions is beneficial to emotion recognition performance.

Guo (2012) found that observers spent more time fixating the eye region for all expressions and intensities and Vaidya et al. (2014) showed that eye fixations were the best predictor of emotion recognition performance for subtle facial expressions. As seen in Figure 33, Vaidya et al.'s ideal observer analysis revealed that, for subtle expressions, visual information lies within the eyes across high and low frequencies, however, they also suggest that the mouth region contains informative visual information at higher spatial frequencies. We found that initially fixating on an eye led to a greater number of neutral responses compared to the mouth and the brow regions in our study, which is contradictory to Vaidya et al.'s ideal observer analysis, which suggested that the eye region is informative for discriminating happy, disgusted, fearful and surprised expressions from neutral poses. Several reasons might be behind this contradictory finding: While the eye region might be useful for discriminating subtle expressions from neutral, it might not be as informative when distinguishing expressions from each other. As can be seen in Figure 33, there is minimum difference between the eye regions of the expressions studied by Vaidya et al. (2014). Another reason might be owing to the limited number of intensities studied by Vaidya et al. They studied only one low intensity for each expression and this was in the range of 20 to 40%. Blais, Roy, Fiset, Arguin and Gosselin (2012) suggest that observers prefer to utilise the mouth since the mouth has a highly dynamic motion range compared to the eyes or other facial features. Since we have used 3 different intensity levels and the different intensities were presented within the same block, this suggested dynamic change might have been more apparent to our observers leading them to make use of the mouth region compared to the eyes. In accordance with this suggestion, here we show that, regardless of expression intensity, initially fixating on the mouth region improved emotion recognition for fearful, surprised and disgusted facial expressions. This finding corroborates the results from the two experiments described in Chapter 2 and further shows that, contrary to the initial hypothesis of the current study, informative facial features contribute to emotion recognition performance of both prototypical and lower intensity facial expressions similarly. Additionally, analysis of the frequency of neutral responses showed that even at the lowest intensity level, initially fixating at the mouth region led to fewer neutral responses for fearful and surprised facial expressions.

While these results suggest that initially fixating informative features improve emotion recognition accuracy of emotions regardless of intensity, we did not find results supporting our hypothesis that initially fixating informative features contribute more to recognition of lower intensity expressions. In this experiment, same intensities (30%, 50% and 70%) were used for all expressions. However, the thresholds for recognizing the expressions of each of the six basic emotions differ (Calvo, Avero, Fernandez-Martin & Recio, 2016; Hess et al, 1997) Calvo et al. (2016) showed that happy expressions were recognizable above threshold at 20% intensity; sad, surprised and angry expressions needed to be at 40% intensity and fear needed to be at 50% intensity. The lowest intensity used in this experiment (30%) is very close to the threshold intensity for anger and surprised facial expressions and the mid-level (50%) intensity matches the threshold intensity for fearful expressions according to the results of Calvo et al. (2016). Therefore, the lack of an effect of intensity on the contribution of informative facial features to emotion recognition might be due to mid- and low-level intensities already providing unambiguous facial information for the recognition of each expression. Future research can look into increasing the range of lower level intensities used in order to pinpoint at what intensity foveal processing of an informative feature starts contributing to emotion recognition. Additionally, in this study, we used the same intensity levels for different facial expressions. The differing thresholds for different facial expressions might be partially attributed to the changes in facial morphology (Calvo et al., 2016; Maher, Ekstrom & Chen, 2014). Therefore, it is possible that, at the same intensity level, the facial morphological changes for one expression can be more informative compared to others. This can further be supported by our finding that there is no difference in recognition accuracy for disgust expressions of 50 and 70% intensity, suggesting that a more modest facial morphological change is sufficient for disgust recognition. Therefore, future research should also look into choosing intensities that produce comparable morphological changes in the facial features of each of the facial expressions.

One major limitation of this study is the lack of ideal observer analysis. One assumption made here is that the informative facial features are constant across all intensities for each expression. However, it is possible that different facial features become informative for each expression at different intensities. Compare the ideal observer analysis from Smith et al. (2005) and Vaidya et al. (2014) for example: While the ideal observer analysis for the prototypical expressions in both suggest that informative features are spread across different regions of the face, the analysis of subtle facial expressions in Vaidya et al. (2014) suggest that informative features are localized to the eyes. Though we did not find that initial fixation on the eyes improved emotion recognition for any of the lower intensity expressions, there are several differences in the initial fixation effects between this study and studies reported in Chapter 2. One example is disgust recognition when the mouth is initially fixated. While in Chapter 2, we reported that initially fixating the mouth improved disgust recognition compared to all other initial fixation locations, in this experiment this effect was limited to improved recognition only compared to the brow. Additionally, we found that the eyes, as well as the mouth, improved disgust recognition compared to the brow. This suggests that mouth region might be less informative for disgust expressions at lower intensities and this suggestion can be supported by Schurgin et al. (2014)'s finding that observers shifted from looking more at the eyes towards looking more at the upper nose and lip with increasing expression intensity.

Alternatively, recognition of lower intensity facial expressions might be more reliant on holistic processing, therefore; require integration of visual information from several features (Guo et al., 2012). Contrary to our initial suggestion that informative features of low intensity expressions will benefit from processing of high spatial frequencies, the ambiguity of each facial feature might lead observers to seek out visual information from several features before responding. In this case, forcing fixation on an informative feature would not aid emotion recognition, however; it would be interesting to investigate whether there is a *pattern* of eye movements that would lead to improved emotion recognition. For example, for lower intensity disgust expressions it might be beneficial to fixate on the eyes followed by the mouth compared to fixating on the eyes followed by the brow. Future research can look into this further.

Contrary to our expectations, we did not find that initially fixating the brow region improved emotion recognition accuracy. However, we found that initially fixating the brow region led to the least number of neutral responses for angry faces (even though this did not reach statistical significance). It is therefore possible that the brow region of angry faces contributes to the discrimination of angry faces from the neutral faces. Even though the similarity of the furrowed brow region between anger and disgust in the current study might mean that initially fixating this region might not provide any distinctive information as to the identity of the expression, the brow region seems to aid the discrimination of angry expressions from neutral expressions. Support for this can be found in Schurgin et al. (2014) study. They show that observers looked more within the nasion of neutral faces when examining angry and neutral faces when instructed to indicate whether faces showed any expression. Additionally, Atkinson and Smithson (2020) found that initially fixating the brow region of angry faces improved emotion recognition accuracy and this was accompanied by a reduction of neutral responses given to angry expressions, in two experiments in which both angry and neutral faces were presented.

Fearful facial expressions were misclassified as surprised, similar to our findings in Chapter 2 and further extending this finding by showing that subtle expressions of fear and surprise are still confused but their confusion is not alleviated until an intensity of 50% is surpassed. Even though this misclassification seems to be reduced by initially fixating the mouth region at 70% intensity, no apparent benefit of initial fixation is observed for lower intensities. Du and Martinez (2011) showed that the confusion of fear as surprise is high at mid to high image resolution but is very low in low image resolution. A parallel could be drawn between this finding and ours: The percentage of misclassification between fear and surprise increases with increasing expression intensity. One common aspect of increasing intensity and increasing image resolution is less ambiguity in facial features. Since these two expressions share more AUs with each other than distinctive ones (Roy-Charland, Perron, Beaudry & Eady, 2014), it is possible that a high degree of muscle activation of a distinctive AU (for example, the lip stretcher action unit which is distinctive to fear and early attention to the region affected by the activation of this AU (i.e., the mouth) is necessary for the misclassification to be resolved. Additionally, disgust was most often misclassified as anger and this was consistent across all intensities and is in line with Guo et al.'s (2012) findings that among expressions of 40% and higher intensities, disgust is most often misclassified as anger. This misclassification seems to be more prominent when initially fixating the brow region and this might be due to the similarity of this region in both expressions (Jack, Garrod & Schyns, 2014). Disgust was confused less with anger when initially fixating at the mouth region for all intensities. This is somewhat different to the misclassification pattern of fear and surprise where the resolution of the misclassification was affected by intensity.

In line with our analysis of saccade direction in Chapter 2, we wanted to know whether reflexive saccades are directed to informative facial features in lower intensity facial expressions as well. If the saccade direction is purely affected by the featural information pertaining to each expression, we would expect there to be differences between lower intensity and higher intensity expressions – in the case of this study between 30% and 50% and 70% expressions – since featural information is more ambiguous in lower intensity expressions. However, the saccade direction analysis showed only that the reflexive saccades tend to seek out upper facial features compared to lower facial features regardless of expression and intensity. This finding agreed with the findings from the experiments reported in Chapter 2. Furthermore, the saccade path analysis revealed that the reflexive saccades made by our observers were most similar to a saccade path targeting the brow region regardless of expression and intensity. In line with our previous findings (Chapter 2), saccades from either eye tended to target the brow or the opposite eye.

There are several processes that can affect the direction of early saccades. As discussed in the Discussion section of Chapter 2, the centre-of-gravity effect might be leading the early saccades towards the centre of the face image (Bindemann et al., 2009). We did present some evidence supporting this in Chapter 2, with caution. Peterson and Eckstein (2012) showed that observers prefer fixating just below the eyes on briefly presented faces. They also suggested that, when taking into account the capabilities of the foveated visual system, this location provides the observer with the optimal amount of visual information to resolve face-related tasks. It is, therefore, possible that these early saccades are targeting this optimum location just below the eyes. While we found several expression specific effects in the experiments reported in Chapter 2, these were not consistent across the two experiments therefore we cannot reliably argue for these effects without further replication.

The only consistent result we have received across the three studies reported in this and the previous chapter is the higher proportion of saccades upwards from the mouth compared to downwards from the eyes. The tendency of saccades to move upwards from the mouth more often might be a face-specific effect rather than reflecting the seeking out of informative facial features. Support for this can be found in studies of macaque social communication by Guo, Robertson, Mahmoodi, Tadmor and Young (2003) and Guo (2007). Guo et al. (2003) showed that when presented with face images of other monkeys and humans, the macaques direct their first fixation more often towards the eyes compared to other facial features such as the mouth and the nose and the first fixation to the eyes comes later when faces are scrambled (locations of facial features shuffled) or inverted. Guo (2007) further showed that macaques direct their first fixation towards the eye region even when the face image is scrambled to make the eyes not visible. These studies suggest that early fixations to the eyes might be driven by the top-down knowledge of the location of the eyes rather than its perceptual or social saliency. Some research on human observers suggests that the initial tendency to look at the eyes might be automatic and cannot be suppressed (Laidlaw & Kingstone, 2017; Laidlaw, Risko & Kingstone, 2012; Thompson, Foulsham, Leekam & Jones, 2019). Laidlaw et al. (2012) showed that when participants were asked to not look at the eyes and not look at the mouth, fixations and dwell times on the eyes were longer at the eyes for 'don't look at the eyes' condition compared to fixations and dwell times on the mouth for 'don't look at the mouth' condition. These results were true for fixations taking place both within 100ms of viewing time and after. Thompson et al. (2019) carried out a face recognition task where participants needed to identify a target face they had seen from a pair of target and foil faces. On half of the trials, participants were prompted to look at the eyes or at the mouth to aid their detection of the target face. On unprompted trials, they found that first fixations by participants were made more often to the eyes compared to the mouth. More importantly, there was no difference between initial fixations to the eyes when participants were prompted to look at the eyes or the mouth. On the other hand, they found that over the course of the 700ms the face remained on the screen, participants spent more time looking within the prompted region. The researchers, therefore, suggest that despite the possibility of voluntarily suppressing fixations to the eye region, there is an automatic process that directs attention to the eyes which is difficult to suppress. If directing early attention (i.e., first fixation, first saccade, etc) is an automatic process brought about by the face image, then early saccades (ones happening immediately after face onset) might not be reliable as a measure of what facial features are being sought out in a face image. However, this interpretation needs to be approached with caution since the previous studies mentioned did not investigate saccade behaviour for expressive faces and the current thesis did not investigate saccade behaviour on non-expressive faces.

Another eye movement parameter we were interested in was latencies of reflexive saccades from each initial fixation location for each expression at each intensity. According to the saccade generation model by Findlay and Walker (1990), if a fixated location requires effortful visual processing, the eye will remain at that location for longer. On the other hand, if there is relatively less visual information to be processed at a given location, the eye will move to the next fixation target sooner. According to this model, we expected the saccade latencies to differ across expression intensities: More specifically, we expected shorter saccade latencies from informative initial fixation locations such as the eyes, brow and the mouth since the visual information regarding the identity of expressions at these regions will be more ambiguous at lower intensities. While we failed to find an effect of intensity on saccade latencies, we found that reflexive saccades from the cheeks had shorter latencies compared to the brow and the mouth and the reflexive saccades from the eyes had shorter latencies compared to the brow (as reported in the Appendix). The saccade latencies from the cheeks were partly in line with our previous experiments and are also in line with our suggestion that the cheeks are relatively less informative compared to the other fixation locations on the face. Contrary to the results from Chapter 2, on the other hand, we did not find any difference in latencies between the cheeks and the eyes and the latencies from the eyes were shorter compared to latencies from the brow. A tentative suggestion from these results might be that the eye region is less informative for lower intensity expressions compared to peak intensity expressions.

A limitation of this study is that we opted to use a morphing technique to create lower intensity/subtle facial expressions. While the morphing technique is a valid way of obtaining low intensity facial expressions and has been used in the literature many times, the ecological validity of the created face images might not be very high. As Matsumoto and Hwang (2014) state, the linear morphing of an expressive face into a neutral face might not accurately and realistically represent the formation of subtle facial expressions in real life since it might not align with the anatomical movement of facial muscles.

Both in Chapter 2 and this chapter, we have shown that initially fixating on informative facial features of certain facial expressions conveys an advantage in emotion recognition accuracy. Especially, we have shown that initially fixating on the mouth region of both prototypical, high intensity facial expressions and lower intensity facial expressions improves recognition of fear, surprise and disgust and reduces misclassification of fear as surprise and disgust as anger. Furthermore, despite not obtaining a significant result, we show that initially fixating on the brow region reduces the misclassification of anger as neutral. Throughout this thesis, our reasoning for obtaining such results has been that fixating on an emotion-informative facial feature such as the mouth or the brow, led to improvements due to the processing of high spatial frequency information from these facial regions. However, since we did not measure the spatial frequency content of the studied faces, this suggestion is yet untested. In Experiment 4, which will be reported in the next chapter, we will investigate the effect of either masking or stripping away the high spatial frequency information from these informative facial features on emotion recognition performance to more directly test our hypothesis that processing high spatial frequency information is contributing to emotion recognition performance.

Chapter 4

Contribution of high spatial frequencies of informative facial features to facial expression recognition and/or discrimination.

Introduction

Facial expressions have features or regions that are informative for their accurate recognition (Smith & Schyns, 2009; Smith, Cottrell, Gosselin & Schyns, 2005; Smith & Merlusca, 2014; Wegrzyn, Vogt, Kireclioglu, Schneider & Kissler, 2017) and observers spend more time looking at these facial features when allowed to view faces freely (Eisenbarth & Alpers, 2011; Schurgin et al., 2014). In accordance with this, in the previous chapters, we presented evidence that a single, brief fixation on an informative feature can improve emotion recognition for some expressions, compared to a single, brief fixation on a different feature: fixating on the mouth led to higher recognition accuracy for disgust compared to fixating other locations and reduced the misclassification of disgust as anger. Initially fixating the mouth also improved emotion recognition accuracy for fear and surprise by reducing the misclassification of fear as surprise. This effect of initially fixating the mouth was still the present when these expressions were presented at lower intensities (with the use of a morphing technique). Using the same brief-fixation paradigm, Atkinson and Smithson (2020) recorded an improvement of emotion recognition performance for angry faces when a single fixation was enforced on the central brow region as compared to an eye or cheek. In line with this result, we found in Chapter 3 that initially fixating the brow led to a reduction in the number of neutral responses to angry faces of lower intensities. These findings supplement the results of studies indicating the presence of informative facial regions (e.g., Smith et al., 2005; Smith & Merlusca, 2014; Wegrzyn et al., 2017) by showing that when facial expressions are fully visible, foveal processing of certain informative regions leads to an improvement in recognition of the facial expression.

In contrast to the *Bubbles* studies where only certain facial regions/features were available for emotion recognition, in all the studies reported in this thesis, while fixation was on a non-informative feature (e.g., a cheek), the informative feature (e.g., mouth for disgust) was still visible in the extrafoveal visual field. Despite the ability or opportunity to process the face and all its features holistically, only initially fixating on an informative feature lead to recognition improvement. Therefore, we suggest that fixating, and thus foveal processing of, an informative feature conveys additional recognition advantage compared to extrafoveal processing of this same informative feature. The main difference in the visual information obtained in these two cases is owing to the capabilities of the foveated visual system: While the fovea is capable of processing high spatial frequency information obtaining finer detail from the fixated feature, with increasing eccentricity from the fovea, the ability of the visual system to process high spatial frequency information declines, therefore, obtaining coarser information from the rest of the face. There is evidence in the literature suggesting that facial expression recognition is possible through the processing of a range of spatial frequencies. For example, using the *Bubbles* technique (described in detail in Chapter 1), Smith and Merlusca (2014) separated images of facial expressions into five non-overlapping spatial frequency bands and quantified how much information was used from each band and what region of the face this information resided. They have found that for expressions of anger, fear and disgust, the information available in the face images was spread across the spatial frequency bands. The wrinkles around the eyes, nose and the mouth were informative for disgust recognition across a range of SFs when the disgust expressions were presented among a set of faces displaying five other emotions or a neutral expression. However, the informative SF spectrum for anger and fear changed depending on the expression combinations used, especially in the lower spatial frequency bands. High spatial frequency information from the eyes was consistently informative for fear recognition (when compared against only neutral, happy and neutral, angry and neutral and to all other expressions) and lower spatial frequencies from the various regions of the mouth also became informative. The furrowed brow and the eyebrows were informative for anger recognition across SF bands, with the brow region chosen by the experiments reported here being informative at the highest spatial frequencies. On the other hand, when compared against only neutral faces, observers made use of a broader range of information from across the angry face at lower spatial frequencies.

Similarly, Smith and Schyns (2009) used the *Bubbles* technique and quantified what information was used for the recognition of basic expressions and neutral faces within 4 non-overlapping spatial frequency bands. They showed that a higher proportion of visual information was being used at lower spatial frequencies for happy, surprised, disgusted and angry faces whereas the recognition of fearful, sad and neutral faces depended more on the use of a higher proportion of high spatial frequencies. According to their usage of high or low spatial frequency information, Smith and Schyns (2009) identified happy, surprised, disgusted and angry expressions as distal and fear, neutral and sad expressions as proximal expressions, meaning that the first group will be better recognised across a range of distances while the second group will be better recognised at shorter distances. Furthermore, Bayle, Schoendorff, Hénaff and Krolak-Salmon (2011) showed that it is possible to detect (recognised as emotional) fear and disgust at eccentricities up to 40 and 35 degrees respectively, which indicates the presence of informative low spatial frequency information even at these extreme eccentricities. Since regardless of whether initial fixation was on an informative feature or not, the participants in our brief-fixation experiments were very good at recognising the underlying expression, the contribution of extrafoveal information to emotion recognition in our studies cannot be disregarded. However, we observed relatively better recognition when the fixated region was informative compared to when it was non-informative. Therefore, we argue that directly fixating on an informative feature contributes to emotion recognition through the processing of high spatial frequencies by the fovea. In the following experiment, we test this hypothesis directly by either low-pass filtering or masking the initially fixated facial features and comparing emotion recognition accuracy to expression recognition of the original image.

Several studies support the idea that visual detail or high spatial frequencies contribute to facial emotion recognition performance. Du and Martinez (2011) administered a 7-alternative forced choice emotion recognition task in which participants were shown images of all six basic expressions along with neutral poses at 5 gradually decreasing resolutions for 500ms. They showed that while the recognition of happy, surprised and neutral faces was less affected by the reduction of resolution, anger, fear, sadness and disgust showed significant drops in recognition accuracy at relatively higher resolutions. Recognition of angry expressions was impaired with decreasing resolution and at the mid-resolution used in Du and Martinez's (2011) study, they started becoming confused with neutral expressions. A visual inspection of the images implies that loss of informative features with decreasing resolution, i.e., the furrows in the brow region, might have led to this misclassification. Visual detail within the brow region that is only present at higher resolutions might be necessary to detect angry expressions. Guo, Soornack and Settle (2018) conducted a similar experiment where they presented participants with facial expressions of all the six basic emotions and asked them to categorize and rate the intensity of each expression. They used 4 gradually declining image resolution levels. In addition to the emotion recognition task, Guo et al. (2018) also recorded the eye movement patterns specific to each expression at each resolution. Emotion recognition results were mostly in line with Du and Martinez (2011) where happiness and surprise were less affected by a decrease in resolution while anger, fear, sadness and disgust were susceptible to this manipulation at relatively higher resolutions. In addition to facial expression recognition performance, the decrease in resolution also affected gaze allocation to facial features: While the proportion of fixations directed to the eye region decreased with decreasing image resolution, the proportion of fixations within the nose region increased.

While not directly relevant to the study at hand, since they did not study the contribution of high spatial frequency information to emotion recognition, the studies discussed in the previous paragraph can inform our study. The impairment in emotion recognition performance caused by degrading the visual information in the image of a facial expression provides support for the contributory effect of visual detail from facial features to emotion recognition from faces. More directly relevant to the study at hand, Guo et al. (2018) further examined the recognition and eye movement patterns for facial expressions that contained 4 non-overlapping and decreasing spatial frequency bands. Decrease in the high end of the spatial frequency spectrum led to a decrease in emotion recognition for all expressions except for happy expressions: while observers showed similar performance with faces containing the broad SF content and SF bandwidth of 30-15 cycles/image, there was a drop in emotion recognition at 15-7.5 cycles per image and a further drop with 7.5-3.8 cycles per image. Additionally, increasing face blur led to an increase in the duration of individual fixations directed at the face suggesting an increase in difficulty to discriminate facial expressions. Like decreasing image resolution, there was a decrease in the proportion of fixations at the eye region and an increase in the proportion of fixations in the nose region. This might suggest that, the decrease in visual detail due to high spatial frequency filtering made fixating individual facial features redundant, leading the participants to endorse a more holistic face viewing strategy.

Besides spatial frequency filtering, presentation of facial expression images in the periphery further supports the contributory effect of foveating the face to emotion recognition. Calvo, Nummenmaa and Avero (2010) found that all facial expressions of emotion could be recognized above levels expected by chance when presented extrafoveally (the inner edge of the face 2.5 degrees of visual angle away from a central fixation) for 150ms. Participants were asked to fixate a central point on each trial. While participants were free to move their eyes, a gaze contingent mask blocked central vision, therefore preventing foveal processing of the face if participants moved their eyes to the location of the face. Since high spatial frequency information from the face would not be accessible at the location used by Calvo et al. (2010), this study provides strong support for the recognition of facial expressions using low spatial frequencies alone. However, the study cannot inform us about what facial feature provided this informative content and whether there was a difference between a central fixation and the extrafoveal presentation which is a more direct way of understanding the relative importance of foveal versus extrafoveal processing of features within a face for emotion recognition.

Calvo, Fernández-Martín and Nummenmaa (2014) used a similar task to Calvo et al. (2010) and added a central presentation and a peripheral presentation (6 degrees away from fixation) to the parafoveal presentation (2.5 degrees away from fixation). Participants fixated a central fixation cross which roughly corresponded to the nose when faces were presented centrally and was located 2.5 degrees or 6 degrees away from the inner edge of the face when faces were presented parafoveally or peripherally, respectively. All faces were presented for 150ms. The additional distances revealed that for all expressions except happy faces there was a gradual decline in emotion recognition performance with increasing eccentricity. Therefore, when faces are viewed centrally, the high spatial frequencies obtained by foveal examination contribute to emotion recognition performance, and their gradual removal leads to respective gradual decline in emotion recognition, possibly due to the decline in visual detail that can be obtained from facial features. In an additional experiment, Calvo et al. (2014) presented either only the eyes (including the brow and eyebrows) or only the mouth centrally and peripherally (6 degrees away from fixation). When the eyes were presented centrally and peripherally, the recognition of anger and surprise was better compared to other expressions. This recognition advantage for anger might have resulted from the presence of the informative furrowed brow region; however, this should be interpreted along with the finding that peripheral presentation of the eye region led to a decrease in emotion recognition for all expressions. Therefore, it is reasonable to assume that foveal processing of the eye region will lead to higher emotion recognition for anger compared to a condition where HSFs are reduced. However, since the eye region contained the region of the eyes and eyebrows as a strip it is difficult to speculate whether it is the furrowed brow that might have led to a recognition advantage for angry faces. Furthermore, when the mouth region was presented peripherally, the recognition of fearful expressions dropped below chance levels. Since enforced fixation on the mouth region led to increased emotion recognition accuracy in the experiments reported in this thesis, the high spatial frequency information at the mouth region appears important for the recognition of fearful expressions. A further comparison by Calvo et al. (2014) of how central and peripheral presentation of the whole face or a face part affect emotion recognition performance showed that while the sole presentation of the mouth in the periphery impairs the recognition of anger and fear more than peripheral presentation of the whole face and only eyes, there was no difference between the impairment caused by the peripheral versus central presentation of the eyes and the whole face. Therefore, some information preserved within the eye region might aid peripheral recognition of fearful and angry faces.

While existing evidence suggests that different expressions show different susceptibilities to spatial frequency content modulation, a decline in emotion recognition accuracy with decreasing HSF information is equivocally reported. While the above studies gave us an insight into how decline in spatial frequency content of a whole face affects emotion recognition, the contribution of spatial frequency decline of a specific facial feature is rarely researched. As previously discussed, when a non-informative facial feature was fixated in our previous studies, the informative facial feature was still processed in the extrafoveal regions. However, when these informative regions were fixated - when they were processed foveally - we recorded an increase in emotion recognition performance for certain expressions. Therefore, it seems that foveating an informative feature and processing the feature extrafoveally leads to a difference in emotion recognition performance.

In this study, we investigate whether the improvement in emotion recognition when fixating an informative facial feature can be explained by the extrafoveal blurring of the informative feature. We use the brief-fixation paradigm used in our previous studies with several changes:

First, in this experiment, we use the brow, the mouth and the midpoint between these features as the initial fixation locations. The selection of these locations was informed by the results of our previous studies which showed that initially fixating on the mouth improved emotion recognition for fear, surprise and disgust and initial fixation on the brow reduced (albeit not significantly) the number of neutral responses to angry faces. Second, as well as the original images of facial expressions of anger, fear, surprise and disgust, we used images in which the initial fixation locations were masked by either blurring, to eliminate the high spatial frequencies, or occluded. The size of the masks was chosen to cover the fovea respective to the enforced initial fixation location: The centre of the mask was centred on the coordinates of the initial fixation location. To answer our research question, the blurring simulated the extrafoveal blurring expected when the non-informative facial feature was fixated; more specifically, the blurring of the mouth simulated the extrafoveal blurring of the mouth when the brow was fixated, and the blurring of the brow simulated the extrafoveal blurring of the brow when the mouth was fixated. As well as the blurred fixation locations, in separate trials we occluded the fixated location in order to prevent foveal processing of the fixated feature and force the participant to rely on the extrafoveal processing of the non-occluded facial features. If the emotion recognition advantage we have observed previously was caused by high spatial frequencies processed when foveating the informative feature, we would expect to find a decrease in emotion recognition when the informative feature is blurred compared to the original image. On the other hand, if the improvement is simply due to the foveal processing of the informative facial feature regardless of spatial frequency content, we would expect to find no recognition difference between the original image and the blurred image but a decrease in recognition when the informative feature is occluded. In the light of Atkinson and Smithson's (2020) finding of improved anger recognition at initial fixation on the brow and our finding of reduced neutral responses for anger when fixating the brow region, we expected that blurring and/or masking the initially fixated brow will lead to a decline in anger recognition. We also expected that blurring and/or masking the initially fixated mouth region will lead to emotion recognition decline for fear, surprise and disgust since we previously showed a contribution of initially fixating this feature to recognition of these expressions. However, if the improvement in emotion recognition performance was not due to processing of high spatial frequency information, it is possible that we will not observe any difference in recognition performance.

Methods

Participants

A total of 40 participants took part in the experiment but data from only 36 participants were analysed due to technical problems with some of the participants. The 36 participants had a mean age of 19.6 (SD= 3.88) years and comprised 35 females and 1 male. All participants had normal or corrected-to-normal vision. Sample size was selected to be similar to the sample size of studies with similar design. Participants gave written consent to take part in the experiment and were rewarded participant pool credit for their participation. The study was approved by the Durham University Psychology Department Ethics Sub-committee.

Stimuli

12 facial identities were used from the Radboud Face Database (Langner et al., 2010). All faces were of Caucasian adults (6 females, 6 males) with frontal pose and gaze. Angry, fearful, surprised, disgusted and neutral poses of each identity were selected for the experiment leading to a total of 60 images. The same subset of expressions used in Experiment 2 and Experiment 3 reported in Chapters 2 and 3 were used – anger, fear, surprise and disgust. The reason behind choosing these expressions is explained in more detail in Chapter 2. The same subset of expressions was used here in order to draw links between the results of this study and the results

Figure 34. Examples of images used in the experiment. The first row shows the mask condition for the initial fixation locations of brow, midpoint and mouth and the second row shows the low-pass filter condition for the same initial fixation locations. While the blurring resulting from the low-pass filtering of the fixation location is not obvious in this figure, the images presented to the participants during the experiment were larger, making this blurring more apparent.

Figure 35. The three fixation locations (brow, center and mouth) superimposed on an example angry expression.

reported in previous Chapters. We included the neutral face images in this study in order to emphasize the importance of the brow region for discriminating angry faces from neutral poses. All images were in colour. For all the facial identities and facial expressions two types of masks were created: a low-pass filter mask which filtered out the high spatial frequencies (low-pass filter condition) and a mask which occluded the initial fixation location (occluded condition). Both masks covered three degrees of visual angle around the brow, the mouth or the midpoint between the two. The mask overlaid onto the colour images used the average colour of the area to be masked. The low-pass filter used mimicked the decline in visual acuity at the target region when the other main initial fixation location was fixated. More specifically, the filter at the mouth region mimicked the decline in visual acuity expected at the mouth region when an observer is fixating on the brow region, and the filter at the brow region mimicked the decline in visual acuity if fixation were at the centre of the mouth. To generate the face images with the SF-filtered features, we used components of Gielser and Perry's (Giesler & Perry, 1998; Perry & Giesler, 2002) Space Variant Imaging System toolbox (http://svi.cps.utexas.edu/software.shtml) to generate images of faces that were blurred to match the variable blurring induced by the retina for a given fixation location on the face. We then overlaid the resulting face image with the corresponding original image using alpha blending in MATLAB so that the resultant stimulus image showed the unaltered face with the target region blurred. In addition to the masked images, the original, unmasked versions of the same images were also used. An example of masks used in the experiment can be seen superimposed onto an angry expression can be seen in Figure 34 and examples of mask conditions for the remaining expressions can be seen in the Appendix (Figure A11 and A12). Finally, an example of the position of the new initial fixation locations can be seen in Figure 35.

Design

A within-subjects design was used for the experiment with *expression* of the face (anger, fear, surprise, disgusted and neutral), *initial fixation location* (brow, midpoint and mouth) and *mask* condition (original image, low-pass filtered and occluded) as independent variables. There were 4 blocks of 135 trials each. Within each block, each expression was presented at each initial fixation location with each mask 3 times. The order of presentation was randomised for each participant.

Procedure

The procedure of this experiment was similar to that of the previous brief-fixation experiments reported in this thesis. The experiment was executed and controlled using the MATLAB programming language with the Psychophysics Toolbox extensions. To control stimulus presentation and to measure gaze behaviour, we used an EyeLink 1000 desktop-mounted eye-tracker (SR Research, Mississauga, ON, Canada), recording at 500 Hz. Each stimulus block was started with the default nine-point calibration and validation sequences. Recording was binocular but only data from the left eye was analysed since data from the left eye started the trial in a gaze-contingent manner. Default criteria for fixations, blinks, and saccades implemented in the EyeLink system were used. Each trial started with a fixation cross located at one of 25 possible locations on the screen. This was to make both the exact screen location of the fixation cross and the to-be-fixated facial feature unpredictable. These 25 possible locations for the fixation cross were at 0, 25, or 50 pixels left or right and up or down from the center of the screen. These fixation-cross positions were pseudo-randomly ordered across trials. Faces were presented in a gaze-contingent manner on each trial: The participants needed to fixate within 0.6 degrees of visual angle of the fixation cross for 6 consecutive eye-tracking samples following which a face showing one of the four target expressions was presented. The location of the fixation cross corresponded to one of the informative (brow or mouth) or non-informative (midpoint) initial fixation locations for each of the presented faces. The initial fixation locations were reduced from the previously reported studies since we have found that the brow might reduce misclassification of anger as neutral and that the mouth was informative for the recognition of fearful, surprised and disgusted facial expressions. Each face was presented for 82.4 ms (7 monitor refreshes) on a monitor with an 85Hz refresh rate. Following the face presentation, the participant pressed a key on a QWERTY keyboard to indicate their answer. The row of number keys near the top of the keyboard were used, with 4 for anger, 5 for fear, 8 for surprise, 9 for disgust and 0 for neutral. The keys were labelled A, F and Su, D and N from left to right and the order of these keys remained the same for each participant. Participants pressed the A and F keys with the left and the Su, D and N keys with the right hand. This configuration was chosen to optimize the reach of the participants to the keyboard from either side of the chinrest. Participants were again asked to memorise the keys to avoid looking down on the keyboard during responding. The next trial started once a valid response had been registered for the presented face.

Results

Emotion Recognition Accuracy

The average reaction time for this task was 1.36s (SD=2.07). The average reaction time for each condition and a visualization of the reaction time data distribution can be seen in the Appendix, Table A11 and Figure A13. Responses made within 200msec were removed from further analysis since they might indicate automatic responding rather than result from genuine perceptual process. In total, 0.004% (8 trials in total) of all trials were removed.

The unbiased hit rates are summarised in Figure 36. A three-way repeated measures ANOVA was carried out on the unbiased hit rates using expression (anger, fear, surprise, disgust and neutral), initial fixation (brow, mouth and centre/midpoint) and mask condition (original, low-pass filtered and occluded) as factors. Where the assumption of equal variances was violated, the Greenhouse-Geisser corrected values are reported. There was a main effect of expression ($F_{(2.36, 82.65)}$= 80.65, p< .001, η_p^2= .70) which indicated that the neutral faces were better

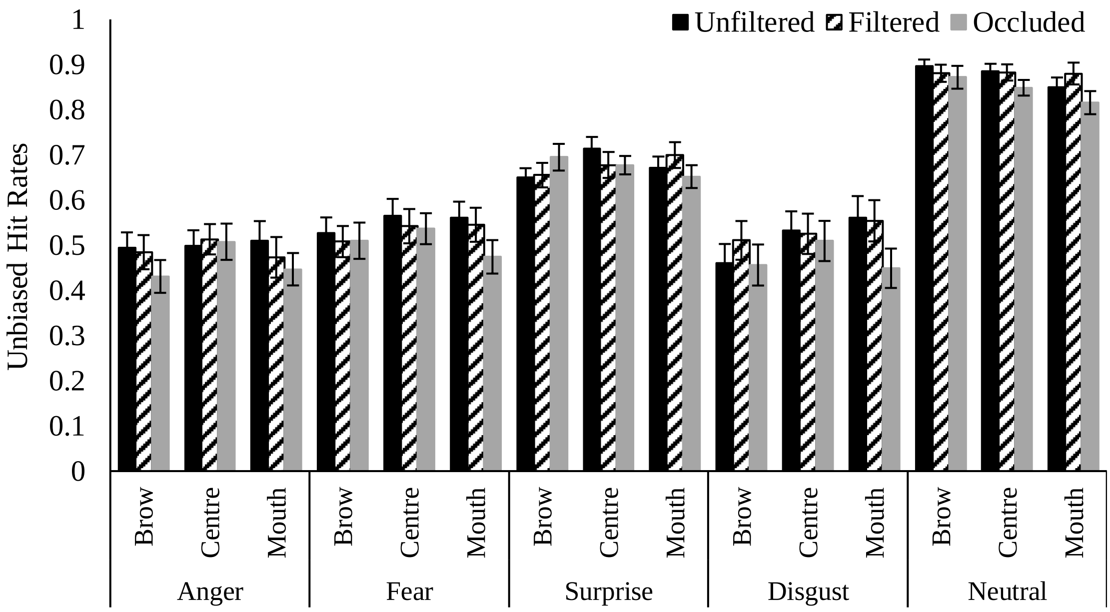

Figure 36. The unbiased hit rates for the five expressions at each initial fixation with each mask condition.

recognised compared to all other expressions (all ps< .001) and surprised expressions are better recognised compared to all other expressive faces (all ps< .001). There was also a main effect of the mask condition ($F_{(2, 70)}$= 7.26, p= .001, η_p^2= .17) which indicated that occluding the initial fixation locations led to lower accuracy compared to the original (p= .002) and low-pass filtered faces (p= .036). The main effect of initial fixation was not significant ($F_{(1.69, 59.20)}$= 1.55, p= .22, η_p^2= .04). The interaction between emotion and initial fixation ($F_{(5.64, 197.47)}$= 2.02, p= .069, η_p^2= .055) did not survive Greenhouse-Geisser correction. There was no interaction between emotion and mask condition ($F_{(4.64, 162.42)}$= 1.28, p= .28, η_p^2= .035) or initial fixation and mask condition ($F_{(3.65, 127.60)}$= 2.10, p= .09, η_p^2= .06). The three-way interaction did not survive Greenhouse-Geisser correction ($F_{(8.92, 312.11)}$= 1.71, p= .087, η_p^2= .05).

To further investigate the non-significant trend for the three-way interaction, two-way repeated measure ANOVAs were carried out separately for each of the three initial fixation locations. This allows us to compare the effect of type of masking on a single initial fixation location and whether this varied with expression.

For the brow region, there was a main effect of expression ($F_{(2.46, 86.12)}$= 72.52, p< .001, η_p^2= .67) and an interaction between expression and mask condition ($F_{(4.20, 146.91)}$= 2.44, p= .046, η_p^2= .07). The effect of mask condition did not reach significance ($F_{(2, 70)}$= 0.40, p= .67, η_p^2=.01). The interaction was followed up by conducting separate one-way ANOVAs comparing emotion recognition accuracy for each facial expression across mask conditions. For angry expressions, the main effect of mask condition approached significance ($F_{(2, 70)}$= 2.86, p= .064, η_p^2= .08). A visual inspection of the unbiased hit rate data implies that occluding the brow region leads to lower anger recognition compared to low-pass filtering and the original image. Pairwise comparisons revealed no difference between mask conditions. In order to test our specific hypothesis relating to the effect of masking of the brow region for the recognition of angry expressions, planned comparisons were carried out to test whether anger accuracy differed when participants fixated the low-pass filtered or occluded brow relative to the brow of the original face image. There was no difference in accuracy between fixation of the low-pass filtered and unaltered brow ($t_{(35)}$= 0.509, p= .307). Anger accuracy when participants fixated the occluded brow was significantly lower than when they fixated the unaltered brow ($t_{(35)}$= 2.19, p= .018). The results regarding the informative brow region of angry expressions partially support our hypothesis. As expected, there was a decline in anger recognition accuracy when the emotion-informative brow region was occluded. However, contrary to our expectation, filtering out the high spatial frequency information from the emotion-informative brow region did not affect anger recognition accuracy. There was no effect of mask condition for the brow region for fearful faces ($F_{(2, 70)}$= 0.40, p= .67, η_p^2= .011), surprised faces ($F_{(1.56, 54.56)}$= 2.86, p= .08, η_p^2= .076), disgusted faces ($F_{(2, 70)}$= 2.65, p= .08, η_p^2= .070) or neutral faces ($F_{(2, 70)}$= 0.65, p= .52, η_p^2= .018).

For the fixation location corresponding to the mid-point between the brow and the mouth, there was a main effect of expression ($F_{(2.45, 85.85)}$= 57.08, p< .001, η_p^2= .62). There was no main effect of mask condition ($F_{(2, 70)}$= 0.61, p= .55, η_p^2= .02) nor an interaction between expression and mask condition ($F_{(5.24, 183.32)}$= 0.65, p= .67, η_p^2= .02).

For initial fixation at the mouth, there was a main effect of expression ($F_{(2.79, 97.53)}$= 60.77, p< .001, η_p^2= .64) and a main effect of mask condition ($F_{(2, 70)}$= 12.19, p< .001, η_p^2= .26). There was no interaction between expression and mask condition ($F_{(5.72, 200.14)}$= 1.48, p= .19, η_p^2= .04). Since we predicted that filtering and/or occluding the mouth region of the fearful, surprised, and disgusted expressions would impair recognition accuracy, we compared emotion recognition accuracy for these expressions across mask conditions by one-way ANOVAs. The main effect of mask condition for fear ($F_{(2, 70)}$= 6.36, p= .003, η_p^2= .15) showed that, in line with our hypothesis, fixation on the occluded mouth region led to lower recognition accuracy compared to fixation on the mouth of the original image (p= .002) and fixation on the low-pass filtered mouth (p= .029). On the other hand, contrary to our expectations, there was no difference between fixation on mouth of the original image and fixation on the low-pass filtered mouth. There was no effect of mask condition for surprised expressions ($F_{(2, 70)}$= 2.20, p= .12, η_p^2= .059) which was contrary to our expectation that occluding and/or low-pass filtering of the emotion-informative mouth region would impair recognition of surprised expressions. The main effect of mask condition for disgust ($F_{(2, 70)}$= 9.38, p< .001, η_p^2= .21) showed a decrease in emotion recognition accuracy when fixation was on the occluded mouth area compared to both fixation on the mouth of the original image (p= .001) and the low-pass filtered mouth (p= .014) as suggested by the hypothesis that mouth is emotion-informative for disgust. There was no difference between fixation on the mouth of the original image and fixation on the low-pass filtered mouth, which denies our suggestion that high spatial frequency information at the emotion-informative mouth region contributes to disgust recognition. The main effect of mask

Table 8. Misclassifications for the brow region. The first column represents the presented expressions while the rows represent the % of times each response label was used when the target expression was presented.

	Anger	**Fear**	**Surprise**	**Disgust**	**Neutral**
Unfiltered					
Anger	71.18	8.17	1.29	**16.99**	**2.37**
Fear	1.28	65.95	**24.84**	7.49	0.43
Surprise	0.86	8.99	85.44	3.21	1.50
Disgust	**30.77**	1.07	2.35	64.74	1.07
Neutral	3.86	0.21	0.43	1.72	93.78
Filtered					
Anger	67.45	9.21	2.78	**15.20**	**5.35**
Fear	2.15	63.01	**26.45**	7.74	0.65
Surprise	0.86	7.07	86.94	3.64	1.50
Disgust	**27.14**	0.21	2.14	68.80	1.71
Neutral	3.00	0.43	0.64	0.43	95.50
Masked					
Anger	63.25	10.68	1.50	**15.81**	**8.76**
Fear	1.08	64.52	**24.52**	9.03	0.86
Surprise	0.64	7.91	87.82	2.35	1.28
Disgust	**31.33**	1.72	3.00	63.73	0.21
Neutral	3.87	0.65	0.65	1.29	93.55

condition for neutral faces was almost significant ($F_{(2,70)}$= 3.10, p= .051, η_p^2= .081). This main effect indicated that fixation on the occluded mouth region led to lower neutral expression recognition compared to fixation on the mouth of the original face image.

Confusion matrices

To investigate whether there were systematic misclassifications between the expression categories investigated, we have tabulated confusion matrices separately for each initial fixation location and mask condition. One-way repeated measures ANOVAs were conducted as response option as a factor to compare percentage of emotion categories used as response options for each of the expressions for each initial fixation location separately. Then, the effect of mask condition on the most common misclassifications was measured using a one-way repeated measures ANOVAs with mask condition as a factor.

Table 8 shows the misclassifications at the brow region. When the brow region was initially fixated, there was a main effect of response for anger ($F_{(1.75, 61.40)}$=13.57, p<.001, η_p^2= .28), fear ($F_{(1.25, 43.66)}$=53.41, p<.001, η_p^2= .60), surprise ($F_{(1.93, 67.56)}$=18.15, p<.001, η_p^2= .34) and disgust ($F_{(1.19, 41.66)}$=47.11, p<.001, η_p^2= .57). Angry expressions were more often misclassified as disgust compared to surprise and neutral (both ps< .001) and least often as surprised. There was no effect of mask condition on misclassification of anger as disgust

Table 9. Misclassifications for the central region. The first column represents the presented expressions while the rows represent the % of times each response label was used when the target expression was presented.

	Anger	**Fear**	**Surprise**	**Disgust**	**Neutral**
Unfiltered					
Anger	68.31	8.78	2.57	**13.49**	**6.85**
Fear	2.78	68.38	**20.09**	7.69	1.07
Surprise	0.00	5.56	89.96	2.78	1.71
Disgust	**24.84**	0.64	3.00	70.45	1.07
Neutral	1.07	0.43	0.64	1.72	96.14
Filtered					
Anger	69.53	8.37	1.72	**13.30**	**7.08**
Fear	1.72	67.10	**21.72**	7.96	1.51
Surprise	0.86	6.64	86.08	3.64	2.78
Disgust	**27.41**	0.86	2.57	68.52	0.64
Neutral	0.86	0.00	0.64	0.43	98.07
Masked					
Anger	69.81	7.49	2.36	**14.78**	**5.57**
Fear	2.58	65.67	**22.32**	7.73	1.72
Surprise	0.21	6.41	88.46	3.21	1.71
Disgust	**27.41**	1.71	3.85	66.17	0.86
Neutral	2.79	0.21	0.00	1.07	95.92

($F_{(2, 70)}$=.22, p=.80, η_p^2= .006). There was an effect of mask condition on the misclassification of anger as neutral ($F_{(1.58, 55.37)}$=7.18, p=.003, η_p^2= .17). Masking the brow region led to higher misclassification of anger as neutral compared to unfiltered brow. Fear was most often misclassified as surprise (ps< .001) and surprised expressions were most often misclassified as fear (anger: p< .001; disgust: p= .015; neutral: p< .001). Finally, disgusted expressions were most often misclassified as anger (ps< .001).

Table 9 shows the misclassifications at the central region. When center location was initially fixated, there was a main effect of response for angry ($F_{(1.82, 63.69)}$=7.21, p=.002, η_p^2= .17), fear ($F_{(1.50, 52.53)}$=37.38, p<.001, η_p^2= .52), surprise ($F_{(2.33, 81.70)}$=10.22, p<.001, η_p^2= .23) and disgust ($F_{(1.18, 41.16)}$=54.56, p<.001, η_p^2= .61). Angry expressions were least often misclassified as surprise (fear: p= .042; disgust: p< .001; neutral: p< .001). Fear was most often misclassified as surprise (all ps< .001) and more often misclassified as disgust compared to anger (p= .002) and neutral (p< .001). Surprised expressions were more often misclassified as fear compared to anger (p< .001) and neutral (p= .017) and as disgust compared to anger (p= .018). Finally, disgusted expressions were most likely misclassified as anger (all ps< .001). Since we did not have specific expectations of which expression center region would be informative for, no further analysis was conducted as to whether there was an effect of mask condition on these misclassifications.

Table 10. Misclassifications for the mouth region. The first column represents the presented expressions while the rows represent the % of times each response label was used when the target expression was presented.

	Anger	**Fear**	**Surprise**	**Disgust**	**Neutral**
Unfiltered					
Anger	65.52	7.92	1.71	**16.70**	**8.14**
Fear	2.58	68.88	**19.74**	7.73	1.07
Surprise	0.21	8.55	85.47	4.06	1.71
Disgust	**19.70**	0.86	3.85	74.95	0.64
Neutral	1.28	0.00	0.85	1.28	96.58
Filtered					
Anger	63.50	8.86	1.94	**15.98**	**9.72**
Fear	3.63	67.74	**19.44**	8.33	0.85
Surprise	0.64	7.91	86.97	3.63	0.85
Disgust	**22.22**	1.07	1.50	74.57	0.64
Neutral	1.50	0.43	1.50	1.93	94.64
Masked					
Anger	67.31	5.56	2.35	**11.75**	**13.03**
Fear	4.06	57.48	**26.71**	7.91	3.85
Surprise	0.64	8.15	87.77	2.79	0.64
Disgust	**32.40**	1.07	5.58	60.30	0.64
Neutral	2.78	0.43	0.64	1.28	94.86

Finally, Table 10 shows the misclassifications for the mouth region. When mouth was fixated, there was a main effect of response for angry ($F_{(2.25, 78.77)}$=7.56, p=.001, η_p^2= .18), fear ($F_{(1.47, 51.45)}$=38.67, p<.001, η_p^2= .53), surprise ($F_{(1.49, 52.14)}$=23.06, p<.001, η_p^2= .40) and disgust ($F_{(1.17, 41.07)}$=52.83, p<.001, η_p^2= .60). Anger was misclassified as disgust and neutral more compared to surprise (both *p*s< .001). Fear was most often misclassified as surprised (all *p*s< .001) and surprised expressions were misclassified as fear (anger: p< .001; disgust: p= .017; neutral: p< .001). Surprised expressions were also more often misclassified as disgust compared to anger (p< .001) and neutral (p= .012). Disgust was misclassified most often as anger (all *p*s< .001). There was a main effect of mask condition on the misclassification of fear as surprise ($F_{(2, 70)}$=4.30, p=.017, η_p^2= .11). When the mouth region was masked, the misclassification of fear as surprise was more frequent compared to when the mouth was filtered (p=.039). There was no effect of mask condition on the misclassification of surprise as fear ($F_{(2, 70)}$=.46, p=.63, η_p^2= .013). The effect of mask condition for the misclassification of disgust as anger ($F_{(1.43, 49.88)}$=12.16, p<.001, η_p^2= .26) showed that the misclassification of disgust as anger was more frequent when the mouth region was masked compared to the filtered (p< .001) and original image (p< .001).

Discussion

This study aimed to investigate whether the contributory effect of initially fixating an informative feature resulted from processing of high spatial frequency information at the fixated feature. To test this, we presented participants with expressions of anger, fear, surprise, and disgust, along with emotionally neutral faces. A single fixation was enforced at one of three locations on the face: the brow, the mouth or the midpoint between the two. These initially fixated features were untouched, low-pass filtered or occluded in order to establish the effect of the presence of high spatial frequency information at the fixated informative feature. If foveating an informative facial feature led to emotion recognition improvement over the extrafoveal presentation of this feature due to the processing of high spatial frequencies, we expected to see reduced emotion recognition for the filtered faces compared to the original faces. On the other hand, if the emotion recognition improvement was due to early processing of the informative feature, regardless of spatial frequency content, we expected no difference between the original and the filtered face but a decline when the informative region was occluded.

As previously suggested by Smith et al. (2005) and Merlusca and Smith (2014), the brow region contains useful high spatial frequency information for anger recognition. Furthermore, despite not reaching significance, we reported in the previous chapter a decline in the frequency of neutral responses to lower intensity angry expressions when initial fixation was at the brow region. Therefore, we hypothesised that initially fixating on the brow region leads to an emotion recognition advantage for anger via the processing of high spatial frequencies at the brow region. However, we found no difference in emotion recognition performance between enforced fixation on the untouched and filtered brow for angry faces. However, we found that occluding the brow region when at fixation led to a decline in anger recognition compared to the untouched brow. While this refutes our first hypothesis regarding the informativeness of high spatial frequencies from the fixated brow region for anger, it agrees with our alternative hypothesis that the emotion recognition advantage of initially fixating the brow results from fixating the brow regardless of spatial frequency content. While we found effect of masking of brow region for other expressions, the reduction in emotion recognition accuracy was specific to anger. Therefore, we can argue that the visual processing of the brow region in central vision contributes specifically to anger recognition.

From the findings of the previously reported brief-fixation experiments in this thesis, we established that foveal processing of the central mouth region led to improved emotion recognition for disgust and for fear and surprise to some extent. Therefore, if this improvement was due to processing of HSFs by the fovea at fixation, we expected a decline in emotion recognition for fear, surprise and disgust when the mouth at fixation was blurred. However, if the foveal processing of the informative feature, regardless of spatial frequency content, leads to the recognition improvement then we expected a decline in recognition accuracy when the region was occluded compared to the original, untouched faces.

We found no effect of masking condition for fixation on the mouth of surprised faces. Previous research showed that the recognition of surprised expressions was the most resistant to image resolution and blurring manipulations (Calvo et al., 2014, 2010; Du & Martinez, 2011; Guo et al., 2019). We found no difference in emotion recognition accuracy between the original or low-pass filtered faces for fearful and disgusted expressions when initial fixation was on the mouth. However, similar to our finding for anger, occluding the mouth region led to a drop in recognition accuracy for fearful, disgusted and neutral faces. While the effect of occluding the mouth on emotion recognition appears to apply to almost all expressions studied here, the confusion matrices show that occluding the mouth led to an increase in the misclassification of disgust as anger and fear as surprise. These effects were asymmetrical, meaning that surprised expressions were not misclassified as fearful to the same extent. The same was true for angry expressions misclassified as disgust. This is in line with our previously reported results suggesting that initially fixating on the mouth reduced these misclassifications. However, since the accuracy difference between the original and the low-pass filtered face was minimal, the reduction of this misclassification is unlikely to be caused solely or perhaps even at all by the processed HSF at the mouth region. The effect of masking the mouth region on fear and disgust recognition should be evaluated along with our finding that occluding the mouth region also led to a decline in recognition of neutral and angry faces. While in our previous experiments we found the mouth region to be relatively more informative for fearful and disgusted face, the non-specificity of the masking effect in this experiment might mean that rather than being informative for these specific expressions, the mouth region might be important for the recognition of the expressions involved in this experiment overall. However, since occluding the mouth region increased misclassifications of fear and disgusted expressions, it is reasonable to argue that visual processing of this facial feature contributes to the discrimination of these expressions from surprised and angry expressions respectively.

Overall, we did not find that high spatial frequency information at the fixated region leads to the previously observed emotion recognition improvement. Therefore, it is possible that despite a lack of HSF information at the foveated region, the lower SFs at the fixated location (as well as in the extrafoveal regions) was adequate for improved expression recognition. However, this does not explain why in our previous studies we observed improved emotion recognition when initially fixating an informative feature and not when this feature was in the periphery since mid- to low SF information from the facial feature deemed informative would have still been present and accessible when this feature was not fixated. For example, we previously found that enforced fixation on the mouth region improved emotion recognition accuracy for disgust compared to enforced fixation on an eye, the central brow and a cheek. In the present experiment, we find that there is no difference in emotion recognition accuracy between conditions in which the fixated mouth region was either untouched or low-pass filtered, which means that HSF information obtained at this region is not the key reason for this improvement. Similar mid- to low- SF information from the mouth region would have been available to the retina when the brow region was foveated; however, the emotion recognition improvement was only present when the mouth was foveated. Therefore, it is reasonable to speculate that fixating an informative feature aids emotion recognition via mechanisms other than the processing of HSFs.

One such mechanism might be owing to the phenomenon of crowding in the periphery. Crowding is the phenomenon where an object in the periphery becomes harder to recognize due to the presence of distractor objects nearby. The phenomenon of crowding has been widely researched for letter recognition. One such study from Bouma (1970) suggests that the presence of a non-target letter/object within half the viewing distance around the area of the target object reduces the identification of the target through interference. While there is an abundance of evidence indicating that crowding in the periphery happens between letters and objects, there are only a few studies investigating this phenomenon within objects, or between the features of a single object; for example, how parts of the face, such as the eyes, nose, the hair, crowd another feature, for example the mouth. Martelli, Majaj and Pelli (2005) used words and face caricatures to investigate whether the features within a face crowd each other. In a first experiment, they showed that while a mouth or a letter was identified when it was presented within a face or a word (respectively) compared to when presented alone, at increasing eccentricities the corresponding context (a face or a word) impairs the identification of the letter and the expression of the mouth. To further investigate whether this decline in part recognition in the periphery was due to crowding, Martelli et al. (2005) presented words and caricatured faces at eccentricities between 0 and 12 degrees and at different sizes. The spacing between the target part and the nearby parts was also manipulated. The results indicate that the distance between the target part and the nearby parts that would enable unimpaired part recognition increases with increasing eccentricity – in other words, the more peripherally a target is presented, the larger the spacing between the target and surrounding features need to be. Therefore, the results of the Martelli et al. (2005) study suggest that crowding affects the part-based recognition of faces through interference from nearby features in the periphery; however, the results also suggest that this interference is absent when the target part is presented within context foveally. Liu, Montaser-Kouhsari and Xu (2014) have also investigated whether facial features will lead to self-crowding, and whether this would in turn reduce the discriminability of the facial expression of the self-crowded face. In their experiment, they used facial caricatures similarly to Martelli et al. (2005) and created eight different crowding conditions in order to simulate crowding of the mouth by other internal features and external contours. The conditions they used can be seen in Figure 37. The mouth of the facial caricatures belonging to each crowding condition represented an expression in the happy-sad continuum (see Figure 37*j*) and the observers were required to indicate whether the mouth of the caricature images expressed sadness or happiness. The proportion of the sad responses given to each category was plotted as a sigmoidal curve where the slope of this curve represented the discriminability of the mouth. The results indicate that the discriminability of the expression of the mouth region was reduced with the crowding of the mouth region by the combination of internal and external contours, hair and only a delineating line around the features.

Combined, the results from the studies by Martelli et al. (2005) and Liu et al. (2014) suggest that individual facial features along with the external contours of a face can crowd and therefore impair the identification of single facial features. In the studies described within this paper, when the initial fixation is at the mouth region, the brow, which is the informative region for angry faces, is viewed at an eccentricity of 6.97 degrees. According to the laws of crowding, the presence of another part within 3.49 degrees around the brow might lead to crowding,

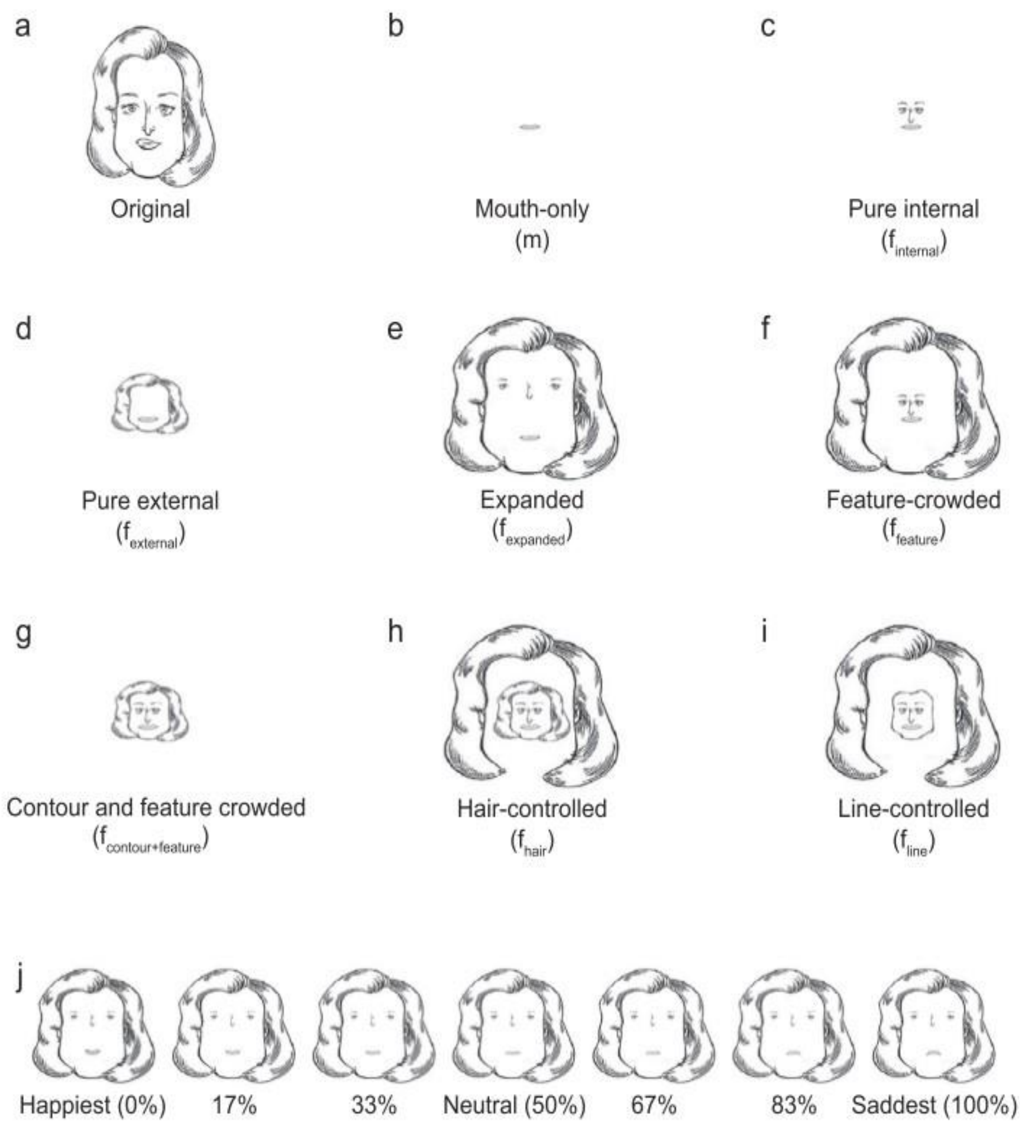

Figure 37. Image from Liu et al. (2014) summarising their experimental conditions.

therefore, interfering with the processing of this region and thus with the recognition of the emotion. In fact, the centre-to-centre distance between the brow region and either eye is 3.02 degrees of visual angle in our experiment. Therefore, it is possible that the flanking eyes interfere with the perception of the brow region in the periphery. Similarly for the mouth region, when initial fixation is at the brow, the presence of the cheeks or the face contour or the face line within 3.49 degrees around the mouth might lead to crowding making the perception of this feature less efficient and therefore resulting in a decline in emotion recognition accuracy which can only be resolved when the informative facial feature is fixated.

We find that occlusion of the initially fixated brow region reduced emotion recognition accuracy for angry expressions only. While the occlusion of the initially fixated mouth reduced recognition accuracy of all expressions except for surprised, it had a specific effect on the misclassification of fear as surprise and disgust as anger. Schurgin et al. (2014) reported similar results: When participants were asked to indicate whether a face expressed anger or no emotion, under conditions in which the eye region was occluded with a continuous horizontal strip that covered both the eyes and the eyebrows, accuracy of participants was worse compared to when the region was visible. Similarly, when the mouth region was occluded, participants were poorer at recognising expressions of happiness and disgust compared to when mouth was visible. Therefore, despite showing reasonable levels of recognition accuracy, in conditions where the informative features were occluded, the decline recorded both in the current study and in Schurgin et al.'s (2014) study indicate that the visibility of informative features contributes to emotion recognition.

A reason for the decline in emotion recognition performance when the informative facial features were occluded might be due to the inability of the participants to move their eyes to further explore the face image due to the brief presentation duration in this study (82.4msec) and in the Schurgin et al. (2014) study (100ms). Studies suggest that the optimal viewing point within a face is slightly below the eyes since a fixation at this region allows optimal extraction of visual information considering the limitations of our foveated visual systems (Hsiao & Cottrell, 2008). Van Belle, de Graef, Verfaillie, Rossion and Lefèvre (2010) showed that when foveal vision is blocked with a mask roughly the size of a facial feature, participants made more central eye movements. Furthermore, a gaze-contingent masking of foveal vision abolished the cultural differences in eye movements between Eastern Asian and Western Caucasian observers (Miellet et al., 2012). Miellet et al. (2012) showed that Western Caucasian observers, whose preferential eye movement strategy when viewing faces is fixating on individual facial features, shifted their fixations towards the centre of the face with increasing blindspot size during a face recognition task. Therefore, it is possible that when the access to an informative facial feature is limited or blocked, observers compensate with a shift in eye movement strategy in order to achieve normal task performance. Since the face presentation in our study was too brief to allow any eye movements, the participants could not resort to a compensatory strategy to obtain visual information from other facial features/regions leading to a decline in expression recognition performance compared to the original and low-pass filtered face conditions.

The mask size in this experiment was chosen to cover the area on the face that would be projected onto the fovea when the intended initial fixation location is fixated. This way, we could directly investigate the contribution of the foveal processing of the fixated facial features in studies reported in Chapters 2 and 3. However, the occluding mask and the blurring filter used do not always fully cover the mouth region of facial expressions. This is especially an issue for expression for which initially fixating the mouth was found to be informative – fear, surprise and disgust. For surprised expressions, the masks used in this experiment was not large enough to cover the open mouth (see Figure A11 and Figure A12) which allowed the participants to recognize the expression regardless of the blurring/occlusion. For fearful expressions, due to the action of the lip stretcher, the mask leaves the sides of the stretched lips visible which could have held informative value for fear recognition. For disgusted expressions, while the mask size is large enough to cover the mouth region, it is possible that the participants used information retained around the mouth region such as the nasolabial folds. This limitation could have been overcome by choosing a larger mask which would cover the whole facial feature. However, the research question to be addressed with this manipulation was whether occlusion or filtering of the part of the image that would fall on the fovea given the initial fixation location would impair emotion recognition compared to the fully available image. Therefore, the choice for the size of the masks used was guided by the size of the area that would fall on the fovea rather than facial feature size.

The drop in recognition accuracy for unfiltered angry expressions compared to Experiment 2 reported in Chapter 2 can be explained by the addition of the neutral response option. After disgusted and fearful expressions, angry expressions were most often misclassified as neutral. Neutral was not a response option in experiments reported in Chapter 2 and the addition of this response option might have reduced the accuracy of anger recognition by allowing participants to label angry expressions as neutral. Anger recognition accuracy in this experiment is closest to accuracy for angry expressions at 70% intensity reported in Chapter 3 where neutral was also a response option. For expressions of fear, disgust and surprise, the unbiased hit rates for unfiltered expressions are comparable to those in Chapter 2. Similar to the findings reported in the previous chapters and contrary to our expectations, initially fixating the brow region of angry expressions did not lead to improved recognition compared to the mouth or the center (See '*Planned Comparisons for Unfiltered Expressions*' in Appendix). The contribution of initial fixation on the unfiltered mouth to recognition of disgust expressions was as expected: Initial fixation on the mouth led to higher disgust recognition compared to the brow but not the center. However, contrary to results reported in the previous chapters, initially fixating the unfiltered mouth region did not lead to improved emotion recognition for fearful and surprised expressions compared to the unfiltered brow or center. In fact, initial fixation on the center led to numerically higher surprise recognition accuracy compared to fixating on the mouth. Even more strikingly, there was no difference in recognition accuracy for fear and surprise between initial fixations at mouth and brow. In the previous chapters, it was consistently reported that initially fixating on the mouth region led to higher fear and surprise recognition compared to upper facial features (i.e. eyes and brow). Here, while fear recognition was numerically higher when mouth was initially fixated compared to the brow, there was no difference in surprise accuracy between the two initial fixation locations. This is an unexpected result and is difficult to reconcile this with findings reported in Chapter 2. The two main differences between methods reported in Chapter 2 and here are: (1) The addition of neutral faces. (2) The addition of the center fixation location. As shown by Smith et al. (2005), both fearful and surprised expressions transmit visual information useful for expression recognition in both the eye and mouth regions. When given the opportunity to look at the center of the face, it is possible that participants integrated mid- to low spatial frequency information from the eye and mouth regions into a holistic representation of the expression which made the center of the face more informative compared to the mouth. It is more difficult to reconcile the discrepancy in results regarding the improvement in fear and surprise recognition when initially fixating the mouth compared to the brow. An examination of the percentage accuracy for fear recognition when initially fixating the brow does not suggest a substantial difference between Chapter 2 and results reported here. There is, however, a drop in fear recognition when the mouth is initially fixated. Recognition of surprised expressions is lower compared to Chapter 2 both at initial fixation on the brow and the mouth.

Neutral faces had the highest accuracy among all expressions in this study: They were rarely labelled as any other expression and neutral response option was very rarely chosen for any of the other facial expressions (see the low % values in Tables 8, 9 and 10). This might be because the decision to respond ‘neutral’ to a face image is a result of a binary expressive/non-expressive decision. On the other hand, once an expression is detected in a face image, a further decision of what expression is being depicted should be made. With multiple emotion response options available, there is more room for participants to make errors in response to expressive compared to non-expressive face images.

In conclusion, our results suggest that the contribution of foveal processing of informative facial features to emotion recognition found in the previous experiments reported in this thesis is not due to processing of high spatial frequency information from the fixated region as initially hypothesized. While filtering the high spatial frequencies from the fixated facial feature did not lead to impaired emotion recognition performance compared to the original images, our results suggest that occluding these informative features hindered emotion recognition, emphasizing their contribution to recognition. We suggested that crowding of the informative features in the periphery might be one of the reasons leading to the observed emotion recognition improvement when informative features are initially fixated. Future studies are necessary to test this suggestion formally, as well as seeking further potential explanations.

Chapter 5

General Discussion

Contribution of foveal processing of informative facial features to emotion recognition

For *angry expressions*, we hypothesized that a single fixation of the central brow region in the brief-fixation paradigm will improve emotion recognition, relative to a single fixation on a less informative region such as a cheek or an eye, via the processing of informative HSF information from the brow. Previous studies showed that the brow region is informative for the recognition of angry expressions, especially the mid-to-high SF information from the furrow lines (Smith et al., 2005; Smith & Merlusca, 2015; Smith et al., 2018). Furthermore, using the same brief-fixation paradigm as ours, Atkinson and Smithson (2020) showed that initially fixating the central brow region led to improved anger recognition accuracy associated with a decrease in misclassification of angry expressions as neutral. The results reported in Chapter 2 suggest no consistent evidence to support our hypothesis regarding foveal processing of the brow region and anger recognition. Whereas in Experiment 1 we found that initially fixating on the brow improved anger recognition compared to initially fixating the cheeks, in Experiment 2 we found that initially fixating the brow led to *reduced* accuracy compared to initially fixating the cheeks and the mouth.

Despite some evidence for the informativeness of the brow region compared to the cheeks when expressions of anger, fear, surprise and sadness were used, the finding of reduced anger recognition accuracy compared to fixating the mouth and the cheeks with a different combination of expressions (using disgust instead of sadness) suggests that the brow region is not consistently informative for anger recognition. One possible reason for this can be the combination of expressions used in the experiment. Smith and Merlusca (2014) showed that the informativeness of facial regions changes depending on the subset of facial expressions of the six basic emotions that are used in the experiment. In Experiment 2, where fixating on the brow region impaired emotion recognition accuracy, one of the four expressions was disgust – an expression often misclassified as anger (Du & Martinez, 2011; Ekman, Friesen, Sullivan, et al., 1987; Guo, 2012; Widen et al., 2004). Jack, Garrod and Schyns (2014) suggest that this confusion is due to the activation of the nose wrinkler and lip funneler during the temporal unfolding of these two expressions, and the confusion is resolved by the later activation of the upper lip raiser. Since the nose wrinkler activation is also associated with a weak activation of the brow lowerer AU, it is possible that initially fixating on the brow will have led to increased confusions between these expressions resulting in lower accuracy for anger.

On the other hand, initially fixating on the brow reduced the number of neutral responses given to angry expressions (albeit not significantly), as reported in Chapter 3. Lower intensity angry expressions received the highest number of neutral responses compared to fear, surprise and disgust. Therefore, it is possible that the brow region is informative for the discrimination of angry expressions from neutral expressions. This interpretation is in line with previous findings by Atkinson and Smithson (2020). Additionally, when asked to classify angry expressions of varying intensities among angry and neutral expressions, Schurgin et al. (2014) reported that observers spent longer fixating the nasion of neutral faces compared to other facial regions. The nasion region used in Schurgin et al.'s (2014) study is the region between the eyes and the eyebrows. The brow ROI used in the studies reported in this thesis partially comprises the nasion as well as the glabella. In line with the suggestion that the brow region might play an informative role in the discrimination of anger from neutral expressions, we also found that both masking and, to a lesser extent, filtering of the initially fixated brow region led to an increase in the misclassification of angry expressions as neutral.

Overall, while initially fixating the brow region did not consistently improve anger recognition accuracy, there is some evidence for its informativeness in discriminating anger from neutral expressions when foveally processed. Due to the differences between foveal and extrafoveal visual processing discussed in detail in Chapter 1, we suggested that the recognition improvement is due to the processing of high spatial frequency information by the fovea at the fixated region. Directly testing this hypothesis in Chapter 4 showed that while filtering the high spatial frequencies from the fixated brow region did not lead to a decline in anger recognition compared to the original (unaltered) image, occluding the brow region did. This result indicates that while HSF information from the brow region might not be contribute to anger recognition over mid- to low spatial frequencies, the visibility of this region contributes to it significantly.

The results of Experiments 2 and 4 also indicate that HSF information from the mouth region might contribute to anger recognition. The informativeness of both the furrowed brow region and the mouth region for anger recognition was established (Schyns, Petro & Smith, 2009; Smith et al., 2018). Smith et al. (2018) showed that both regions contained useful information for anger recognition across a range of SF bands. It should be noted, however, that masking of the central mouth region in Experiment 4 also led to a decline in emotion recognition for fear and disgust. Along with the finding that initially fixating the mouth region led to higher emotion recognition accuracy overall in Experiment 2, it is probable that the contribution of the mouth region to emotion recognition is not specific to anger. In line with this suggestion, using the *Bubbles* technique, Blais, Roy, Fiset, Arguin and Gosselin (2012) showed that the mouth region was most informative when discriminating between the six basic expressions as well as pain and neutral expressions. In a re-analysis of Smith et al.'s (2005) data, Blais et al. (2012) showed further that the mouth region remained the most informative region to discriminate expressions. While masking of the central mouth led to decline in emotion recognition for multiple expressions (i.e., anger, fear and disgust), the decline in emotion recognition when the brow region is masked in Experiment 4 is unique to anger recognition. This indicates a more specific contribution of the brow region to the recognition of anger.

For *fearful expressions*, we found that a single, brief fixation on the central mouth region led to an increase in fear recognition accuracy compared to a single, brief fixation on the eyes or on the brow for prototypical, full intensity expressions (Experiments 1 and 2). Initially fixating the central mouth region also improved emotion recognition accuracy of lower intensity expressions compared to the brow (Experiment 3). At the lowest intensity (30%), initially fixating the mouth led to fewer neutral responses to fearful faces compared to initially fixating the eyes and the brow. These results indicate that foveating the mouth region contributes to fear recognition over and above its extrafoveal processing. Therefore, we suggested that HSF from the mouth region is important for fear recognition. However, filtering of high spatial frequency information from the mouth region did not reduce fear recognition accuracy compared to the original image, refuting our suggestion that HSF at the mouth region is necessary for fear recognition. On the other hand, occluding the mouth region led to a reduction in fear accuracy compared to the original image. Therefore, similar to our finding for anger, even though foveal processing of the mouth region contributes to fear recognition, this is not due to the processing of HSF from this region.

Most previous studies speak to the importance of the eye region for recognition of fear. Smith et al. (2005) showed that observers make use of HSF information from the wide-open eyes of fearful expressions for fear recognition. Smith et al. (2018) additionally showed that both the eye region and the mouth region are used more compared to a baseline (hair) region when recognising fearful expressions. The informativeness of these regions, however, were SF-dependent: The eye region was more informative at higher spatial frequencies compared to the mouth and both regions were equally informative at lower spatial frequencies. Furthermore, Adolphs, Tranel, Damasio and Damasio (1994) described bilateral amygdala damaged patient, SM, who had a selective impairment of fear recognition. In a follow-up study using the *Bubbles* technique, Adolphs et al. (2005) found that while the control subjects made use of high spatial frequency information from the eye region, SM did not. As well as not utilising HSF from the eye region, Adolphs et al. (2005) also reported that SM manifested a lack of spontaneous fixation to the eye region during an emotion recognition task with unaltered face images. When asked explicitly to look at the eye region, SM's impaired fear recognition was reversed indicating that a failure to fixate and use HSF information from the eye region led to SM's fear recognition impairment.

Considering these previous results from the literature, it is reasonable to expect that initial fixation of the eye region in the brief-fixation paradigm would contribute to fear recognition more compared to initial fixation of other facial features. However, a large number of studies show that fearful and surprised expressions are confused due to their configural similarities (Du & Martinez, 2011; Ekman et al., 1987; Gagnon et al., 2010; Guo, 2012; Roy-Charland et al., 2014, 2015a; Young et al., 1997). The ideal observer analysis by Smith et al. (2005) also showed that fearful and surprised facial expressions are transmitted with a large overlap in their expression signal: Both fear and surprised expressions contain informative visual information at both the eyes and the mouth. In contrast to the ideal observer, human observers disambiguated this overlapping signal by relying more on the mouth region for surprised recognition and on the eyes for fear recognition. Research shows that the configural similarity of fear and surprised expressions leads to their misclassification and that looking at the distinctive AU (i.e., the lip stretcher of the mouth) resolves this misclassification (Chamberland et al., 2017; Roy-Charland et al., 2014, 2015a). Since the lip stretcher AU is a distinctive property of fearful expressions, initially fixating at the mouth region might have become more important for fear recognition in experiments reported in this thesis. This suggestion is supported by our finding that initially fixating the mouth region led to a reduction of misclassification of fear as surprised and masking of the initially fixated mouth led to the highest misclassification of fear as surprise. Therefore, we can conclude that visual processing of the mouth region, regardless of spatial frequency content, contributes to the resolution of the fear-surprise misclassification.

For *surprised* expressions, we found that a single, brief fixation on the central mouth region led to improved emotion recognition accuracy for both prototypical and lower intensity expressions, compared to a single, brief fixation on other regions (Experiments 1, 2 and 3). Additionally, initially fixating the mouth region reduced the number of neutral responses given to surprised expressions at 30 and 50% intensity (Experiment 3). The informativeness of the mouth region for recognition of surprised expressions agrees with the results from Smith et al.'s (2005) *Bubbles* study and extends it to show that fixating on the mouth region contributes to emotion recognition when the whole face is fully visible to the observers. While we expected that initial fixation on the mouth would lead to improvement in emotion recognition due to processing of HSF by the fovea at fixation, neither masking nor filtering of the fixated mouth altered surprise recognition compared to the original image (Experiment 4). It is possible that when the fixated mouth region is filtered or occluded, the recognition of surprised expressions proceeds through processing of the other available facial features. An alternative explanation behind this observation might be due to the size of the masks used in Experiment 4. Due to the activity of the jaw drop action unit, the mouth region of surprised expressions was larger than the size of our mask. Therefore, while we were careful to choose images where our masks covered the mouth region of surprised faces to the same degree as they did for the other expressions, the masks still did not fully cover the mouth region for some of the face images. It is possible that the remaining visible part of the mouth region was enough for participants to recognise surprised expressions with high accuracy.

For *disgusted* expressions, a single, brief fixation on the central mouth region led to higher emotion recognition accuracy compared to a single, brief fixation on the eyes or on the brow (Experiment 2). For lower intensity expressions, fixation on the mouth led to greater accuracy for disgust compared to fixation on the eyes but fixation of the brow also led to greater disgust accuracy than did fixation of the eyes (Experiment 3). Therefore, while foveal processing of the mouth region contributed to the recognition of prototypical disgust expressions, foveal processing of both the mouth and the brow contributed to disgust recognition when it was expressed at lower intensities. The informativeness of the mouth is in line with the findings of Smith et al. (2005) which showed that the observers found information around the mouth region to be useful when categorising disgusted expressions. Additionally, Schurgin et al. (2014) showed that participants spent longer fixating on the upper lip region for disgusted expressions. Furthermore, our finding of decreased disgust recognition accuracy and increased misclassification as anger when the initially fixated mouth region was masked (Experiment 4) also points to the contributory role of foveal processing of the mouth region for disgust recognition.

Overall, for the expressions we studied, we found that fixating on certain facial features led to improved emotion recognition. While some of the facial features that led to improved recognition were in line with what was suggested in the literature, some were novel findings. Due to the nature of our paradigm, when an informative feature was not fixated, it was available extrafoveally. Since specifically fixating on an informative feature was what led to recognition improvement, we suggested that differences between foveal and extrafoveal visual processing should explain this difference. Specifically, we focused on the ability of the fovea to resolve higher spatial frequencies which declines gradually with increasing eccentricity from the fovea. However, filtering of the fixated informative facial features (central brow or mouth) to exclude the highest spatial frequencies did not significantly reduce emotion recognition accuracy for anger, disgust, fear or surprise, compared to unfiltered versions of these faces (Experiment 4). It is possible that the presentation duration in our brief fixation experiment was too short to allow processing of high spatial frequencies in the first place. The visual system is suggested to combine spatial frequency information from the visual image starting from coarse information carried by low spatial frequencies to fine details carried by high spatial frequencies (Marr, 1982; Watt, 1987; Bullier, 2001). Goffaux et al. (2011) have shown that face-responsive brain regions, especially the right fusiform gyrus, responded more strongly to images consisting of LSFs compared to images with only HSFs when images were presented only for 75ms. In contrast, when presented for 150ms, this pattern reversed, and face-responsive regions were more responsive to HSF face images compared to LSF face images. Since we used a presentation duration of 82.4ms in our experiments, it is possible that recognition of facial expressions was relatively more dependent on LSF content. This could explain why there was no difference in emotion recognition accuracy when initially fixating visible informative features compared to filtered informative features. While there is useful visual information in lower spatial frequencies for emotion recognition, this fails to explain why foveal but not extrafoveal processing of features improved recognition.

We need to consider a reason other than differences in spatial frequency processing to explain the differences in emotion recognition performance resulting from foveal versus extrafoveal processing of informative facial features. One candidate explanation is *crowding*. While crowding is typically measured *between* objects, Martelli, Majaj and Pelli, (2005) and Liu, Montaser-Kouhsari and Xu, (2014) showed that features within a face also crowd each other. Martelli et al. (2005) showed that when presented centrally, observers identified the expression of a mouth better when it was presented within the whole face compared to when it was presented in isolation. This is in line with the findings from the part/whole face paradigm (Tanaka & Farah, 1993; Tanaka & Sengco, 1997), which show that the recognition of a face part is better within the face context than when presented alone. This is taken to indicate holistic processing of faces. In contrast, Martelli et al. (2005) showed that when seen at eccentricities of up to 8 degrees, the expression of the mouth was not as well identified when it was seen with the rest of the face. Martelli et al. (2005) also showed that this drop in identification performance for the face part was due to crowding of the mouth by the other internal facial features. The identification of a face part presented in the periphery improved when the distances between the facial features (mouth-to-cheek distance) were increased in caricatured faces. The distance between the mouth and the cheeks needed to improve part identification increased with increasing eccentricity. Liu et al. (2014) further showed that not only internal *facial features* but also face *contour* leads to crowding impairing the identification of the mouth expression when faces were presented peripherally. As discussed in the Discussion section of Chapter 4, while we presented faces centrally, we manipulated initial fixation location. This meant that when not initially fixated, the informative facial feature was seen in extrafoveal vision. Due to the size of our face images, it is possible that an informative facial feature (e.g., the brow) might be crowded by adjacent facial features (e.g., the left and right eye). Therefore, identification of the details of the informative facial feature which would contribute to emotion recognition would be better when this is processed centrally, where crowding is not present.

Future studies could investigate the contribution of informative facial features to emotion recognition in crowded and uncrowded faces. A possible way of testing this is to repeat the brief-fixation paradigm with larger face images. The increased image size will increase the spacing between facial features therefore reducing the crowding of single facial features. If emotion recognition accuracy in uncrowded faces were to improve compared to crowded faces when informative features are in the periphery, this would provide some evidence that crowding impairs the recognition of expressions when the informative feature is not processed foveally. However, larger images will confound the effect of feature size with the effect of spatial distance between facial features. It is known that increasing the size of objects in the periphery equates the peripheral recognition of the target object to its central recognition. A more effective way of manipulating crowding within the face would be to increase the distance between the facial features while keeping their size the same, similar to Martelli et al. (2005) and Liu et al. (2014). However, both of these studies used caricature faces instead of real-life face images. Manipulating the spatial distance between facial features within a real human face image presents several problems: (1) Changing the configural relations between facial features compromises the ecological validity of images. (2) since informative facial information sometimes lies in-between what can be identified as a 'facial feature', for example the brow region used in our studies, it might not be possible to *un-crowd* all facial features/regions that are deemed informative.

Seeking out informative facial features

In addition to examining the contribution of foveal versus extrafoveal processing of informative features to emotion recognition, we wanted to investigate whether observers would seek out informative facial features when these were not initially fixated. Across all three of our experiments (Experiments 1 & 2 in Chapter 2; Experiment 3 in Chapter 3), we consistently found that observers made a higher proportion of reflexive saccades upwards from the mouth and the cheeks compared to downwards from the eyes and brow. This finding is in line with previous findings in the literature (Boll & Gamer, 2014; Gamer & Büchel, 2009; Kliemann et al., 2012; Kliemann et al., 2010; Scheller et al., 2012). Gamer and Büchel (2009) found higher amygdala activation for fearful expressions when initial fixation was on the mouth compared to when the initial fixation is on the eyes. This result led to the suggestion that amygdala plays a role in orienting attention to the informative feature of fearful facial expressions (i.e., the eye region). In a study comparing the emotion recognition abilities of individuals with ASD and neurotypical individuals, Kliemann et al. (2012) also corroborated the finding of higher amygdala activation when initial fixation was on the mouth compared to when it was on the eyes in neurotypical subjects. Scheller et al. (2012), Boll and Gamer (2014), Kliemann et al. (2010) and Kliemann et al. (2012) also recorded higher percentages of fixation changes upwards from the mouth than downwards from between the eyes and Kliemann et al. (2010) further reported that the propensity of the ASD subjects to change their fixations upwards from the mouth was associated with higher emotion recognition.

Similar to those previous studies, we also found that the proportion of upwards compared to downwards reflexive saccades was affected by the emotion expressed by the face. More specifically, for the first experiment (reported in Chapter 2), we found that there were fewer saccades going downwards from the eyes for sad expressions compared to angry, fearful and surprised expressions. For the second experiment (also reported in Chapter 2), we found that while the percentage of saccades going upwards from the mouth was higher for the angry and disgusted expressions, there was no difference between the proportions of upwards compared to downwards reflexive saccades for fearful and surprised expressions. The expression effects on reflexive saccades, however, were not consistent across our experiments. The effects of expression reported in this thesis do not agree with previous findings. Gamer and Buchel (2009), Scheller et al. (2012), Boll and Gamer (2014); Kliemann et al. (2010) and Kliemann et al. (2012) have all found that the fixation changes upwards from the mouth compared to downwards from the eyes were stronger for fearful and neutral faces and this effect was substantially reduced or absent for happy expressions. The modulation of the probability of saccades being directed upwards from the mouth versus downwards from the eyes by facial expression suggests that these early saccades are influenced by the distribution of informative facial features of each expression. However, all these studies reported saccade orientation/direction data for a limited set of expressions - namely fear, happy and neutral (and anger for Gamer & Buchel, 2009; Boll & Gamer, 2014). Fearful and happy expressions have their informative facial features on opposite halves of the face (Smith et al., 2005) – the eyes for fear and the mouth for the happy faces. This might have accentuated the difference between the saccade directions in these studies. In the experiments reported in this thesis, we used additional facial expressions like surprise and disgust. As discussed previously, while previous studies emphasized the importance of eye region for fear recognition, we found that the mouth was more informative for fear recognition, possibly due to the confusion of this expression with surprise. Additionally, we found that initially fixating on the mouth region also improved emotion recognition accuracy for surprised and disgusted expressions. In contrast to the suggestion that reflexive saccades seek out informative facial features, we did not find increased tendency to saccade towards the mouth for any of the expressions. Additionally, Scheller et al. (2012) found higher percentage of fixation changes upwards from the mouth region not only for emotion recognition, but also for identification and gender discrimination. It is possible therefore, that the initial orientation of eye movements towards the upper face is not solely related to seeking out of informative facial features of expressions.

Many factors exert their influence on early saccades. Bindemann, Scheepers, and Burton, (2009) showed that saccades taking place within the first 250ms of face presentation during a face recognition task target the centre of the face image. This is called the *centre-of-gravity effect* and was most apparent for faces presented in profile views: When these faces were being viewed, early saccades targeted a featureless location on the face, which on a frontal face will coincide with the eyes and the nose. While it can be argued that face recognition and facial expression recognition rely on different facial cues, Guo and Shaw (2015) showed that during an emotion recognition task the fixations on the nose decreased monotonically when the viewpoint changed from frontal to mid-profile to profile view. In Guo and Shaw's (2015) study, while the fixations were mainly directed towards the eyes followed by the nose then mouth, the decrease in the number of fixations on the nose indicates a tendency for fixations to target the centre of the face. Indeed, our additional analysis of the reflexive saccades reported in the Appendix showed that these saccades ended within the nose region more than any other region (i.e., eyes, brow, mouth). Therefore, there is evidence to suggest that the reflexive saccades in our experiments were influenced by the *centre-of-gravity effect*.

The centre-of-gravity effect alone is not sufficient to explain our findings, however. If the reflexive saccades were only targeting the centre of the face, then we would expect an equal percentage of saccades upwards from the mouth and downwards from the eyes. The higher percentage of upwards saccades from the mouth indicate that the upper face provides more visual information that needs to be processed early on. Several studies showed that early saccades starting from peripheral fixation locations are directed towards a featureless region just below the eyes during a face recognition task (Hsiao & Cottrell, 2008; Or et al., 2015; Peterson & Eckstein, 2012) with a considerable downward shift during an emotion recognition task (Peterson & Eckstein, 2012). These researchers suggest that observers choose this particular location as a target of early saccades since it constitutes the *centre of information* for the face image. At this location, the visual information across all directions of the face space is equal when the variability in the spatial resolution capabilities of the retina is taken into account. This finding was mainly based on face recognition tasks and it is possible that the centre of information for facial expression recognition might be different. As can be seen in Figure 38, while the preferred landing location for the initial fixation location for the emotion recognition task in Peterson and Eckstein (2012) was lower than the preferred landing location for gender and identity tasks, it appears to be closer to the upper face half. Therefore, instead of seeking out the informative features of facial expressions, the first saccades in our experiments might be seeking out this suggested centre of information. As mentioned previously, the saccade end location analysis reported in the Appendix suggested that the first saccades in Experiments 1 and 2 had a tendency towards the centre of the face. However, since we used the nose ROI as a proxy for the centre of the face, it is possible that the reflexive saccades in our experiments ended within the subregion of the nose suggested by Hsia and Cottrell (2008). Future studies could investigate this possibility by using more comprehensive ROIs including centre of the face (Bindemann et al., 2009) and the centre of information (Hsiao & Cottrell, 2008) as well as the internal features themselves.

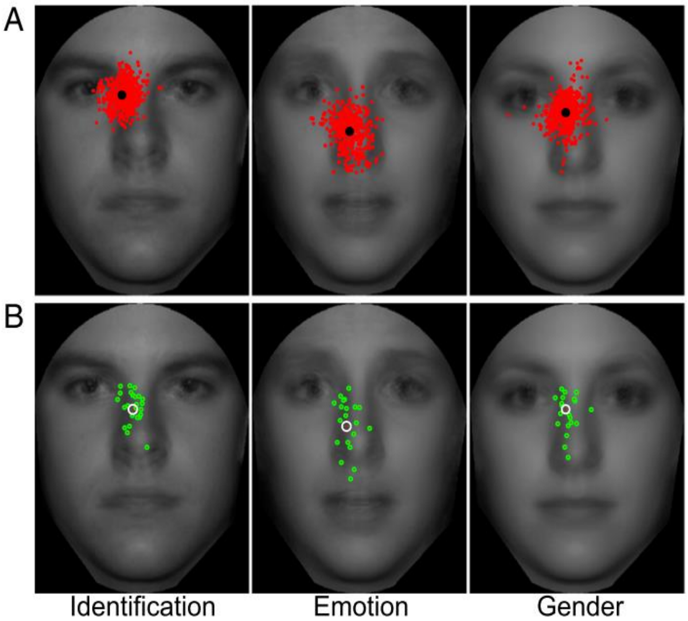

Figure 38. Image from Peterson & Eckstein (2012). The black dot in the top row and the white circle in the bottom row represent the averaged preferred fixation location of observers in this study.

In addition to these '*centre-of-gravity*' and '*centre-of-information*' explanations of our reflexive saccade data, research also suggests that fixation to the eyes during early visual scanning might be guided by top-down knowledge about face configuration (in monkeys - Guo, 2007; Guo, Robertson, Mahmoodi, Tadmor, & Young, 2003) and is automatic regardless of task demands (Laidlaw et al., 2012; Thompson et al., 2019). Laidlaw et al. (2012) showed that participants made more errors than expected by chance when they were instructed to not look at the eye region and Thompson et al. (2019) found that regardless of whether participants were instructed to look at the eyes or the mouth, the first fixations of participants targeted the eye region more frequently than the mouth region. Additionally, Thompson et al.'s (2019) instruction to look at either the eyes or the mouth was accompanied by the comment that the instruction will be useful in resolving the following recognition task. Therefore, if early eye movements are influenced by factors described above and are only partly under volitional control, these eye movements might not provide a pure measure of what observers are seeking out following initial fixation on a facial feature.

We further investigated the approximate path the reflexive saccades followed in order better to describe what these saccades might be targeting. Overall, we found that reflexive saccades were more strongly directed towards the brow compared to other saccade targets. This finding is in line with our saccade direction findings showing higher percentage of saccades upwards from the mouth and the cheeks compared to downwards from the eyes and the brow. However, in this more precise saccade path analysis, we show that these saccades target the brow more than they are targeting the eyes themselves. Among the previous studies which used a similar paradigm (Gamer & Buchel, 2009; Scheller et al., 2012; Boll & Gamer, 2014; Kliemann et al., 2010; Kliemann et al., 2012), only Gamer and Buchel (2009) shifted the face image horizontally in order to make either eye fall on the initial fixation location. In the rest of the studies, when the face image was shifted downwards, the region between the two eyes ended up becoming the initial fixation location instead of either eye individually. While these studies suggest that reflexive saccades upwards from the mouth are targeting the informative eye region for fearful faces, the definition of the eye region is quite broad. Our saccade direction analysis and saccade path analysis combined indicate that while the reflexive saccades are indeed directed upwards from the mouth and cheeks, they might not be seeking out the eyes themselves. It is possible that these saccades are targeting the centre-of-information – i.e., the region just below the eyes – suggested by Hisao and Cottrell (2008). Arizpe, Kravitz, Yovel and Baker (2012) also show that first few fixations from peripheral locations target the region just below the left eye most frequently.

We also analysed the paths of reflexive saccades individually from each initial fixation location. This analysis showed that saccades initiated from either eye more strongly targeted the brow or the other eye while saccades starting from either cheek were more strongly directed towards the mouth. An interesting pattern also emerged from our data indicating that saccades initiated from either eye more strongly targeted the opposite cheek compared to the other cheek and the mouth. This finding is partially in line with Arizpe et al. (2012). Arizpe et al. (2012) reported that the first fixation to the face from a peripheral location ended at a location on the face that was close to the initial fixation location. They further report that the second through to the fifth fixations targeted the side of the face opposite to the initial fixation location. While we find this pattern of fixations targeting the opposite side of the initial fixation location from the initial fixation location, it is likely due to the fact that the first fixation in our experiments was always on the face image. Since the aim of eye movements in visual processing is to collect information about the image, this behaviour of targeting the opposite side of the face fits in with Renninger, Verghese and Coughlan's (2007) suggestion that observers select their saccade target in order to reduce *local uncertainty.* More specifically, since at initial fixation, visual information from the fixated side of the face was sampled and processed, observers plan their next saccade to gather information from the side of the face they haven't processed yet.

Limitations and Suggestions for Future Research

Our results suggest that when initially fixating the informative feature of an expression in central view, this improves emotion recognition. This could be potentially useful in everyday situations where facial expressions are usually fleeting and sampling visual information from the expression-relevant facial feature would allow better recognition compared to expression-irrelevant facial features. However, there are several limitations that need to be considered that might restrict the generalisation of the results described here.

Several researchers showed that observers can recognize facial expressions at increasing viewing distances (Guo, 2013; Hager & Ekman, 1979; Schyns et al., 2009). The equivocal finding from these studies show that recognition accuracy declines with increasing viewing distance. With increasing viewing distances, faces will cover smaller degrees of visual angle which makes several facial features fall on the fovea at the same time and ultimately makes the whole face fall on the fovea. This makes sampling multiple facial features with fixations redundant. This is supported by the findings of Guo (2013). Guo (2013) found that with decreasing image size – simulating increasing viewing distances - fixations were directed more to the center of the faces regardless of facial expression. Therefore, the findings of this thesis do not generalize to faces seen at far distances.

Additionally, these findings only apply to faces that are seen at frontal views. In naturalistic social interactions, faces can be from different views as well as the frontal view therefore it would be beneficial to understand whether the findings of this thesis apply to facial poses other than frontal poses. As opposed to the frontal view of the face, mid-profile and profile views present less visual information that is relevant for expression recognition. Therefore, it is plausible that observers rely on different visual information in profile views compared to frontal views. Guo and Shaw (2015) showed that facial expressions seen at frontal, mid-profile and profile view could be recognized with high accuracy and fixation patterns specific to individual expressions were preserved. However, they also show a decline in recognition for expressions where informative facial features are majorly occluded by the viewpoint alteration. For example, the recognition of sad and disgusted expressions which have expression-informative features within the brow (fully occluded in profile view) and central mouth regions (fully occluded in profile view) declined more in mid-profile to profile views compared to other expressions. Additionally, Skowronski, Milner, Wagner, Crouch and McCanne (2014) showed that the identification of angry and sad expressions is impaired when shown in profile view and this is especially emphasized when the expressions are presented for 100ms. This can be due to the fact that this exposure duration is too short to make multiple fixations and since the informative feature was not immediately visible, one fixation was not adequate to obtain task-relevant information.

Furthermore, research shows that dynamic facial expressions are better recognized compared to static facial expressions. Blais, Fiset, Roy, Regimbald and Gosselin (2017) showed that there are differences in the eye movement patterns for static and dynamic expressions: Eye movements for dynamic expressions were more central compared to static expressions which attracted more fixations towards the eyes and mouth. Therefore, it is possible that individual fixations on informative features might contribute less to the recognition of dynamic expressions. This, in turn, makes the findings reported in this thesis not generalizable to dynamic expressions.

The initial fixation locations used in these experiments have symmetrical locations. However, visual performance across the vertical and horizontal meridians is not homogenous. As well as the previously mentioned decrease in the capability of high spatial frequency processing with increasing eccentricity, it is shown that visual performance along the horizontal meridian of the visual field is better than visual performance along the vertical meridian (Carrasco, Talgar & Cameron, 2001; Virsu & Rovamo, 1979). Additionally, visual performance is shown to be better in the lower visual field compared to the upper visual field (Carrasco et al., 2001; Liu, Heeger & Carrasco, 2006; Talgar & Carrasco, 2002). Crucially, Liu et al. (2006) showed that the asymmetry of visual processing between upper and lower visual fields was more pronounced for high spatial frequency stimuli. These well-established differences might mean that spatial frequencies processed from the lower facial features when fixating an upper face feature will not be equal to the spatial frequencies that could be processed from the upper facial features when a lower facial feature is initially fixated.

The non-homogenous performance for visual tasks along the vertical meridian might also play a role in determining the saccade direction of the initial saccades we investigated here. If relatively higher spatial frequencies can be processed from the lower visual filed compared to the upper visual field, it is possible that our participants prefer to saccade towards the upper visual filed in order to obtain more detailed visual information regardless of expression content present at this location. Future research can test this theory by comparing the direction of saccades from upper and lower facial features for upright compared to inverted facial expressions. If saccades in both inverted and upright faces are preferentially directed upwards from the respective lower facial features, it is possible that these initial saccades are seeking to optimise visual sampling of the whole stimulus instead of seeking expression-informative facial features.

As briefly touched upon in the Discussion section of Chapter 3, one major limitation of the studies reported in this thesis is the lack of ideal observer analysis. We have used the facial regions deemed informative by Smith et al. (2005) as informative/non-informative initial fixation locations for the expressions presented. However, it should be noted that there are several aspects of task and face images themselves which might affect the informativeness of certain facial features for the target expressions. As Merlusca and Smith (2014) showed, informativeness of certain facial features/regions is affected by the comparison categories. Therefore, it is possible that the regions we found to be informative in this thesis might not generalize to recognition of the same expressions presented with alternative expression categories. This can be further supported by the fact that using a similar paradigm, several studies run by Gamer and colleagues (Gamer & Buchel, 2009; Scheller et al., 2012; Gamer et al., 2013) did not find initial fixation at the mouth to be more informative compared to the eyes. Additionally, contrary to Atkinson and Smithson (2020), we did not find that initial fixation on the brow improved emotion recognition in Chapters 2 and 3. In addition to comparison categories, the intensity of the expression can alter the informativeness of facial features. While the model observer analysis for prototypical expressions of fear, surprise, happy and disgust produce patterns of available visual information across different facial regions for these expressions, the model observer for subtle expressions suggest that informative facial regions are restricted to the eye region (Vaidya et al., 2014). For future studies, it would be beneficial to feed expressions of varying intensities into a model observer to extract the available visual information at different intensities and investigate whether visual sampling by observers follow the change in the pattern of available visual information. Another aspect of facial images that might affect the informativeness of facial features is whether the mouth is open or not: For example, in the experiments reported here where expressions of angry, fearful, surprised and disgusted faces were used, angry expressions were the only face images with a closed mouth. Therefore, the distinctive closed mouth might have acted as the informative feature rather than the brow region. Angry expressions come in two variants: open-mouth - showing teeth - and closed mouth - pressed lips. Different expression databases contain different variants of the angry expressions; therefore, the angry expression used in emotion recognition experiments vary depending on the database used. The NimStim database includes open and closed mouth variants of all basic emotions (except surprise) and the validation of this database showed that there is a difference in the recognition of these two variants: Where angry, fearful and happy expressions are recognized better from their open-mouth variants, sad expressions were recognized better from their closed-mouth variant (Tottenham et al., 2009). Kestenbaum and Nelson (1990) showed that 7-month-old infants discriminate expressions showing teeth from expressions that are not. Additionally, Horstmann, Lipp and Becker (2012) showed that open-mouth angry expressions are rated as more intense compared to closed-mouth angry expressions and the visibility of the teeth in the open mouth angry expressions facilitates the visual search for angry expressions. Therefore, it is reasonable to suggest that whether the expression is posed with an open or closed mouth would impact the informativeness of the mouth region. For example, Vaidya et al. (2014) and Smith et al. (2005) differ in the informative regions they suggest for disgust. While Smith et al. (2005) uses the open mouth disgust variant which shows this region to be informative, Vaidya et al. (2014) uses the closed mouth disgust variant and emphasizes the eye region as informative. Therefore, the generalizability of results reported in this thesis to expressions taken from different face databases is limited by these factors.

Conclusion

In summary, it was found that foveal processing of some informative facial features improved emotion recognition, for example foveal processing of the mouth region improved emotion recognition for fearful, surprised and disgusted expressions. However, contrary to our initial expectations, this improvement was not due processing of high spatial frequency information at the foveated feature. Additionally, in line with findings in the literature, it was found that reflexive saccades, saccades that happen after face offset, are more often directed upwards from lower facial features than downwards from upper facial features. Previous studies such as Gamer and Buchel (2009) interpreted this finding as observers seeking out diagnostic facial features, such as eyes, for facial expressions such as fear. Using a more precise measure to map out the paths of these reflexive saccades, it was found that the paths of these saccades are most similar to a saccade path that is directed towards the brow. Therefore, while foveal processing of informative features improves emotion recognition, when these features were present in the extrafoveal visual field, hconclusive evidence that they are sought out. Instead, our data suggest that these early, reflexive saccades are directed upwards towards the brow. This region can be the featureless region suggested to be the *centre-of-information* suggested by Hsiao and Cottrell (2008) or the centre of the face image as suggested by (Bindemann et al., 2009).

The findings of this thesis can inform the field of facial expression perception in several ways: (1) While fixations to the eyes dominate the visual sampling of faces and facial expressions as demonstrated in Chapter 2 and in previous research, indicative of a holistic strategy, this is more likely to be a face-specific strategy which applies to majority of face-related tasks. The fixations that are of functional value – contributory to the task of emotion recognition – are likely to be ones landing on distinctive facial features such as the mouth and the brow. (2) The reflexive saccades that were previously suggested to be seeking-out informative features of facial expressions more likely reflect an automatic and involuntary strategy of looking upwards/at the upper visual field/upper face due to a top-down knowledge of the face configuration and are not best placed to measure what facial features are sought out when recognizing an expression. (3) We showed that there are certain facial features that improve emotion recognition. There are differences in the preferred initial fixation locations among observers and our findings suggest that observers with a preference to sample information from informative features would perform better at emotion recognition. For example, those observers who prefer to fixate the lower face would perform better in disgust, fearful and surprised recognition while those who prefer to fixate upper facial features would be better at recognizing angry expressions, especially when exposure duration is limited.

Appendix

Experiment 1

Reaction Time Descriptive Statistics

Brief Fixation Paradigm

The average reaction times for all conditions of the brief fixation paradigm can be seen in Table A1. A distribution of the reaction time data can also be seen in Figure A1. As can be seen in the figure, the distribution has a very strong right tail. The strongest outliers are a result of technical issues during the experiment which led to a noisy distribution of RT data which would make analysis of this data unreliable. Therefore, no further analysis is conducted on this data.

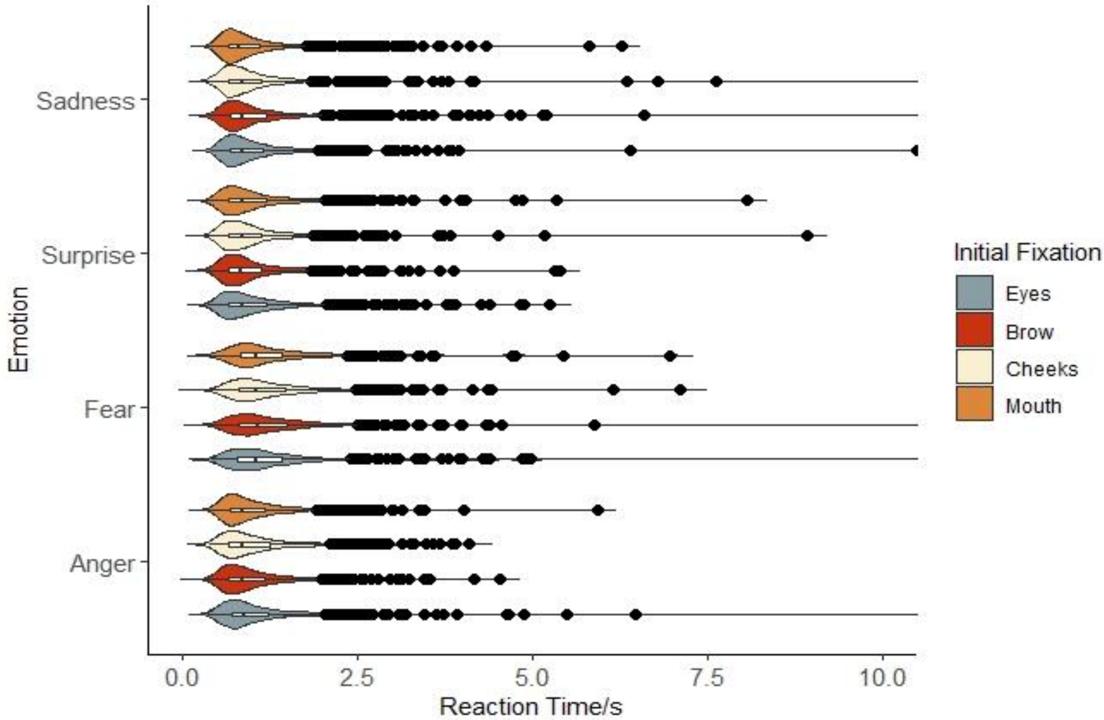

Figure A 1. The distribution of reaction times for each expression at each initial fixation location. The tail of the distribution is limited at 10 seconds to emphasize the shape of the distribution.

Table A 1. Average reaction times and standard deviations for each condition of the brief fixation paradigm.

	Eyes	**Brow**	**Cheek**	**Mouth**
Anger	1.09(0.32)	1.00(0.24)	1.05(0.28)	1.00(0.22)
Fear	1.25(0.35)	1.25(0.32)	1.23(0.28)	1.21(0.25)
Surprise	1.03(0.32)	.98(0.24)	0.99(0.25)	1.04(0.29)
Sad	1.06(0.24)	1.06(0.27)	1.05(0.29)	0.99(0.23)

Long Presentation Paradigm

The average reaction times for all conditions of the long presentation paradigm can be seen in Table A2. Similar to the brief fixation paradigm, the distribution of the reaction times in Figure A2 suggests the presence of strong outliers. Therefore, no further analysis is conducted on the reaction time data.

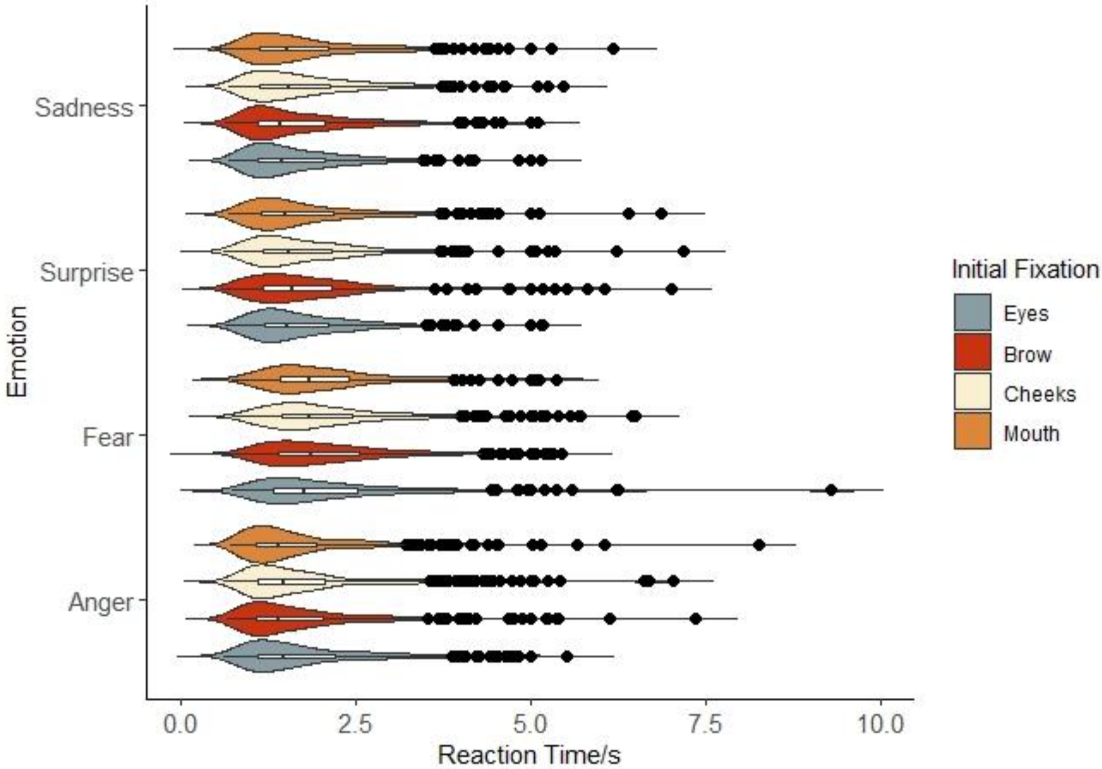

Figure A 2. The distribution of reaction times for each expression at each initial fixation location. The tail of the distribution is limited at 10 seconds to emphasize the shape of the distribution.

Table A 2. Average reaction times and standard deviations for each condition of the long fixation paradigm.

	Eyes	**Brow**	**Cheek**	**Mouth**
Anger	1.78(0.52)	1.70(0.44)	1.80(0.57)	1.69(0.46)
Fear	2.04(0.54)	2.10(0.49)	2.10(0.46)	2.02(0.44)
Surprise	1.71(0.36)	1.73(0.48)	1.80(0.46)	1.76(0.48)
Sadness	1.67(0.42)	1.67(0.40)	1.73(0.43)	1.71(0.48)

Percentage Accuracy Analyses

Brief Fixation Paradigm

There was a main effect emotion ($F(_{3, 96)}$=16.01, p<.001, $\eta_p{}^2$=.33) and a main effect initial fixation ($F_{(2.53, 81.04)}$=3.10, p=.004, $\eta_p{}^2$=.09). Fear was least accurately recognized compared to all other expressions (all ps< .001) and initially fixating the mouth led to higher recognition accuracy compared to the cheeks (p= .016). The interaction between expression and initial fixation was not significant ($F_{(9, 288)}$=1.24, p=.27, $\eta_p{}^2$=.04). The descriptive statistics for this analysis can be seen in Table A3.

Table A 3. Mean accuracy and standard deviation for each experimental condition

	Eyes	**Brow**	**Cheeks**	**Mouth**
Anger	0.84(0.11)	0.87(0.11)	0.82(0.13)	0.86(0.12)
Fear	0.70(0.16)	0.69(0.17)	0.71(0.18)	0.75(0.17)
Surprise	0.86(0.14)	0.89(0.13)	0.88(0.14)	0.89(0.10)
Sad	0.86(0.08)	0.87(0.08)	0.86(0.09)	0.87(0.10)

A significant chi-squared test showed that the use of each emotion response was different from what would be expected if all response options were used equally ($\chi^2(3)$= 131.55, p< .001). Participants used ‘fear’ (M=80.94, SD= 18.30) less than all other responses (anger: M=95.33, SD= 17.40; surprise: M=107.85, SD=19.54; sad: M=99.88, SD=14.89) and used surprise most often. Since the average percentage accuracy measure does not take this bias into account, we will use the unbiased hit rates to measure emotion recognition performance.

Long Presentation Paradigm

There was a main effect emotion ($F_{(1.83, 64.12)}$=24.35, p<.001, $\eta_p{}^2$=.41). Fear was least accurately recognized (all ps< .001) compared to all other expressions. The main effect initial fixation ($F_{(3,105)}$=.03, p=.99, $\eta_p{}^2$=.001) and the interaction did not reach significance ($F_{(6.03, 211.13)}$=.01, p=.37, $\eta_p{}^2$=.03). The descriptive statistics for this analysis can be seen in Table A4.

Similar to the brief fixation experiment, a significant chi-squared test showed that the use of each emotion response was different from what would be expected if all response options were used equally ($\chi^2(3)$= 73.88, p< .001). Again, the fear response (M=41.08, SD= 9.63) was used less compared to all other responses (anger: M=46.97, SD= 5.59; surprise: M=53.72, SD=9.45;

Table A 4. Mean accuracy and standard deviation for each experimental condition

	Eyes	**Brow**	**Cheeks**	**Mouth**
Anger	0.89(0.13)	0.91(0.09)	0.89(0.11)	0.92(0.09)
Fear	0.75(0.20)	0.73(0.22)	0.75(0.18)	0.74(0.19)
Surprise	0.92(0.12)	0.91(0.10)	0.9(0.10)	0.91(0.10)
Sad	0.91(0.09)	0.93(0.08)	0.94(0.08)	0.91(0.09)

sad: M=49.31, SD=5.41). and surprise responses were used most frequently. Since the mean accuracy measure does not take this bias into account, we will use the unbiased hit rates to measure emotion recognition performance.

Saccade Latency of the Reflexive Saccades

As discussed in the Introduction, the generation of a saccade is determined by visual information that maintains a fixation at a certain point or visual information that triggers movement to another target (Findlay & Walker, 1999). The latency of a saccade can be determined by the amount of perceptual processing required at the point of fixation (e.g. Arizpe et al., 2012), therefore; we investigated whether the initial fixation locations influenced the latencies of reflexive saccades. If perceptual processing of the fixated region affects the temporal aspects of saccade generation, we expect saccades from the cheeks to have shorter latencies compared to the other initial fixation locations since the cheeks contain relatively less useful visual information for recognition accuracy. A 4 × 4 repeated measures ANOVA was conducted to compare the average latencies of the initial saccades for each expression at each initial fixation location. Descriptive statistics can be seen in Figure A3. There was no main

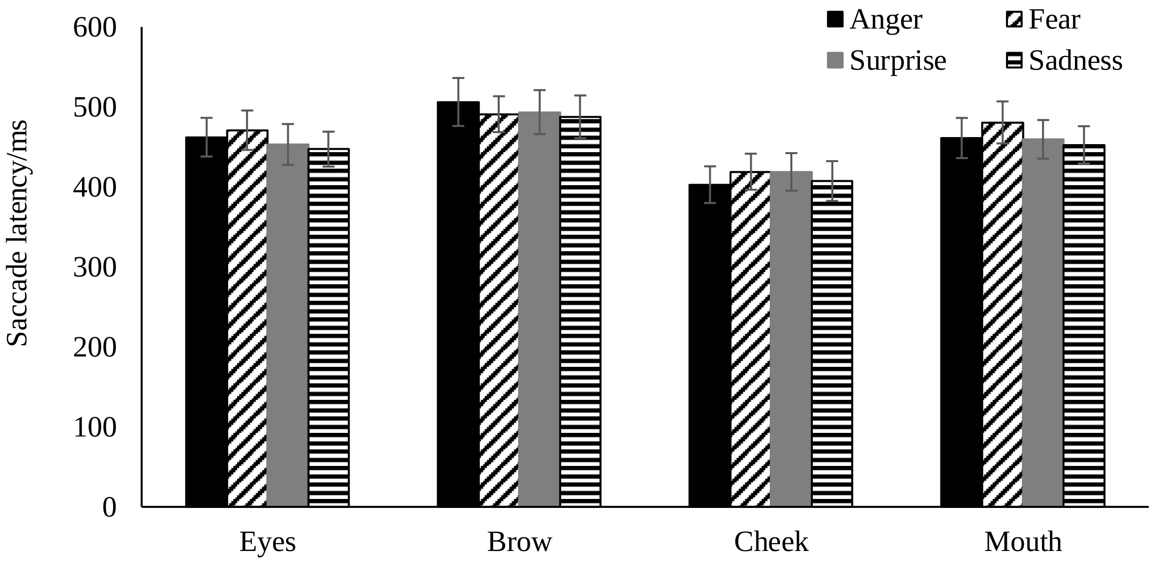

Figure A 3. Mean saccade latencies (in milliseconds) of reflexive saccades from each initial fixation location, separated by expression. Error bars indicate standard error of the mean.

effect of emotion ($F_{(2.29, 66.51)}$=1.58, p= 0.21) and the emotion × fixation location interaction was also not significant ($F_{(5.16, 149.58)}$=0.59, p= 0.71). However, there was a main effect of fixation location ($F_{(2.23, 66.64)}$=8.21, p< 0.001, η_p^{2}=.22). In line with our expectations, the main effect indicated that the saccade latencies from initial fixation on the cheeks were shorter compared to all the other initial fixation locations as expected (cheeks vs eyes: p=.002; cheeks vs brow: p=.001; cheeks vs mouth: p=.008).

First Saccade End Location

We wanted to investigate whether our observers' reflexive saccades in the brief fixation paradigm tended to target the central feature of the face (i.e., the nose) as Bindemann et al., (2009) suggested that early saccades are driven by the centre-of-gravity effect in face recognition. With this analysis, we aimed to investigate whether the reflexive saccades made by the observers were influenced by the centre-of-gravity effect in our study as suggested by Bindemann et al. (2009). A 4 × 4 repeated measures ANOVA was conducted on the total

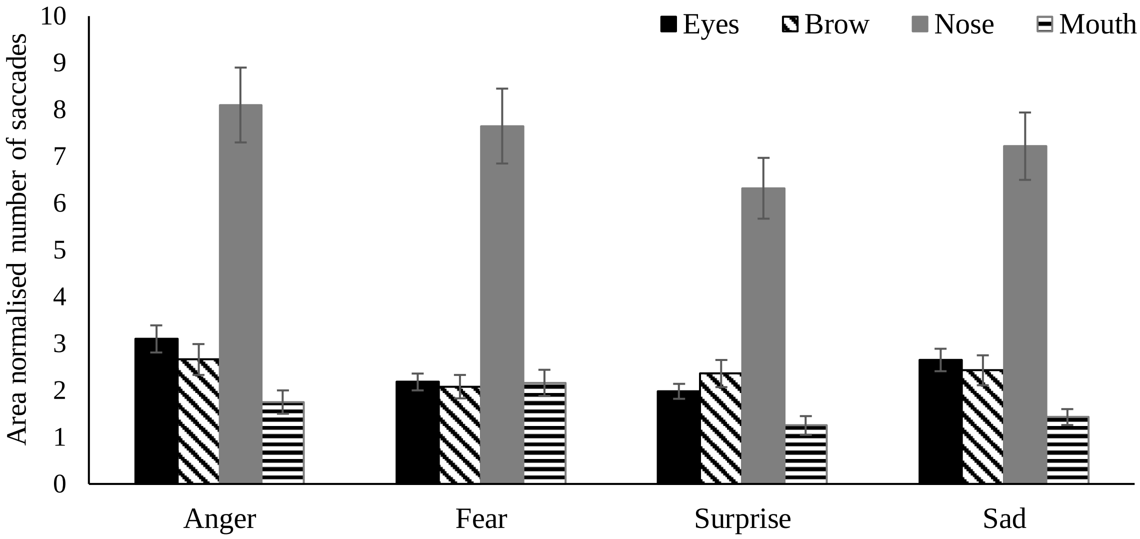

Figure A 4. The average number of reflexive saccades ending in one of the four regions of interest (eyes, brow, nose and mouth) for each expression. Error bars indicate standard error of the mean.

number of saccades ending in the eyes, brow, nose and mouth ROIs for each emotion. The number of saccades were area normalized by dividing them by the percentage of the whole face area covered by the relevant ROI. For the purposes of this analysis, we will accept the nose region as the centre of the face; however, it should be noted that the definition of the nose in this study comprises the area between the bridge and apex of the nose. There was a main effect of emotion ($F_{(3,96)}$=38.61, p< 0.001, η_p^2= .55) and a main effect of ROI ($F_{(1.46,\ 46.68)}$=43.33, p< 0.001, η_p^2= .58). The main effect of emotion indicated that significantly more saccades were made for angry faces compared to all other expressions (all *p*s< .001) and least number of saccades were made for fearful faces (all *p*s< .001). The main effect of ROI showed that significantly more saccades ended in the nose region compared to all other ROIs (all *p*s< .001) and more saccades ended in the eyes compared to the mouth (p= .01). There was also an emotion × ROI interaction ($F_{(6.00,\ 191.89)}$=4.64, p< 0.001. η_p^2= .13) (Figure A4).

To investigate this interaction, planned comparisons were carried out to compare the number of saccades ending in informative facial features compared to relatively non-informative features. This resulted in three one-tailed, paired t-tests for each expression. Where the data failed to meet the assumption of normality, Wilcoxon signed ranks tests were used instead. The results were accepted as significant at Bonferroni-corrected p-value of 0.017. For angry expressions, there were more saccades ending in the nose region compared to the brow region (Z= -4.66, p<.001). There was also a trend towards higher number of saccades ending in the brow region compared to the mouth region, however; this comparison did not reach the corrected significance value (Z= -2.06, p=.020). For fearful expressions, comparing the saccades ending in the mouth region to other ROIs did not reveal any significant differences (mouth ~ eyes: Z= -30, p=.38; mouth ~ brow: Z= -.06, p=.48) except that there were significantly more saccades ending in the nose region compared to the mouth (Z= -4.64, p<.001). There were significantly fewer saccades ending in the mouth region for surprised expressions compared to the eyes (Z= -3.26, p<.001), brow (Z= -2.74, p<.006) and the nose (Z= -5.01, p<.001).

Experiment 2

Reaction Time Descriptive Statistics

Brief Fixation Paradigm

The distribution of reaction times can be seen in Figure A5. Similar to Experiment 1, due to the presence of strong outliers, this data is not analysed further. The average reaction times for all conditions of the brief fixation paradigm can be seen in Table A5.

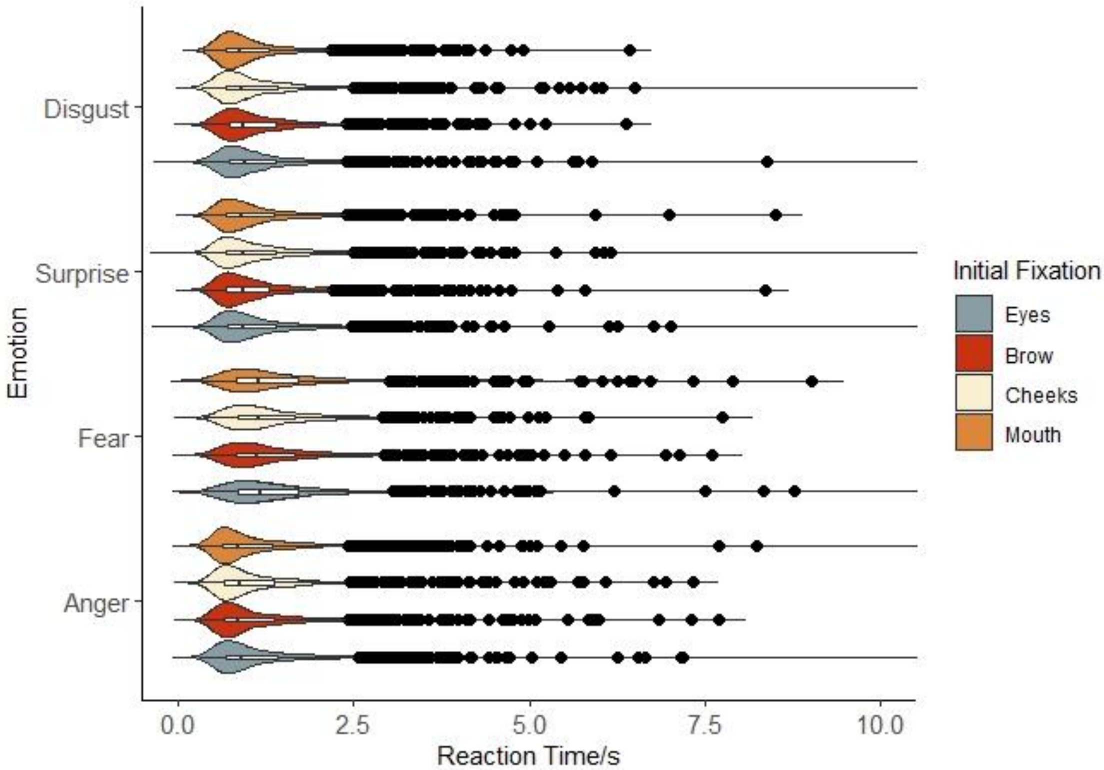

Figure A 5. The distribution of reaction times for each expression at each initial fixation location. The tail of the distribution is limited at 10 seconds to emphasize the shape of the distribution.

Table A 5. Average reaction times and standard deviations for each condition of the brief fixation paradigm.

	Eyes	Brow	Cheeks	Mouth
Anger	1.25(0.40)	1.19(0.40)	1.18(0.39)	1.20(0.45)
Fear	1.53(0.83)	1.39(0.35)	1.40(0.34)	1.44(0.38)
Surprise	1.21(0.35)	1.16(0.36)	1.22(0.39)	1.18(0.41)
Disgust	1.24(0.43)	1.19(0.30)	1.24(0.44)	1.14(0.33)

Long Presentation Paradigm

The distribution of reaction times for the long presentation paradigm can be seen in Figure A6 and the average reaction times for all conditions can be seen in Table A6. The reaction time data will not be further analysed due to the presence of strong outliers.

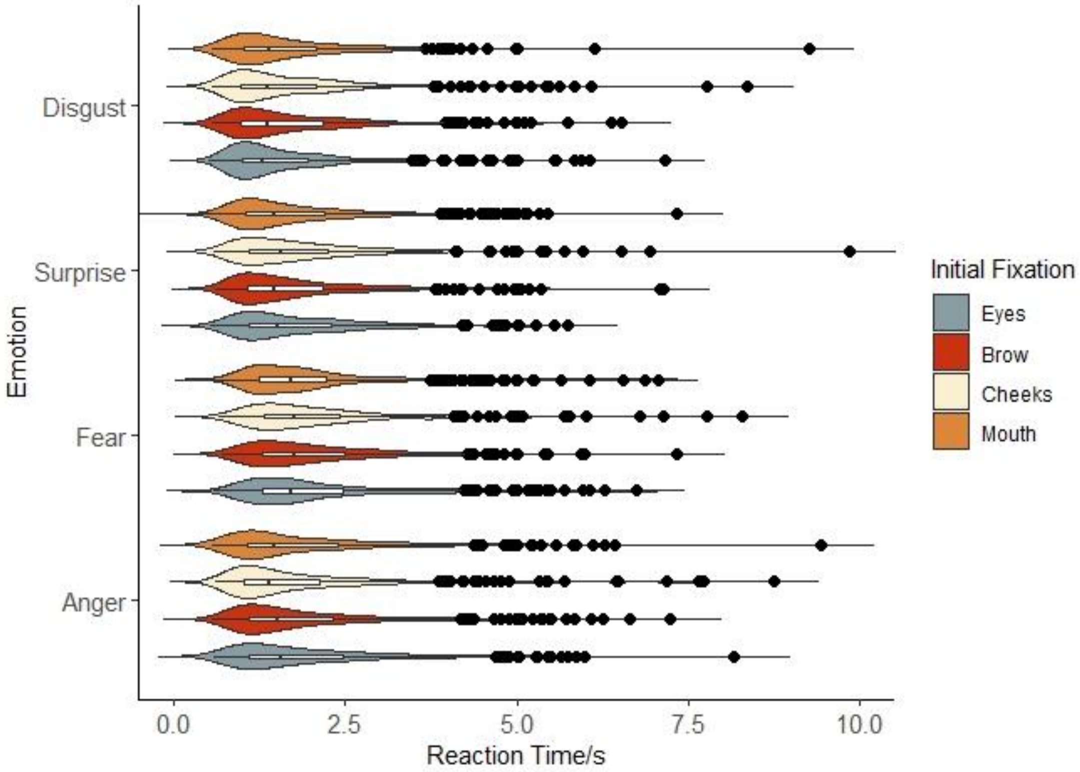

Figure A 6. The distribution of reaction times for each expression at each initial fixation location. The tail of the distribution is limited at 10 seconds to emphasize the shape of the distribution.

Table A 6. Average reaction times and standard deviations for each condition of the long fixation paradigm.

	Eyes	Brow	Cheeks	Mouth
Anger	1.93(0.61)	1.91(0.66)	1.78(0.65)	1.90(0.63)
Fear	2.05(0.61)	2.05(0.54)	2.06(0.62)	1.97(0.53)
Surprise	1.83(0.59)	1.79(0.61)	1.85(0.58)	1.79(0.65)
Disgust	1.65(0.65)	1.71(0.59)	1.69(0.70)	1.68(0.59)

Percentage Accuracy Analyses

Brief Fixation Paradigm

There was a main effect expression ($F_{(3, 117)}$=12.01, p<.001, η_p^2=.24) and main effect initial fixation ($F_{(3, 117)}$=8.97, p<.001, η_p^2=.19). The interaction approached, however, failed to reach significance ($F_{(6.65, 259.40)}$=1.88, p=.08, η_p^2=.05). The descriptive statistics for this analysis can be seen in Table A7.

Table A 7. Mean accuracy and standard deviation for each experimental condition

	Eyes	**Brow**	**Cheek**	**Mouth**
Anger	0.80(0.18)	0.81(0.20)	0.82(0.20)	0.81(0.22)
Fear	0.70(0.21)	0.68(0.21)	0.72(0.19)	0.72(0.21)
Surprise	0.85(0.10)	0.88(0.10)	0.88(0.09)	0.90(0.09)
Disgust	0.74(0.17)	0.71(0.21)	0.76(0.18)	0.79(0.19)

A significant chi-squared test showed that the use of each emotion response was different from what would be expected if all response options were used equally ($\chi^2(3)$= 173.21, p< .001). The fear response (M=83.25, SD= 19.03) was used the least, followed by disgust (M=90.38, SD=19.29) and anger (M=102.18, SD= 21.27). Surprise response (M=108.13, SD=21.47) was used most frequently. Since the mean accuracy measure does not take this bias into account, we will use the unbiased hit rates to measure emotion recognition performance.

Long Presentation Paradigm

There was a main effect emotion ($F_{(2.34, 89.03)}$=6.63, p=.001, η_p^2=.15) which showed that surprised expressions were better recognized compared to fearful (p= .002) and disgusted (p= .001) expressions. Neither the main effect initial fixation ($F_{(2.45, 93.27)}$=.16, p=.89, η_p^2=.004). nor interaction reached significance ($F_{(9, 342)}$=1.43, p=.17, η_p^2=.036). Table A8 shows that descriptive statistics.

Table A 8. Mean accuracy and standard deviations for each experimental condition

	Eyes	**Brow**	**Cheek**	**Mouth**
Anger	0.81(0.21)	0.85(0.19)	0.85(0.19)	0.83(0.18)
Fear	0.79(0.16)	0.78(0.19)	0.78(0.18)	0.78(0.18)
Surprise	0.91(0.13)	0.88(0.12)	0.91(0.12)	0.90(0.12)
Disgust	0.78(0.20)	0.78(0.18)	0.78(0.18)	0.82(0.16)

A significant chi-squared test showed that the use of each emotion response was different from what would be expected if all response options were used equally ($\chi^2(3)$= 40.80, p< .001). Fear response (M=43.05, SD=8.63) was used the least followed by than anger (M= 50.49, SD= 9.83) and surprise (M=51.00, SD= 8.43). Since the mean accuracy measure does not take this bias into account, we will use the unbiased hit rates to measure emotion recognition performance.

Saccade Latency of the Reflexive Saccades

Saccade latency can be an indication of how much visual information is being processed at the point of fixation before a saccade is initiated according to the saccade generation model of Findlay (1999). Therefore, we investigated the mean saccade latencies from each of the initial fixation locations for the first saccades, as we did for Experiment 1. We expected the saccade latencies from the cheeks to be shorter compared to other initial fixation locations since we expect cheeks to have less diagnostic visual information compared to the eyes, brow and mouth. A 4 × 4 repeated measures ANOVA with emotion and initial fixation location as factors was carried out (Figure A7).

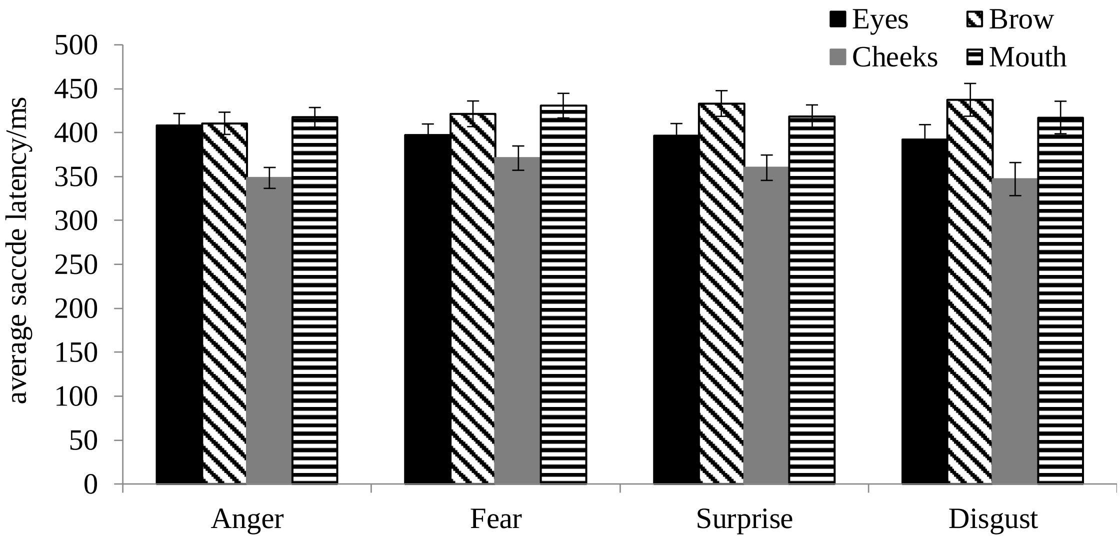

Figure A 7. Mean saccade latencies (in milliseconds) of reflexive saccades from each initial fixation location, separated by expression. Error bars indicate standard error of the mean.

There was no main effect of emotion ($F_{(3, 111)}$= 1.34, p=.27) but there was a significant main effect of initial fixation location ($F_{(1.80, 66.83)}$= 10.72, p< .001, η_p^2=.23). This main effect indicated that the saccade latencies from the initial fixation location at one of the cheeks were significantly shorter than the saccade latencies from the eyes, the brow and the mouth (all *p*s< .001). Additionally, first saccade latencies were shorter from the eyes compared to the brow. There was also a significant interaction ($F_{(5.98, 221.42)}$= 2.68, p= .02, η_p^2=.07).

To investigate this interaction, separate one-way ANOVAs were carried out for each initial fixation location separately. There was no main effect of emotion on saccade latency for initial fixation at the eyes ($F_{(3,111)}$= 1.51, p=.22) or mouth ($F_{(3,111)}$= 0.71, p= .55). The significant main effect of emotion at the brow ($F_{(3,111)}$= 3.38, p= .02, η_p^2= .084) indicated that first saccade latencies for the angry faces were shorter compared to disgusted faces (p= .003). Finally, the main effect of emotion for initial fixation at the cheeks ($F_{(3, 111)}$= 5.26, *p*= .002, η_p^2= .12) indicated that first saccade latencies were longer for fear compared to anger (*p*= .027) and disgust (*p*= .002).

To investigate whether the saccade latencies from the facial feature deemed informative for each expression was longer compared to relatively non-informative features, planned comparisons were carried out for each expression. This led to 3 one-tailed paired samples t-tests where the distribution of mean saccade latencies were normal and 3 Wilcoxon signed ranks tests where the distribution was not normal.

For angry facial expressions, comparing the saccade latencies from the brow region to other initial fixation locations revealed that the latencies from the cheeks are significantly shorter compared to latencies from the brow (*Z*= -4.07, *p*< .001). No other comparison was significant (brow vs eyes: Z= -.56, p= .29; brow vs mouth: *Z*= -30, *p*= .39).

For fearful expressions, comparing latencies from the mouth region to other initial fixation locations showed that saccade latencies from the mouth region were shorter compared to the cheeks ($t_{(37)}$= 3.46, *p*<.001). Despite a large numerical difference between the eyes and the mouth, this comparison did not reach significance ($t_{(37)}$= 1.79, *p*=.04). There was no significant difference between the mouth and the brow ($t_{(37)}$= .49, *p*=.32).

For surprised facial expressions, there was only a significant difference between the mouth and the cheeks showing significantly shorter latencies for cheeks compared to the mouth ($t_{(37)}$= 3.78, *p*<.001).

For disgusted facial expressions, initially fixating on the cheek region led to significantly shorter latencies compared to initially fixating the mouth region (*Z*= -3.98, *p*< .001). Despite a large difference in latencies between the mouth and the brow regions implying that latencies from the brow region was longer compared to the mouth, this did not reach significance (*Z*= -1.68, *p*< .047).

First Saccade End Locations

To investigate whether our observers demonstrated a tendency to direct their fixations towards the centre of the briefly presented faces, we compared the mean frequency of first saccades ending in the eye, brow, nose and cheek regions of interest for each emotion. Since Bindemann et al. (2009) showed that the early saccades in face recognition are directed to the geometric centre of the face stimuli, it is important to investigate whether the reflexive saccades in this experiment might also be affected by this centre-of-the gravity effect, as we did for Experiment 1. Therefore, instead of the actual centre of the faces, it is more appropriate to indicate that this analysis will compare the mean frequency of first saccades that end within the central feature (i.e., the nose) of the presented faces. A 4 × 4 repeated measures ANOVA with emotion and

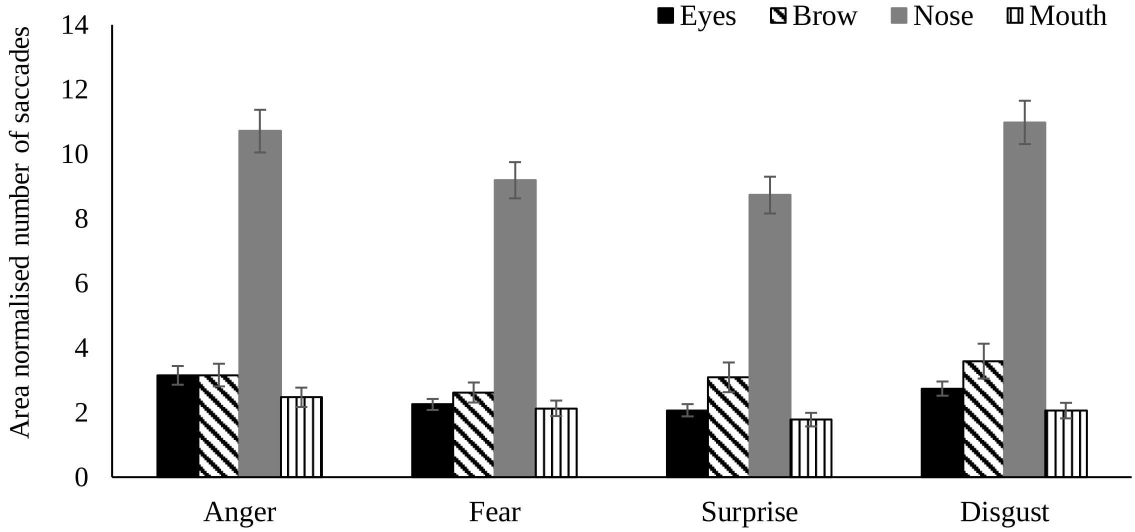

Figure A 8. The average number of reflexive first saccades ending within each of the four regions of interest (eyes, brow, nose and mouth) for each expression. The values are area normalized. Error bars indicate standard error of the mean.

region of interest as factors was conducted. Graphical illustration of the descriptive statistics can be seen in Figure A8. There was a significant main effect of emotion ($F_{(3,111)}$= 56.09, p< .001, η_p^2= .60) and a significant main effect of region of interest ($F_{(2.09, 77.43)}$= 86.06, p< .001, η_p^2= .70). The main effect of emotion indicated that more saccades were made for angry and disgusted expressions compared to fearful and surprised expressions (all *p*s< .001). The main effect of region of interest reflected the fact that the first saccades ended in the nose region significantly more frequently than in the eyes, brow or the mouth ROIs (all *p*s<.001). There was also a significant interaction between emotion and ROI ($F_{(5.42, 200.56)}$= 6.53, p< .001, η_p^2= .15). Planned comparisons were carried out to further investigate this interaction. A Bonferroni-corrected *p*-value of 0.017 was used for multiple comparisons and where the data distribution failed the assumption of normality, Wilcoxon signed ranks test were used instead of paired samples t-tests. For angry expressions, in contrast to our expectations, reflexive saccades did not end in the brow region significantly more compared to other initial fixation locations (brow vs eyes: Z=-.20, p=.43; brow vs mouth: Z=-.88, p=.19). For fearful expressions, similarly, the first saccades tended to end significantly more in the nose region (Z=-5.33, p< .001) compared to the mouth. No other comparison reached significance (mouth vs eyes: Z= -.62, p= .27; mouth vs brow: Z=.37, p=.19). For surprised expressions, first saccades ended in the nose region (Z= -5.32, p< .001) and brow region (Z=2.37, p=.009) significantly more compared to the mouth, there was no difference between the eyes and the mouth (Z=-1.02, p=.16). For disgusted expressions, saccades ended in the nose region significantly more compared to the mouth (Z=-5.37, p< .001). Despite not reaching significance, there were a higher number of saccades that ended in the eye (Z=-1.86, p=.031) and brow regions (Z=-1.57, p=.058) compared to the mouth.

Experiment 3

Reaction Time Descriptive Statistics

The distribution for the reaction times can be seen in Figure A9 and the average reaction times for all experimental conditions of Experiment 3 can be seen in Table A9. Reaction time data is not further analysed due to the presence of strong outliers.

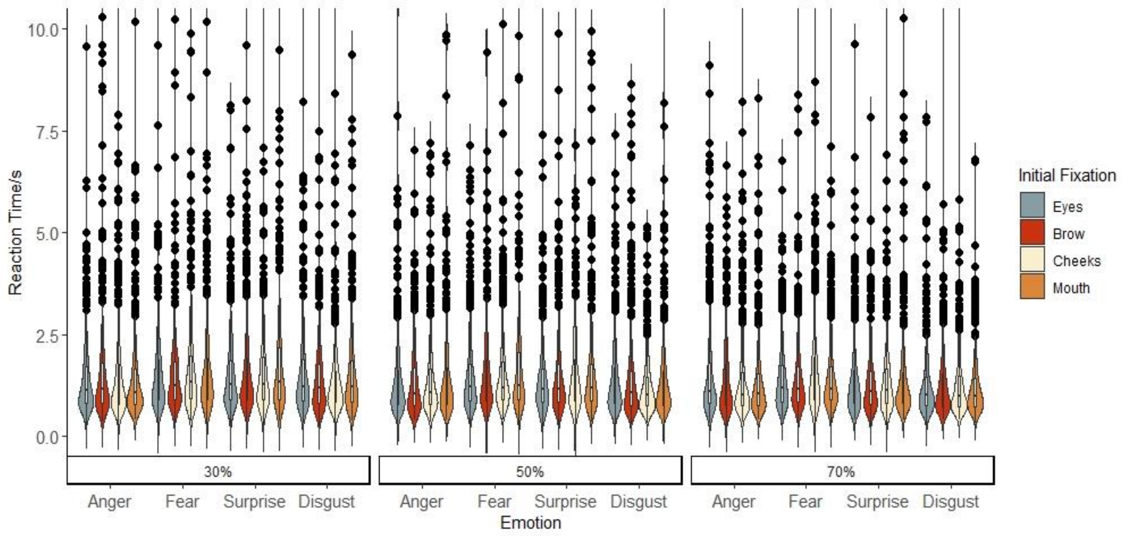

Figure A 9. The distribution of reaction times for each expression at each initial fixation location. The tail of the distribution is limited at 10 seconds to emphasize the shape of the distribution.

Table A 9. Average reaction times and standard deviations for each condition of the long fixation paradigm.

		Eyes	Brow	Cheek	Mouth
Anger	30	1.43(0.44)	1.57(0.67)	1.55(0.56)	1.43(0.63)
	50	1.38(0.46)	1.38(0.46)	1.42(0.51)	1.43(0.58)
	70	1.51(0.47)	1.37(0.42)	1.37(0.50)	1.37(0.47)
Fear	30	1.57(0.45)	1.51(0.42)	1.67(0.61)	1.67(0.52)
	50	1.53(0.48)	1.55(0.50)	1.56(0.40)	1.60(0.52)
	70	1.44(0.36)	1.46(0.44)	1.68(0.61)	1.53(0.53)
Surprise	30	1.57(0.48)	1.64(0.65)	1.64(0.54)	1.79(0.61)
	50	1.53(0.90)	1.51(0.58)	1.56(0.62)	1.55(0.49)
	70	1.41(0.43)	1.35(0.41)	1.41(0.46)	1.38(0.53)
Disgust	30	1.57(0.51)	1.48(0.44)	1.41(0.49)	1.54(0.53)
	50	1.39(0.44)	1.43(0.60)	1.23(0.33)	1.38(0.54)
	70	1.27(0.38)	1.31(0.36)	1.28(0.32)	1.23(0.37)

Percentage Accuracy Analyses

There was a main effect emotion ($F_{(2.58, 102.98)}$=5.13, p=.004, η_p^2=.11). Fear was recognised significantly less accurately compared to surprise (p= .009). The main effect initial fixation ($F_{(2.36, 94.41)}$=10.33, p<.001, η_p^2=.21) showed that initial fixation at the mouth led to significantly higher accuracy compared to all other initial fixation locations (eyes and brow p< 0.001; cheeks: p=0.001). The main effect intensity ($F_{(1.50, 59.98)}$=472.55, p<.001, η_p^2=.92) was as expected with 70% intensity leading to the highest and 30% intensity leading to the lowest recognition accuracy. There was an interaction between emotion and initial fixation location ($F_{(9, 360)}$=2.58, p=.007, η_p^2=.06).

There was no effect of initial fixation location for anger ($F_{(2.57, 102.70)}$=0.30, p=.80, η_p^2=.007). The effect of initial fixation for fear ($F_{(2.45, 98.06)}$=4.80, p=.006, η_p^2=.11) showed that initially fixating the mouth led to higher fear recognition compared to the eyes (p=.002) and the brow (p=.006). Furthermore, the main effect of initial fixation for surprise ($F_{(3, 120)}$=4.04, p=.009, η_p^2=.092) showed that initially fixating the mouth led to higher surprise recognition compared to the eyes (p=.004) and the cheeks (p=.006). Finally, initial fixation effect for disgust ($F_{(3, 120)}$=8.31, p<.001, η_p^2=.17) showed that initial fixation on the mouth led to higher disgust accuracy compared to eyes (p=.005), brow (p<.001) and cheeks (p=.004). Additionally, fixating the eyes led to higher disgust accuracy compared to the brow (p=.021). There was also an interaction

Table A 10. The average accuracy and standard deviation for each expression at each initial fixation and intensity.

		Initial Fixation Location			
Expression	**Intensity**	**Eyes**	**Brow**	**Cheek**	**Mouth**
Anger	**30**	0.35(0.16)	0.37(0.15)	0.36(0.17)	0.34(0.17)
	50	0.57(0.20)	0.56(0.19)	0.57(0.17)	0.58(0.19)
	70	0.63(0.20)	0.65(0.15)	0.66(0.18)	0.64(0.17)
Fear	**30**	0.29(0.19)	0.29(0.15)	0.31(0.17)	0.36(0.20)
	50	0.53(0.21)	0.51(0.21)	0.57(0.20)	0.57(0.18)
	70	0.60(0.25)	0.60(0.27)	0.61(0.25)	0.67(0.20)
Surprise	**30**	0.31(0.18)	0.37(0.20)	0.35(0.20)	0.43(0.22)
	50	0.64(0.21)	0.67(0.21)	0.63(0.19)	0.67(0.23)
	70	0.78(0.17)	0.77(0.20)	0.77(0.21)	0.78(0.23)
Disgust	**30**	0.45(0.17)	0.44(0.16)	0.45(0.16)	0.50(0.18)
	50	0.64(0.18)	0.56(0.23)	0.61(0.03)	0.66(0.20)
	70	0.59(0.21)	0.55(0.21)	0.59(0.22)	0.65(0.22)

between emotion and intensity ($F_{(4.17, 166.88)}$=16.19, p<.001, η_p^2=.29). All expressions followed the same pattern as the main effect of intensity except for disgust where there is no difference between 50 and 70% intensities (p=.452). There was no interaction between initial fixation and intensity ($F_{(6, 240)}$=0.74, p=.622, η_p^2=.02. There was no three-way interaction ($F_{(18, 720)}$=.98, p<.478, η_p^2=.02). The descriptive statistics for the mean percentage accuracy can be seen in Table A10.

Similar to the previous two studies reported in Chapter 2, a significant chi-squared test showed that the use of each emotion response was different from what would be expected if all response options were used equally ($\chi^2(4)$= 641.63, p< .001). The neutral response option (M=13.65, SD= 6.42) was the least frequently used option while other response options were used with similar frequency (anger: M=22.23, SD= 5.74; fear: M=20.02, SD= 5.93; surprise: M=22.52, SD= 5.64; disgust: M=21.58, SD= 4.90).

Saccade Latency Analysis

Faster saccade latencies should be observed for lower intensity expressions overall. According to the saccade generation theory of Findlay & Walker (1999), the amount of visual information at the fixation location decides when the eye moves. Since there is less task-relevant visual information to be processed for the lower intensity expressions, we expected that even at the informative facial features such as the eyes, brow and mouth, saccades would be made to another location quicker in order to obtain further visual information to resolve the emotion recognition task. A 4×3×3 repeated measures ANOVA was run on the average saccade latencies. Descriptive statistics of this analysis can be seen in Figure A10. In contrast to our

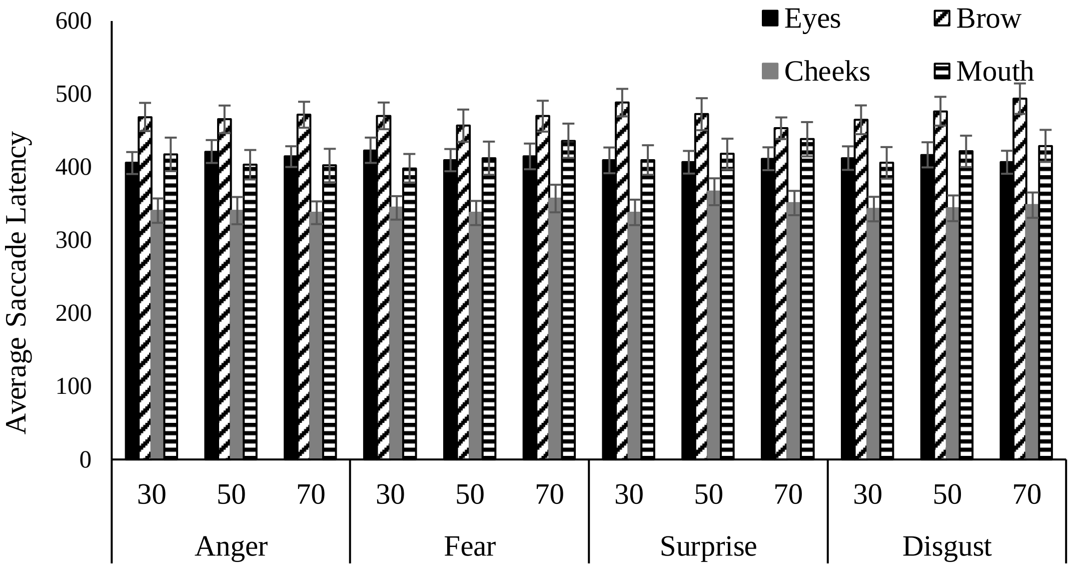

Figure A 10. Mean saccade latencies (in milliseconds) of reflexive saccades from each initial fixation location, separated by expression. Error bars indicate standard error of the mean.

prediction that intensity will have an effect on saccade latencies, neither the main effect of intensity ($F_{(3, 117)}$= 0.34, p=.71) nor any of its interactions were significant. There was also no main effect of expression ($F_{(3, 117)}$= 1.01, p=.39). The significant main effect of initial fixation ($F_{(1.93, 75.14)}$= 9.83, p<.001, η_p^2= .20) indicated that the saccade latencies from initial fixation on the cheeks was shorter compared to the brow and the mouth (ps< .001) which was in line with our findings from the first two studies reported in Chapter 2. Furthermore, saccade latencies from the brow region were longer compared to eyes (p=.002) and cheeks (p< .001).

Experiment 4

Stimuli examples

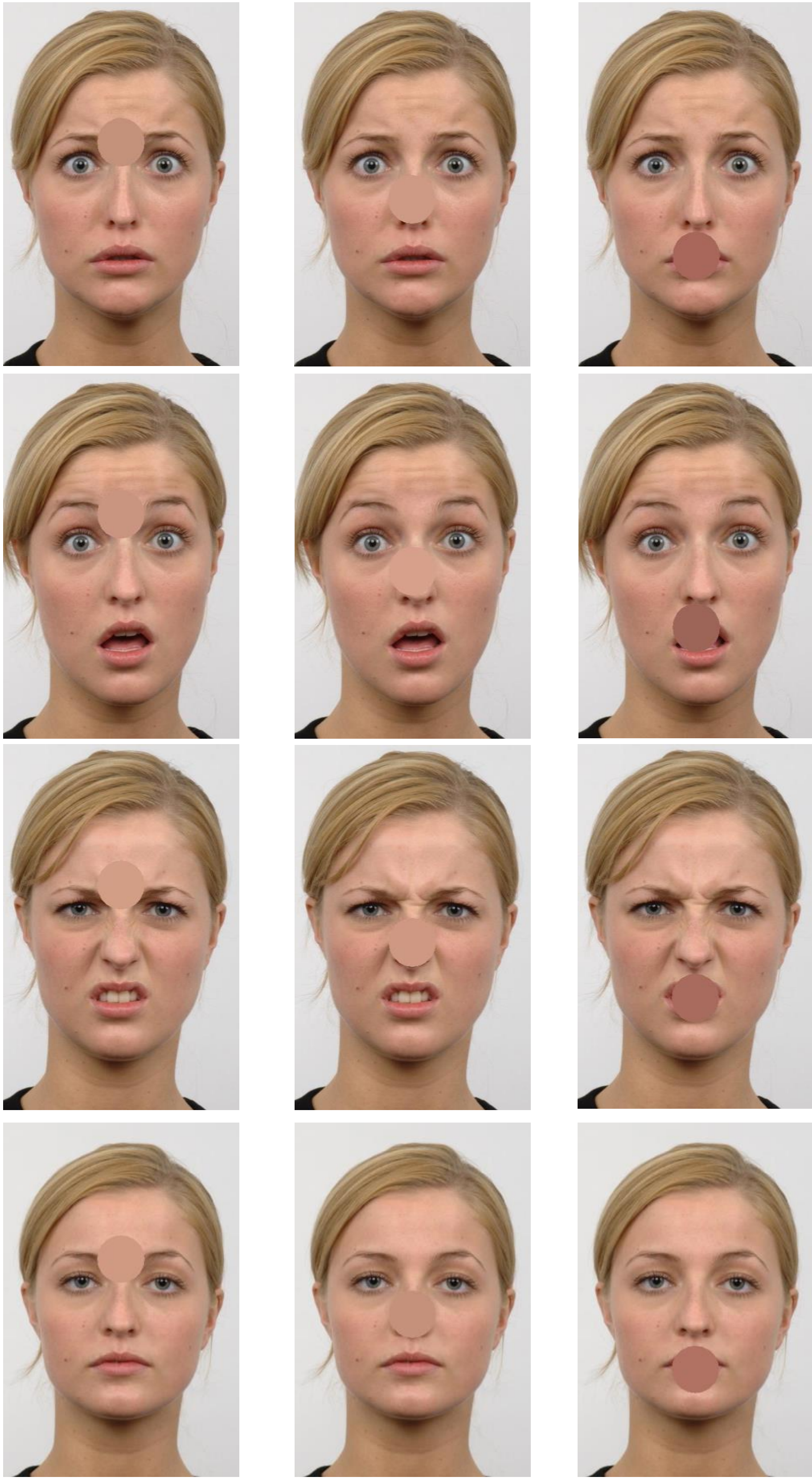

Figure A 11. Examples of expressions of fear, surprise, disgust and neutral with the initial fixation locations of brow, centre and mouth masked.

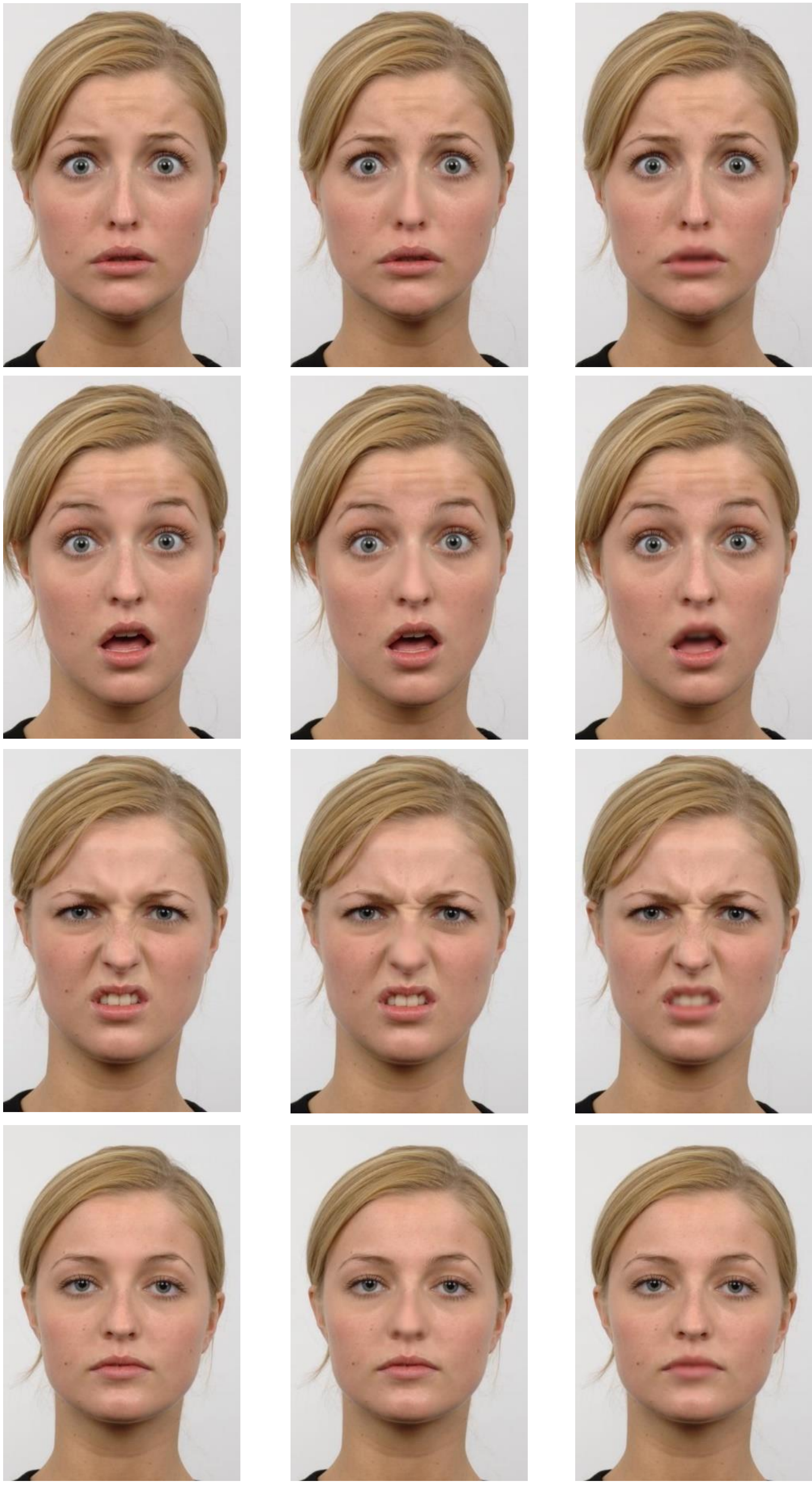

Figure A 13. Examples of expressions of fear, surprise, disgust and neutral with the initial fixation locations of brow, centre and mouth filtered.

Reaction Time Descriptive Statistics

The distribution of the reaction times for Experiment 4 is visualised in Figure A13 and the average reaction times can be seen in Table A11. There are no further analyses for reaction time data due to the presence of strong outliers.

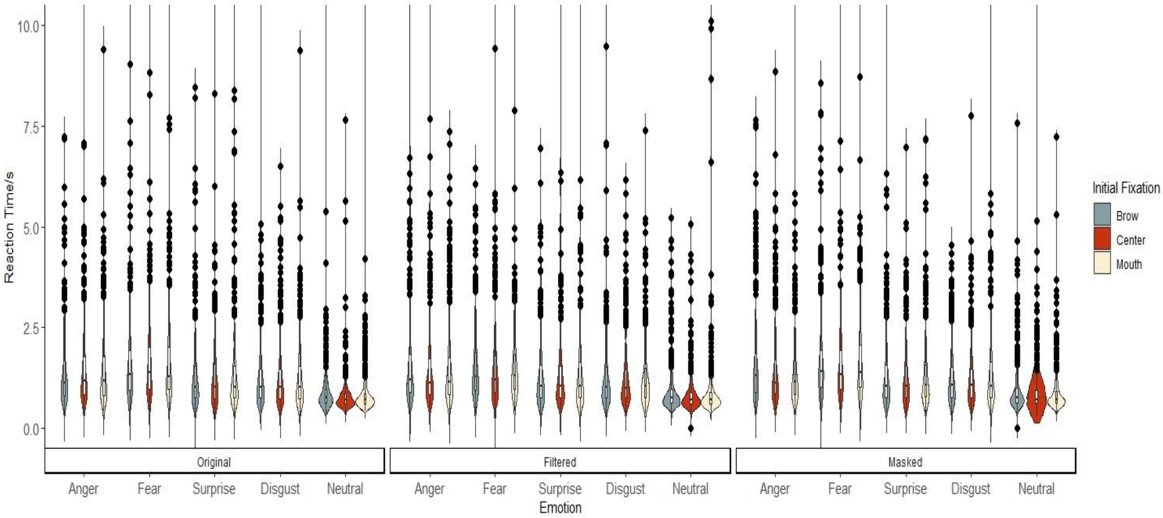

Figure A 15. The distribution of reaction times for each expression at each initial fixation location. The tail of the distribution is limited at 10 seconds to emphasize the shape of the distribution.

Table A 11. Average reaction time and standard deviation for each expression at each initial fixation location and mask condition.

		Brow	Centre	Mouth
Anger	Unfiltered	1.42(0.39)	1.52(0.42)	1.50(0.40)
	Filtered	1.67(0.64)	1.47(0.50)	1.50(0.40)
	Occluded	1.65(0.47)	1.45(0.41)	1.43(0.43)
Fear	Unfiltered	1.72(0.60)	1.75(0.74)	1.66(0.50)
	Filtered	1.60(0.44)	1.69(1.02)	1.56(0.41)
	Occluded	1.66(0.48)	1.62(0.45)	1.69(0.44)
Surprise	Unfiltered	1.34(0.39)	1.35(0.60)	1.43(0.52)
	Filtered	1.31(0.40)	1.31(0.40)	1.37(0.43)
	Occluded	1.32(0.42)	1.28(0.46)	1.35(0.39)
Disgust	Unfiltered	1.30(0.51)	1.26(0.42)	1.26(0.39)
	Filtered	1.29(0.34)	1.26(0.39)	1.27(0.40)
	Occluded	1.25(0.32)	1.27(0.40)	1.35(0.44)
Neutral	Unfiltered	1.13(1.14)	0.85(0.20)	0.89(0.33)
	Filtered	1.00(0.28)	0.89(0.30)	0.91(0.36)
	Occluded	0.98(0.28)	1.39(3.03)	0.91(0.23)

Percentage Accuracy Analyses

There was a main effect emotion ($F_{(2.85, 99.80)}$=39.73, p<.001, η_p^2=.53). Neutral expressions were recognised significantly better compared to all other emotions (all ps< .001). Surprised expressions were recognised better compared to all but neutral expressions (all ps<.001). The main effect initial fixation ($F_{(1.53, 53.63)}$=3.57, p=.047, η_p^2=.093) showed that initial fixation on the centre led to a higher emotion recognition compared to the brow (p= .003). In the unbiased hit rates analysis, this is not significant. The main effect mask ($F_{(2, 70)}$=9.57, p<.001, η_p^2=.215) showed that occluding led to the lowest recognition compared to the original and filtered images (original image: p=.001; filtered image: p=.005). There was no interaction between emotion and initial fixation ($F_{(8, 280)}$=.75, p=.65, η_p^2=.02). There was an interaction between emotion and mask ($F_{(5.98, 209.16)}$=4.31, p<.001, η_p^2=.11) but no interaction between initial fixation and mask (only approaching significance; $F_{(4, 140)}$=2.21, p=.071, η_p^2=.059). These interactions were not significant in the analysis with unbiased hit rates. There was also a significant three-way interaction which was not significant when analysis was conducted on unbiased hit rates ($F_{(9.85, 344.69)}$=3.37, p<.001, η_p^2=.09).

Table A 12. The average accuracy and standard deviation for each expression at each mask condition and initial fixation location. (UN= Unfiltered, F= Filtered, O= Occluded).

		Eyes	Brow	Mouth
Anger	**UN**	0.70(0.19)	0.68(0.20)	0.65(0.25)
	F	0.66(0.22)	0.70(0.17)	0.63(0.23)
	O	0.62(0.20)	0.69(0.21)	0.66(0.21)
Fear	**UN**	0.64(0.19)	0.67(0.22)	0.68(0.19)
	F	0.62(0.22)	0.66(0.22)	0.67(0.22)
	O	0.64(0.24)	0.65(0.19)	0.57(0.22)
Surprise	**UN**	0.85(0.10)	0.90(0.10)	0.85(0.10)
	F	0.86(0.13)	0.86(0.13)	0.87(0.14)
	O	0.88(0.13)	0.89(0.10)	0.87(0.10)
Disgust	**UN**	0.63(0.27)	0.69(0.25)	0.73(0.24)
	F	0.68(0.25)	0.67(0.25)	0.73(0.26)
	O	0.61(0.28)	0.66(0.25)	0.58(0.24)
Neutral	**UN**	0.94(0.07)	0.96(0.05)	0.97(0.07)
	F	0.96(0.07)	0.98(0.05)	0.94(0.09)
	O	0.93(0.10)	0.96(0.06)	0.95(0.09)

In order to follow-up this interaction, one-way repeated measures ANOVAs were conducted for each expression at each initial fixation location separately with mask as a factor. When the initial fixation was on the brow, there was a main effect of mask for angry expressions ($F_{(2, 70)}$=4.95, p=.010, η_p^2=.12). Occluding the brow region led to a drop in anger recognition accuracy compared to the original image (p=.019). There was no main effect of mask for fearful ($F_{(2, 70)}$=.32, p=.73, η_p^2=.009), surprised ($F_{(2, 70)}$=.70, p=.50, η_p^2=.02) or neutral ($F_{(2, 70)}$=.92, p=.40, η_p^2=.03) expressions when these are seen initially at the brow region. The main effect of mask for disgust expressions when initially fixating the brow ($F_{(1.71, 59.95)}$=2.95, p=.07, η_p^2=.078) was almost significant. This effect was associated with a reduced disgust accuracy when the brow is occluded compared to when it was filtered (p= .026).

There was no effect of mask for anger ($F_{(1, 60.21)}$=.26, p=.74, η_p^2=.007), fear ($F_{(2, 70)}$=.46, p=.63, η_p^2=.013), surprise ($F_{(2, 70)}$=2.01, p=.14, η_p^2=.054), disgust ($F_{(2, 70)}$=.86, p=.43, η_p^2=.024) or neutral ($F_{(2, 70)}$=1.19, p=.31, η_p^2=.033) faces at the centre fixation point.

When the initial fixation was on the mouth, there was a main effect of mask condition for fear ($F_{(2, 70)}$=10.18, p<.001, η_p^2=.23) and disgust ($F_{(1.56, 54.75)}$=18.97, p<.001, η_p^2=.35). Occluding the initially fixated mouth region led to lower fear recognition compared to the original (p< .001) and the filtered (p= .003) mouth. For disgusted expressions, occluding the initially fixated mouth region led to lower recognition accuracy compared to filtered and original mouth (both *p*s <.001). There was no effect of masking condition for the remaining expressions (anger: ($F_{(2, 70)}$=1.07, p=.35, η_p^2=.030); surprise: ($F_{(1.67, 58.44)}$=0.66, p=.50, η_p^2=.018); neutral: ($F_{(2, 70)}$=1.98, p=.15, η_p^2=.054).

Similar to the previous studies reported in Chapters 2 and 3, a significant chi-squared test showed that the use of each emotion response was different from what would be expected if all response options were used equally ($\chi^2(4)$= 262.49, p< .001). Fear responses were used least often (M=16.47, SD=4.15) followed by disgust (M=18.84, SD=4.39). Surprise responses were used most frequently (M=23.30, SD=5.61) followed by neutral (M=21.32, SD=1.89) and anger (M=20.07, SD=4.71).

Planned Comparisons for Unfiltered Expressions

In order to see whether the effect of initial fixation location on emotion recognition accuracy was consistent between Chapter 2 and Chapter 4, planned comparisons were conducted on the unbiased hit rates for unfiltered faces. In Chapter 2, we hypothesized that initially fixating the brow would improve anger recognition accuracy. We also hypothesized and provided evidence that initial fixation on the mouth improved emotion recognition accuracy for fear, surprise and disgust. Therefore one-tailed paired samples t-tests were conducted comparing the unbiased hit rates for anger when initially fixating the brow to other initial fixation locations. Planned comparisons were also conducted comparing unbiased hit rates for fear, surprise and disgust when initially fixating the mouth to other initial fixation locations.

There was no difference between unbiased hit rates for angry expressions when initially fixating the brow compared to the center ($t_{(35)}$=-.18, *p*=.43) or the mouth ($t_{(35)}$=-.51, *p*=.31). Initial fixation on the mouth did not lead to higher fear recognition compared to the brow ($t_{(35)}$=1.08, *p*=.14) or the center ($t_{(35)}$=-.10, *p*=.45). Initial fixation on the mouth also did not lead to better emotion recognition for surprised expressions compared to the brow ($t_{(35)}$=.87, *p*=.19) or the cheeks ($t_{(35)}$=-1.56, *p*=.06). Initial fixation on the mouth improved disgust recognition compared to the brow ($t_{(35)}$=2.87, *p*<.001) but not the center ($t_{(35)}$=.97, *p*=.17).

References

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*(6507), 669–672. https://doi.org/10.1038/372669a0

Adolphs, R., Tranel, D., Hamann, S., Young, A., Calder, A. J., Phelps, E., … Damasio, A. R. (1999). Recognition of Facial Emotion in Nine Subjects with Bilataral Amygdala Damage. *Neuropsychologia*, *37*, 1111–1117.

Adolphs, Ralph, Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature*, *433*(7021), 68–72. https://doi.org/10.1038/nature03086

Arizpe, J., Kravitz, D. J., Yovel, G., & Baker, C. I. (2012). *Start Position Strongly Influences Fixation Patterns during Face Processing : Difficulties with Eye Movements as a Measure of Information Use*. *7*(2). https://doi.org/10.1371/journal.pone.0031106

Atkinson, A. P., & Smithson, H. E. (2013). Distinct Contributions to Facial Emotion Perception of Foveated versus Nonfoveated Facial Features. *Emotion Review*, *5*(1), 30–35. https://doi.org/10.1177/1754073912457226

Atkinson, A.P. and Smithson, H.E. (2020) The impact on emotion classification performance and gaze behavior of foveal versus extrafoveal processing of facial features. *Journal of experimental psychology: Human perception and performance*, *46*(3), 292-312.

Azzopardi, P., & Cowey, A. (1993). Preferential representation of the fovea in the primary visual cortex. *Nature*, *361*(6414), 719–721. https://doi.org/10.1038/361719a0

Banks, M. S., Sekuler, A. B., & Anderson, S. J. (1991). Peripheral spatial vision: limits imposed by optics, photoreceptors, and receptor pooling. *Journal of the Optical Society of America A*, *8*(11), 1775. https://doi.org/10.1364/JOSAA.8.001775

Bannerman, R. L., Hibbard, P. B., Chalmers, K., & Sahraie, A. (2012). Saccadic latency is modulated by emotional content of spatially filtered face stimuli. *Emotion*, *12*(6), 1384–1392. https://doi.org/10.1037/a0028677

Bavelas, J. B., Black, A., Lemery, C. R., & Mullett, J. (1986). "I Show How You Feel". Motor Mimicry as a Communicative Act. *Journal of Personality and Social Psychology*, *50*(2), 322–329. https://doi.org/10.1037/0022-3514.50.2.322

Bayle, D. J., Schoendorff, B., Hénaff, M.-A., & Krolak-Salmon, P. (2011). Emotional Facial Expression Detection in the Peripheral Visual Field. *PLoS ONE*, *6*(6), e21584.

https://doi.org/10.1371/journal.pone.0021584

Beaudry, O., Roy-charland, A., Perron, M., & Cormier, I. (2014). Featural processing in recognition of emotional facial expressions. *Cognition and Emotion*, *28*(3), 416–432. https://doi.org/10.1080/02699931.2013.833500

Benson, P. J., Campbell, R., Harris, T., Frank, M. G., & Tovée, M. J. (1999). Enhancing images of facial expressions. *Perception and Psychophysics*, *61*(2), 259–274. https://doi.org/10.3758/BF03206887

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision*, *9*(2:7), 1–16. https://doi.org/10.1167/9.2.7

Blairy, S., Kleck, R. E., & Hess, U. (1979). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behaviour*, *21*(4), 241-257

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*, *3*(8). https://doi.org/10.1371/journal.pone.0003022

Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia*, *50*(12), 2830–2838. https://doi.org/10.1016/j.neuropsychologia.2012.08.010

Bodenschatz, C. M., Kersting, A., & Suslow, T. (2018). Effects of Briefly Presented Masked Emotional Facial Expressions on Gaze Behavior: An Eye-Tracking Study. *Psychological Reports*. https://doi.org/10.1177/0033294118789041

Boll, S., & Gamer, M. (2014). 5-HTTLPR modulates the recognition accuracy and exploration of emotional facial expressions. *Frontiers in Behavioral Neuroscience*, *8*(JULY), 1–10. https://doi.org/10.3389/fnbeh.2014.00255

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226*(5241), 177-178.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433-436

Burt, D. M., & Perrett, D. I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia*, *35*(5), 685–693. https://doi.org/10.1016/S0028-3932(96)00111-X

Bullier, J. (2001). Integrated model of visual processing. *Brain Research Reviews,* 36, 96-107

Butler, S., Gilchrist, I. D., Burt, D. M., Perrett, D. I., Jones, E., & Harvey, M. (2005). Are the perceptual biases found in chimeric face processing reflected in eye-movement patterns? *Neuropsychologia*, *43*(1), 52–59. https://doi.org/10.1016/j.neuropsychologia.2004.06.005

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. https://doi.org/10.1016/j.visres.2016.05.008

Caldara, R., Zhou, X., & Miellet, S. (2010). Putting culture under the “Spotlight” reveals universal information use for face recognition. *PLoS ONE*, *5*(3), 1–12. https://doi.org/10.1371/journal.pone.0009708

Calder, A. J., Young, A. W., Rowland, D., & Perrett, D. I. (1997). Computer-enhanced emotion in facial expressions. *Proceedings of the Royal Society B: Biological Sciences*, *264*(1383), 919–925. https://doi.org/10.1098/rspb.1997.0127

Calder, Andrew J., Rowland, D., Young, A. W., Nimmo-Smith, I., Keane, J., & Perrett, D. I. (2000). Caricaturing facial expressions. *Cognition*, *76*(2), 105–146. https://doi.org/10.1016/S0010-0277(00)00074-3

Calder, Andrew J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527–551. https://doi.org/10.1037/0096-1523.26.2.527

Calvo, M. G., Avero, P., Fernández-Martín, A., & Recio, G. (2016). Recognition thresholds for static and dynamic emotional faces. *Emotion*, Vol. 16, pp. 1186–1200. https://doi.org/10.1037/emo0000192

Calvo, M. G., Fernández-Martín, A., Gutiérrez-García, A., & Lundqvist, D. (2018). Selective eye fixations on diagnostic face regions of dynamic emotional expressions: KDEF-dyn database. *Scientific Reports*, *8*(1), 1–10. https://doi.org/10.1038/s41598-018-35259-w

Calvo, M. G., Fernández-Martín, A., & Nummenmaa, L. (2014). Facial expression recognition in peripheral versus central vision: Role of the eyes and the mouth. *Psychological Research*, *78*(2), 180–195. https://doi.org/10.1007/s00426-013-0492-x

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods*, *40*(1), 109–115. https://doi.org/10.3758/BRM.40.1.109

Calvo, M. G., Nummenmaa, L., & Avero, P. (2010). Recognition advantage of happy faces in extrafoveal vision: Featural and affective processing. *Visual Cognition*, *18*(9), 1274–1297. https://doi.org/10.1080/13506285.2010.481867

Carrasco, M., Talgar, C. P., & Cameron, E. L. (2001). Characterizing visual performance fields: Effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spatial Vision*, *15*(1), 61–75. https://doi.org/10.1163/15685680152692015

Chamberland, J., Roy-Charland, A., Perron, M., & Dickinson, J. (2017). Distinction between fear and surprise: an interpretation-independent test of the perceptual-attentional limitation hypothesis. *Social Neuroscience*, *12*(6), 751–768. https://doi.org/10.1080/17470919.2016.1251964

Chen, M.Y., & Chen, C.C. (2010). The contribution of the upper and lower face in happy and sad facial expression classification. *Vision Research*, *50*(18), 1814–1823. https://doi.org/10.1016/j.visres.2010.06.002

Chen, J. & Tiddeman, B. (2010). Multi-cue facial feature detection and tracking under various illuminations, *International Journal of Robotics and Automation*, *25*(2)

Corden, B., Chilvers, R., & Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger's syndrome. *Neuropsychologia*, *46*(1), 137–147. https://doi.org/10.1016/j.neuropsychologia.2007.08.005

Coren, S., & Hoenig, P. (1972). Effect of Non-Target Stimuli upon Length of Voluntary Saccades. *Perceptual and Motor Skills*, *34*, 499–508.

Delicato, L. S. (2020). A robust method for measuring an individual's sensitivity to facial expressions. *Attention, Perception, and Psychophysics*. https://doi.org/10.3758/s13414-020-02043-w

Du, S., & Martinez, A. M. (2011). The resolution of facial expressions of emotion. *Journal of Vision*, *11*(13), 24–24. https://doi.org/10.1167/11.13.24

Duncan, J., Gosselin, F., Cobarro, C., Dugas, G., Blais, C., & Fiset, D. (2017). Orientations for the successful categorization of facial expressions and their link with facial features. *Journal of Vision*, *17*(14), 7. https://doi.org/10.1167/17.14.7

Eisenbarth, H., & Alpers, G. W. (2011). *Happy Mouth and Sad Eyes : Scanning Emotional Facial Expressions*. *11*(4), 860–865. https://doi.org/10.1037/a0022758

Ekman, P., & Friesen, W. V. (1976). Measuring Facial Movement. *Environmental Psychology and Nonverbal Behavior*, *1*(1), 56–75. https://doi.org/10.1037/h0030377

Ekman, P., Friesen, W. V, O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., … Tzavaras, A. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, Vol. 53, pp. 712–

717. https://doi.org/10.1037/0022-3514.53.4.712

Ekman, P., Friesen, W. V, Sullivan, M. O., Chan, A., Heider, K., Lecompte, W. A., … Scherer, K. (1987). *Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion*. *53*(4), 712–717. https://doi.org/10.1037/0022-3514.53.4.712

Ekman, P., & Friesen, W. V. (1978). *The Facial Action Coding System (FACS): A technique for the measurement of facial action*. Palo Alto, CA: Consulting Psychologists Press.

Ekman, P., Sorenson, R. E., & Friesen, W. V. (1969). Pan-Cultural Elements in Facial Displays of Emotion. *Science*, *164*(3875).

Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research*, *22*(8), 1033–1045. https://doi.org/10.1016/0042-6989(82)90040-2

Findlay, J. M., & Gilchrist, I. D. (1997). Spatial scale and saccade programming. *Perception*, *26*(9), 1159–1167. https://doi.org/10.1068/p261159

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences*, *22*(4), 661–674. https://doi.org/10.1017/S0140525X99002150

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Gagnon, M., Gosselin, P., Hudon-ven der Buhs, I., Larocque, K., & Milliard, K. (2010). Children’s recognition and discrimination of fear and disgust facial expressions. *Journal of Nonverbal Behavior*, *34*(1), 27–42. https://doi.org/10.1007/s10919-009-0076-z

Gamer, M., & Büchel, C. (2009). Amygdala activation predicts gaze toward fearful eyes. *Journal of Neuroscience*, *29*(28), 9123–9126. https://doi.org/10.1523/JNEUROSCI.1883-09.2009

Gamer, M., Schmitz, A. K., Tittgemeyer, M., & Schilbach, L. (2013). The human amygdala drives reflexive orienting towards facial features. *Current Biology*, Vol. 23. https://doi.org/10.1016/j.cub.2013.09.008

Goffaux, Valerie, Peters, J., Haubrechts, J., Schiltz, C., Jansma, B., & Goebel, R. (2011). From coarse to fine? spatial and temporal dynamics of cortical face processing. *Cerebral Cortex*, *21*(2), 467–476. https://doi.org/10.1093/cercor/bhq112

Goffaux, Valérie, & Rossion, B. (2006). Faces are “spatial” - Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human*

Perception and Performance, *32*(4), 1023–1039. https://doi.org/10.1037/0096-1523.32.4.1023

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research*, *41*(17), 2261–2271. https://doi.org/10.1016/S0042-6989(01)00097-9

Gosselin, F., & Schyns, P. G. (2002). RAP: A new framework for visual categorization. *Trends in Cognitive Sciences*, Vol. 6, pp. 70–77. https://doi.org/10.1016/S1364-6613(00)01838-6

Griffiths, S., Jarrold, C., Penton-Voak, I. S., Woods, A. T., Skinner, A. L., & Munafò, M. R. (2019). Impaired Recognition of Basic Emotions from Facial Expressions in Young People with Autism Spectrum Disorder: Assessing the Importance of Expression Intensity. *Journal of Autism and Developmental Disorders*, *49*(7), 2768–2778. https://doi.org/10.1007/s10803-017-3091-7

Guo, K. (2007). Initial fixation placement in face images is driven by top-down guidance. *Experimental Brain Research*, *181*(4), 673–677. https://doi.org/10.1007/s00221-007-1038-5

Guo, K. (2012). Holistic Gaze Strategy to Categorize Facial Expression of Varying Intensities. *PLoS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0042585

Guo, K. (2013). Size-invariant facial expression categorization and associated gaze allocation within social interaction space. *Perception*, *42*(10), 1027–1042. https://doi.org/10.1068/p7552

Guo, K., Robertson, R. G., Mahmoodi, S., Tadmor, Y., & Young, M. P. (2003). How do monkeys view faces? - A study of eye movements. *Experimental Brain Research*, *150*(3), 363–374. https://doi.org/10.1007/s00221-003-1429-1

Guo, K., & Shaw, H. (2015). Face in profile view reduces perceived facial expression intensity: An eye-tracking study. *Acta Psychologica*, *155*, 19–28. https://doi.org/10.1016/j.actpsy.2014.12.001

Guo, K., Soornack, Y., & Settle, R. (2019). Expression-dependent susceptibility to face distortions in processing of facial expressions of emotion. *Vision Research*, *157*(October 2017), 112–122. https://doi.org/10.1016/j.visres.2018.02.001

Hager, J. C., & Ekman, P. (1979). Long-distance of transmission of facial affect signals. *Ethology and Sociobiology*, *1*(1), 77–82. https://doi.org/10.1016/0162-3095(79)90007-4

He, P., & Kowler, E. (1989). The role of location probability in the programming of saccades:

Implications for "center-of-gravity" tendencies. *Vision Research*, *29*(9), 1165–1181. https://doi.org/10.1016/0042-6989(89)90063-1

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*(1), 98–106.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*(4), 241–257. https://doi.org/10.1023/A:1024952730333

Horstmann, G., Lipp, O. V., & Becker, S. I. (2012). Of toothy grins and angry snarls-open mouth displays contribute to efficiency gains in search for emotional faces. *Journal of Vision*, *12*(5), 1–15. https://doi.org/10.1167/12.5.7

Hsiao, J. H., & Cottrell, G. (2008). Two Fixations Suffice in Face Recognition. *Psychological Science*, *19*(10), 998–1006.

Innes, R. B., Burt, D. M., Birch, Y. K. & Hausmann, M. (2016). A leftward bias however you look at it: revisiting the emotional chimeric face task as a tool for measuring emotion lateralization. *Laterality,* 21(4-6): 643-661.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural Confusions Show that Facial Expressions Are Not Universal. *Current Biology*, *19*(18), 1543–1548. https://doi.org/10.1016/j.cub.2009.07.051

Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current Biology*, *24*(2), 187–192. https://doi.org/10.1016/j.cub.2013.11.064

Jack, R. E., & Schyns, P. G. (2015). The Human Face as a Dynamic Tool for Social Communication. *Current Biology*, *25*(14), R621–R634. https://doi.org/10.1016/j.cub.2015.05.052

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*(6). https://doi.org/10.1037/xge0000162

Jonas, J. B., Schneider, U., & Naumann, G. O. H. (1992). *Count and density of human retinal photoreceptors*. *230*, 505–510.

Kelly, D. J., Miellet, S., & Caldara, R. (2010). Culture shapes eye movements for visually homogeneous objects. *Frontiers in Psychology, 1*(APR), 1–7. https://doi.org/10.3389/fpsyg.2010.00006

Kennedy, D. P., & Adolphs, R. (2010). Neuropsychologia Impaired fixation to eyes following amygdala damage arises from abnormal bottom-up attention. *Neuropsychologia*, *48*(12), 3392–3398. https://doi.org/10.1016/j.neuropsychologia.2010.06.025

Kestenbaum, R., & Nelson, C. A. (1990). The recognition and categorization of upright and inverted emotional expressions by 7-month-old infants. *Infant Behavior and Development*, *13*(4), 497–511. https://doi.org/10.1016/0163-6383(90)90019-5

Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical Reflexive Gaze Patterns on Emotional Faces in Autism Spectrum Disorders. *Journal of Neuroscience*, *30*(37), 12281–12287. https://doi.org/10.1523/JNEUROSCI.0688-10.2010

Kliemann, Dorit, Dziobek, I., Hatri, A., & Heekeren, H. R. (2012). The Role of the Amygdala in Atypical Gaze on Emotional Faces in Autism Spectrum Disorders. *The Journal of Neuroscience*, *32*(28), 9469–9476. https://doi.org/10.1523/JNEUROSCI.5294-11.2012

Kliemann, Dorit, Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical reflexive gaze patterns on emotional faces in autism spectrum disorders. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(37), 12281–12287. https://doi.org/10.1523/JNEUROSCI.0688-10.2010

Laidlaw, K. E. W., & Kingstone, A. (2017). Fixations to the eyes aids in facial encoding; covertly attending to the eyes does not. *Acta Psychologica*, *173*, 55–65. https://doi.org/10.1016/j.actpsy.2016.11.009

Laidlaw, K. E. W., Risko, E. F., & Kingstone, A. (2012). A new look at social attention: Orienting to the eyes is not (entirely) under volitional control. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1132–1143. https://doi.org/10.1037/a0027075

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the radboud faces database. *Cognition and Emotion*, *24*(8), 1377–1388. https://doi.org/10.1080/02699930903485076

Leonards, U., & Scott-Samuel, N. E. (2005). Idiosyncratic initiation of saccadic face exploration in humans. *Vision Research*, *45*(20), 2677–2684. https://doi.org/10.1016/j.visres.2005.03.009

Liu, P., Montaser-Kouhsari, L., & Xu, H. (2014). Effects of face feature and contour crowding in facial expression adaptation. *Vision Research*, *105*, 189–198. https://doi.org/10.1016/j.visres.2014.10.014

Liu, T., Heeger, D. J., & Carrasco, M. (2006). Neural correlates of the visual vertical meridian asymmetry. *Journal of Vision*, *6*(11), 1294–1306. https://doi.org/10.1038/jid.2014.371

Loftus, G. & Harley, E. (2005). Why is it easier to identify someone close than far away? *Psychonomic Bulletin and Review*, *12*(1), 43-65

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. San Francisco, CA: W.H. Freeman

Maher, S., Ekstrom, T., & Chen, Y. (2014). Greater perceptual sensitivity to happy facial expression. *Perception*, *43*(12), 1353–1364. https://doi.org/10.1068/p7806

Marneweck, M., Loftus, A., & Hammond, G. (2013). Psychophysical measures of sensitivity to facial expression of emotion. *Frontiers in Psychology*, *4*(FEB), 1–6. https://doi.org/10.3389/fpsyg.2013.00063

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, *5*(1), 6. https://doi.org/10.1167/5.1.6

Matsumoto, D., & Hwang, H. C. (2014). *Judgments of Subtle Facial Expressions of Emotion*. *14*(2), 349–357. https://doi.org/10.1037/a0035237

McKone, E. (2009). Holistic processing for faces operates over a wide range of sizes but is strongest at identification rather than conversational distances. *Vision Research*, *49*(2), 268–283. https://doi.org/10.1016/j.visres.2008.10.020

Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision, 14*(7). https://doi.org/10.1167/14.7.6

Miellet, S., He, L., Zhou, X., Lao, J., & Caldara, R. (2012). When East meets West: Gaze-contingent Blindspots abolish cultural diversity in eye movements for faces. *Journal of Eye Movement Research*, *5*(2), 1–12. https://doi.org/10.16910/jemr.5.2.5

Neath-Tavares, K. N., & Itier, R. J. (2016). Neural processing of fearful and happy facial expressions during emotion-relevant and emotion-irrelevant tasks: A fixation-to-feature approach. *Biological Psychology*, *119*, 122–140. https://doi.org/10.1016/j.biopsycho.2016.07.013

Neath, K. N., & Itier, R. J. (2014). Facial expression discrimination varies with presentation time but not with fixation on features : A backward masking study using eye- tracking. *COGNITION AND EMOTION*, *28*(1), 115–131. https://doi.org/10.1080/02699931.2013.812557

Or, C. C. F., Peterson, M. F., & Eckstein, M. P. (2015). Initial eye movements during face identification are optimal and similar across cultures. *Journal of Vision*, *15*(13), 12. https://doi.org/10.1167/15.13.12

Oruç, I., & Barton, J. J. S. (2010). Critical frequencies in the perception of letters, faces, and novel shapes: Evidence for limited scale invariance for faces. *Journal of Vision*, *10*(12), 1–12. https://doi.org/10.1167/10.12.20

Palermo, R., & Coltheart, M. (2004). Photographs of facial expression: Accuracy, response times, and ratings of intensity. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 634–638. https://doi.org/10.3758/BF03206544

Papinutto;, M., Lao;, J., Ramon;, M., Caldara;, R., & Miellet, S. (2017). The Facespan—the perceptual span for face recognition. *Journal of Vision June*, *17*(5), 1–12. https://doi.org/doi:10.1167/17.5.16

Parkinson, B. (2005). Do facial movements express emotions or communicate motives? *Personality and Social Psychology Review*, *9*(4), 278–311. https://doi.org/10.1207/s15327957pspr0904_1

Pelli, D. G. (1997). The VideoToolbox software for visual psychysics: Transforming numbers into movies, *Spatial Vision 10,* 437-442

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*(10), 1129–1135. https://doi.org/10.1038/nn.2187

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *PNAS*, *109*(48). https://doi.org/10.1073/pnas.1214269109

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., … Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine*, *40*(11), 1919–1929. https://doi.org/10.1017/S0033291709992364

Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? Eye movements reduce local uncertainty. *Journal of Vision, 7*(3), 1–17. https://doi.org/10.1167/7.3.6.Introduction

Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research*, *21*(3), 409–418. https://doi.org/10.1016/0042-6989(81)90169-3

Rosenholtz, R. (2016). Capabilities and Limitations of Peripheral Vision. *The Annual Review of*

Vision Science, *2*, 437–457. https://doi.org/10.1146/annurev-vision-082114-035733

Roy-Charland, A., Perron, M., Beaudry, O., & Eady, K. (2014). Confusion of fear and surprise: A test of the perceptual-attentional limitation hypothesis with eye movement monitoring. *Cognition and Emotion*, *28*(7), 1214–1222. https://doi.org/10.1080/02699931.2013.878687

Roy-Charland, A., Perron, M., Young, C., Boulard, J., & Chamberland, J. A. (2015a). The Confusion of Fear and Surprise: A Developmental Study of the Perceptual-Attentional Limitation Hypothesis Using Eye Movements. *Journal of Genetic Psychology*, *176*(5), 281–298. https://doi.org/10.1080/00221325.2015.1066301

Roy-Charland, A., Perron, M., Young, C., Boulard, J., & Chamberland, J. A. (2015b). The Confusion of Fear and Surprise: A Developmental Study of the Perceptual-Attentional Limitation Hypothesis Using Eye Movements. *Journal of Genetic Psychology*, *176*(5), 281–298. https://doi.org/10.1080/00221325.2015.1066301

Saumure, C., Plouffe-Demers, M.-P., Estéphan, A., Fiset, D., & Blais, C. (2018). The use of visual information in the recognition of posed and spontaneous facial expressions. *Journal of Vision*, *18*(9), 21. https://doi.org/10.1167/18.9.21

Scheller, E., Büchel, C., & Gamer, M. (2012). Diagnostic features of emotional expressions are processed preferentially. *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0041792

Schurgin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion recognition in faces. *Journal of Vision*, *14*(13), 14–14. https://doi.org/10.1167/14.13.14

Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). *Transmission of Facial Expressions of Emotion Co- Evolved with Their Efficient Decoding in the Brain : Behavioral and Brain Evidence*. *4*(5). https://doi.org/10.1371/journal.pone.0005625

Sekiguchi, T. (2011). Individual differences in face memory and eye fixation patterns during face learning. *Acta Psychologica*, *137*(1), 1–9. https://doi.org/10.1016/j.actpsy.2011.01.014

Şentürk, G., Greenberg, A. S., & Liu, T. (2016). Saccade latency indexes exogenous and endogenous object-based attention. *Attention, Perception, and Psychophysics*, *78*(7), 1998–2013. https://doi.org/10.3758/s13414-016-1136-1

Shanok, N. A., Jones, N. A., & Lucas, N. N. (2019). The Nature of Facial Emotion Recognition Impairments in Children on the Autism Spectrum. *Child Psychiatry and Human*

Development, *50*(4), 661–667. https://doi.org/10.1007/s10578-019-00870-z

Skowronski, J. J., Milner, J. S., Wagner, M. F., Crouch, J. L., & McCanne, T. R. (2014). Pushing the boundaries of human expertise in face perception: Emotion expression identification and error as a function of presentation angle, presentation time, and emotion. *Journal of Experimental Social Psychology*, *50*(1), 166–174. https://doi.org/10.1016/j.jesp.2013.10.001

Smith, F.W., & Schyns, P. G. (2009). Smile through your fear and sadness. *Psychological Science*, *20*(10), 1202–1209. https://doi.org/10.1111/j.1467-9280.2009.02427.x

Smith, Fraser W., & Rossit, S. (2018). Identifying and detecting facial expressions of emotion in peripheral vision. *PLoS ONE*, *13*(5), 1–15. https://doi.org/10.1371/journal.pone.0197160

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*(3), 184–189. https://doi.org/10.1111/j.0956-7976.2005.00801.x

Smith, M. L., Grühn, D., Bevitt, A., Ellis, M., Ciripan, O., Scrimgeour, S., … Ewing, L. (2018). Transmitting and decoding facial expressions of emotion during healthy aging: More similarities than differences. *Journal of Vision*, *18*(9), 1–16. https://doi.org/10.1167/18.9.10

Smith, M. L., & Merlusca, C. (2014). *How Task Shapes the Use of Information During Facial Expression Categorizations*. *14*(3), 478–487. https://doi.org/10.1037/a0035588

Song, H., Chui, T. Y. P., Zhong, Z., Elsner, A. E., & Burns, S. A. (2011). Variation of cone photoreceptor packing density with retinal eccentricity and age. *Investigative Ophthalmology and Visual Science*, *52*(10), 7376–7384. https://doi.org/10.1167/iovs.11-7199

Sowden, P.T., Schyns, P.G. (2006). Channel surfing in the visual brain. *Trends in Cognitive Sciences, 10,* 538-545.

Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007). Analysis of face gaze in autism using “Bubbles.” *Neuropsychologia*, *45*(1), 144–151. https://doi.org/10.1016/j.neuropsychologia.2006.04.027

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, *11*(5), 1–82. https://doi.org/10.1167/11.5.13.Contents

Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008).

Expressing fear enhances sensory acquisition. *Nature Neuroscience*, *11*(7), 843–850. https://doi.org/10.1038/nn.2138

Talgar, C. P., & Carrasco, M. (2002). Vertical meridian asymmetry in spatial resolution: Visual and attentional factors. *Psychonomic Bulletin and Review*, *9*(4), 714–722. https://doi.org/10.3758/BF03196326

Tanaka, J. W., & Farah, M. J. (1993). Parts and Wholes in Face Recognition. *The Quarterly Journal of Experimental Psychology Section A*, *46*(2), 225–245. https://doi.org/10.1080/14640749308401045

Tanaka, J. W., Kaiser, M. D., Butler, S., & Grand, R. Le. (2012). Mixed emotions : Holistic and analytic perception of facial expressions Mixed emotions : Holistic and analytic perception of facial expressions. *Cognition and Emotion*, *26*(6), 961–977. https://doi.org/10.1080/02699931.2011.630933

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*(5), 583–592. https://doi.org/10.3758/BF03211301

Thompson, S. J., Foulsham, T., Leekam, S. R., & Jones, C. R. G. (2019). Attention to the face is characterised by a difficult to inhibit first fixation to the eyes. *Acta Psychologica*, *193*(January), 229–238. https://doi.org/10.1016/j.actpsy.2019.01.006

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., … Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

Vaidya, A. R., Jin, C., & Fellows, L. K. (2014). Eye spy: The predictive value of fixation patterns in detecting subtle and extreme emotions from faces. *Cognition*, *133*(2), 443–456. https://doi.org/10.1016/j.cognition.2014.07.004

Van Belle, G., de Graef, P., Verfaillie, K., Rossion, B., & Lefèvre, P. (2010). Face inversion impairs holistic perception: Evidence from gaze-contingent stimulation. *Journal of Vision*, *10*(5). https://doi.org/10.1167/10.5.10

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research*, *37*(3), 475–494.

https://doi.org/10.1007/BF00236818

Watt, R.J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America A,* 4, 2006-2021

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS ONE*, *12*(5), 1–15. https://doi.org/10.1371/journal.pone.0177239

Widen, S. C., Russell, J. A., & Brooks, A. (2004). *Anger and Disgust: Discrete or Overlapping categories?*

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology*, *19*(6), 739–749. https://doi.org/10.1037/0894-4105.19.6.739

Yang, N., Shafai, F., & Oruc, I. (2014). Size determines whether specialized expert processes are engaged for recognition of faces. *Journal of Vision*, *14*(8), 1–12. https://doi.org/10.1167/14.8.17

Yarbus, A. L. (1967). *Eye Movements and Vision*. New York: Plenum Press.

Yin, R.K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81,* 141-145

Young, A.W., Hellawell, D., & Hay, D. C. (1987), Configural information in face perception. *Perception, 16,* 747-759

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, *63*(3), 271–313. https://doi.org/10.1016/S0010-0277(97)00003-6

Yu, D. and, & Chung, S. T. L. (2011). Critial orientation for face identification in central vision loss. *Optometry and Vision Science*, *88*(6), 724–732. https://doi.org/10.1021/nl061786n.Core-Shell