University of Glasgow | Institute of Neuroscience & Psychology

Understanding age-related differences in the content and temporal dynamics of visual information processing

Hannah Lois Gilman

MA, MSc

Submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

February 2019

School of Psychology

Institute of Neuroscience and Psychology

College of Science and Engineering

University of Glasgow

Abstract

The N170, a negative amplitude peak occurring at approximately 170 ms post-stimulus onset, is an event-related potential (ERP) component observed during electroencephalography (EEG) recordings that preferentially responds to faces compared to other objects (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Bötzel, Schulze, & Stodieck, 1995). EEG research has suggested that the N170 may be modulated by the eye region, however this has received much debate (e.g. Bentin et al., 1996; Eimer,1998; Taylor, Itier, Allison, & Edmonds, 2001). Most recently, Rousselet, Ince, Rijsbergen, & Schyns, (2014) used Gaussian apertures ('bubbles') (Gosselin & Schyns, 2001) in a reverse correlation experiment to demonstrate that increased visibility of the contralateral eye leads to larger and earlier N170s in a face versus noise detection task. However, these results may be explained by the phenomenon of 'left gaze bias' – a preferential looking towards the left visual field. To understand if contralateral eye sensitivity can be explained by a non-feature specific attentional bias to the left, in the first study (Chapter 2) we investigated contralateral eye sensitivity to faces of different image sizes in a face versus noise detection task. Using reverse correlation and Mutual Information (MI) we found that contralateral eye sensitivity is size tolerant, suggesting that contralateral eye sensitivity does reflect feature encoding rather than a general left attentional bias. Next we wanted to address whether eye coding precedes other feature encoding in a more heterogeneous face set. The traditional 'bubbles' technique relies on stimuli being spatially aligned i.e. the eyes, nose, mouth of all images in the stimulus set to be in comparable positions for averaging bubble-masks across stimuli. To overcome this limitation, we used an adaption of *BubbleWarp* (Gill, DeBruine, Jones, & Schyns, 2015) a new technique outlined in Chapter 3, to retrospectively 'warp' Gaussian bubble masks to an average face image. Using this new technique, in Chapter 4, we tested the assumption that contralateral eye sensitivity preceded sensitivity to other facial features, specifically the mouth, in a gender and expressive versus non-expressive (EXNEX) categorisation task in young adult participants. Using MI onset analysis, we found idiosyncratic differences in MI onsets suggesting preferential encoding of the eye before other facial features for ~65 % of participants. This revealed that whilst there is an eye

coding preference, there is not a constraint to encoding the contralateral eye before other facial features. Aging is marked by a decline in processing speed (Salthouse, 1996) and previous work has suggested that whilst older adults process the contralateral eye in face versus noise detection tasks, this processing is weaker and delayed compared to younger adults (Jaworska, 2017). In Chapter 5 using the same task as in Chapter 4, we quantified age-related differences in feature processing speed by calculating 50 % integration times in younger and older participants. There was a ~20 ms delay in eye encoding for older compared to younger adults. We found a 9 ms delay in mouth encoding in the gender task and no differences in mouth processing speed in the EXNEX task. This suggests that there was not a general, uniform delay in processing speed of all facial features across tasks. Overall, our results demonstrate for the first time that 1) contralateral eye sensitivity is tolerant to changes in stimulus size, stimulus set, task demands and age, 2) contralateral eye sensitivity preferentially precedes sensitivity to the mouth but is not a prerequisite in gender or EXNEX categorisation tasks and 3) older adults process the same facial feature information as younger adults, but feature coding is not uniformly delayed compared to younger adults.

Table of Contents

List of Tables

List of Figures

List of Supplementary Figures

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Guillaume Rousselet, for all of your support throughout my Masters and PhD. Thank you for your unique blend of humour and patience. Your enthusiasm at a new figure or analysis has kept me believing in this project even during the worst of times. Your zealousness for robust methods still inspires me, and I am in no doubt that working with you has succeeded in challenging me to push myself and what I can achieve.

I would like to thank Dr. Nicola van Rijsbergen for going above and beyond both professionally and personally, during the course of this PhD and previously. No words will ever express how grateful I am for your encouragement and support in all its forms.

To all my former-PhD lab colleagues - Dr. Katarzyna Jaworska, Dr. Fei Yi and Dr. Magdalena Bieniek – I could not have asked to have followed in better footsteps or been part of a more supportive lab group. I would also like to thank all the students I've worked with over the years - Victoria Nicholls, Yiqiong Yang, Josh Hamilton, Emma Robb, Morika Georgieva, Marsella, Abbas Ali & Rory Boyle – for their company and assistance in collecting data. And of course, my thanks go to all of the individuals who participated in these studies, especially the older adults who continue to give up their time for research and who are the most remarkable, diverse and humorous group of participants I've had the pleasure of spending time with in the lab.

I'd also like to thank Dr. Robin Ince for his patience in answering all my mutual information questions, and Dr. Oliver Garrod and Dr. Daniel Gill for your assistance with implementing Bubble-Warp. I would also like to thank Prof. Phillippe Schyns and his lab group for accommodating me in lab meetings.

My thanks go to the staff at Glasgow University Counselling and Psychological Services for the speed and quality of the support I received at times of difficulty, and in particular to Dr. NJ – thank you.

And finally Jessie – you never signed up for a PhD, but you've endured this journey with me anyway, thank you. I promise we can go camping again now.

Chapter 1 - Introduction

Faces are important social stimuli for achieving mutual understanding and communicating a wide range of signals, such as threat and aggression. We engage in many basic face processing tasks automatically in naturalistic settings, such as judging the age, sex and emotional state from others faces. However, our understanding of the computational stages that visual processing undergoes from initial sensory input to decision making, and how these information processing stages are affected by healthy ageing remains elusive.

Of particular interest to psychophysicists (and others) is the posterior lateral electro and magneto encephalographic response (EEG / MEG) occurring in the range of 140 to 200 ms subsequent to the presentation of a face stimulus. This response – the N170 event-related potential (ERP) in EEG and its MEG equivalent the M170 – has been a subject of considerable debate (Earp & Everett, 2013). Specifically arguments have concentrated on whether there is specialist face processing in the brain. This has resulted in arguments over whether the N170 is face *sensitive* (i.e. is modulated *more* but not *exclusively* by faces than other objects) or face *specific* (i.e. is modulated *exclusively* by faces but not objects). Assuming the N170 is at least face *sensitive,* until recently little has been understood about the specific information, i.e. *what* is being processed during the time window of the N170, and its underlying mechanism, nor how that mechanism may change with ageing observers.

This thesis will expand upon work suggesting that the N170 predominantly encodes information about the contralateral eye (Rousselet, Ince, van Rijsbergen, & Schyns, 2014). The main contributions of this work are:

1. Demonstrating that contralateral eye sensitivity is tolerant to changes in stimulus size, stimulus set, task demands and age.

2. Testing whether encoding of the contralateral eye is task invariant i.e. whether contralateral eye encoding is the first stage of visual processing in a range of categorisation tasks (face versus texture; male versus female; expressive versus non-expressive).

3. Testing whether the timing of contralateral eye encoding is consistent relative to tasks involving another facial feature, namely the mouth.

4. Quantifying age-related differences in the strength and timing of feature encoding

Firstly, we will outline why faces are important stimuli and why difficulties processing faces have negative social consequences for older adults. Next, we will assess the evidence of the existence of the N170 and its neurological underpinnings. We will then address how the N170 in older adults' compares to that in younger adults and how these differences may reflect a general phenomenon of change in the visual processing system with ageing. Lastly, we will address evidence that the N170 is sensitive to and modulated by the presence of the contralateral eye in younger and older adults.

Faces are important social stimuli

Faces are important social stimuli with high evolutionary importance. Faces are dynamic tools for achieving mutual understanding and communicating social signals, such as threat and aggression (Zhang, 2018). Every day we judge the identity, age, sex, personality traits and emotional state of faces we encounter. Those who experience difficulty in face processing, for example due to developmental prosopagnosia, experience negative psychosocial consequences such as anxiety and the avoidance of social situations (Dalrymple et al., 2014; Yardley, McDermott, Pisarski, Duchaine, & Nakayama, 2008), suggesting that face processing is an integral component of our social interactions with others.

Age-related differences in face processing have important social consequences for older adults

Difficulties in processing faces as we age, such as reduced ability in older adults' recognition of facial identities (Boutet, Taler, & Collin, 2015; Konar, Bennett, & Sekuler, 2013; Meinhardt-Injac, Persike, & Meinhardt, 2014), assessment of sex and age (Carbon, Grüter, & Grüter, 2013) and emotion recognition (Ruffman, Henry, Livingstone, & Phillips, 2008; Sullivan, Campbell, Hutton, & Ruffman, 2017; Sullivan & Ruffman, 2004) may have similar negative psychosocial

consequences as those seen in developmental prosopagnosia. Difficulty with emotion recognition may be particularly problematic for older adults, as an inability to respond appropriately to social cues may lead to social isolation, reduced quality of life and may even increase physiological decline with ageing due to loneliness (Hawkley & Cacioppo, 2007).

Face processing

Faces may be processed qualitatively differently from other objects. For example, there may be a double-dissociation between specific impairments in the recognition of faces (prosopagnosia) and non-face objects (object agnosia) following lesion to the temporal lobe (Henke, Schweinberger, Grigo, Klos, & Sommer, 1998). It has also been suggested that faces are processed more "holistically" than other non-face objects - which are processed in a more "part-based" fashion (Piepers & Robbins, 2012). "Holistic" processing suggests that the combined processing of two or more of the basic features of a face (eyes, nose mouth) results in "emergent features" of the face becoming apparent – i.e. properties that cannot be derived from purely the properties of the individual facial features. Emergent features necessarily suggest that the interrelations between facial features are important (i.e. configural information). Configural information can be understood at two levels – *first-order relational properties* (i.e. the basic configuration of a face, such as 2 eyes above a central nose and mouth) and *second-order relational properties* (i.e. variations in the inter-feature spacing and positioning between features) (Piepers & Robbins, 2012).

The "two-streams hypothesis" of visual perception differentiates the dorsal-stream of vision as primarily encoding spatial relationships between objects, from the ventral-stream which encodes features related to object recognition. Within the ventral stream, complex objects, including faces, can be represented in a distributed or clustered manner. For example, the fusiform gyrus cluster for face processing. However, the dorsal stream may also hold behaviourally relevant information for face processing as well (Jeong & Xu, 2016).

There may also be differences between processing of familiar compared to unfamiliar faces. Familiar and unfamiliar faces may be represented differently, such as evidence from prosopagnosia patients demonstrating a dissociation

between familiar and unfamiliar face processing across a range of tasks (Johnston & Edmonds, 2009). However, it has been suggested that whilst the familiarity of the face may modify later ERP components, early stages of face processing may not be modulated by face familiarity (Johnston & Edmonds, 2009).

The N170 is a 'face selective' component

Event-related potentials (ERPs) are stimulus time-locked EEG changes in voltage. ERPs are the summed activity of post-synaptic potentials produced by pyramidal neurons (discussed later in this chapter: *The N170 neuronal sources)*. Early ERPs are thought to reflect physical properties of the stimulus, whilst later components are thought to reflect information processing (Sur & Sinha, 2009). Several ERPs components have been shown to be modulated by faces. For example, the N100 (a negative potential at around 100 ms post stimulus onset) may by modulated by affect, such as a larger N100 for fearful faces (Luo, Feng, He, Wang, & Luo, 2010). The P300 (positive voltage at approximately 300 ms) on the other hand is thought to reflect greater attention and stimulus evaluation. Much focus of the face processing literature however has ben directed to the N170.

The N170, a negative amplitude peak occurring at approximately 170 ms post-stimulus onset, is one of the earliest ERP components observed during EEG recordings that preferentially responds to faces (Bentin et al., 1996; Bötzel et al., 1995) (the P1 occurring before the N170 may also be face sensitive (Itier & Taylor, 2004)). The N170 is typically right lateralised, with larger amplitude over the right than left hemisphere, though there are individual differences to this pattern. The N170 has been shown to typically have larger amplitude for faces compared to other non-face objects (Rousselet, Husk, Bennett, & Sekuler, 2008) and varies in latency compared to other object categories (Nemrodov, Anderson, Preston, & Itier, 2014). The N170 response is also larger when objects are perceived as containing a face, even when a face is not there (so called 'face-in-things' stimuli) than for objects where no face is perceived (Proverbio & Galli, 2016). The N170 component and it's MEG equivalent the M170, are apparent even for schematic faces (Bentin, Sagiv, Mecklinger, Friederici, & von Cramon, 2002).

The N170 does not appear to be specific to human faces, with animal faces eliciting the N170 in human observers as well, though with a delayed peak (Carmel & Bentin, 2002; Rousselet, Mace, & Fabre-Thorpe, 2004). The N170 is elicited by human faces even when faces are not attended (Carmel & Bentin, 2002), though may diminish with repeated presentations of unattended face stimuli (Heisz, Watter, & Shedden, 2006a, 2006b; Mercure, Kadosh, & Johnson, 2011).

The N170 may be best understood as a late variant of the N1 component that occurs in the EEG signal in response to objects (Rossion & Jacques, (2011), though see Itier & Taylor, (2004) who argue the N170 is qualitatively different from the N1). As such, the N170 may be an enhanced N1 in response to faces. The magnitude and timing of the N170 in response to faces is idiosyncratic across participants, stimuli and task (Rossion & Jacques, 2011) and may also be influenced by methodological design, such as whether the same electrode is used for analysis across all participants, or whether electrode selection is optimised on a participant-by-participant basis. The increase in the amplitude for faces compared to objects could reflect alternate phenomena – either an increase in neural activity to faces compared to objects *or* a greater consistency in single-trial latency leading to decreased inter-trial jitter (Rossion & Jacques, 2011). Differences in the latency of the N170 may reflect the time taken to activate face representations, or the speed of accumulation of evidence at the neuronal level.

The N170 and individual differences

Individual differences in the latency and amplitude of ERP waveforms are rife, but inter-participant variability in ERPs is to be expected in EEG research (Luck, 2005). However, between-participants idiosyncrasies are consistent across sessions within participants, suggesting that these differences between participants are predominantly stable (Luck, 2005). These inter-participant differences may reflect anatomical differences e.g. in the folding pattern of the cortex affecting the size of the ERP component at scalp electrodes depending on its location – i.e. whether a generator of an ERP component is in the sulci or gyri (Luck, 2005). As a consequence of large individual variation, group results in the EEG literature should always be treated with caution in the absence of

supporting individual participant data and effect sizes. In this thesis we detail throughout the wide inter-participant variability in ERPs.

The N170 bottom-up / top-down debate

The underlying processing mechanism of the N170 is still heavily debated. Some authors suggest that the N170 reflects a bottom-up (i.e. stimulus driven) structural encoding of face features, proceeding the conscious detection of a face (Eimer, 2000; Sagiv & Bentin, 2001). However, others suggest a top-down (i.e. task or experience driven) mechanism, wherein the N170 may be modulated by task requirements, or experience - such as when geometrical shapes not normally processed as face components are primed to be interpreted as face patterns (Bentin & Golland, 2002; Kato et al., 2004). This suggests that the N170 may, in part, be modulated by an individual's subjective experience of perceiving a stimulus as a face, rather than objective reality. As such, the N170 may reflect a complex mixture of bottom-up and top-down influences.

The N170 neuronal sources

Identifying the neuronal sources underpinning the N170 is still controversial. This is in part due to the problem of estimating the location and distribution of the neuronal sources responsible for producing a given pattern of electrical potentials recorded on surface level scalp electrodes (the inverse problem). The EEG signal is primarily thought to reflect a summation of the excitatory and inhibitory potentials of the long perpendicular dendrites of cortical pyramidal neurons (Kirschstein & Köhling, 2009). Whilst individual neuronal action potentials are too small and fleeting to be recorded at the scalp, post-synaptic potentials have a longer duration and the summation of these can be recorded via EEG (Kirschstein & Köhling, 2009). Pyramidal neurons create electrical dipoles between the body of the neuron and the dendrites, with excitation/inhibition depending upon the distance of the excitatory postsynaptic potential in relation to the neuronal body (Jackson & Bolger, 2014). The EEG signal consists of the summation of these multiple, potentially counter-directional dipoles, so long as enough of the dipoles are in parallel and are synchronously active (Jackson & Bolger, 2014). The net summed positivity or negativity recording by any single given electrode will depend upon its relative

position to the positive or negative end of the (summed) dipoles, as well as the orientation and distance of the dipoles from the electrode (Jackson & Bolger, 2014). The polarity of the measured signal does not reflect the current within the neuron, as a positive or negative EEG signal depends upon whether the inhibitory or excitatory postsynaptic potential is closer or further from the scalp (Jackson & Bolger, 2014). Thus, the challenge for localising the source on the N170 is to try to reconstruct and distinguish between the multiple possible sources producing a recorded EEG signal.

Those who have attempted to conduct source analysis on the N170 suggest the involvement of the superior temporal sulcus (STS) region (Itier & Taylor, 2004). This result is consistent with the conclusions of combined EEG-fMRI (Nguyen & Cunnington, 2014) and EEG-MEG (Burra, Baker, & George, 2017) studies, which benefit from combining the high temporal resolution of EEG with the spatial resolution afforded by other brain imaging techniques. Whilst the STS is thought to be involved in the processing of 'changeable' aspects of faces, such as eye gaze direction and expression (Haxby, Hoffman, & Gobbini, 2000), a wider face-perception network has also been identified. This diffused face-perception network is thought to include the Fusiform Face Area (FFA) (Deffke et al., 2007), which may be involved in processing 'invariant' aspects of faces (such as facial structure and identity) (Haxby et al., 2000), and the Occipital Face Area (OFA) (Pitcher, Walsh, & Duchaine, 2011). Overall, the N170 may be best understood as originating from a combination of areas (rather than a single area) within the face-processing network (Dalrymple et al., 2011).

The N170 in healthy ageing

The timing and amplitude of the N170 may be affected by the process of healthy ageing. We will limit our discussion of ageing to 'healthy older adults', who for the purposes of this thesis can be assumed to refer to individuals over the age of 60 years old who demonstrate typical levels of cognitive abilities as measured by the Mini Mental State Examination (Folstein, Folstein, & McHugh, 1975), Montreal Cognitive Assessment (Nasreddine et al., 2005) or similar. However, as a note of caution, whilst these participants may pass these cognitive screening tests, white matter loss has been documented in the absence of neurological and neurodegenerative disease, with an estimated 37 % loss in posterior white

matter volume between the ages of 46 and 92 years (Piguet et al., 2009). This suggests that even in apparently very healthy individuals age-related cellular changes are already occurring. Similarly, pathological changes, for example in amyloid deposits and neurofibrillary changes (Braak & Braak, 1991), as well as potentially functional brain changes in cerebral blood flow (Beason-Held et al., 2013) occur before the beginnings of behaviourally evident cognitive decline. Therefore, a 'healthy' older adult sample cannot be guaranteed on the basis of behavioural screening tests alone.

Compared to younger adults, the N170 in older adults is typically delayed and larger (Daniel & Bentin, 2012; Gazzaley et al., 2008; Nakamura et al., 2001; Rousselet et al., 2009; Wiese, Schweinberger, & Hansen, 2008). The N170 in healthy older adults may also be less lateralised. Whilst in younger adults the N170 is predominantly stronger in the right than left hemisphere, older adults may demonstrate a reduction in interhemispheric asymmetry due to increased activity in the left hemisphere (Daniel & Bentin, 2012).

Delays in the timing of the N170 peak in older compared to younger adults have been interpreted as reflecting age-related differences in face processing speed with age. However, comparing N170 peaks across age groups should be treated with caution, as the N170 may not reflect the same neuronal process in older as younger adults if, for example, visual processing has slowed to such an extent that later time windows in older adults become functionally equivalent to the N170 in younger adults (Rousselet et al., 2009).

Older adults however also experience decline on a range of non-face tasks. For example in house/letter/face versus texture tasks, older adults were consistently slower than younger adults, as measured using behavioural reaction times, as well has having a delayed and enhanced N170 peak for all object categories compared to younger adults (Jaworska, 2017). Thus it is important to consider whether age-related differences in the N170 are face specific or general across all object processing tasks.

Age-related differences in visual processing speed and recruitment of brain areas

Differences in the N170 in healthy ageing may reflect a wider pattern of age-related changes in visual processing in general, such as reduced processing speed, dedifferentiation (i.e. reduced neural specialisation) and compensatory mechanisms in brain activity. As such, we will now discuss the phenomenon of age-related decline of efficiency in visual perception more generally.

Most notable is the argument for a general decline in processing speed with ageing (Salthouse, 1996). Salthouse (1996) argues that a decrease in the processing speed of simple processing operations results in increased impairment for more complex operations, the so called 'complexity effect'. A slowing of neuronal processes may begin as early as the 20s (Salthouse, 2010). Recent attempts to quantify the reduction in processing speed using brain imaging methods are consistent with the idea that ageing begins during this decade, with evidence of an increasingly delayed noise sensitivity in older adults of approximately 1 ms per year from age 20 onwards (Rousselet et al., 2010). The authors also suggested that there is a qualitative 'shift' from a 'younger' to and 'older' pattern of brain activity at 47±2 years, suggesting that ageing occurs throughout the adult lifespan. Whilst ageing may also affect fine motor skills, age-related differences in processing speed based upon pen and paper tasks endure even when controlling for age related differences in motor dexterity (Ebaid, Crewther, MacCalman, Brown, & Crewther, 2017).

A decline in processing speed with ageing may reflect either a general age-related slowing as suggested by Salthouse (1996), or be a consequence of dedifferentiation or compensatory neural recruitment (Grady et al., 1994) which we shall now discuss.

If older adults' brains are trying to compensate for deficits by recruiting additional brain areas not seen in younger adults and/or increase in recruitment of brain areas seen in younger adults, then this compensatory recruitment of brain areas may be an adaptive mechanism by which older adults try to maintain task performance. In face detection tasks, compensation may be achieved by altering the balance of brain areas recruited (Burianová, Lee, Grady, &

Moscovitch, 2013). For example, older adults may recruit the Prefrontal Cortex (PFC) more than younger adults to compensate for reduced activity in the visual processing regions. This 'posterior-anterior shift with ageing' (PASA) model (Davis, Dennis, Daselaar, Fleck, & Cabeza, 2008) is suggested to reflect a compensatory mechanism (however, an opposite pattern of activity was identified using a longitudinal rather than cross-sectional design (Nyberg et al., 2010)). Additionally, a 'hemispheric asymmetry reduction in older adults' (HAROLD) model (Cabeza, 2002), has suggested that increased bilateral recruitment by older adults is also compensatory (Cabeza, Anderson, Locantore, & McIntosh, 2002). However, there is the suggestion that compensatory activity may only apply when there are low levels of cognitive load, as according to the CRUNCH model (compensation-related utilisation of neural circuity) additional recruitment of neuronal resources by older adults is no longer effective at higher levels of cognitive load (Reuter-Lorenz & Campbell, 2008). Whilst differences in the brain areas recruited for a particular task may not always equate to increased task performance for older adults, teasing apart whether performance would be *more* impaired without this increased activity, and whether this increased activity is really reflecting a compensatory mechanism is difficult (Grady, 2012).

An alternative view is to ascribe recruitment of additional brain areas by older adults as a reflection of a dedifferentiation of neuronal networks. For example, a reduction in neural specificity to faces compared to other categories (Burianová et al., 2013) would suggest that there is a lack of specificity in the neural processing of older adults. Again, teasing apart whether the recruitment of bilateral brain areas reflects a compensatory or non-beneficial mechanism is difficult when only comparing cross-sectional data.

Changes in visual processing with age may also reflect the consequences of changes in brain structure due to the decline in grey and white matter in the brain (Giorgio et al., 2010). Age-related changes in grey and white matter have been directly implicated in the slowing of processing speed (Eckert, Keren, Roberts, Calhoun, & Harris, 2010). White matter integrity has been related to the speed of performance in older adults, so could explain declines in visual processing speed (Kerchner et al., 2012). However, yet another approach is to

suggest that age-related declines in vision are due to functional rather than structural changes (Andersen, 2012). For example, functional differences in neurons and neuronal communication in the visual system, such as reduced levels of gamma aminobutyric acid (GABA) or acetylcholine (ACh) may play a role (Hasselmo & Sarter, 2011; Pitchaimuthu et al., 2017).

Older adults may also have a reduced ability to ignore task irrelevant information in the early stages of visual processing that may interact with processing speed (Gazzaley et al., 2008) due top-down modulation deficits (Kalkstein, Checksfield, Bollinger, & Gazzaley, 2011). Greater processing of task-irrelevant information could contribute to the decline in processing speed of task-relevant information in older compared to younger adults.

Age-related macular degeneration and visual processing

Ageing is also marked by degeneration of the eyes and visual function, including macular degeneration, glaucoma and cataracts (Owsley, 2011). Hence, one explanation for changes in visual processing may be due to deterioration of optical functions, such as visual acuity (Gittings & Fozard, 1986) and contrast sensitivity (Owsley, Sekuler, & Siemsen, 1983). For example, age-related differences in visual acuity and contrast sensitivity could result in delayed processing speeds. However, evidence suggests that visual acuity and contrast sensitivity alone cannot account for age-related difficulties in face recognition (Boutet et al., 2015), and other optical factors such as senile miosis and individual differences in pupil size cannot account for ageing differences and inter-subject variability in processing speed (Bieniek, Frei, & Rousselet, 2013), suggesting that differences in visual processing with ageing cannot be explained by low-level factors alone.

Information Processing during the N170

We have seen that the N170 is a face-sensitive component of the ERP time course that is evident in both younger and older adults, though is delayed with ageing – but what is the information *content* driving this N170 response in younger and older adults? Henceforth we refer to the key information content,

i.e. the facial features that an individual uses to resolve a face processing task as the *diagnostic information* required for the task. To understand the diagnostic information being processed during the time window of the N170 in younger and older adults, we first need to review the image sampling techniques available for answering this question.

Using full images of face stimuli is problematic when seeking to understand what information is being processed during the N170 as it is unclear what particular facial features are driving neuronal responses – any single feature or combination of features may potentially be diagnostic. Instead, we need to isolate single facial features to understand the relative contribution that their visibility has on modulating the N170. Whilst face images can be cropped to reveal particular features in isolation (such as cropping a face into rectangles containing only the eyes or only the mouth or nose – for an example see Bentin et al., (1996) experiment 4) isolating single facial features from the context of a face and comparing the pattern of the N170 is problematic – removal of the face context may change the shape of the N170 compared to full images of faces (Daniel & Bentin, 2012).

An alternative to using full or cropped face images is to use reverse correlation techniques. Multiple methods are available, for example one method used in psychophysics would be to present pure noise-based stimuli (i.e. white or correlated noise) with instructions to detect a face or object within the noise. Such designs have been used to understand 'internal representations' of face stimuli (Smith, Gosselin, & Schyns, 2012). Another method would be filtering face stimuli with noise (i.e. embedding a face 'base image' overlaid with noise) to introduce random variance to the stimulus on each trial and averaging stimuli resulting in a particular behavioural judgement. For example such an approach has been used to understand the internal representations of trustworthy, happy and angry faces and the diagnostic facial features for these representations (éthier-Majcher, Joubert, & Gosselin, 2013).

Using noise-based stimuli, the diagnostic regions for identifying the illusory presence of a face in noise-only stimuli has been used to indicate that face-perception in the absence of a face relies on an eye/nose/mouth pattern (Rieth, Lee, Lui, Tian, & Huber, 2011). However, these approaches are limited

when trying to elucidate how particular features may modulate brain responses, as the visibility of single features are still not isolated.

An alternative approach is to sample full face images through a number of apertures, so as to reveal a random subset of facial features on each trial. The aim of such techniques is to establish the relative importance or contribution of particular facial features to e.g. behavioural and/or brain responses, whilst maintaining the face context (unlike when cropping facial features as described above). This technique was first introduced by Haig (1985) in which an image could be segmented into a square checkerboard, with a select number of squares revealed on each trial. This technique was then expanded on in a variant known as the Bubbles technique (Gosselin & Schyns, 2001). This principled quantitative technique involves sampling an image through Gaussian apertures called 'bubbles'. These bubbles are placed in random locations trial by trial, approximating a uniform sampling of all image regions across trials. In the case of images of faces, using this technique allows different face regions and their combinations to be sampled on each trial and can be used to resolve what information content from a visual display is diagnostic for e.g. correct assignment of an image in a categorisation task.

The N170 is sensitive to the contralateral eye in younger and older adults

Previous EEG research has suggested that the N170 may be modulated by the eye region. However, results have been inconsistent. For example, Bentin et al. (1996) suggested that the N170 elicited by isolated images of eyes is larger than the N170 elicited by images of whole faces. On the other hand, Eimer (1998) suggested no difference in N170 amplitude between faces with and without eyes & eyebrows, though found that when the eyes & eyebrows were absent from the face the N170 was delayed. Taylor, Itier, Allison, & Edmonds (2001) meanwhile found a delayed N170 latency for eyes viewed in isolation compared to full face images. It is difficult to draw conclusions about the role of the eyes in modulating the N170 when comparing whole face images to cropped images of eyes where the face context has been removed. Whilst Eimer (1998) made an attempt to preserve the face context in eyeless & eyebrowless faces by only removing these features from the face, their stimuli resulted in a salient blank

expanse over the region where the eye and eyebrow was removed. In an attempt to remove eye information whilst maintaining the face context, Jaworska (2017) replaced the eye regions with varying amounts of phase noise. Whilst not directly comparing 'no-eyes' with full faces images, she suggested that weaker eye visibility was associated with small amplitude and delayed latency of the N170 compared to trials where eye visibility was greater.

In a study using the aforementioned bubble technique combined with EEG recording, Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin (2003) suggested that the N170 amplitude increased with increased presence of the eye region in a gender and expressive versus non-expressive categorisation task. More recent evidence has also suggested that the N170 is eye-sensitive, as the increased presence of the contralateral eye leads to larger and earlier N170s in a face versus noise detection task in younger adults (Rousselet, Ince, Rijsbergen, & Schyns, 2014). The authors suggest that the presence of the contralateral eye was encoded during the rising part of the N170, with maximum sensitivity before the N170 peak. Increased visibility of the contralateral eye also decreased N170 latencies, resulting in increased reaction times. In contrast, the latter part of the N170 is sensitive to the transfer of facial features across hemispheres, such as the encoding of the other eye from the other hemisphere (Ince et al., 2016). Sensitivity to other facial features, such as the mouth, are suggested to follow subsequently (M. L. Smith, Gosselin, & Schyns, 2007).

Rousselet et al. (2014) suggest that N170 sensitivity to contralateral eye visibility reflects a bottom-up, data driven model of face processing. However, top-down interpretation of a particular configuration of a stimulus as representing the eyes may also be an important factor for the N170. The N170 to an 'isolated eyes' schematic (two dots with no face context) elicited a very similar N170 to schematic faces, but only after participants were primed to interpret the stimuli as eyes (Bentin & Golland, 2002). This may suggest that whether a particular shape or feature is inferred as representing the contralateral eye may be important in modulating the N170 component. This has implications for whether the presence of a face is first inferred by detecting the contralateral eye, or whether inferring the presence of a face then leads to processing of the contralateral eye.

Secondly, the results of Rousselet et al. (2014) could reflect a generic, non-feature specific orientating towards the left due to a 'left gaze bias'. In an illusory face detection task during which only noise stimuli were presented, Rieth et al. (2011) observed a upper-left bias for face detection. This suggests that observers may orientate to this location in an automatic fashion, rather than this orientating being driven by the presence of the left eye. We outline this argument in more detail in Chapter 2.

Recently, Jaworska (2017) expanded the results of Rousselet et al. (2014) to older adults, showing that older adults processed the same facial features (i.e. the eyes) in a face detection task, but that the processing of the eyes is weaker and delayed in older adults. When comparing single-trial fluctuation of the N170 between younger and older adults, it was shown that increased visibility of the contralateral eye increased the amplitude on the N170 in younger and older adults. There was also a reduction in the lateralisation of the N170 in older adults, with eye visibility modulating the N170 in the right hemisphere in young but not in older participants.

Aims of the thesis

This thesis will address the following questions:

- Can the N170 contralateral eye sensitivity be understood as arising from a general non-feature specific orientating to the left?

- Is the N170 contralateral eye sensitivity tolerant to changes in stimulus size, stimulus set and task demands?

- Does encoding of the presence of the contralateral eye precede sensitivity to another facial feature (the mouth) in a male versus female (GENDER) and expressive versus non-expressive (EXNEX) categorisation tasks?

- Do healthy older adults process the same facial features as younger adults and how does the strength and timing of feature sensitivity compare in older compared to younger adults?

In Chapter 2 we present results from a face versus noise detection task where we parametrically manipulated the size of face stimuli to demonstrate that the N170 contralateral eye sensitivity results previously reported (e.g. Rousselet et al. 2014) cannot be explained as arising from a general non-feature specific orientating to the left due to a left gaze bias. In addition, we will demonstrate that the N170 contralateral eye sensitivity is tolerant to changes in stimulus size for all but very small face images.

In Chapter 3, we will outline a novel approach, *Bubble-Warp* (Gill et al., 2015), to utilising Gaussian bubbles for faces images in which facial features have not been spatially aligned. Using a new stimulus set and *Bubble-Warp* in Chapter 4 we will demonstrate that visibility of the contralateral eye *and* mouth modulates single-trial N170 responses in both a GENDER and EXNEX categorisation task in younger adult participants. We will also demonstrate that the relative timing of onset of sensitivity towards the eyes and mouth is idiosyncratic, with processing of the contralateral eye preceding processing of the mouth in only 60 – 70 % of participants depending on task.

In Chapter 5 we quantify age related differences in the timing and strength of feature sensitivity in the GENDER and EXNEX categorisation tasks. We will demonstrate that older adults rely on the same facial features as younger adults, though with increased dependence on visibility of the mouth for resolving the EXNEX task. We will demonstrate that eye processing is 20 -23 ms slower in older compared to younger adults, whilst mouth processing is not delayed in the EXNEX task and only delayed by 9 ms in the GENDER task. We will show that our results suggest that feature encoding is not uniformly delayed in ageing and has consequences for theories suggesting a general decline in processing speed in ageing.

Chapter 2 - N170 sensitivity to contralateral eye area is scale tolerant

Introduction

The N170 ERP is an integrated measure of cortical activity that varies in amplitude and latency across trials and is primarily associated with the extraction of visual information about faces (and other objects). A larger N170 amplitude has been observed for faces compared to other objects (see Chapter 1: *Introduction*) – an effect which cannot be explained by differences in low level visual properties, such as differences in amplitude spectra between object categories (Rousselet, Husk, Bennett, & Sekuler, 2008) (despite claims by Thierry, Martin, Downing, & Pegna (2007) that the N170 is sensitive to 'inter-stimulus perceptual variability' which they claim is greater for faces than other object categories - see Bruno Rossion & Jacques (2008) for a critique).

Quantification of the coding function of the N170 during face processing has shown that the presence of pixels around the eye contralateral to the recording electrode modulates single-trial variability in N170 latency and amplitude at lateral-occipital electrodes (Rousselet, Ince, van Rijsbergen, & Schyns, 2014). This result is consistent with previous studies using a different face set and tasks (M. L. Smith, Gosselin, & Schyns, 2004). N170 sensitivity to the contralateral eye has been suggested to indicate that the first processing step indexed by the N170 is the coding of the contralateral eye area, followed by cross-hemispheric transfer of the ipsilateral eye (Ince et al., 2016). Contralateral eye sensitivity is particularly strong for right hemisphere lateral-occipital electrodes, where visibility of the contralateral left eye modulates N170 latency variability, with increased visibility of the left eye resulting in an earlier peak of the N170 (Ince et al., 2016). Together, this would suggest that the early part of the N170 predominantly reflects a feature-specific encoding mechanism i.e. sensitivity to the contralateral eye.

However, alternative explanations for apparent sensitivity to the left contralateral eye are possible. Rather than this result indicating a feature-specific encoding of the left eye, it is possible that sensitivity to this area of the image may rather reflect a generic, non-feature specific orientating towards the

upper left quadrant of the stimulus space in relation to a fixation cross. Under this assumption, previous results could be interpreted as indicating a general bias in the allocation of visual attention to an area relative to an experimental fixation cross (e.g. the upper left quadrant) and not as reflecting any specific feature-related encoding mechanism.

This alternative explanation has roots in research suggesting that human viewers routinely demonstrate a left gaze bias (LGB) i.e. a higher probability of first gaze directed at the left hemiface (from the viewers perspective) and/or longer looking times towards the left hemiface. LGB occurs when viewing faces, regardless of task demand, gender, familiarity or facial expression (Guo, Smith, Powell, & Nicholls, 2012) and does not appear to be related to handedness or eye dominance (Leonards & Scott-Samuel, 2005), though may be related to habitual scanning directions used in language (Heath, Rouhana, & Ghanem, 2005). LGB is present in 6 month old infants (Guo, Meints, Hall, Hall, & Mills, 2009), and is not limited to human faces, with human viewers exhibiting a LGB towards monkey, dog and cat faces as well (Guo, Tunnicliffe, & Roebuck, 2010). Whilst a LGB is stronger for upright faces than inverted faces (Leonards & Scott-Samuel, 2005), there is evidence of a tendency to inspect the left side of inverted faces first (Guo et al., 2009). One potential limitation of LGB research is the tendency to present stimuli centrally, with trials beginning with fixation on the centre of the face. However, LGB has been demonstrated even when faces are presented parafoveally (Hui-Wen Hsiao & Cottrell, 2008), though LGB was eliminated when faces were presented to the left of the screen (Luke & Pollux, 2016).

One explanation for LGB during face viewing relies on the assumption of a biologically based face asymmetry, in which the left hemiface (from the viewers' perspective) could contain more diagnostic information than the right. For example, there may be hemifacial asymmetries in the expression of emotion, with greater physical movement of the left than right hemiface (Nicholls, Ellis, Clement, & Yoshino, 2004) that affects perceptions of trustworthiness (Okubo, Ishikawa, & Kobayashi, 2018) and the recognition and expression of emotions (Nicholls, Ellis, Clement, & Yoshino, 2004). Thus LGB may have developed through experience as a behaviourally adaptive, task-driven bias

due to greater diagnosticity of the left compared to right hemiface for socially relevant tasks. If LGB is task-driven, then changing the cognitive demands of a task should result in differential gaze patterns relevant for the task at hand. Whilst a LGB has been demonstrated across a range of facial expressions tasks (Guo et al., 2012), there are notable changes in whether the LGB is orientated to the upper or lower part of the face with changing task demands. For example, there is a gaze bias for the upper face during identity tasks and a bias for the lower face during expression tasks (Malcolm, Lanyon, Fugard, & Barton, 2008) suggesting that LGB may vary with task demand.

Alternatively, LGB may be a result of cortical lateralisation of 'face-sensitive' brain regions i.e. the typically larger activation of the right than left hemisphere Fusiform Face Area (Kanwisher, McDermott, & Chun, 1997). LGB may arise because the left hemiface is projected onto the 'face-sensitive' right hemisphere and so is automatically attended to more readily than the right hemiface (as opposed to the left hemiface being attended to because it provides more diagnostic information than the right hemiface). Whilst lateralisation of face processing may develop in later childhood and may be linked to reading acquisition and the development of left lateralisation for written words (Ventura, 2014), there is evidence that a LGB is evident in young children (Guo et al., 2009; Racca, Guo, Meints, & Mills, 2012). Unpicking whether LGB is consequence of cause of right lateralised face processing areas is yet to be elucidated.

If left eye sensitivity in face detection tasks is due to a generic, automatic orientating to the left side of faces, then we would expect this bias to apply to other categories of objects as well, potentially including textures. However, when comparing faces and textures in a face discrimination task, there was no apparent sensitivity to the equivalent location of the left eye on texture trials (Rousselet et al., 2014). This has led to the suggestion that, at least in a face detection task, left eye sensitivity cannot be accounted for by a general, non-feature specific bias in the allocation of attention to eye location. However, faces are highly symmetric compared to textures. This symmetry could provoke a lateralised attentional bias that is specific for symmetric face images and not evident in non-symmetrical landscapes or patterns (Leonards & Scott-Samuel,

2005) and potentially suggests that LGB may not be provoked by texture images, making textures a poor experimental control for understanding if apparent contralateral eye sensitivity in face detection tasks is feature specific. An alternative approach would be to compare face processing to other more symmetrical object categories such as houses or cars, which also elicit a LGB (Levine, Banich, & Koch-Weser, 1984). However, images of front-view houses and cars may be perceived to be 'face-like' due to configural cues, such as windows in a house or the headlights of a car being perceived as 'eye-like', making it difficult to tease apart feature specific left eye sensitivity from a general non-feature specific orientation to the left.

An alternative approach to comparing face processing to textures or other objects would be to use images of faces at a range of different sizes. This preserves the face symmetry, whilst also introducing spatial uncertainty by interleaving face images of different sizes, so that participants cannot use a single face template or allocate attention to a very specific area on the screen. If participants have a general left gaze bias that is not specific to any particular facial feature, participants should attend to a similar area of the stimulus space across trials, regardless of the feature present (e.g. if participants always attend slightly to the upper left of the fixation cross, this area could display the eye for 1 face size, but the nose or forehead for faces of other sizes). If however participants attend to the left eye due to eye coding being the first step in facial processing, then participants should 'track' the eye and direct attention to this feature, regardless of its relative position to a centrally presented fixation cross. If tracking of the eye occurs across varying face sizes, this would suggest that left eye sensitivity is feature specific, rather than due to a general orientating to the left due to LGB.

Does changing the size of facial stimuli affect the N170 differentially? This question has not been formally investigated in depth in the psychophysics literature. One attempt at addressing this question has focused on the 'vertex positive potential' (VPP) (Jeffreys, 1989). The VPP is a positive peak in the ERP time course to face stimuli that occurs around 150-200 ms post stimulus onset on midline central and parietal electrodes. The VPP reflects the same brain process as the N170, with both components accounted for the same dipole configuration

(Joyce & Rossion, 2005). Early work using line drawings of faces has considered the effect of stimulus size on the VPP, and has suggested that changes to the size of face stimuli produced potentials of similar amplitude, for all but very small images where the face was less distinct (Jeffreys, 1989; Jeffreys, Tukmachi, & Rockley, 1992).

However, more recent research using naturalistic photographs of faces has suggested that changing face size may modulate the amplitude of the N170, with a larger N170 observed for large compared to small high resolution face images, but reversed for low resolution images (i.e. larger N170 for smaller than larger faces) (Mercure, Dick, Halit, Kaufman, & Johnson, 2008). This may be due to changes in the spatial frequency content of the images that occurs when modifying stimulus sizes - reducing the size of a stimulus leads to an increase the cycle/degree of spatial frequencies due to compression, whilst simultaneously reducing cycles/image causing some high spatial frequency details to be lost (Mercure et al., 2008). High resolution large images of faces, and low resolution small images of faces in the Mercure et al. (2008) study, shared the same spatial frequency as measured by cycles/image which may explain the pattern of their results.

In the current experiment, we set out to resolve whether left eye sensitivity can be understood by a generic spatial attention to the left, or whether left eye sensitivity is feature specific. Participants completed a face versus noise texture discrimination task. Images were presented at 4 different sizes (see Figure 1), with trial by trial randomisation of image size and category. For face trials, this resulted in variation of the distance of the left eye from the central fixation cross, increasing spatial uncertainty. As a consequence, participants could not use a single specific face-template or allocate attention to a very specific area on the screen. If the N170 contralateral eye sensitivity reflects a generic non-feature specific orientating towards the left, we would expect a systematic left bias across face trials, not specific to the eye region. For example, the same spatial location would be attended to, regardless of the underlying facial features at that location. If however there is a systematic eye-specific feature processing, we can expect that (regardless of stimulus size), it is predominantly the eye area as a feature that is modulating behaviour and EEG. Our design also

allowed us to secondly investigate the previously untested assumption that the encoding of the contralateral eye is tolerant to changes in size of the stimulus.

We will demonstrate that contralateral eye sensitivity is size tolerant and that as face size increases, sensitivity is specifically modulated by the pupil/iris of the contralateral eye. We will replicate and expand the group results of Rousselet et al. (2014) demonstrating the modulation of the N170 in face but not noise trials, but we will also demonstrate that at the individual level there is large idiosyncrasies in N170 modulation.

Materials and Methods

Participants

The study comprised 6 participants: 4 women, median age = 23 years (min 21, max 26). Participants did not report any eye condition, history of mental illness, or were currently taking psychotropic medications, suffered from any neurological condition, had diabetes, or had suffered a stroke or a serious head injury. Volunteers were also excluded from participation if they had not had their eyes tested within the last 3 years, in order to minimise the chances that volunteers did not know of an underlying eye condition. Participants' contrast sensitivity and visual acuity was assessed in the lab (Table 1). Contrast sensitivity was assessed using the Pelli-Robson chart (Pelli & Robson, 1988). Visual acuity at 40 cm and 63 cm were assessed using the Colenbrander mixed contrast card set (Colenbrander & Fletcher, 2004), and 6 m assessed using the Bailey-Lovie Chart (Bailey & Lovie, 1980). All participants had normal or corrected-to-normal visual acuity and normal contrast sensitivity of 1.95 log units (Mäntyjärvi & Laitinen, 2001). Participants gave informed consent to participate in the study and were compensated £6/hr for their participation. The experiment was approved by the local ethics committee.

	LC 40	HC 40	LC 63	HC 63	LC 600	HC 600	CS
P1	95 *[0.10]*	100 *[0.00]*	98 *[0.04]*	104 *[-0.08]*	97 *[0.06]*	105 *[-0.10]*	1.95
P2	97 *[0.06]*	105 *[-0.10]*	95 *[0.10]*	107 *[-0.14]*	97 *[0.06]*	105 *[-0.10]*	1.95
P3	96 *[0.08]*	105 *[-0.10]*	100 *[0.00]*	103 *[-0.06]*	103 *[-0.06]*	105 *[-0.10]*	1.95
P4	94 *[0.12]*	102 *[-0.04]*	100 *[0.00]*	105 *[-0.10]*	100 *[0.00]*	108 *[-0.16]*	1.95
P5	103 *[-0.06]*	105 *[-0.10]*	99 *[0.02]*	105 *[-0.10]*	105 *[-0.10]*	110 *[-0.20]*	1.95
P6	93 *[0.14]*	104 *[-0.08]*	95 *[0.10]*	105 *[-0.10]*	89 *[0.22]*	96 *[0.08]*	1.95

Table 1: Visual Acuity and Contrast Sensitivity Scores For each of 6 participants independently, visual acuity and Contrast sensitivity (CS) scores are given. Visual acuity scores are reported for low contrast (LC) and high contrast (HC) charts presented at 40 cm, 63 cm, and 6 m viewing distance, expressed as raw visual acuity scores (VAS). The corresponding logMAR scores are presented below in italics, where higher values indicate poorer vision and negative values represent normal vision (logMAR score of 0 corresponds to 20/20 vision). Contrast Sensitivity (CS) scores are expressed in log units.

Stimuli

Stimuli were greyscale images of faces and textures (Gold, Bennett, & Sekuler, 1999). Face stimuli were 10 greyscale images of front-view faces (5 female, 5 male). Images were cropped within a common oval frame to remove external features (hair, ears) and pasted on a uniform grey background (Gold et al., 1999). Textures were created by randomising the phase of the face images (0% phase coherence) (see Rousselet et al. 2014). As a result of phase randomization textures lacked the local edge characteristics of faces, so that all face features were theoretically diagnostic, i.e. any feature was theoretically sufficient to detect faces. This ensured we were not biasing participants to use a particular face feature for accurate face detection. All stimuli had an amplitude spectrum set to the mean amplitude of all faces.

Throughout the experiment we used 4 image sizes (Figure 1, Panel A), corresponding to an entire image size of 82 x 82 pixels (size 1), 164 x 164 pixels (size 2), 247 x 247 pixels (size 3) and 331 x 331 pixels (size 4). This equated to 3, 6, 9, and 12 degrees of visual angle respectively. The face oval of each image size was smaller, equating to 2.1˚x1.5˚ (size 1), 4.3˚ x 3˚ (size 2), 6.4˚ x 4.5˚

(size 3) and 8.6° x 6° (size 4) of visual angle. Our size 3 variant was closest in size to the stimuli used in previous work, in which the face oval was 7.0° x 4.9° of visual angle (Rousselet et al., 2014). Each of the face images were presented at each of the different image sizes.

On each trial, stimulus information was revealed through two-dimensional Gaussian apertures. Gaussian aperture size was kept constant across all stimulus sizes (s.d. = 10) with only the number of apertures varying among stimulus sizes: 1 aperture for the smallest stimulus size, increasing to 4, 9 and 16 apertures respectively as stimulus size increased. The increase in the number of apertures with increasing stimulus size ensured that we approximately matched the average sampled area of the stimulus across all stimulus sizes (Figure 1, Panel B). Gaussian apertures were randomly located across the stimulus space, with the constraint that the centre of each aperture was at a unique position and that the centre of the aperture remained within the common oval frame. In the rest of this chapter we will refer to these masks with Gaussian apertures as ‘bubble masks’, and apertures as ‘bubbles’.

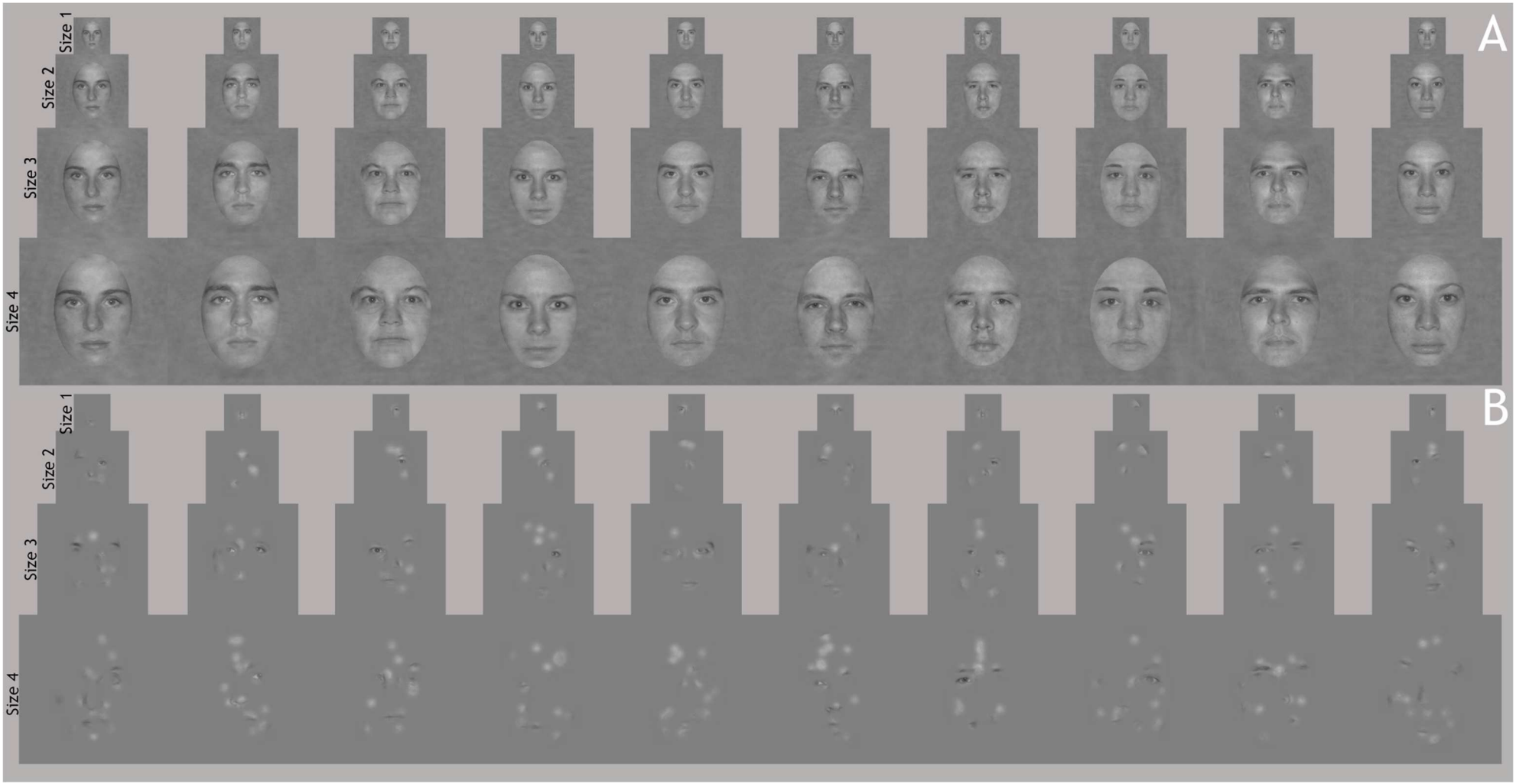


Figure 1: Example face Stimuli. Example stimuli presented without bubbles (Panel A) and with bubbles (Panel B). Faces images consisted of 10 identities (5 female, 5 male). Images were presented at 4 different images sizes (size 1 = smallest, size 4 = largest).

Procedure

During the experiment participants sat in a sound attenuated booth and rested their head on a chin rest. Viewing distance measured from the chin rest to the monitor screen was 80 cm.

Each participant completed 6 experimental sessions, with each session following the same procedure. Stimuli were displayed on a Samsung SynchMaster 1100Mb monitor (600 x 8000 pixels, height and width: 30 x 40cm, 21 x 27° of visual angle, refresh rate: 85 Hz, bits per pixels: 32). Participants were given experimental instructions, including a request to minimise eye blinks and body movements during each block. Participants were asked on each trial to decide whether they had seen a face or noise texture as quickly and as accurately as possible; they pressed '1' for a face stimulus and '2' for a texture stimulus on the numerical pad of a keyboard, using the index and middle fingers of their right hand. At the end of every block they received feedback on their overall performance (median reaction time and mean percentage correct), and, after Block 1, on their performance overall across all blocks completed thus far on that individual experimental session. Median reaction times and mean percent correct remained on the screen until participants pressed a key to move on to the next block.

At the beginning of each session, participants completed 1 practice block of 80 trials with full images (without bubble masks). Participants then performed 14 blocks of 80 trials per block of bubbled images. Each block consisted of 10 face images presented once at each of the 4 stimulus sizes (i.e. 40 face trials) and 10 texture trials of each of the 4 stimulus sizes (i.e. 40 texture trials). Face and noise texture trials and size of stimulus were randomised within each block. Each experimental session consisted of 1200 trials in total, including 80 practice trials and 1120 experimental bubbled trials. In each session participants completed 140 face and 140 texture trials of each stimulus size. Across the 6 experimental sessions participants completed 480 practice trials without bubble masks and 6720 experimental trials with bubble masks, consisting of 840 face trials and 840 texture trials at each of the four stimulus sizes.

Each trial began with a small black fixation cross displayed at the centre of the monitor for a random time interval of about 500-1000 ms. This was followed by either a face or noise texture image presented for ~ 82 ms. A blank grey screen followed the stimulus presentation and remained until the participant's response. Participants were required to make a response before the next trial began. After a response there was an inter-trial stimulus interval of 800 ms before the next trial began. The fixation cross, stimulus, and blank response screen were all displayed on a uniform grey background with mean luminance 43 cd/m^2.

EEG Recording and Pre-Processing

EEG data were recorded at 512 Hz using an active electrode amplifier system (BIOSEMI, Amsterdam, the Netherlands) with 128 electrodes mounted on an elastic cap. Four additional flat electrodes were placed on the outer canthi and below each of the eyes. Electrode offsets were kept between ±20 μV.

EEG data were pre-processed using Matlab 2013b and the open-source toolbox EEGLAB version 13. Data were band-pass filtered between 1 Hz and 40 Hz using a non-causal fourth order Butterworth filter. A second dataset was created by pre-processing data with a fourth order Butterworth filter – high-pass causal filter at 1 Hz and low-pass non-causal filter at 40 Hz. Data was then re-sampled to 500 Hz.

Data from the two datasets were epoched between -300 to 1000 ms around stimulus onset. Baseline correction was performed using the average activity between -300 to 0 ms only for the high-pass causal filter data set. For the non-causal filtered dataset, the channel mean was removed from each channel instead.

Noisy electrodes and trials were then detected by visual inspection of the non-causal dataset and rejected from the two datasets on a participant-by-participant basis. The reduction of blink and eye-movement artefacts was performed using ICA, as implemented in the *infomax* algorithm from EEGLAB. ICA was performed on the non-causal filtered dataset, and then applied to the causal filtered dataset on a participant-by-participant basis, in order to remove

the same components from both datasets. Components representing blinks and eye-movements were identified by visual inspection of their topographies, time-courses and amplitude spectrum. After rejection of artefactual components, data epochs were removed based on an absolute threshold value larger than 100 μV and the presence of a linear trend with an absolute slope larger than 75 μV per epoch and R^2 larger than 0.3.

Finally, we calculated spherical spline current source density (CSD) waveforms using the CSD toolbox. CSD waveforms were computed using parameters 50 iterations, m=4, lambda =10-5.

Electrode Selection

Detailed analyses were performed on a subset of electrodes. We pre-specified three clusters of posterior electrodes at the left, midline and right hemisphere (Figure 2). Our central posterior electrode cluster (CE) comprised 11 electrodes from Cz down the vertical midline, including CPz, Pz, POZ and Oz. Our left poster electrode cluster (LE) comprised 19 electrodes including P07 and its neighbouring electrodes. Our right posterior electrode cluster (RE) comprised 19 electrodes, including PO8 and its neighbouring electrodes.

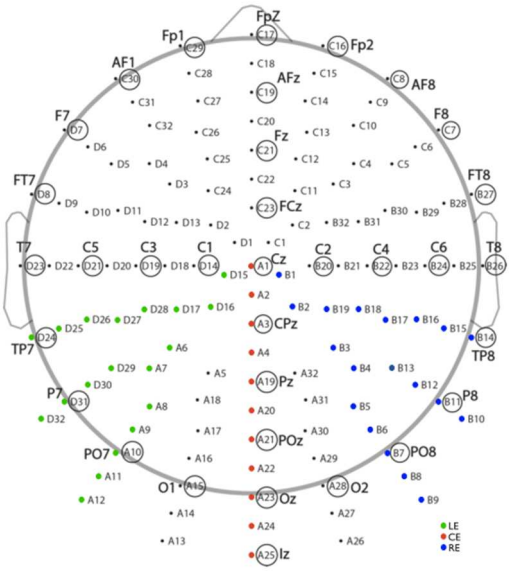


Figure 2: Electrode Selection for analysis consisted of three clusters of electrodes. A posterior right electrode cluster (RE) consisted of 19 electrodes including PO8 (shown in blue). A posterior left electrode cluster (LE) consisted of 19 electrodes including PO7 (shown in green). A midline electrode cluster consisted of 7 electrodes including Pz and Oz (shown in red). Image reproduced from Jaworska (2017).

Statistical Analyses

Statistical analyses were conducted using Matlab 2013b.

Mutual Information

We used Mutual Information (Thomas & Joy, 2006) to quantify the dependence between stimulus features and behavioural responses, and stimulus features and brain responses. Mutual Information (MI) is a non-parametric measure that quantifies (in bits) the difference in entropies and reduction in uncertainty about one variable after observation of another (i.e. sampled pixels and behaviour; sampled pixels and brain signal responses). One of the benefits of using MI is that it is a direct measure of effect size on a common meaningful scale, and as such direct comparisons across neural responses can be made. MI can be used to study the selectivity of neural and behavioural responses to external stimuli in single trials and is sensitive to non-linear effects (Schultz, Ince, & Panzeri, 2015). Several tools for computing MI are available through an open source toolboxes, including the 'gcmi' (Gaussian copula mutual information) toolbox (Ince et al., 2017) used for analysis described in this thesis (https://github.com/robince).

We calculated several MI quantities in single participants. For each MI calculation, we combined data from each of the 6 experimental sessions. For face and texture trials independently, we calculated for each of four stimulus sizes:

- MI(PIX,**RT**) to establish the relationship between image pixels and reaction times for faces and textures at each image size. We copula normalised RT's using the *copnorm* function. We then computed MI using the *info_gg* function to calculate MI between Gaussian copula normalised RTs and Gaussian bubble mask values. MI was calculated with bias correction for the entropy of Gaussian variables (Ince, Giordano, et al., 2016; Ince, Petersen, Swan, & Panzeri, 2009).

- MI(PIX, **CORRECT**) to establish the relationship between image pixels and correct responses for faces and textures at each image size. We computed

MI using the *info_gd* function to calculate MI between discrete responses (correct versus incorrect) and Gaussian bubble mask values.

- MI(PIX, [**ERP**, **grad**]) to establish the relationship between image pixels and ERPs. We calculated *bivariate* MI, which considers the recorded voltage at each time point together with the temporal gradient of the ERP. Including the temporal gradient results in a smoothing effect, where artificial dips in MI due to the bimodal ERP crossing the zero line are smoothed out (Ince et al., 2016). We copula normalised ERPs using the *copnorm* function and computed MI using the *info_gg* function to calculate MI between Gaussian copular normalised ERP and ERP gradient, and Gaussian bubble mask values. MI analysis was computed between -300 ms pre-stimulus onset to 400 ms post stimulus onset. We calculated MI independently at each time point, pixel and electrode, for each participant and image size using the non-causal dataset.

Mutual Information Classification Images

We refer to MI between pixels and behaviour, or pixels and ERPs as 'classification images'. These images reveal the pixels associated with modulations of the particular response being calculated.

Mutual Information Timecourses

We calculated how MI values between pixels and brain responses were modulated over time. For each of three pre-specified posterior electrode clusters (LE, RE, CE), for each participant, we took the maximum MI value across all pixels and electrodes within each cluster for each time point between -300 ms pre- stimulus onset to 400 ms post stimulus onset.

Feature of Interest Analysis

MI is directionless, in that higher mutual information values can reflect either the presence *or* the absence of a feature in modulating responses. Using a reverse analysis, we quantified by how much changing the amount of information about the presence of specific image features (e.g. the left and right eye) modulated brain responses. To this end, we created 'feature of

interest' masks for the left and right eye using the average image of all face stimuli. We created the masks by taking the classification image of the group averaged MI(PIX, [**ERP, grad**]) to face trials for the largest (i.e. Size 4) faces. Then, for the right and left eye separately, we centred a circle (radius = 15 pixels) on the pixel that showed the maximum MI value for each eye at the group level. The same mask was used for each stimulus size, scaled appropriately (Figure 3).

Figure 3: Feature of interest masks We created feature of interest masks using the average of all face stimuli (background face) for the contralateral left and right eye by centring an ellipse (radius = 15 pixels) on the pixel with the group averaged maximum MI value for each eye independently.

Using the feature of interest masks, we calculated on a trial-by-trial basis the visibility of each eye independently; obtained as a scalar value of the sum of the total pixel visibility within the ellipse of each feature of interest mask (each pixel had a visibility value between 0 – not visible to 1 – completely visible). We then split these visibility values into ten equally populated bins ranging from the lowest (bin 1) to the highest (bin 10) visibility values. We used these feature visibility bins to compare how the degree of feature visibility modulated ERP responses to face and noise trials.

To quantify the effect that visibility of each feature had on brain responses, we selected a single electrode of interest from the LE, CE and RE clusters. We optimised electrode selection for each individual for each image size, by selecting the electrode from each cluster that had the maximum MI value between pixels and ERP and ERP gradient (MI(PIX,[**ERP, grad**])) to face trials within the time window of 80 – 180 ms post-stimulus onset. Selected LE

electrodes were P7/PO7 or their surrounding neighbours. Selected RE electrodes were P8/PO8 or their surrounding neighbours.

To compare how feature visibility modulated average ERP waveforms, for each electrode of interest we sorted single trial ERP's for each image size into 10 equally populated bins based on a scalar value of the sum of the pixel visibility within the ellipse of each feature of interest mask, ranging from the lowest (bin 1) to the highest (bin 10) visibility values. For each bin, we averaged ERPs across all trials.

Results

Behavioural Results

Behavioural results are given here for trials with bubble masks, unless otherwise stated. Practice trials presented without bubble masks revealed the whole face image and were used to familiarise participants with the task. We found high levels of accuracy and speed across all image sizes for all participants, with both accuracy and reaction times being predominantly modulated by visibility of the left eye.

Reaction Times and Percentage Correct

For each image type (face or texture) and each image size (size 1 = smallest, size 4 = largest) we calculated the mean percentage correct and median reaction time for each participant (Table 2; Figure 4).

Participants were generally more accurate and faster in correctly detecting the presence of a face for sizes 2-4 relative to accuracy for size 1, the smallest faces. This may be due to a response bias towards textures that is particularly evident for the smallest faces. A bias towards textures for the smallest faces could reflect participants missing the presentation of the stimulus for example due to the size of the stimulus, or be a consequence of the smallest faces being revealed by a single bubble, where there is only a single area of the image revealed. This may be comparatively more difficult, as if a salient facial feature such as an eye, nose or mouth is not revealed by the bubble, participants may report detecting a texture only, whereas with the larger face sizes multiple

bubbles are used, so several areas of the image are revealed, any one of which may reveal a salient facial feature for accurate face detection.

		Participant	Size 1	Size 2	Size 3	Size 4
Accuracy (%)	Face	P1	71	92	98	96
		P2	81	96	94	96
		P3	66	78	81	77
		P4	83	94	93	94
		P5	79	96	99	98
		P6	88	94	94	97
	Texture	P1	98	96	98	98
		P2	89	95	93	96
		P3	98	99	100	100
		P4	87	91	98	97
		P5	96	99	99	99
		P6	98	95	99	98
Reaction time (ms)	Face	P1	457	409	405	415
		P2	348	303	300	295
		P3	425	422	406	414
		P4	366	339	329	320
		P5	441	397	375	379
		P6	371	335	330	324
	Texture	P1	468	422	456	467
		P2	360	359	340	333
		P3	356	324	312	307
		P4	389	362	344	347
		P5	415	402	398	392
		P6	361	346	350	333

Table 2: Behavioural results Behavioural results (mean accuracy and median reaction times) are given at each image size (size1 = smallest, size 4 = largest) and each participant separately, for face and textures.

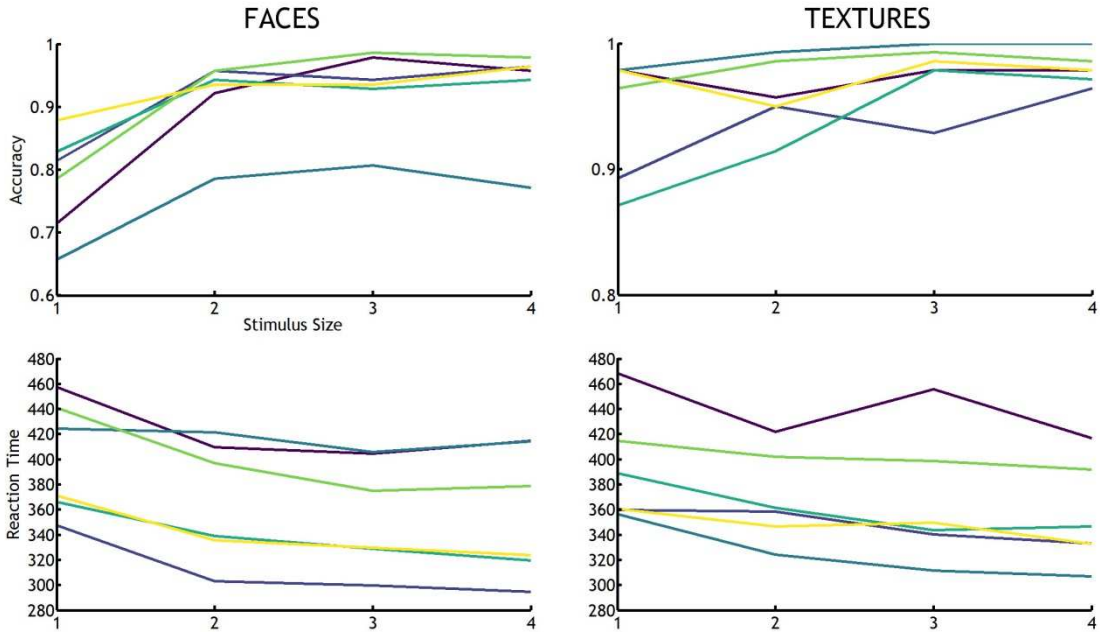


Figure 4: Behavioural Results. Behavioural results for face trials (left) and texture trials (right) for four stimulus sizes. Accuracy (top row) is given in percentage points (0-1). Reaction times (bottom row) are given in ms. Each line represents an individual participant.

Behavioural MI Classification Images

To determine what image features are associated with behavioural responses, we looked at mutual information between pixels and reaction times, MI(PIX, **RT**), and pixels and correct/incorrect responses, MI(PIX, **CORRECT**).

Individual classification images for MI(PIX, **RT**) are presented in Figure 5. In general, visibility of the left eye (and occasionally the right eye) modulated reaction times across all images sizes and participants, for face but not texture trials. In texture trials, there was some sensitivity to the edge of the face oval, which may be due to a visible edge between the face oval texture and background texture being salient.

For size 3 images (closest in size to stimuli used in Rousselet et al. 2014) all participants showed sensitivity to the left eye (from the viewers perspective) on face but not texture trials. This is consistent with the findings of Rousselet et al. (2014).

For larger (size 4) faces all participants showed sensitivity to the left eye in face but not texture trials (though MI is weaker for some participants than for size 3 faces). Additionally, 3/6 participants showed clear sensitivity to the right eye for large faces, which for participant P5 was stronger than MI values to the left eye. Compared to size 3 faces, the eye hotspot shrinks to a more concentrated area around the iris/pupil. This may be due to the relative size of the bubble compared to the size of the image. With a comparatively smaller bubble size compared to image size, it is possible to see more fine-tuned sensitivity to the eye, with a less distributed MI cluster around the eye than in size 3 faces.

For smaller size 2 and size 1 faces, we again see sensitivity to the left eye in face but not texture trials. With reduced stimulus size, the hotspot around the eye is more distributed, with a large spread of MI values around the eye region due to the bubble to image size ratio. This is particularly the case for our smallest faces, where the hotspot covers a large region around the left eye and expanding to the nose, with the hotspot drifting away from the pupil. Higher MI values around the edge of the face oval in texture trials across all image sizes

could be a low level edge detection effect, as there is a visible edge between the face oval texture and background texture.

The relative width and spread of the hotspot demonstrated on each image size may be an effect of the comparative bubble size for each image. We maintained the same bubble size across images, meaning that for large images a single bubble revealed a much smaller portion of the image than a single bubble revealed for the smallest faces. To demonstrate what facial features were revealed by a single bubble for each image size, for each participant and stimulus size we centred a single bubble on the pixel showing the maximum MI value in face trials (Figure 6). As image size increases, the bubble shrinks to be centred upon the pupil/iris area for all participants. For the smallest faces, our MI hotspot drifted towards the centre of the face, rather than being centred on the pupil. This may be a result of the ratio between the size of the bubble and the size of the image, as centring a single bubble on pixels more towards the centre of the face revealed the right eye and nose, in addition to the left eye. Higher overall MI values for the smallest faces may be a consequence of having a single area of the face sampled, rather than more distributed sampling of the face as in the largest face trials, where bubbles may not reveal concurrent areas (see Figure 2).

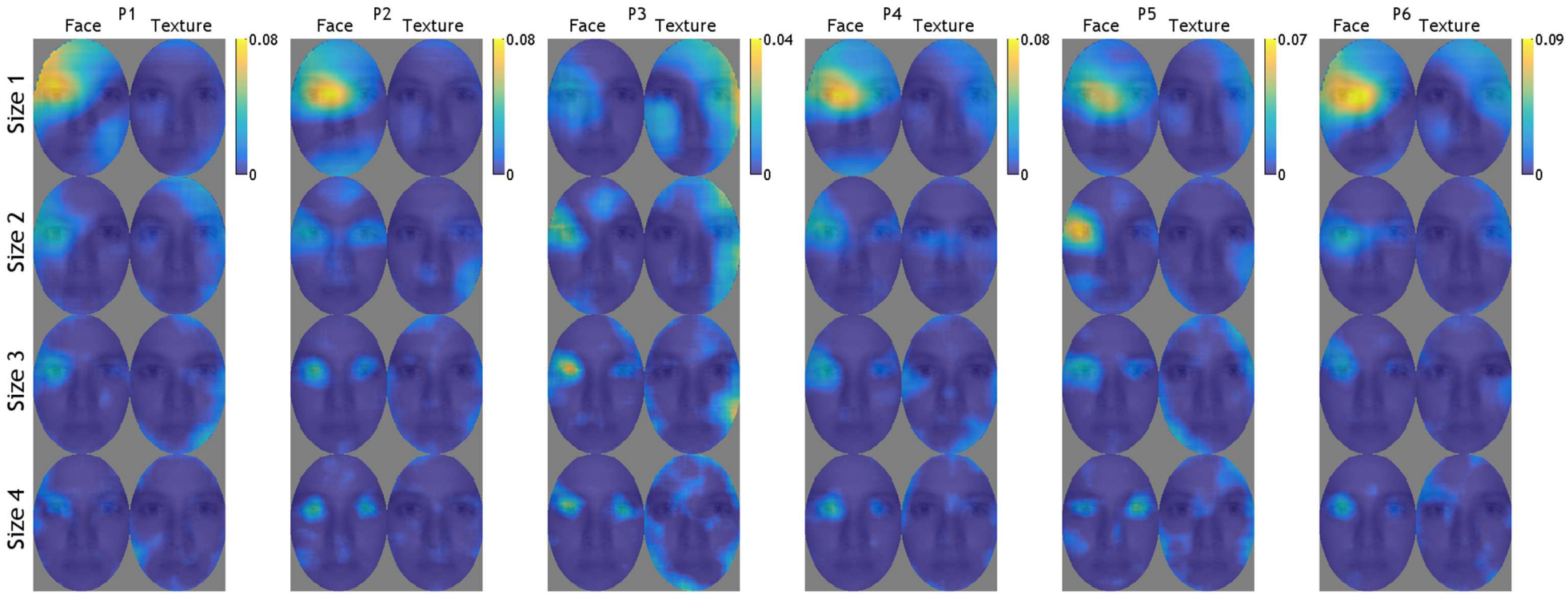


Figure 5: Mutual Information Reaction Time classification images We calculated MI(pix, **RT**) for each of 6 participants for face and noise textures at each image size (size 1 = smallest, size 4 = largest). We scaled classification images independently for each participant, depicted by the colour bar for each participant.

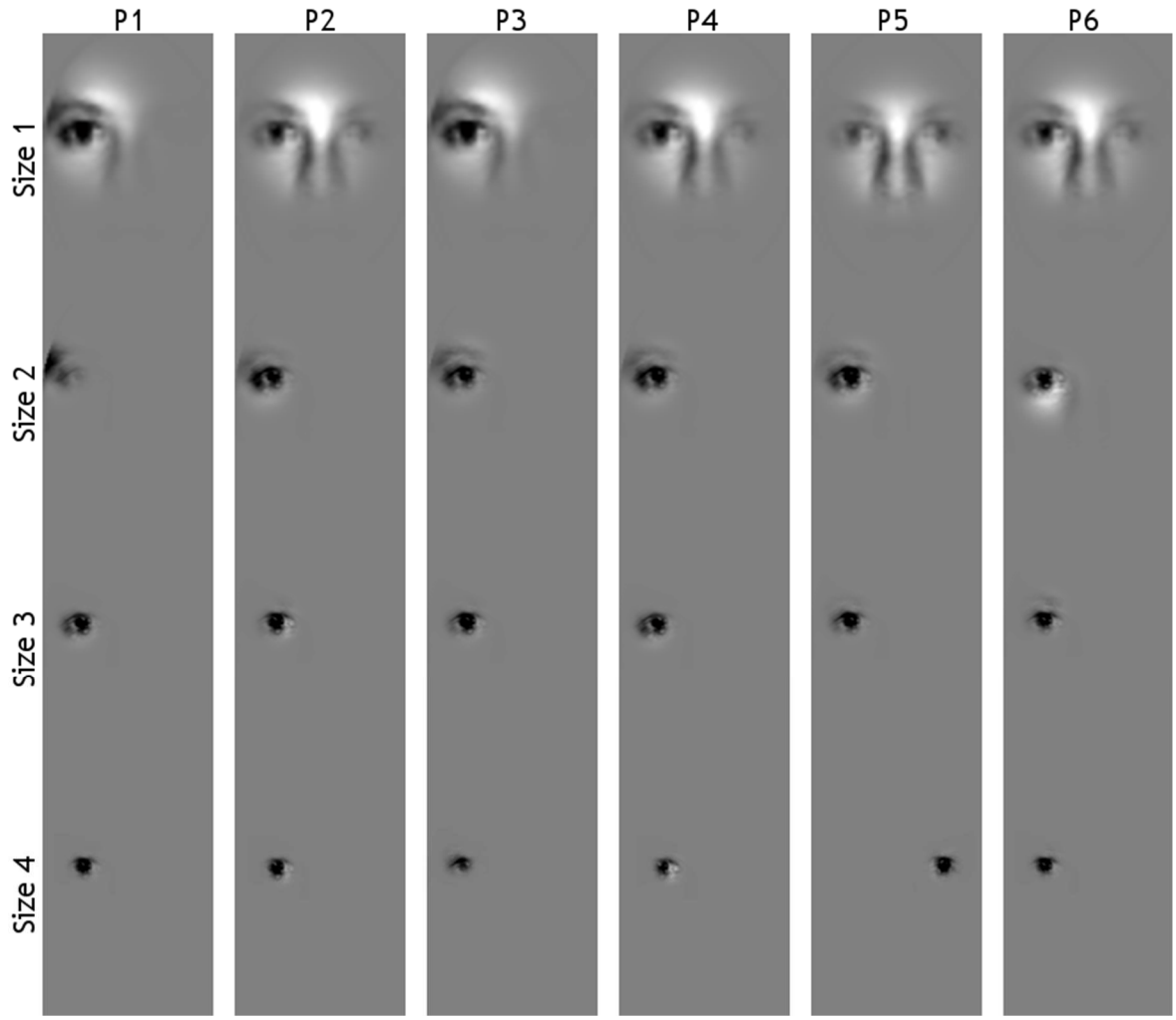


Figure 6: Information associated with speed of response For each of 6 participants and 4 image sizes, we placed a single bubble on the pixel with the highest MI value in face trials as calculated in Figure 5. This reveals the facial features that are revealed by a single bubble, and highlights the increased specificity on the iris/pupil area with increasing stimulus size. For the smallest images where the highest MI pixel was often close to the nose, a hotspot centred on this location still revealed information about one or both eyes for all participants.

For individual classification images for MI(PIX, **CORRECT**) we found weaker sensitivity to the left eye for size 2 – 4 faces (Figure 7). This is similar to the pattern of results observed by Rousselet et al. (2014) who reported sensitivity to the left eye for MI (PIX, **CORRECT**) for only a few participants. In our results, Participant 4 had the clearest example of sensitivity to the left eye, which similar to MI(PIX, **RT**) shrank to the pupil/iris area with increasing image size. For sizes 2 - 4, we observed no sensitivity to any single area in texture trials.

We found the strongest MI values between pixels and correct responses in trials where the smallest (size 1) faces were presented. The location of these hotspots varied among participants, with 4 participants displaying a hot spot towards the left eye/nose area, whereas 1 participant (P2) had the strongest MI for the

forehead and P4 had the strongest MI for an area of the right cheek. A hotspot centred to the left of the nose and below the eye is comparable to the point of first gaze location for face identification identified in some studies (Or, Peterson, & Eckstein, 2015; Peterson & Eckstein, 2013). One explanation for these differences in optimal performance fixation points on the face is individual differences in degradation of contrast sensitivity in the periphery (Peterson & Eckstein, 2013), wherein some individuals may have poorer visibility in upper or lower peripheral vision. For very small (size 1) faces, the ratio between our bubble size and image size may have resulted in bubbles whose centres fell below the left eye to be optimal for processing information about the eye in viewers with better upper than lower peripheral vision for example.

For each participant and stimulus size, we centred a single bubble on the pixel showing the maximum MI value in face trials (Figure 8). As image size increases, the bubble shrinks to be centred upon the pupil/iris area for most, but not all participants. For the smallest faces, our MI hotspot drifted towards the centre of the face, rather than being centred on the pupil. This may be a result of the ratio between the size of the bubble and the size of the image, as centring a single bubble on pixels more towards the centre of the face revealed the right eye and nose, in addition to the left eye. Higher overall MI values for the smallest faces may be a consequence of having a single area of the face sampled, rather than more distributed sampling of the face as in the largest face trials, where bubbles may not reveal concurrent areas (see Figure 8).

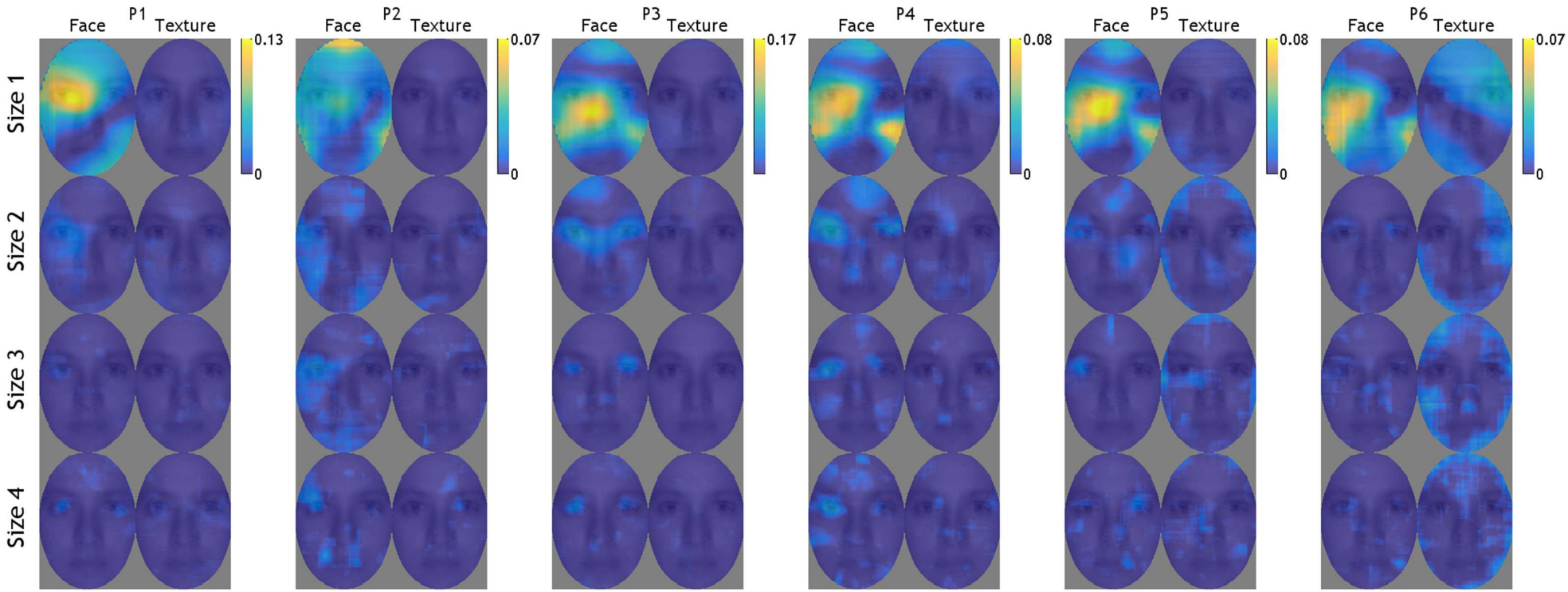


Figure 7: Mutual Information Percentage correct classification images We calculated MI(pix, **CORR**) for each participant (N = 6) for face and noise textures at each image size (size 1 = smallest, size 4 = largest). We scaled classification images independently for each participant, depicted by the colour bar for each participant.

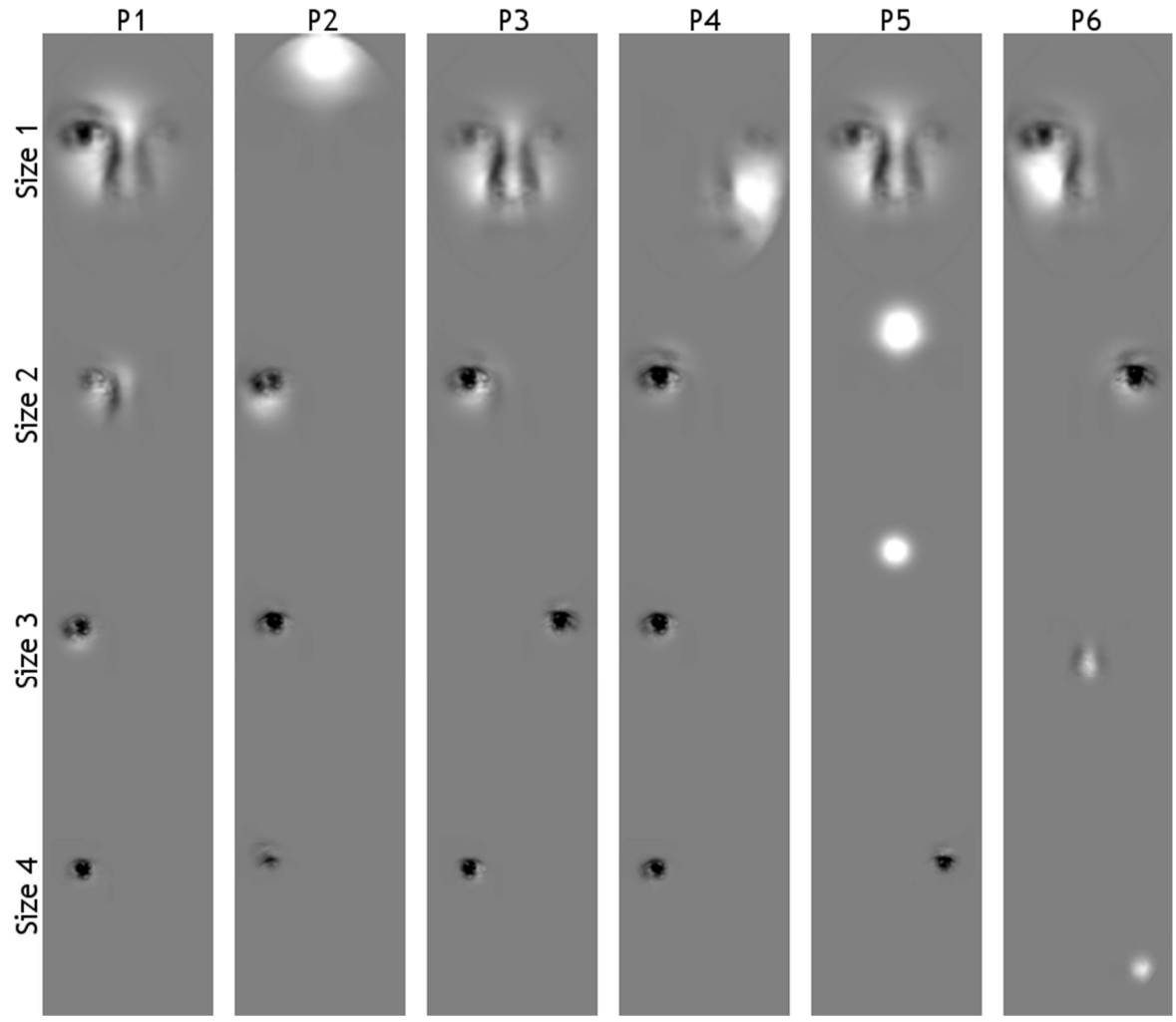


Figure 8: Information associated with correct responses. For each of 6 participants and 4 image sizes, we placed a single bubble on the pixel with the highest MI value in face trials as calculated in Figure 7. This reveals the facial features that are revealed by a single bubble, and highlights the increased specificity on the iris/pupil area with increasing stimulus size (for most participants). For the smallest images where the highest MI pixel was predominantly close to the nose, a hotspot centred on this location still revealed information about one or both eyes for all participants but P2.

In summary, we replicated the results of Rousselet et al. (2014) for size 3 faces (most comparable to the stimulus size in their experiment) finding strong left eye sensitivity in all participants for reaction times for face but not texture trials. Whilst most participants' accuracy was modulated by the left eye region for most image sizes, this association was weaker than for reaction times and more variable across participants and sizes. We found no comparable sensitivity to any single area for texture trials.

With increasing stimulus size, we found a shrinking hotspot of sensitivity between pixels and reaction times revealing sensitivity to the left (and for some participants also the right) iris/pupil area. For small faces the hotspot was larger covering a wide area around the left of the nose and left eye for both reaction

times and accuracy, which may reflect individual differences in optimal viewing patterns i.e. a bubble located centrally on small faces offered visibility of both eyes simultaneously.

ERP Results

We have seen that for all but the smallest images of faces, the left eye region modulated reaction times and correct responses, a pattern not identified in texture trials. As with previous research, this suggests that the left contralateral eye region modulates behavioural responses for face but not texture trials. This supports the suggestion from Rousselet et al. (2014) that behavioural modulation by visibility of the eye region is not due to retinal bias or attention to the eye region, as eye modulation effects were not present in texture trials. In addition, these effects were not specific to one image size, but were apparent for images presented at a variety of sizes, suggesting that this reflects a general processing mechanism used during face detection tasks that is tolerant to changes, at the very least, in face size.

Next, we will consider ERP results. We predicted that we would replicate the results of Rousselet et al. (2014), with maximum MI at left and right lateral occipital electrodes for contralateral eye sensitivity in face but not texture trials, and that this pattern of results would be apparent regardless of image size.

We calculated for every participant for every size separately, the maximum MI at each time point and electrode between pixels and brain responses for every electrode in the right and left hemisphere cluster from –300 ms pre-stimulus onset to 400 ms post stimulus onset. We also computed corresponding classification images displaying the maximum MI across all electrodes for each hemisphere across all time points. We present results for face (Figure 9) and texture (Figure 10) trials separately.

As predicted we found contralateral eye sensitivity for all faces sizes, though this sensitivity was less lateralised and sensitivity was more central for the smallest faces. We found no corresponding sensitivity to the eye region in texture trials in larger textures, though we did find a dispersed sensitivity in the

smallest texture trials. This sensitivity may be explained by increased sensitivity to the *absence* of facial features in texture trials. In the smallest textures the bubble revealed an uninterrupted view of a large portion of the stimulus. In larger textures, identifying the absence of visible face features, particularly the absence of an eye, nose or mouth would require integration of information across several occluders (bubbles) in all but the smallest textures, where the absence of these features would be more readily apparent.

For face trials (Figure 9) maximum MI values were similar for P1-3 between the two hemispheres, and MI was stronger for the right than left hemisphere for P4-6. As with our behavioural results, as image size increased, the MI hotspot shrank to an area consisting of the contralateral pupil/iris area, coupled with a reduction in the strength of MI for most participants. For some participants and face sizes a second later peak is evident, potentially indicating later MI sensitivity to the integration of information across multiple bubble occluders.

For very small (size 1) faces however laterality is reduced, particularly for the left hemisphere where the hotspot is shifted towards the middle of the nose rather than the contralateral (right) eye area. Whilst our results for size 2-4 faces are in line with the results of Rousselet et al. (2014), finding maximum MI at the left and right lateral occipital electrodes for contralateral eye sensitivity in face trials, this result does not hold well for very small faces.

For texture trials (Figure 10) MI values were weaker for all participants and all image sizes. The strongest MI values remained for the smallest (size 1) textures, whilst MI timecourses to the largest size 3 and 4 textures was predominantly flat. Unlike in face trials, there was no hotspot around the pixels corresponding to the iris/pupil for size 2-4 textures – though P1 and P3 had some sensitivity to the area corresponding to the eyebrow for size 2 textures only.

For the smallest textures, some participants demonstrated sensitivity to the area corresponding to the eye (see P1 & P3). This sensitivity was weaker, less eye centred and more spread out, and less lateralised then that seen in faces trials. This pattern of results may be explained by increased sensitivity to the *absence* of facial features in texture trials when the texture images were small and occluded by a single bubble.

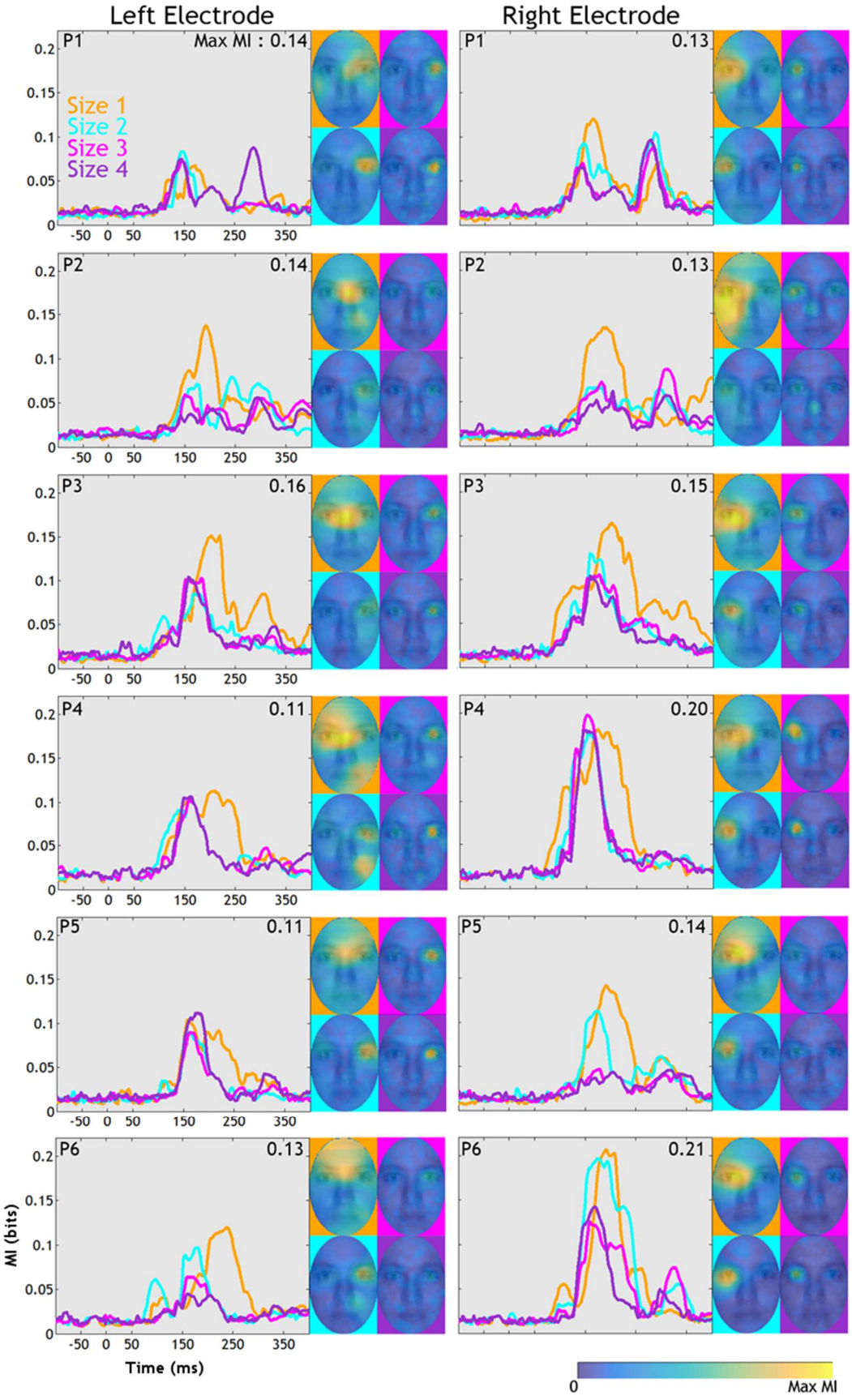


Figure 9: Face MI Timecourse. For each participant (P1-6) and 4 images sizes (different coloured lines), the maximum MI was calculated independently for the left (left column) and right (right column) electrode clusters. Classification images were scaled independently for each participant, and represent the maximum MI at each pixel across all electrodes and time points. Small numbers in the upper right corner of each figure denote the maximum MI value the classification image colour map is scaled to for each participant independently.

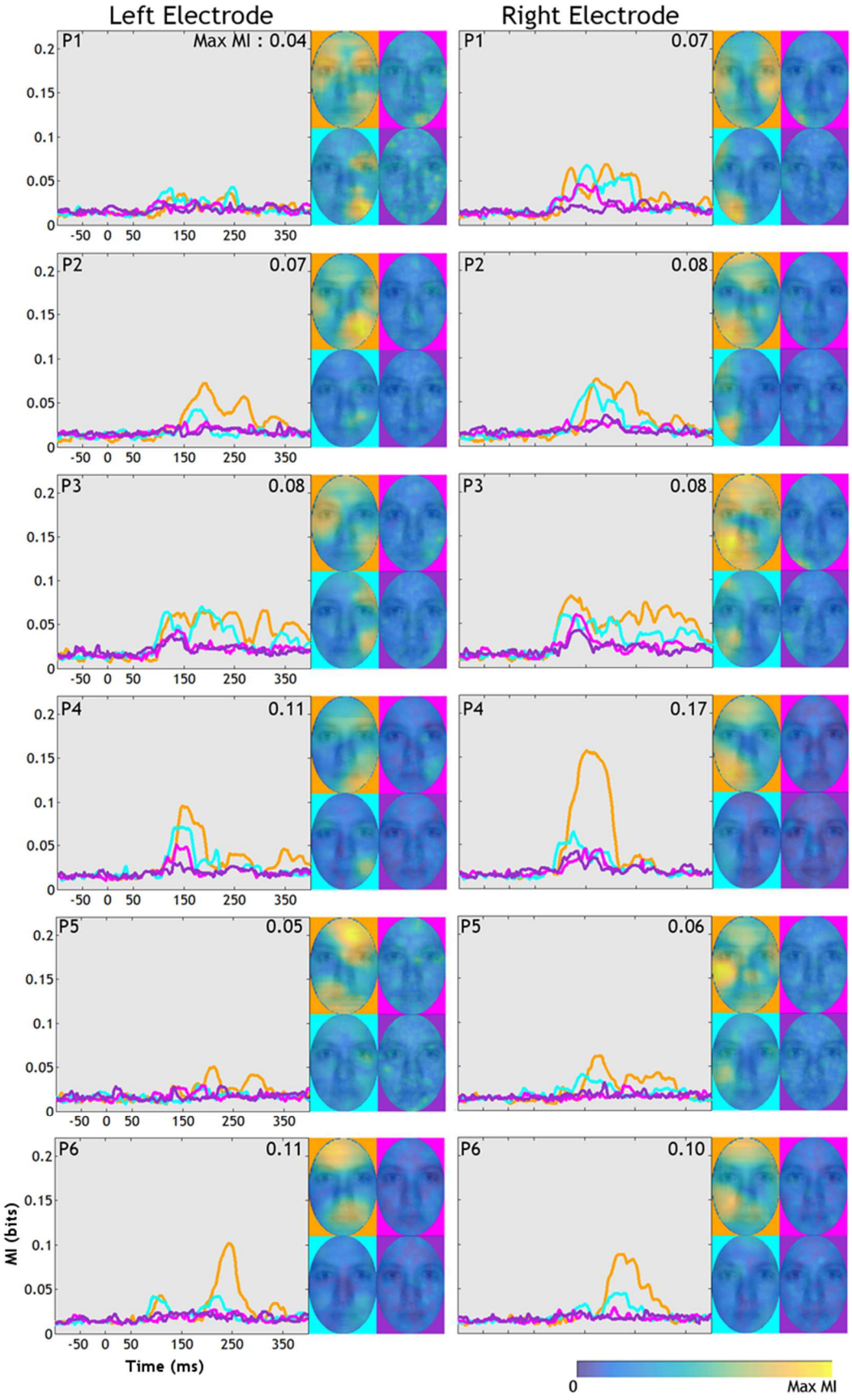


Figure 10: Texture MI Timecourse. For each participant (P1-6) and 4 images sizes (different coloured lines), the maximum MI was calculated independently for the left (left column) and right (right column) electrode clusters. Classification images were scaled independently for each participant, and represent the maximum MI at each pixel across all electrodes and time points. Small numbers in the upper right corner of each figure denote the maximum MI value the classification image colour map is scaled to for each participant independently.

ERPs by Feature Visibility

We have seen that the presence of the left eye modulated reaction times and accuracy for most participants in face but not texture trials at all stimulus sizes. We have also seen that posterior lateral brain signals are modulated by the contralateral eye region, at least in larger faces.

MI does not reveal how the shapes of the ERPs are affected by the presence or absence of the left and right eye. Using a reverse analysis in which the face was split into 16 horizontal bands, Rousselet et al. (2014) demonstrated that increased visibility of the band which included the eyes modulated ERPs for face but not noise trials. For face trials, the N170 at both left and right posterior lateral electrodes was larger and shifted towards the left (i.e. earlier peak latency) as visibility of the eye band increased. We anticipated that we would replicate these results, finding that increased visibility of the eye region would modulate ERPs in face but not texture trials.

To establish how visibility of the left and right eye regions affected the shape of our ERPs in face and texture trials we conducted a reverse analysis isolating the specific EEG modulations associated with the presence of the eye region. Using the feature of interest masks (see *Feature of Interest Analysis*) we calculated independently for each participant and image size, the visibility of each eye region on a trial-by-trial basis, obtained as a scalar value of the sum of pixel visibility within the ellipse of each eye mask. We then split these visibility values into ten equally populated bins ranging from the lowest (bin 1) to the highest (bin 10) values. We then sorted single trial ERPs into 10 bins, based on the eye region visibility in each trial. For face trials, this resulted in ERPs being binned based on visibility of the eye. For texture trials, this resulted in ERPs being binned based on visibility of texture in the areas corresponding to the eye on face trials. We present averaged ERPs within each bin for each image size for each participant independently, for face (Figure 11 - Figure 16) and texture (Figure 17 - Figure 22) trials.

For face trials, increased visibility of the contralateral eye region resulted in larger N170s and a leftward (i.e. earlier peak latency) shift for both the left and right electrode for all participants and image sizes (except P2, size 4, left

electrode where modulation was less clear). ERPs to the smallest (size 1) faces were typically of a smaller amplitude and delayed compared to larger face sizes. A delayed and reduced VPP for smaller faces was also identified by Jeffreys (1989).

In texture trials, increased visibility of the area corresponding to the contralateral eye on face trials also modulated the amplitude and latency of ERPs for some participants (e.g. see Figure 21). Increased texture visibility in the eye area may drive this response due to the presence of structured, high contrast elements in the eye area. An N170 response to noise stimuli has previously been documented in some participants (Rousselet et al., 2009).

Our results from face trials are consistent with those reported by Rousselet et al. (2014) demonstrating that increased visibility of the contralateral eye region modulates single-trial ERPs in face trials for all images sizes. Whilst we found modulation for the smallest faces, this was less lateralised in some participants' and the N170 was generally diminished.

Unlike Rousselet et al. (2014) we did observe modulation of ERPs in texture trials in some individuals. This modulation differed between participants, i.e. in some participants increased visibility of the area corresponding to the eye region increased amplitudes and latencies, whilst in other participants the opposite pattern was identified. However, Rousselet et al. (2014) only present *group-averaged* ERP time courses (calculated across 16 subjects). When calculating mean ERPs for each bin across the 6 subjects in our experiment (see Supplementary 1- Supplementary 2), our results are similar to Rousselet et al. (2014) for image size 3 - the closest in size to the images used in Rousselet et al. (2014) - in that we found no modulation in texture trials with increased visibility of the area corresponding to the eye in face trials. We also find no modulation for the largest (size 4) textures, though we did still find evidence of modulation in size 2 textures (increase in amplitude and earlier N170 with increased visibility). For the smallest textures (size 1) we find a flat early time course with little evidence of an N170, though there may be modulation later in the ERP time course. This suggests there are individual idiosyncrasies in ERP responses to increased visibility of texture information in the areas corresponding to the eye in face trials that are masked by group-average ERP analysis.

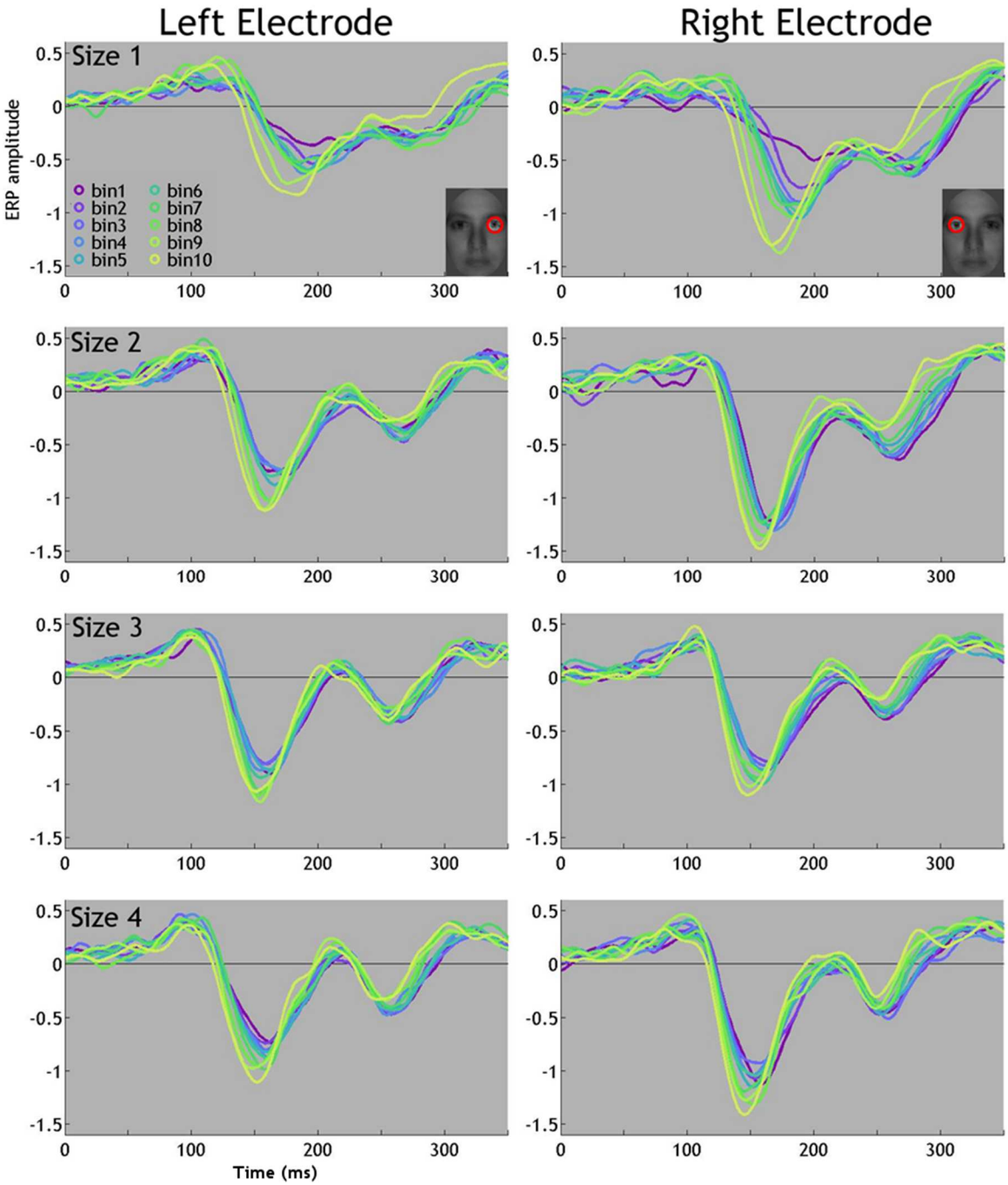


Figure 11: ERP modulation by feature visibility on face trials (P1). Average ERPs during face trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the contralateral eye (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the contralateral eye region, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest faces) to size 4 (largest faces).

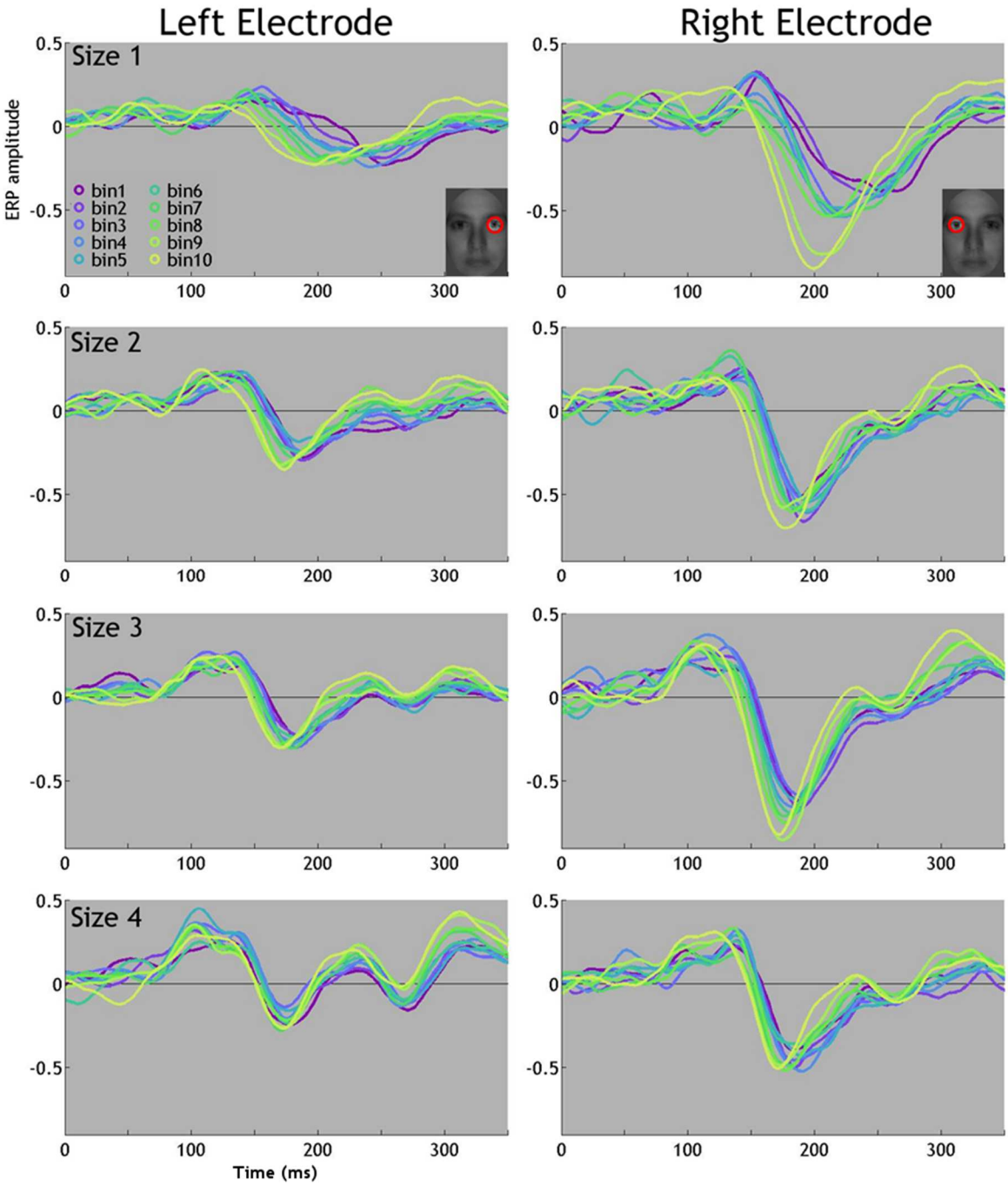


Figure 12: ERP modulation by feature visibility on face trials (P2). Average ERPs during face trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the contralateral eye (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the contralateral eye region, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest faces) to size 4 (largest faces).

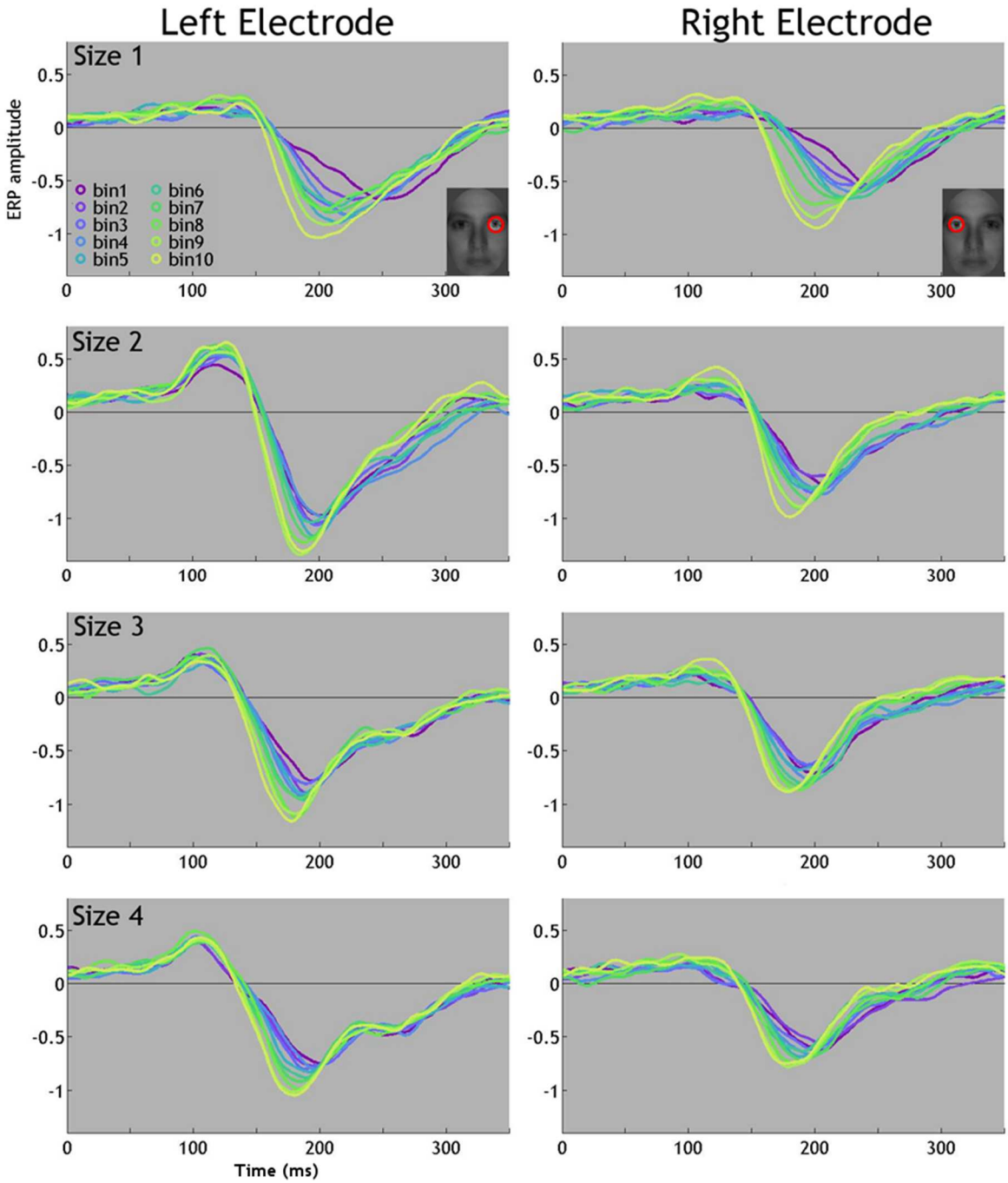


Figure 13: ERP modulation by feature visibility on face trials (P3). Average ERPs during face trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the contralateral eye (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the contralateral eye region, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest faces) to size 4 (largest faces).

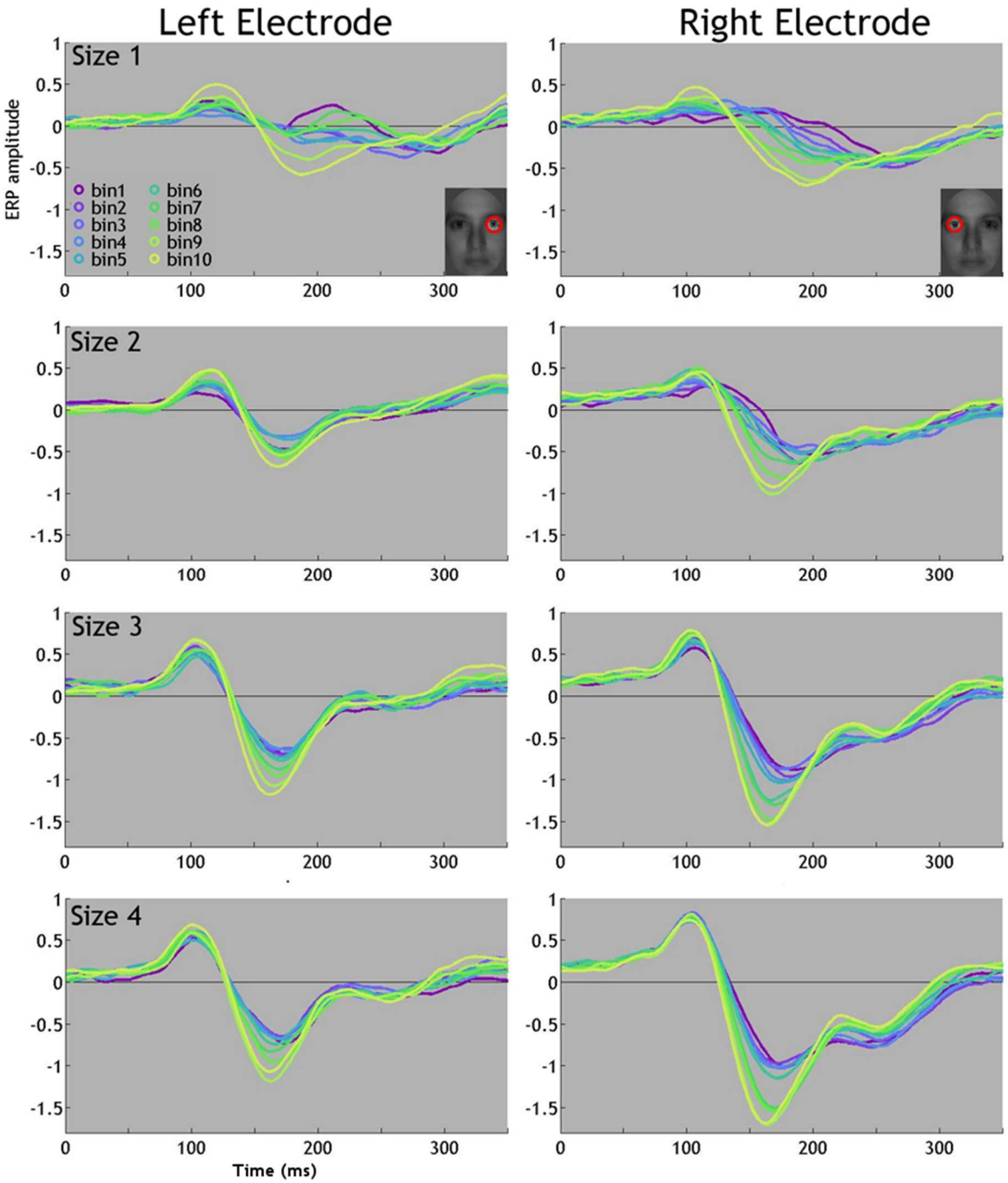


Figure 14: ERP modulation by feature visibility on face trials (P4). Average ERPs during face trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the contralateral eye (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the contralateral eye region, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest faces) to size 4 (largest faces).

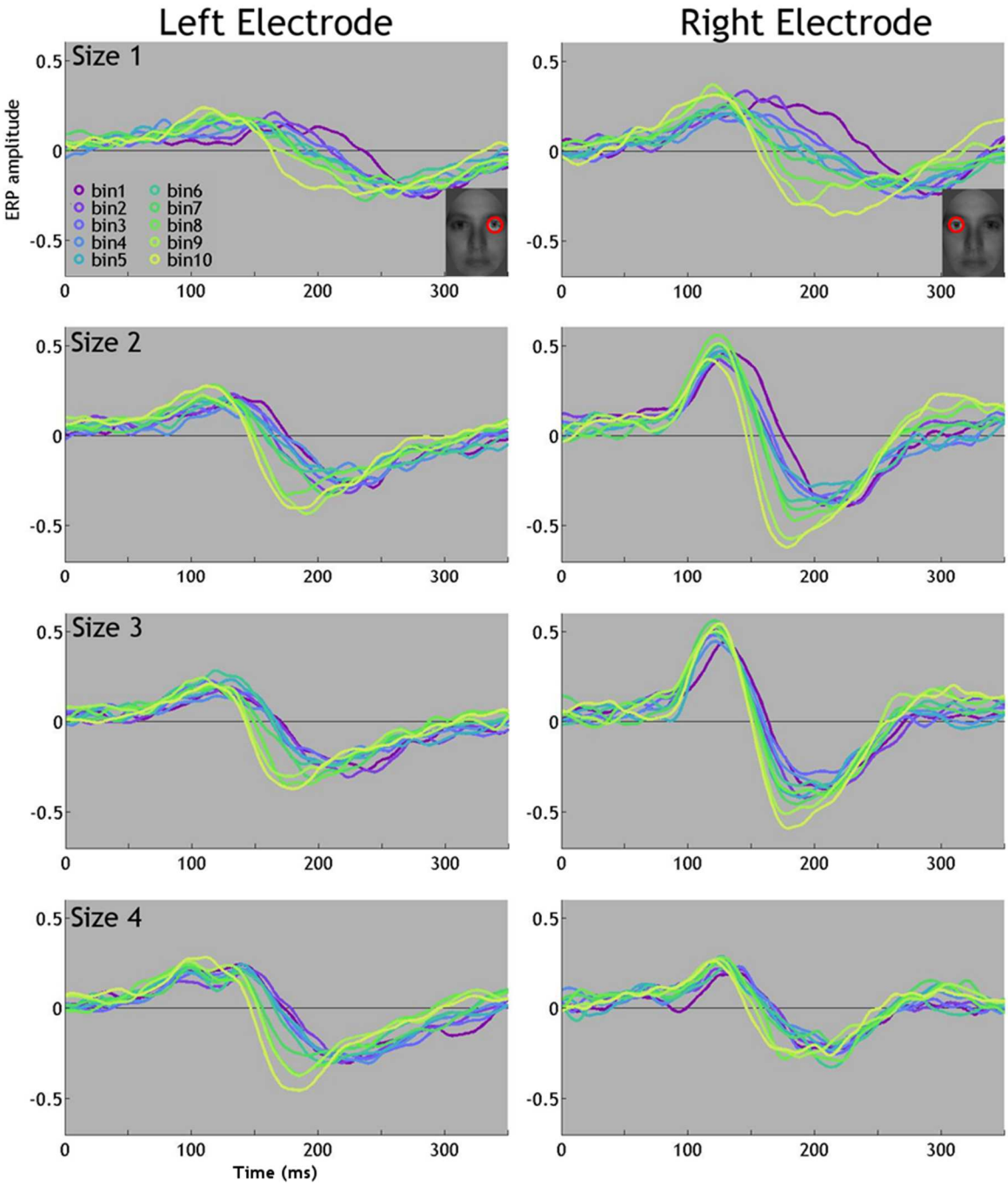


Figure 15: ERP modulation by feature visibility on face trials (P5). Average ERPs during face trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the contralateral eye (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the contralateral eye region, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest faces) to size 4 (largest faces).

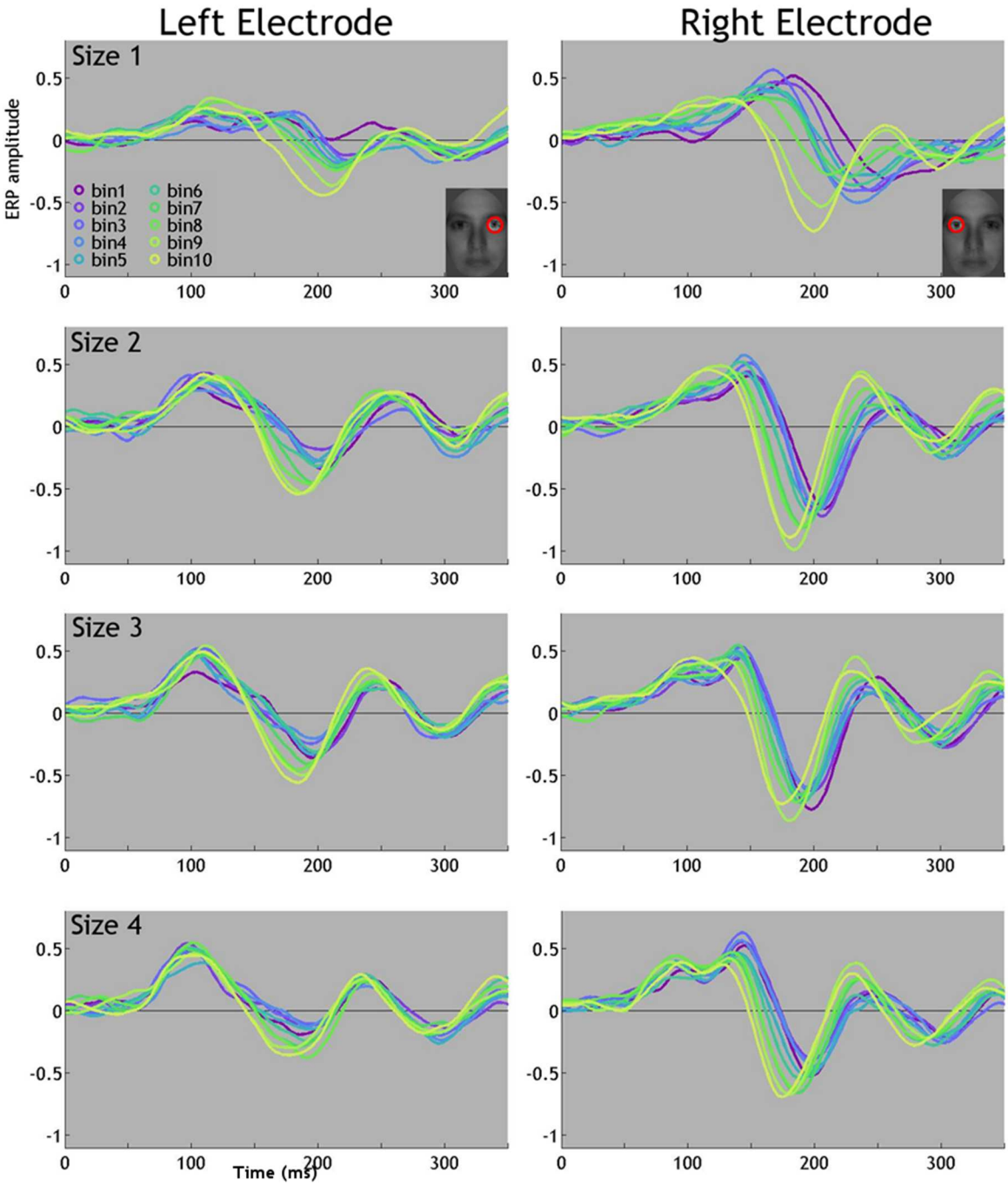


Figure 16: ERP modulation by feature visibility on face trials (P6). Average ERPs during face trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the contralateral eye (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the contralateral eye region, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest faces) to size 4 (largest faces).

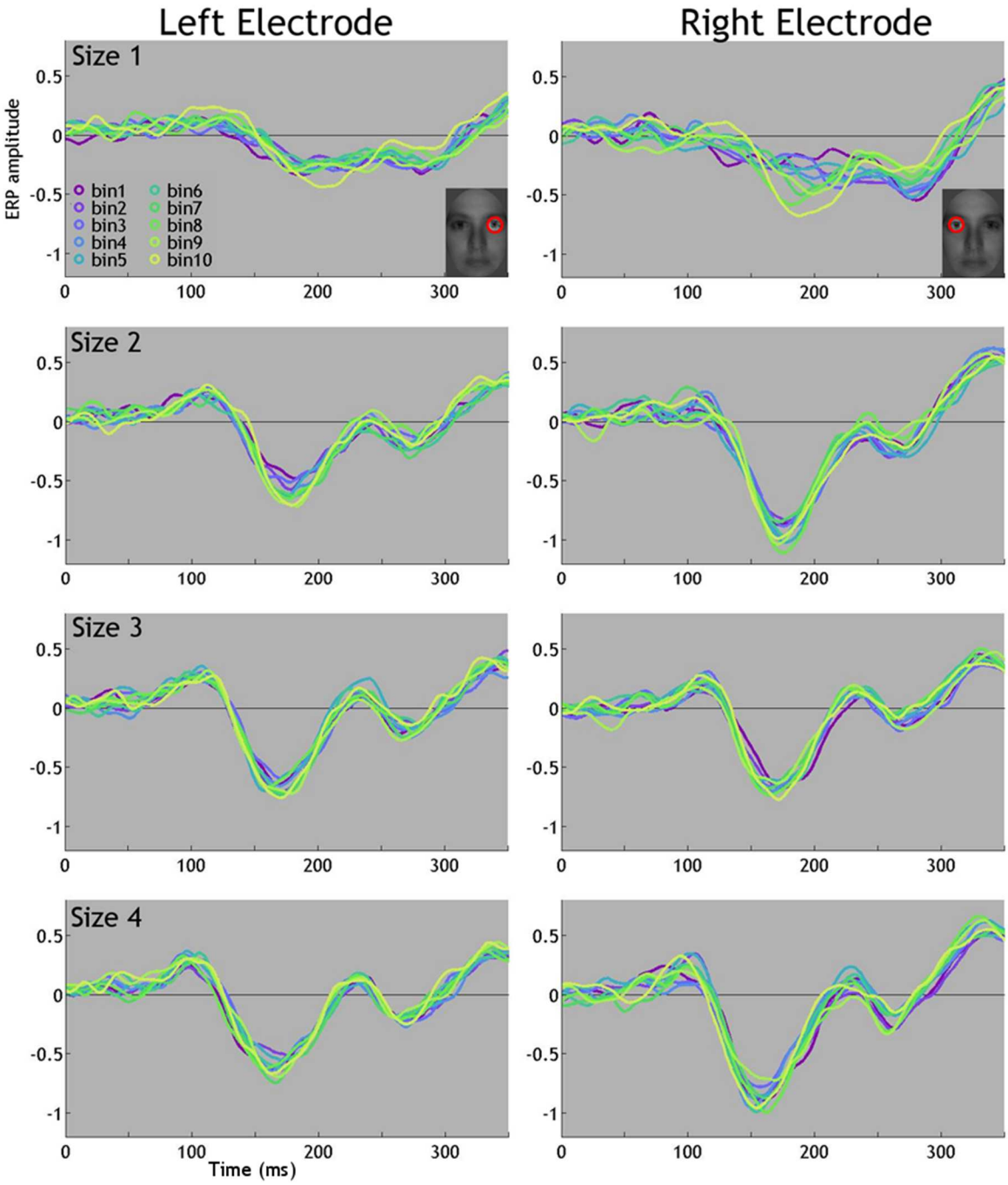


Figure 17: ERP modulation by feature visibility on texture trials (P1). Average ERPs during noise trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the area corresponding to the contralateral eye region on face trials (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the area corresponding to the contralateral eye region on face trials,, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest textures) to size 4 (largest textures).

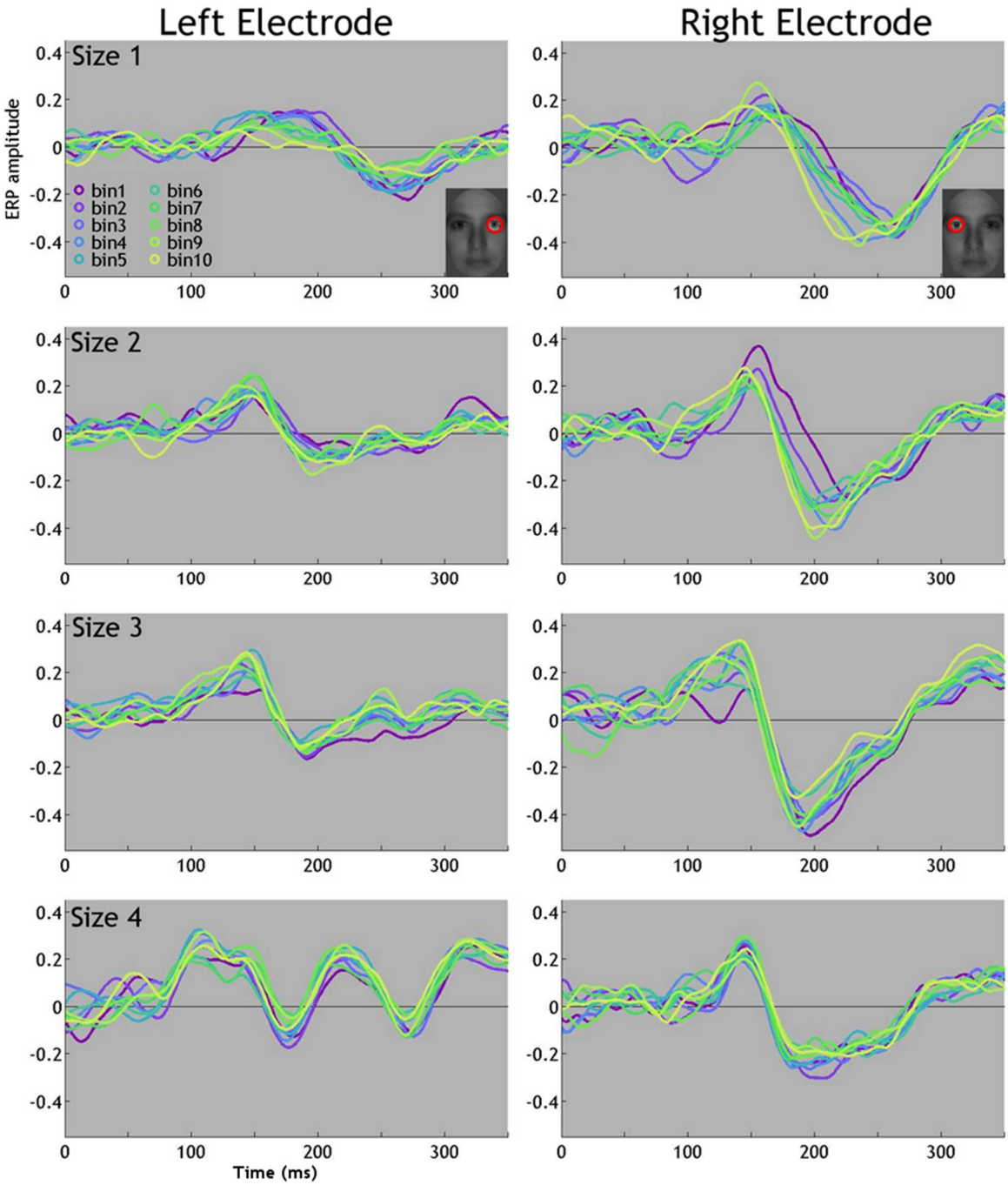


Figure 18: ERP modulation by feature visibility on texture trials (P2). Average ERPs during noise trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the area corresponding to the contralateral eye region on face trials (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the area corresponding to the contralateral eye region on face trials,, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest textures) to size 4 (largest textures).

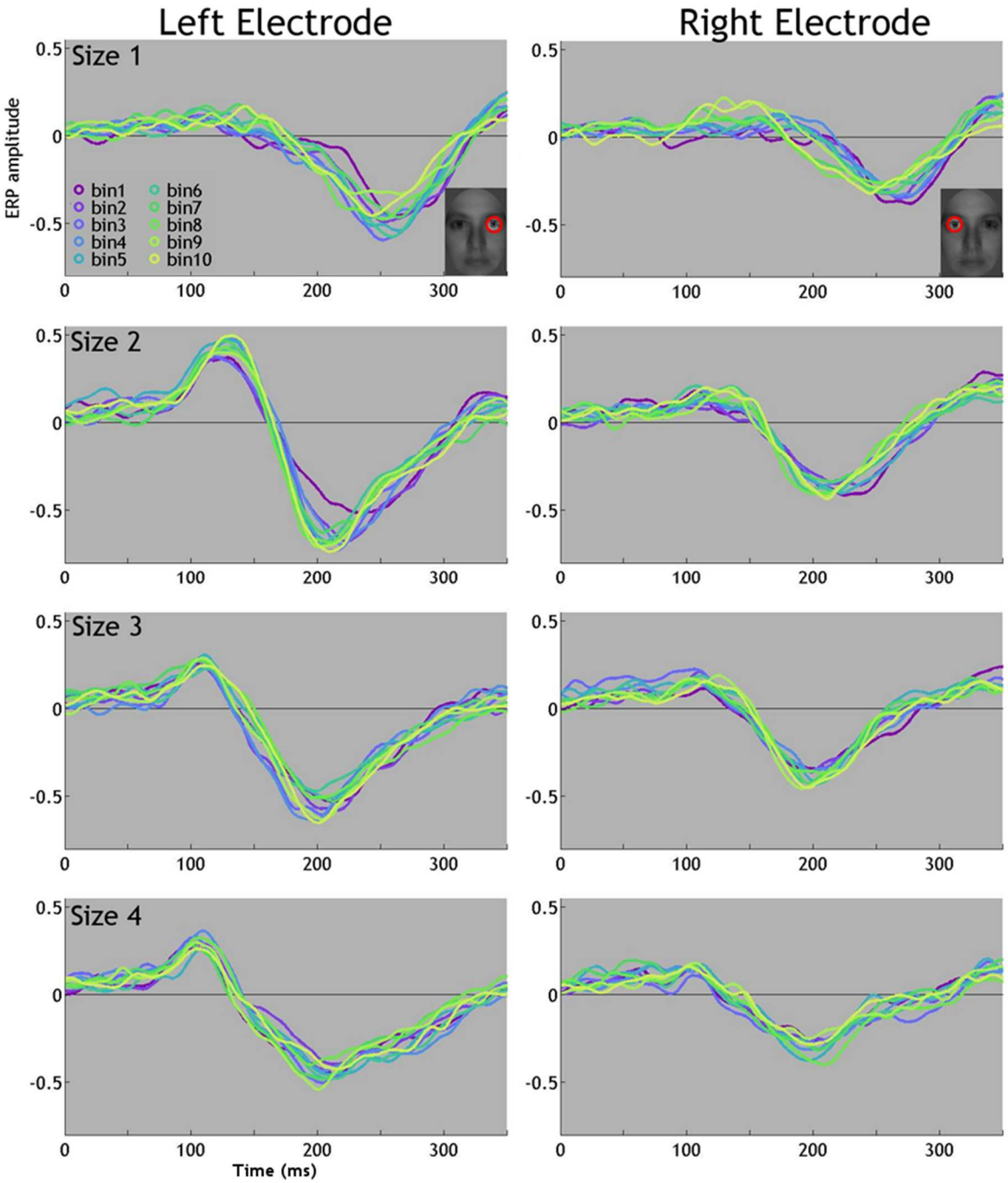


Figure 19: ERP modulation by feature visibility on texture trials (P3). Average ERPs during noise trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the area corresponding to the contralateral eye region on face trials (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the area corresponding to the contralateral eye region on face trials,, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest textures) to size 4 (largest textures).

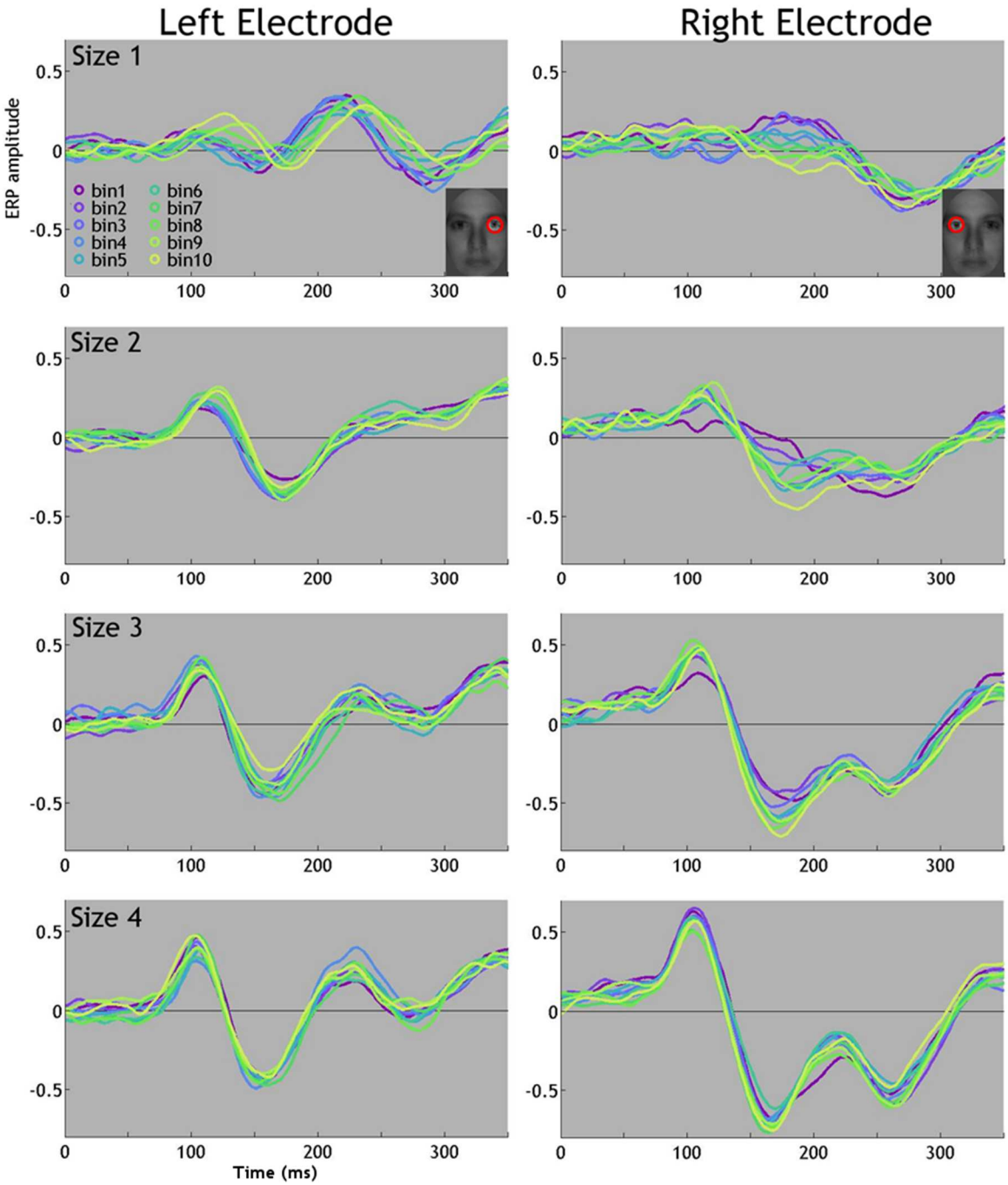


Figure 20: ERP modulation by feature visibility on texture trials (P4). Average ERPs during noise trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the area corresponding to the contralateral eye region on face trials (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the area corresponding to the contralateral eye region on face trials,, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest textures) to size 4 (largest textures).

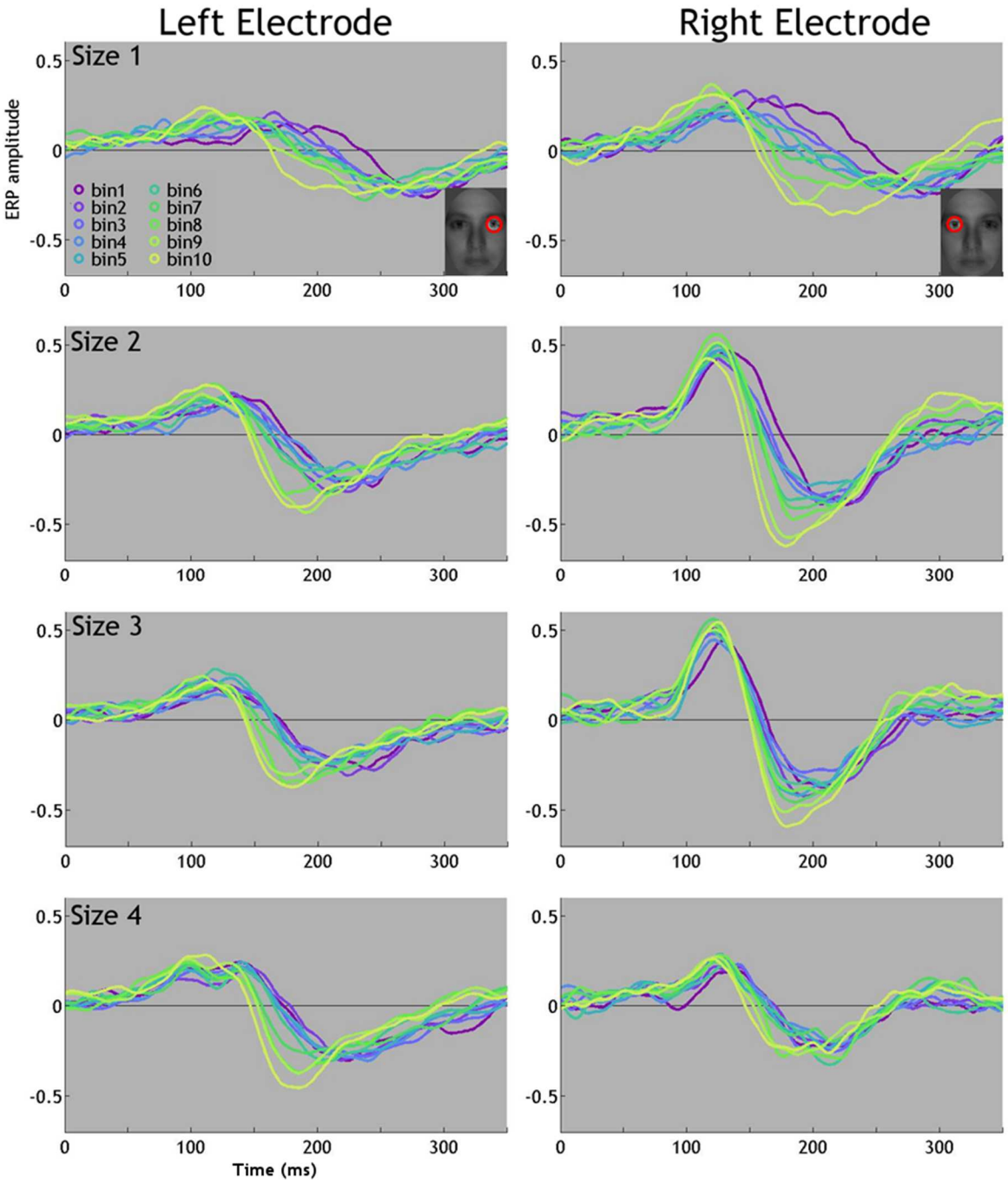


Figure 21: ERP modulation by feature visibility on texture trials (P5). Average ERPs during noise trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the area corresponding to the contralateral eye region on face trials (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the area corresponding to the contralateral eye region on face trials,, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest textures) to size 4 (largest textures).

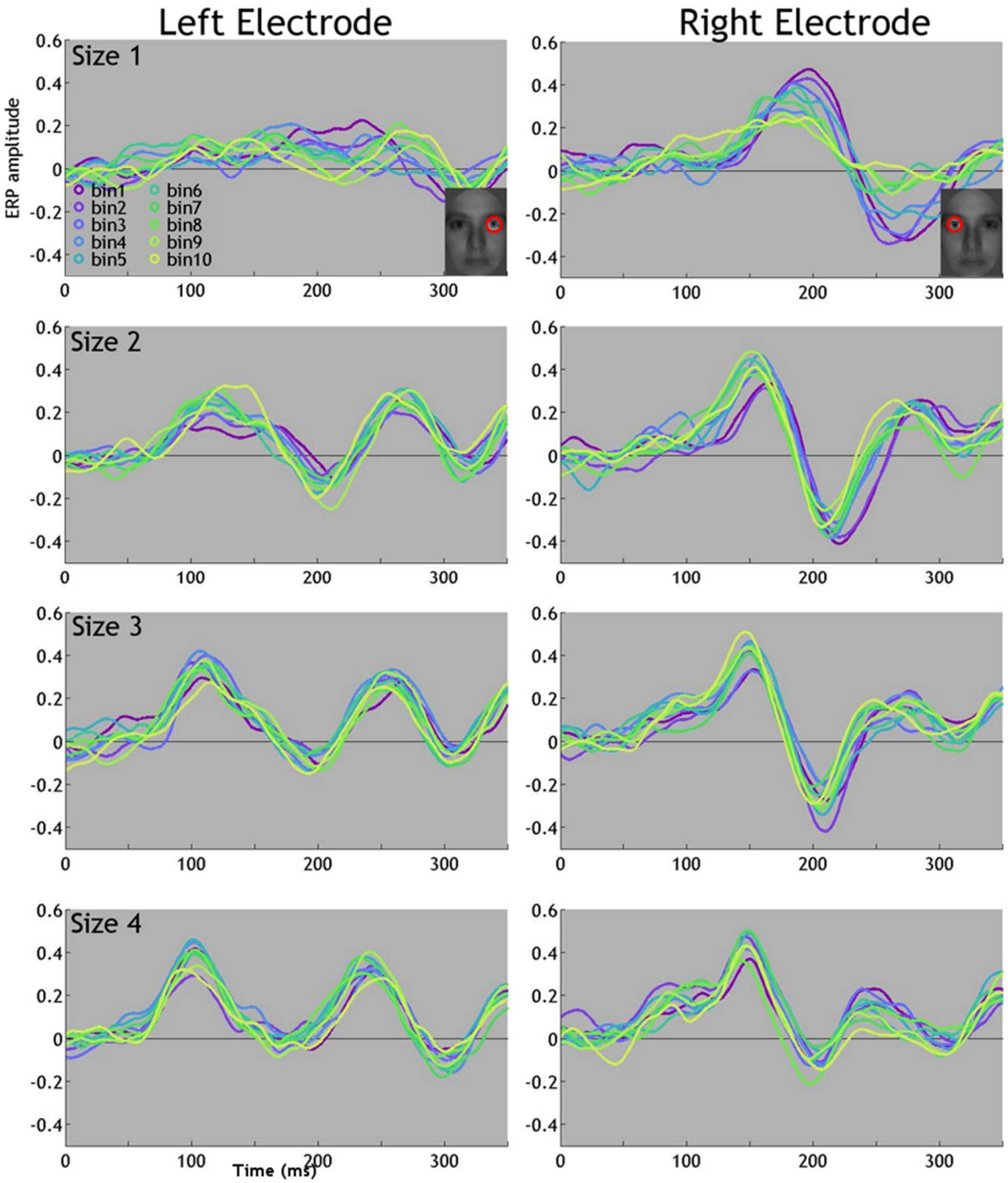


Figure 22: ERP modulation by feature visibility on texture trials (P6). Average ERPs during noise trials at the left and right hemisphere electrodes (left and right columns), grouped by the amount of visibility of the area corresponding to the contralateral eye region on face trials (highlighted by the red ellipse). ERPs were separated into 10 bins based on the degree of visibility of the area corresponding to the contralateral eye region on face trials,, from the least visible (bin 1) to the most visible (bin 10) trials. ERP modulations are presented separately for each image size, from size 1 (smallest textures) to size 4 (largest textures).

Discussion

Previous work (Rousselet et al., 2014) has suggested that the N170 predominantly reflects the encoding of the contralateral eye in face detection tasks, and is not due to retinal bias or allocation of attention to the eye area of the face. In the current experiment using a range of image sizes, we have established that the N170 sensitivity to the contralateral eye area is scale tolerant though less lateralised for very small faces, and cannot purely be explained as a non-feature specific allocation of attention to the left of a central fixation cross.

In the current study we found that participants' reaction times and accuracy in a face versus texture detection task were modulated by the amount of visibility of the left eye in face but not texture trials. We also found brain activity was modulated by the visibility of the contralateral eye for all face sizes, though contralateral eye sensitivity was less lateralised for very small (3 degrees of visual angle) faces. With increasing stimulus size, we found a shrinking hotspot of association between pixels and behavioural responses and pixels and brain responses revealing sensitivity to the iris/pupil area.

Can these results be explained by a left gaze bias? In texture trials, behavioural classification images did not indicate any significant association between pixels and behavioural responses. This suggests that participants did not rely on any one particular area more than another area in identifying texture trials and goes against the explanation of a left gaze bias. Whilst in our experiment changing the stimulus size systematically shifted the eye regions increasingly higher and into the periphery compared to the central fixation cross, Yi (2018) has recently also demonstrated consistent contralateral eye sensitivity to faces that have been vertically aligned so as to present the eyes either above, below or in line with a central fixation cross. Yi's (2018) results also point against a left gaze bias explanation of contralateral eye sensitivity.

On the other hand, modulation of ERP responses to texture trials with increased visibility of the contralateral eye area varied with stimulus size. At the group level, for the largest 2 image sizes there was no modulation with increased visibility of the contralateral eye, replicating the results of Rousselet et al.

(2014). For the smallest 2 textures there was modulation of ERPs with increased visibility of the 'eye' area at the group level. This may reflect increased noise sensitivity towards textures displayed in the eye area for larger textures, which could explain why participants were generally faster and more accurate on larger compared to small trials. Larger textures are displayed with more, comparatively smaller bubbles (compared to 1 large bubble in the smallest textures) which may require information revealed by the various bubbles to be integrated before a face detection decisions can be made. It may be that the *absence* of a readily visible eye and *presence* of a visible texture in the eye location in the larger texture trials drives sensitivity to the 'eye' area in larger texture trials.

In the current experiment we controlled for the percentage of face area revealed in the different sizes of face trials by presenting increasing numbers of bubbles as stimulus size increased, whilst keeping the size of the bubbles the same. We did this to ensure that the *amount* of overall face revealed across the different sized images was the same, so that any differences between image sizes could not be accounted for by differences in the overall amount of information available. However, this approach may have resulted in differences between the face sizes due to the requirement to integrate information across more 'patches' in larger face images compared to in particular the smallest face images where 1 bubble was presented and there was no requirement to integrate information across bubbles. The bubbles manipulation may be akin in essence to perception of partially occluded faces. Studies investigating face amodal completion, where the visual system perceives occluded faces as whole and complete, have demonstrated that amodal completion can modulate the amplitude of ERPs within the time window of the N170 and that the time required for amodal completion increases as a function of the amount of occlusion (Chen, Liu, Chen, & Fang, 2009). In the current experiment it is unclear the extent to which variations in the degree of occlusion and amount of integration across bubbles can explain variations in results between faces of different images sizes. It is possible for example that differences in our results for the smallest face size in which only 1 bubble was shown may be influenced by these factors.

Future studies investigating image size should consider the issue of integration over differing numbers of bubbles as. An alternative approach to varying the number of bubbles across image size would have been to keep the number of bubbles the same e.g. 1 bubble for every face size, and increase the size of the bubble as image size increased to maintain the same percentage of face area revealed across the different image sizes.

We found large individual idiosyncrasies in ERP modulation towards both face and texture trials, with some participants showing consistent modulation of ERPs with increased 'eye' visibility across all texture trials. This individual variation was not apparent when considering group-averaged ERPs. Group-averaged ERPs are problematic as they can cause misleading interpretation of results because group-averaged data can hide reliable inter-subject differences such as those demonstrated in this experiment and elsewhere (Pernet, Sajda, & Rousselet, 2011; Rousselet, Gaspar, Wieczorek, & Pernet, 2011). It is unclear why participants may differ.

One explanation for individual differences could be differences in preferred eye gaze patterns. One weakness of our study was the absence of eye tracking data, as in addition to individual differences in preferred eye gaze patterns, changing the size of the stimuli may in itself change individual eye gaze patterns. Recently Wang (2018) has proposed that when viewing larger images of faces, participants fixated more often and for longer on the eye region than when viewing smaller faces. In the current experiment however, stimuli were presented for a very short time, so participants may not have time to make saccades. However, eye tracking data could be useful in ensuring participants were fixating on the central cross before a trial began.

Our results are consistent with a bottom-up data driven model of face processing as suggested by others (DiCarlo, Johnson, Gross, & Bruce, 1999; Rousselet et al., 2014). Under this model, detecting the presence of the contralateral eye may be the first step in inferring the presence of a face. This may then be followed by the processing of other task-relevant facial features. In the current experiment, detecting the presence of the eye was sufficient to accurately judge the presence or absence of a face. However, in other, more complex face processing tasks, the processing of other additional facial features may be required for

accurate performance – such as the processing of the mouth region in an expressive versus non-expressive detection task (Gosselin & Schyns, 2001). To achieve this aim, we need to test the same participants using the same set of stimuli in a variety of face processing tasks relying on other diagnostic facial features other than just the eye.

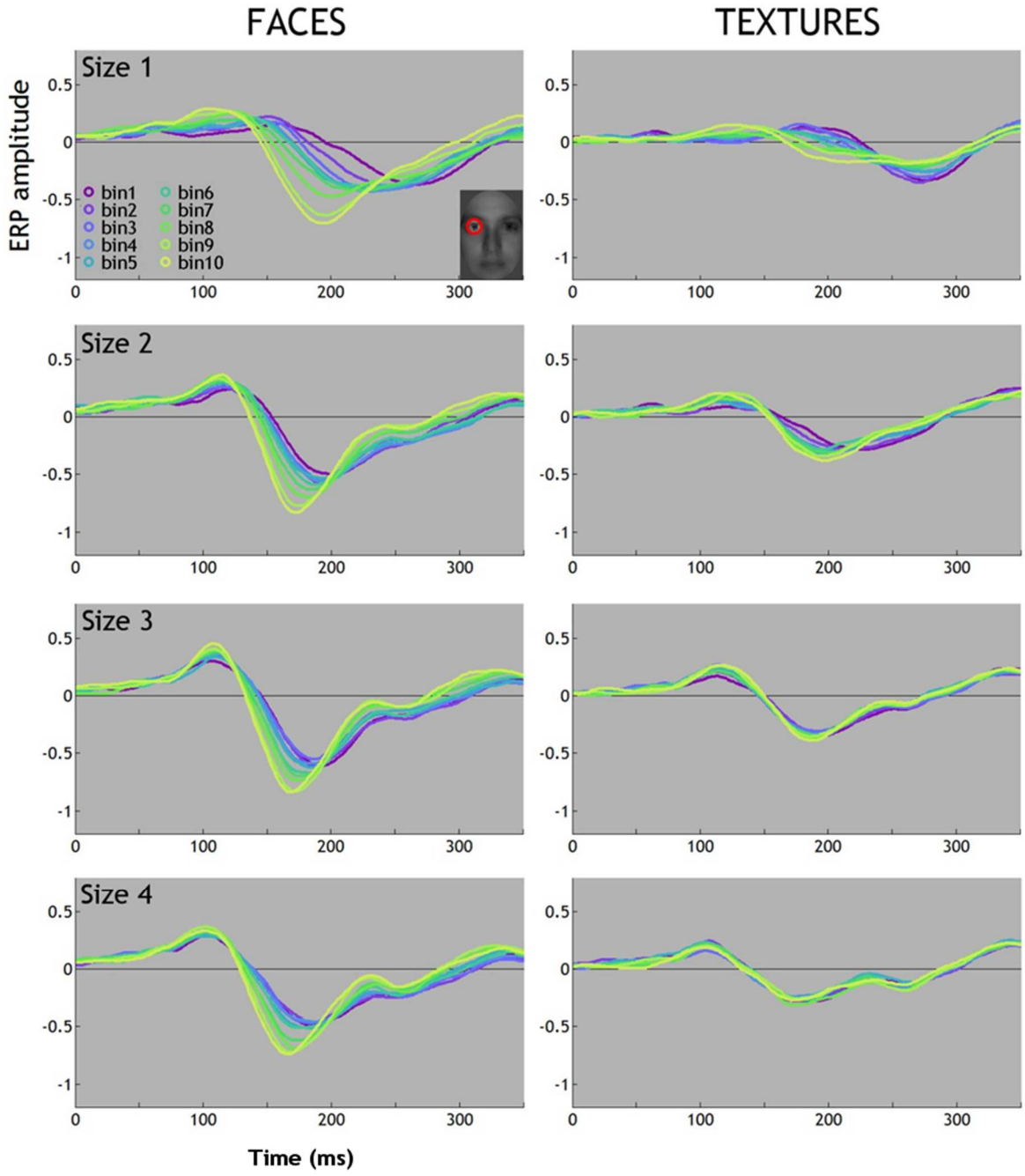
Whilst we have begun to determine the tolerance of the N170 contralateral eye sensitivity to changes in stimulus size, other stimulus related properties are still to be fully investigated, including, for example, different stimulus sets and changes to the size and number of the Gaussian apertures. Additional work from our lab has already demonstrated contralateral eye sensitivity in an alternative stimulus set using expressive faces and a single, large Gaussian aperture (Rousselet, Gilman, Ince, & Schyns, 2014), as well as demonstrating that changing the image contrast delays and reduces contralateral eye sensitivity (Rousselet, Gilman, Ince, & Schyns, 2014; Yi, 2018). Similarly, embedding faces in a more ecologically valid context, such as presenting faces out with a common face oval frame and presenting face images in colour are crucial to formally verify the N170 contralateral eye sensitivity as a global phenomenon.

Controversy exists over whether the N170 is also sensitive to differences in facial emotions. For example, a recent meta-analysis of 57 studies concluded that angry, fearful and happy faces elicit the strongest N170 amplitudes (Hinojosa, Mercado, & Carretié, 2015). If, as we suggest, the N170 reflects the processing of the pupil area, as is seen with increasing image size, this could account for differences in N170 sensitivity to emotional faces, as the region around the eyes is also the most diagnostic area for correct identification of emotion for fearful and angry faces (M. L. Smith, Cottrell, Gosselin, & Schyns, 2005). This is particularly the case for images of fearful faces where the eye is more open, with a larger presentation of the sclera and greater visibility of the pupil. N170 modulation by fearful faces (Batty & Taylor, 2003) thence may in fact be an artefact of greater pupil visibility and thus greater modulation of the N170, rather than processing of the emotion of the face per se. One way in which to formally test such a hypothesis would be to systematically manipulate the eye regions of emotional faces by superimposing ‘fearful’ or ‘non-fearful’ eyes on

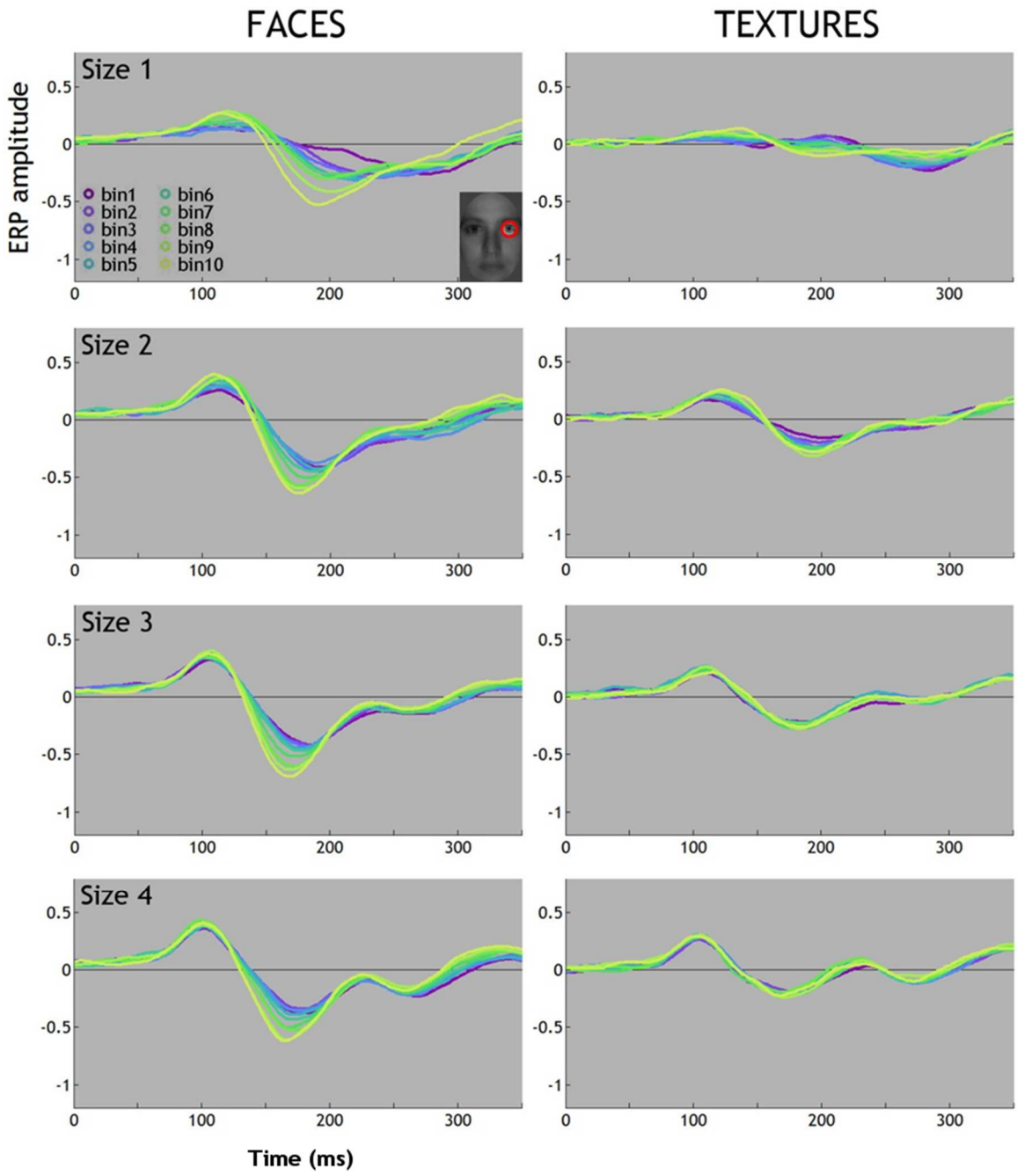
faces displaying fear or non-fearful emotions and test the resulting N170 modulations.

In conclusion, here we show for the first time that N170 contralateral eye sensitivity is scale tolerant. Specifically, as image size increases, we observed a hotspot around the contralateral eye area which shrank to the iris/pupil area for large faces. These results suggest the contralateral eye sensitivity cannot purely be explained as a non-feature specific allocation of attention to the left of a central fixation cross. Instead, contralateral eye sensitivity reflects the encoding of a specific feature – the eye.

Chapter 2 Supplementary Figures



Supplementary 1: Right Electrode Group Averaged ERPs Group averaged ERPs at the right electrode for faces (left) and textures (right) at four images sizes (size = 1 smallest, size 4 = largest)



Supplementary 2: Left Electrode Group Averaged ERPs Group averaged ERPs at the left electrode for faces (left) and textures (right) at four images sizes (size = 1 smallest, size 4 = largest)

Chapter 3 - Methods for Chapter 4 & 5

In Chapter 2, we have seen that the N170 contralateral eye sensitivity is scale tolerant in a face versus texture detection task. Next, we investigated whether the N170 contralateral eye sensitivity result is resistant to further changes in task demand, changes in stimulus properties and age of participants.

In Chapter 4 our aims are twofold: firstly, we quantified the extent to which more complex tasks would alter N170 eye sensitivity, where the eyes varied in terms of behavioural relevancy for completing the task. Secondly, we ascertained whether contralateral eye sensitivity consistently precedes sensitivity to other facial features, even under conditions where other features are more behaviourally task relevant than the eyes. By varying task demands, we tracked not only eye sensitivity as in Chapter 2, but simultaneously tracked sensitivity to another facial feature - the mouth. To achieve these aims, we used the same stimulus set in two different discrimination tasks (expressive versus non-expressive - henceforth '**EXNEX**' - and female versus male – henceforth '**GENDER**') in young participants (see Chapter 4 for more rationale). These tasks have been shown to diverge in the features that are behaviourally relevant for fast and accurate discrimination, namely that the mouth is the most behaviourally relevant feature for the **EXNEX** task, whilst the eyes are more behaviourally relevant for the **GENDER** task (Gosselin & Schyns, 2001; Schyns, Bonnar, & Gosselin, 2002; Schyns et al., 2003; M. L. Smith et al., 2004, 2007).

In Chapter 5, we expand on the results from Chapter 4 to understand the diagnostic information used by older adults; quantify how the N170 contralateral eye sensitivity in the two discrimination tasks outlined above compare in healthy older adults and understand the timing of feature sensitivity in older compared to younger adults (see Chapter 5 for details).

The following methods, pre-processing and statistical analyses are applicable to the experiments reported in both Chapter 4 and Chapter 5. Information on participants is covered separately within each Chapter.

CHOICE OF STIMULI

In Chapter 4 and 5, rather than using grey-scale images of faces, we utilised the MaxPlanck FACES dataset (Ebner, Riediger, & Lindenberger, 2010). This dataset contains naturalistic, coloured images of faces displaying different expressions, that have been validated (Ebner et al., 2010). This dataset was used in an attempt to use more naturalistic stimuli than those often used within the field of facial perception research. Attempts at control over stimuli used in facial perception experiments has resulted in the preponderance of greyscale image of faces being used, which are often cropped within common oval frames where the hair, ears and contour of the face are missing (e.g. Gold et al., 1999), have unrealistic normalised hairstyles added (e.g. Gosselin & Schyns, 2001) or have inter-feature distances normalised (e.g. Dailey et al., 2010) (see Figure 23 for examples). This has led to face perception research being dominated by studies on unrealistic face stimuli. Cropping images of faces (for example into a face oval as in Figure 23; A) may also increase the amplitude on the N170 (Dering, Martin, Moro, Pegna, & Thierry, 2011). We argue that utilising a diverse stimulus set where stimuli are more naturalistic i.e. coloured images of faces with ears, neck, variation in hairstyle and variation in facial feature inter-distances (see Figure 25 for examples), could better reveal the strategies observers use in natural social interactions. For example, skin colour may affect a variety of facial perception tasks including face detection (Bindemann & Burton, 2009) and face recognition (Yip & Sinha, 2002) tasks. Facial colour may also affect recognition of facial emotions (Nakajima, Minami, & Nakauchi, 2017). For a comprehensive understanding of facial perception it is then important to move towards methods of using naturalistic datasets.

PROBLEMS WITH THE GOSSELIN & SCHYNS (2001) STIMULI

When introducing the Bubbles technique, Gosselin & Schyns (2001) demonstrated the utility of the bubbles for understanding task diagnostic information using the GENDER/EXNEX task described previously. Since then, several subsequent studies (Schyns et al., 2002, 2003, M. L. Smith et al., 2004, 2007) have relied on the assumption that in the EXNEX task the mouth is task-diagnostic, whilst in the GENDER task the eyes and mouth are task diagnostic. However, we have identified that there are particular methodological concerns with using the original stimuli used by Gosselin & Schyns (2001) (stimuli are demonstrated in Figure 23, C). Our concerns are thus:

- Poor application of the normalised hairstyle
- Non-FACS coded stimuli & inconsistencies in happy expressions (mouth-open versus mouth-closed)
- Identification of makeup being worn by female but not male models
- Too few identities

Attempts at applying a normalised hairstyle to the stimuli have resulted in inconsistent results. For example, in Figure 23 C, top row, third face from the left the application of the normalised hairstyle has also artificially smoothed the skin on the forehead as well as smoothing and lightening the eyebrows to make the eyebrows less distinctive. Smiling can cause wrinkling of the forehead, as happy expressions can be characterised by raised inner eyebrows, so artificial smoothing of this information may remove diagnostic information.

The stimuli set have not been subjected to validation for example by the Facial Action Coding System (FACS) (Friesen, 1978). This is particularly evident in the discrepancy between open-mouth and closed-mouth happy expressions. For example, in Figure 23 C, third row, fourth face from the left, this model displays a closed-mouth happy expression in contrast to the open-mouth happy expressions of all other models in the stimulus set. This may have incurred an undocumented response bias for this stimulus, as over trials “closed-mouth” may become synonymous “non-expressive” for this stimuli set.

Female models have been identified as wearing eye makeup. For example, in Figure 23 C, first and second row, first face from the left, there is clearly dark lines around the eyes indicative of eyeliner/mascara being worn. The darkened lips may also indicate that lipstick is being worn. This may enhance the contrast of the eye and mouth region in female compared to male faces, making the eye and mouth region diagnostic due to the presence of makeup. As cosmetic makeup may affect the N170 during face perception tasks (Tanaka, 2016) the lack of control in this stimulus set is problematic.

The Gosselin & Schyns (2001) stimulus set consists of 10 identities. Viewed over multiple trials, this could result in familiarisation with the identities of the stimulus set. It is possible that over repeated trials participants may be able to use identity as a way of discriminating GENDER. We suggest that too few identities are available to use in this stimulus set.

Hence, for the purposes of our experiment we adopted a new stimulus set that overcame the aforementioned methodological issues with the Gosselin & Schyns (2001) stimulus set.

Moving towards using naturalistic stimuli whilst using bubbles is problematic however due to the need to average over trials. For example, in Chapter 2, we used Bubbles to successfully demonstrate that in a face versus noise detection task the N170 was predominantly modulated by visibility of the pixel area around the left contralateral eye across a range of stimulus sizes. We achieved this result by using a systematically constrained stimulus set i.e. grey-scale images of faces cropped within a common oval frame, where the eye positions were normalised (see Chapter 2: *Stimuli* or Figure 23A). This approach is typical in perception research as a whole, and historically has been a necessary prerequisite in reverse correlation research due to the need to compare across trials with different underlying stimuli. Without ensuring that the location of facial features between images within the data set are the same (i.e. by computer manipulation of the location and inter-feature distances of the nose, eyes, mouth) averaging across images becomes problematic. This is because we ascertain how visibility of any given pixel modulates brain and behavioural responses by averaging across trials, and then relate significant clusters of pixels back to which feature or features were revealed by those pixels. By virtue of this approach, this requires an underlying assumption that if one cluster of pixels reveals e.g. the eye on one image, the same cluster of pixels should reveal the eye on a different image within the stimulus set. If facial features are not presented in approximately the same pixel location as other stimuli within the stimulus set this could potentially become meaningless. This would be the case when inter-image variability is large, and the bubble sampling is done with small kernels.

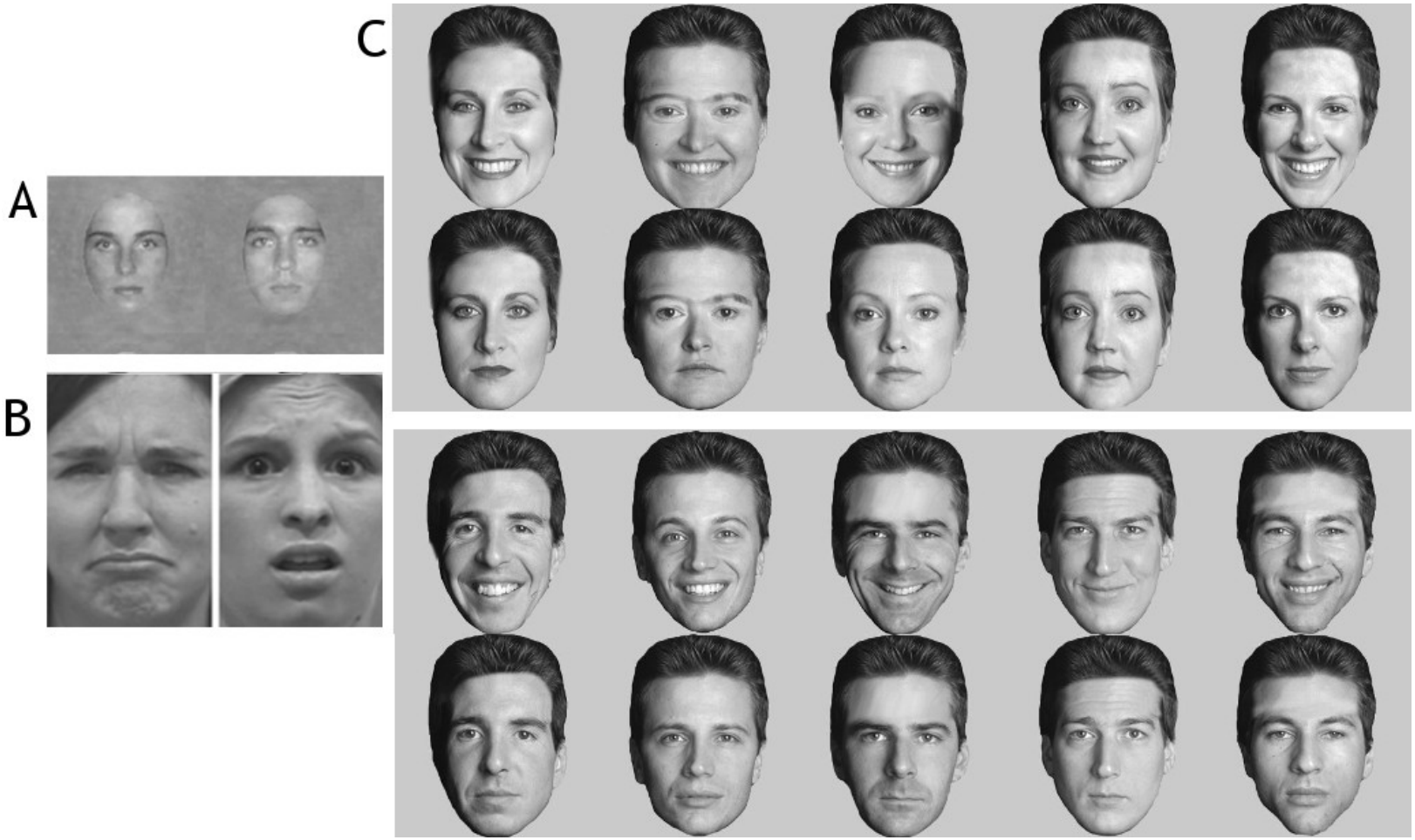


Figure 23: Example face stimuli Examples of face stimuli used in facial perception research. A) Face images cropped within common oval frame with no visible hair or ears (Gold et al., 1999). B) Face images linearly scaled to normalise distance between the eyes and between the eyes and the mouth (Dailey et al., 2010)**.** C) Face images with normalised hairstyles (Gosselin & Schyns, 2001). Note the occurrence of eye makeup on the female models (top two rows) and the discrepancy in the closed-mouth happy expression for 1 male model.

This problem of comparing over trials is more pronounced when using the bubbles technique. When using bubbles, methodological decisions need to be made as to the size of the Gaussian apertures utilised in the study. In a scenario where there is variation within a face dataset (such as in the FACES dataset (Ebner et al., 2010) used in Chapter 4 & 5), using a larger Gaussian aperture may still ensure the same facial feature is revealed across stimuli, even if the exact location of that feature varies among stimuli. For example, with larger apertures the pupil and corner of the mouth can be made visible on two different images despite inter-image variation in the pixel location of those features (see Figure 24A & B). As Gaussian aperture size decreases, i.e. the bubbles become smaller, averaging becomes more problematic as the same facial feature may not be revealed when there is variation in the position of that feature between images. For example, a smaller bubble, centred on the same pixel location as the larger bubble, may now only reveal the pupil and corner of the mouth on one image, but not the other (see Figure 24 C & D).

In Chapter 2, we kept the size of the bubbles constant across image sizes but varied the number of bubbles presented for each image. As image size increased, we argued that the hotspot around the eye region shrank, increasingly revealing sensitivity to the iris/pupil area. We were able to demonstrate this due to having a comparatively smaller bubble (in relation to the total face size), where we could precisely link pixels to a certain feature across all images in the data set. Hence there is a trade-off between reducing the bubble size to achieve a more precise knowledge of the facial features revealed by any given pixel, whilst also keeping bubbles large enough to ensure overlap of facial features to be able to present naturalistic images of faces that vary.

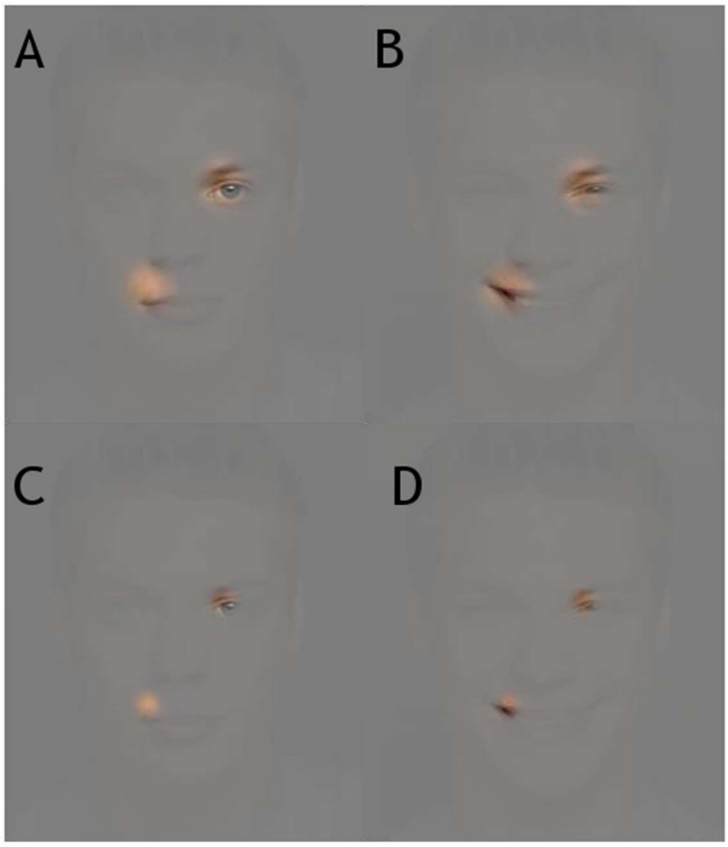


Figure 24: Differential feature content across images Panel A & B: Gaussian apertures centred on the same pixel location across images (s.d = 15) reveal the pupil and corner of the mouth on a 'neutral' (A) and 'happy' (B) face. Panel C & D: Gaussian aperture (s.d = 8) centred on the same pixel locations as A & B. The corner of the mouth is visible in C but not D. The pupil is revealed in C but only partially in D. Figure shows 1 ID only due to licensing agreement restrictions of the dataset.

In Chapters 4 & 5, we address this issue by using naturalistic faces that vary in size, inter-feature distances and hair style. By using a novel technique, *BubbleWarp* (Gill et al., 2015), we are also able to precisely quantify what facial feature is revealed on each trial, and relate this to the same facial feature revealed by a *different* cluster of pixels on an alternate facial image. *BubbleWarp* is based upon a process of delineating images of faces using landmarks representing salient facial locations (e.g. the corner of the eyes, nose and mouth). By computing the average template of all landmarks, we can reconstruct for each individual image within the dataset a classification image representing the facial features revealed by the randomly located Gaussian apertures (bubbles) on each trial (see *Warped Bubble Masks* for details). The advantage of this approach is that:

1) We move beyond merely comparing in pixel space and the assumption that we can infer content from an analysis of pixels as with the traditional bubbles approach (Gosselin & Schyns, 2001), and

2) Due to having no configural limitation we can use a varied heterogeneous dataset of real-world faces that have limited normalisation.

Stimuli

The stimuli used in Chapter 4 and Chapter 5 were coloured pictures of faces from the Max Planck FACES database (Ebner et al., 2010). Faces from 20 identities (10 women, 10 men) were front view photographs, with two images of each identity, where the image was either expressive ('happy') or non-expressive ('neutral'). Face models were of younger adult Caucasians within the age range of 19-31 years. Example of stimuli displayed in Figure 25 showcase 1 exemplar from the dataset. Due to licensing issues more exemplars of individuals cannot be displayed here, so we also present images of the *average* of 1) female faces; 2) male faces; 3) happy faces; 4) neutral faces (generated by averaging across all images of the relevant category used in the experiment).

Images were first delineated using WebMorph (DeBruine, 2018), a web based version of Psychomorph. Faces were delineated using 189 points placed on the face following a standardised template system outlining key facial features, such as the eyes, nose, mouth and hairline. Images were then adjusted so that the centre of each pupil was aligned to the average pupil height across all images. This allowed the images to be aligned to correct for images where the head was tilted and so the fixation cross appeared approximately in the location of the nose. Differences in inter-pupil distances were not altered. Images were then cropped to remove blank edges after pupil height alignment and re-sized to 253 x 294 pixels to reduce image size. Original hairstyle and inter-feature distances were maintained. Stimuli spanned approximately 9.2° x 7.9° of visual angle. This is approximately equivalent to the visual angle of the size 3 faces presented in Chapter 2.

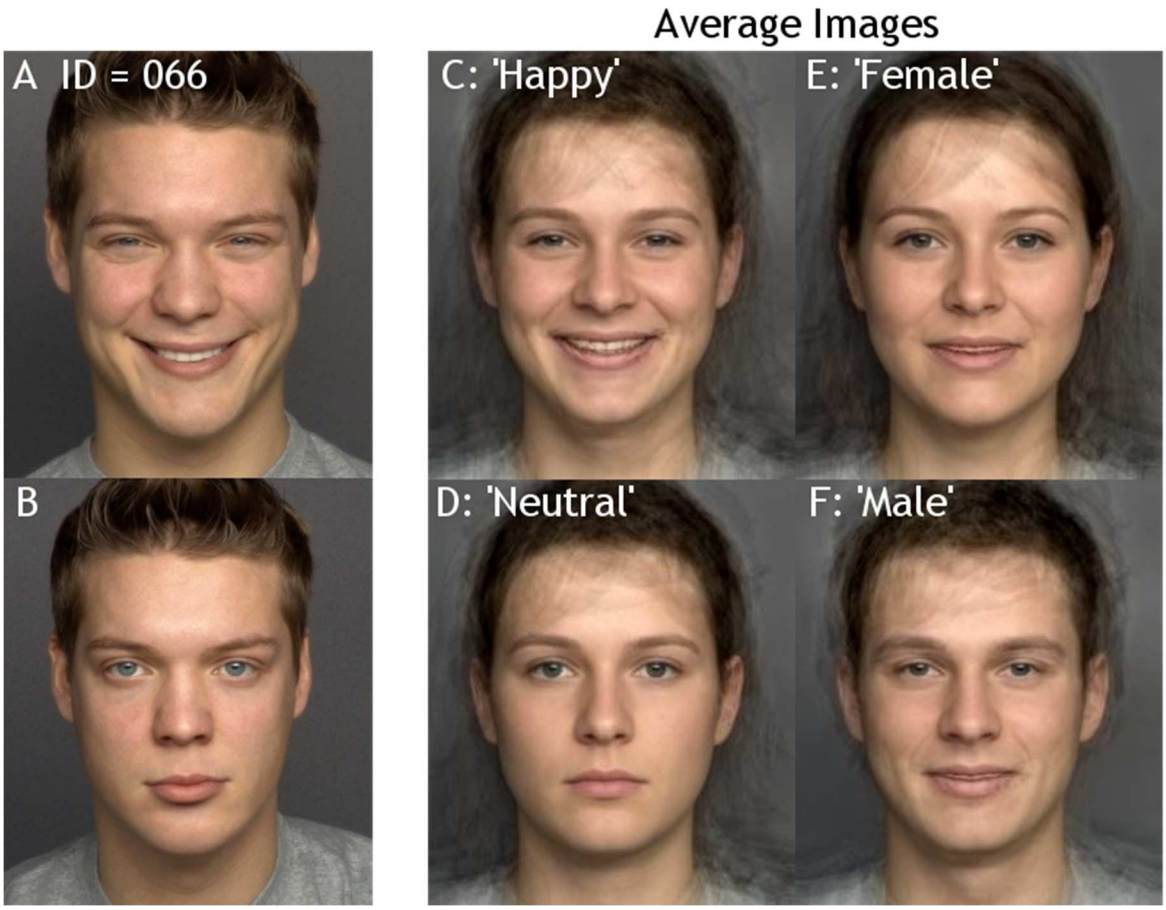


Figure 25: Example Stimuli Example stimuli are shown for 1 individual (FACES ID = 066) presenting 'happy' (A) and 'neutral' (B) expressions. Due to licensing agreement restrictions of the dataset, only 1 original ID can be demonstrated. Instead, we present average images calculated as the average across all images used in the data set presented here for demonstration purposes only. C) Average 'happy' expression across images. D) Average 'neutral' expression across images. E) Average 'female' face and F) Average 'male' face.

On each trial facial information was revealed through 20 two-dimensional Gaussian apertures (s.d. = 15) randomly located across the stimulus space, with the constraint that the centre of each aperture was at a unique position (see Figure 26). Bubbles were not constrained to the face area and could be located anywhere within the image, including non-facial areas such as the hair or shoulders. In Chapter 4 and 5 we will refer to these masks with Gaussian apertures as bubble masks. This sampling strategy approaches a uniform sampling of all face regions across trials, and information sampling was dense enough so that on each trial face features were revealed, but sparse enough to prevent ceiling effects.

Figure 26: Example of bubbled stimuli Examples of stimuli shown in the experiment. One identity is shown displaying a 'neutral' expression (top row) and 'happy' expression (bottom row) overlaid with bubble mask. Each bubble mask is composed of 20 Gaussian apertures (s.d. = 15). This sampling strategy approaches a uniform sampling of all face regions across trials.

Procedure

During the experiments described in Chapter 4 and 5, participants sat in a sound attenuated booth and rested their head on a chin rest. Viewing distance measured from the chin rest to the monitor screen was 45 cm. At the beginning of each of two experimental sessions, participants were fitted with a BIOSEMI (Amsterdam, the Netherlands) head cap with 128 electrodes and 4 additional ocular electrodes placed at the outer canthi and below each eye.

Stimuli were displayed on a VIEWPixx monitor (1920 x 1200 pixels; 22.5 inch diagonal display size; 120 Hz refresh rate). Participants were given experimental instructions including a request to minimise eye blinks and body movements during each block. Participants were asked to categorise images as fast and as accurately as possible: they pressed '1' for expressive and '2' for non-expressive faces (expression task), or they pressed '1' for male and '2' for female faces (gender task) on the numerical pad of a keyboard, using the index and middle fingers of their right hand. At the end of every block they received feedback on their overall performance (median reaction time and mean percentage correct), and, after Block 1, on their performance overall across all blocks completed thus far. Median reaction times and mean percent correct remained on the screen until participants pressed a key to move on to the next block.

Before the main experiment on each session, participants performed a practice block with full images without bubble masks, to help to minimise spatial uncertainty about the stimuli, and for comparison between full images and bubble trials as part of the analyses. After the practice block, participants performed 14 blocks of the images with bubble masks. All 15 blocks had the same structure. They consisted of 80 trials, with 20 face identities; each displayed 2 different emotions (expressive 'happy' or non-expressive 'neutral'). Each image was presented twice in each block. Participants could take a break at the end of each block. Each session consisted of 1200 trials, including 80 practice trials. All participants participated in two experimental sessions, with a different categorisation task in each session. Participants only completed one of the two tasks on each session. Task order was counterbalanced across participants. Each session lasted about 50-70 mins, including breaks, and excluding the time required to apply the EEG electrodes prior to actual testing.

Each trial began with a small black fixation cross displayed at the centre of the monitor screen for a random time interval of ~500-1000 ms. This was followed by an image of a face presented for ~100 ms. A blank grey screen followed stimulus presentation and remained until participant's response. After response there was a post-stimulus interval of 800 ms before the next trial began. The fixation cross, stimulus and blank response screen were all displayed on a uniform grey background with mean luminance = ~46 cm/m^2.

EEG Recording and Pre-Processing

EEG data were recorded at 512 Hz using an active electrode amplifier system (BIOSEMI, Amsterdam, the Netherlands) with 128 electrodes mounted on an elastic cap. Four additional flat electrodes were placed on the outer canthi and below each of the eyes. Electrode offsets were kept between ±20 μV.

EEG data were pre-processed using Matlab 2013b and the open-source toolbox EEGLAB version 13. Data were band-pass filtered between 1 Hz and 40 Hz using a non-causal fourth order Butterworth filter. A second dataset was created by pre-processing data with a fourth order Butterworth filter – high-pass causal filter at 1Hz and low-pass non-causal filter at 40 Hz.

Data from the two datasets were epoched between -300 to 1000 ms around stimulus onset. Baseline correction was performed using the average activity between -300 to 0 ms only for the high-pass causal filter data set. For the non-causal filtered dataset, the channel mean was removed from each channel instead.

Noisy electrodes and trials were then detected by visual inspection of the non-causal dataset and rejected from the two datasets on a participant-by-participant basis. The reduction of blink and eye-movement artefacts was performed using ICA, as implemented in the *infomax* algorithm from EEGLAB. ICA was performed on the non-causal filtered dataset, and then applied to the causal filtered dataset on a participant-by-participant basis, in order to remove the same components from both datasets. Components representing blinks and eye-movements were identified by visual inspection of their topographies, time-courses and amplitude spectrum. After rejection of artefactual components, a fourth order Butterworth 40 Hz low pass filter was performed again and data epochs were removed based on an absolute threshold value larger than 100 μV and the presence of a linear trend with an absolute slope larger than 75 μV per epoch and R^2 larger than 0.3.

Finally we calculated spherical spline current source density (CSD) waveforms using the CSD toolbox. CSD waveforms were computed using parameters 50 iterations, m=4, lambda =10^{-5}.

Electrode Selection

Detailed analyses were performed on a subset of electrodes. We pre-specified three clusters of posterior electrodes at the left, midline and right hemisphere (Figure 27). Our central posterior electrode cluster (CE) comprised 7 electrodes from Pz down the vertical midline, including Oz. Our left posterior electrode cluster (LE) comprised 19 electrodes including P07 and its neighbouring electrodes. Our right posterior electrode cluster (RE) comprised 19 electrodes, including PO8 and its neighbouring electrodes. Our lateral electrode selection excludes those electrodes immediately adjacent to the midline.

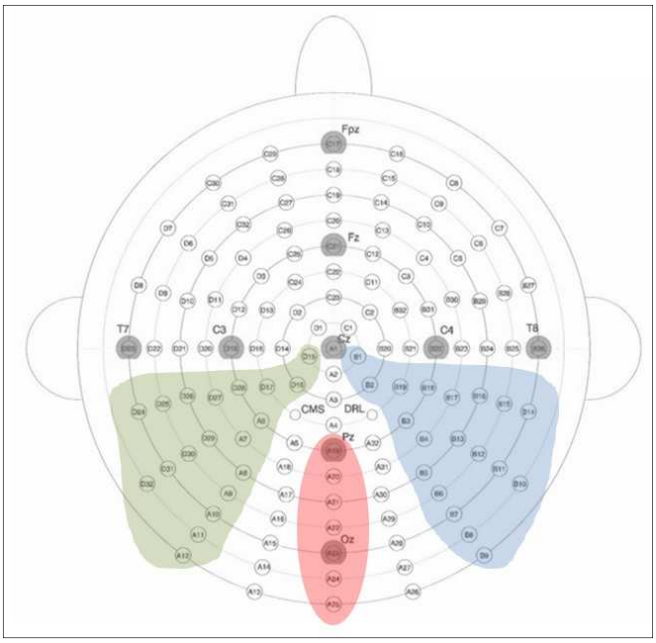


Figure 27: Electrode Selection for analysis consisted of three clusters of electrodes. A posterior right electrode cluster (RE) consisted of 19 electrodes including PO8 (shown in blue). A posterior left electrode cluster (LE) consisted of 19 electrodes including PO7 (shown in green). A midline electrode cluster consisted of 7 electrodes including Pz and Oz (shown in red).

Statistical Analyses

Statistical analyses were conducted using Matlab 2013b and the LIMO EEG toolbox (Pernet, Chauveau, Gaspar, & Rousselet, 2011). Throughout Chapters 4 & 5, square brackets indicate 95 % confidence intervals computed using the percentile bootstrap technique, with 1000 bootstrap samples. Unless otherwise stated, median values are Harrell-Davis estimates of the 2nd quartile (Harrell & Davis, 1982).

Warped Bubble Masks

In Chapter 4 and Chapter 5 we employed the use of the FACES dataset (see *Stimuli*). Due to the varied nature of the stimulus set, revealing the same cluster of pixels of each image did not reveal the same feature content. To allow averaging across the images in our data we re-established the connection between visible pixels and underlying facial feature across images within the dataset using the novel BubbleWarp technique (Gill et al., 2015). To this end, for each trial we retrospectively 'warped' the Gaussian bubble mask presented on that trial to the average face image (calculated as an average of all stimuli within the dataset) so that key facial feature points (such as the corner of the

mouth, edge of the iris) were aligned across trials making Gaussian aperture masks comparable.

This was achieved through the following steps: First, prior to the experiment each stimulus image (see Figure 28; A) was delineated in Webmorph (DeBruine, 2018) using facial feature 'fiducial' points (see *Stimuli* for more details). This provides a 'template' file for each image within the data set, where the co-ordinates for each of the fiducial points is unique (see Figure 28; B). Each template file is tessellated i.e. formed into a 'mesh' of imaginary lines joining the fiducial points together to make triangles. This is achieved through using Delaunay triangulation - a widely used tool in computer graphic algorithms that connects points in a nearest neighbour fashion whilst ensuring that the circumcircle associated with each triangle contains no other point in its interior. Next, we created an 'average' face (see Figure 28; D) from all stimuli within the data set. This is accomplished by moving the fiducial points for each individual image in the dataset to the mid-point location across all images in the data set (see Figure 28; C). The difference between the original template for each image and the average feature positions defines by how much each image would have to change in order to assume the same configuration as the average face. We refer to this as the 'warp-deformation'. In this fashion the original delineation points give 'base' co-ordinates that are 'warped' to achieve the 'target' co-ordinate destination on the average face. It is then possible to 'warp' the original face image to the average face image, so that the configuration of the original face images matches that of the average face image.

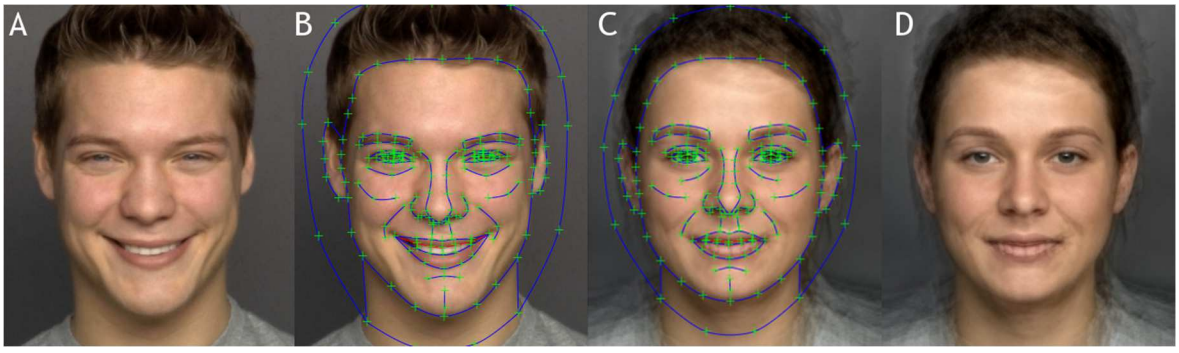


Figure 28: Delineation of stimuli A) Example stimulus set (adjusted to average pupil height and cropped). B) Example stimulus overlaid with 189 delineation points mapping key features of the face. C) 189 delineation points of the average face. D) The average face created by averaging over all images within the dataset

In this fashion, we warped bubble masks (as seen in the experiment by participants) to the average face image, using a custom made Matlab function *pawarp*. For each stimulus we calculated the warp deformation and then applied this calculation to the bubble mask on each trial involving that stimulus.

We refer to the resulting image as the *warped bubble masks*. As a result the bubble masks are equated across all trials, i.e. all bubble masks that revealed the eye of a face image in the experiment, regardless of which x-y coordinate pixels were actually revealed, would be 'warped' to reveal the same features on the average face. For instance, if the corner of the mouth was revealed in a given trial, the bubble mask is warped to ensure that the warped mask revealed the same pixel content (the corner of the mouth) in the averaged face. Throughout Chapters 4 & 5, analyses using bubble masks and pixels refer to analyses computed using the warped bubble masks.

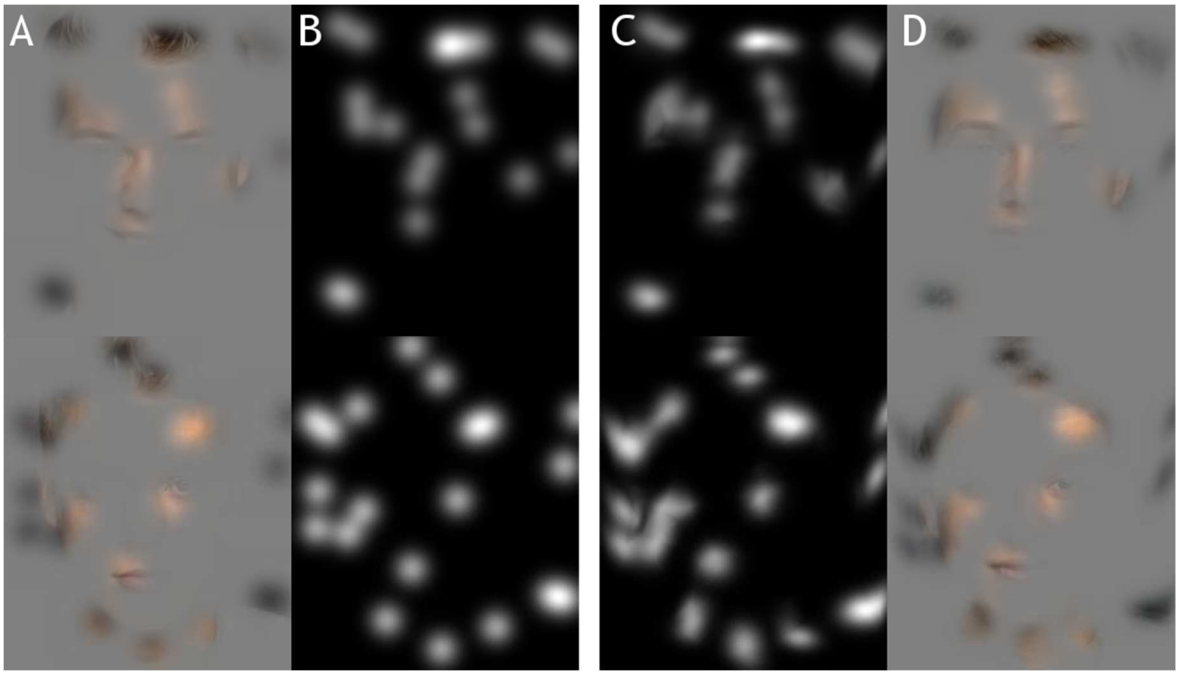


Figure 29: Bubble mask warping Column A) Examples of bubbled stimuli as seen by participants' during the experiment. Column B) corresponding bubble masks overlaying the stimuli seen in A. We 'warp' the original bubble masks (B) to our warped bubble masks (C). In column D) we show the warped bubble mask revealing the same features of the average face.

Mutual Information

We used Mutual Information (Thomas & Joy, 2006) to quantify the dependence between stimulus features and behavioural responses, and stimulus features and brain responses. Mutual Information (MI) is a non-parametric measure that quantifies (in bits) the difference in entropies and reduction in uncertainty

about one variable after observation of another (i.e. sampled pixels and behaviour; sampled pixels and brain signal responses). MI can be used to study the selectivity of neural and behavioural responses to external stimuli in single trials and is sensitive to non-linear effects (Schultz et al., 2015). Several tools for computing MI are available through an open source toolboxes, including the 'gcmi' (Gaussian copula mutual information) toolbox (Ince et al., 2017) used for analysis described in this thesis (https://github.com/robince).

Unlike in Chapter 2, in Chapter 4 and 5 we calculated MI quantities using warped bubble masks (*see Warped Bubble Masks*). We sampled the entire stimulus space, including the edge of the image where no facial features were presented, with the only constraint being that the centre of each bubble was located within the image. Over trials, this constraint leads to a systematic under sampling of the edge of the image corresponding to the size of the radius of the bubble, as there are never bubbles presented in which the centre of the bubble is located outside the edge of the image. Bubble masks specify the visibility of each pixel between 0 (not visible) and 1 (visible). A systematic under sampling of the edge of the image leads to more repeat zero values at the edge of the image than in the rest of the image. Due to the MI calculation toolbox we utilised taking the trial order when there are tied ranks, this can lead to erroneous MI effects in the periphery of the image that represent systematic effects of trial order (rather than a stimulus specific effect). For example a systematic slowing of reaction times over the course of an experiment may result in high MI values at the edge of the image.

Given that we made no prediction about the role of stimulus specific effects at the edge of the image, we minimised any sequential effects in the design by 'binning' our bubble masks prior to calculating MI. We binned the bubble masks into 2 bins using a fixed threshold of 0.05 i.e. so that values equal to or larger than 0.05 equalled 1, whilst values less than 0.05 equalled 0. Using alternative bin thresholds resulted in results consistent with our threshold of choice. This negates the issue of tied ranks in our MI calculations. This resulted in pixels being treated as a discrete rather than gaussian values, as in Chapter 2.

We calculated for each participant separately for the expressive and gender task:

- MI(PIX, **RT**) to establish the relationship between image pixels and reaction times. We copula normalised RT's using the *copnorm* function. We then computed MI using the *mi_gd* function to calculate MI between Gaussian copula normalised RTs and discrete bubble mask values. MI was calculated with bias correction for the entropy of Gaussian variables (Ince et al., 2016; Ince et al., 2009).

- MI(PIX, **CORRECT**) to establish the relationship between image pixels and correct responses. We computed mutual information using the *calc_info* function to calculate MI between discrete responses (correct versus incorrect) and discrete bubble mask values.

- MI(PIX, **[ERP, grad]**) to establish the relationship between image pixels and ERPs. For ERP analysis the warped bubble masks were down-sampled from 253 x 294 pixels to 147 x 127 pixels using the *imresize* function before the bubble masks were binned (see above) to reduce heavy computation time. We calculated *bivariate* MI, which considers the recorded voltage at each time point together with the temporal gradient of the ERP. Including the temporal gradient results in a smoothing effect, where artificial dips in MI due to the bimodal ERP crossing the zero line are smoothed out (Ince et al., 2016). We copula normalised ERPs using the *copnorm* function and computed MI using the *mi_gd* function to calculate MI between Gaussian copula normalised ERP and ERP gradients, and discrete bubble mask values. MI analysis was computed between -300 ms before stimulus onset to 600 ms following stimulus onset. We calculated MI independently at each time point, pixel and electrode, for each participant in each task using the non-causal dataset.

Mutual Information Classification Images

We refer to MI between pixels and behaviour, or pixels and ERPs as 'classification images'. These images reveal the warped pixels associated with modulations of the particular response being calculated.

Mutual Information Frequency of Significant Effects

In order to establish which parts of the classification image showed significant association with the behavioural performance, we performed a permutation test.

For each participant independently, we shuffled response labels (RT / correct) 1000 times. For each iteration, we calculated MI between the warped bubble masks and the randomly shuffled response labels, and saved the maximum MI value across all pixels. This creates a distribution of MI values under the null hypothesis that the variables (pixel MI values and behavioural responses) are statistically independent. To obtain the image pixels associated with the response at the significance level of 0.05, we compared the original MI scores against the 95^{th} percentile of the permutation distribution. For each pixel, we calculated the frequency of significant effects, as the total number of participants whose MI values were significant for each pixel.

Mutual Information Timecourses

We calculated how MI values between pixels and brain responses were modulated over time. For each of three pre-specified posterior electrode clusters (LE, RE, CE), for each participant, we took the max MI value for every pixel from each electrode cluster at each time point between -50 ms pre-stimulus onset to 450 ms post stimulus onset. We also calculated a 95 % confidence interval using the percentile bootstrapping technique with 1000 permutations.

Feature of Interest Analysis

MI is directionless, in that higher mutual information values can reflect either the presence *or* the absence of a feature modulating behavioural responses. Using a reverse analysis we quantified by how much changing the amount of information about the presence of specific image features (e.g. the eyes and the mouth) modulated brain and behavioural responses. To this end we created feature of interest masks for the left eye, right eye and mouth (Figure 30). To create the masks we centred an ellipse on the left eye, right eye and mouth region of the average face (see *Warped Bubble Masks*).

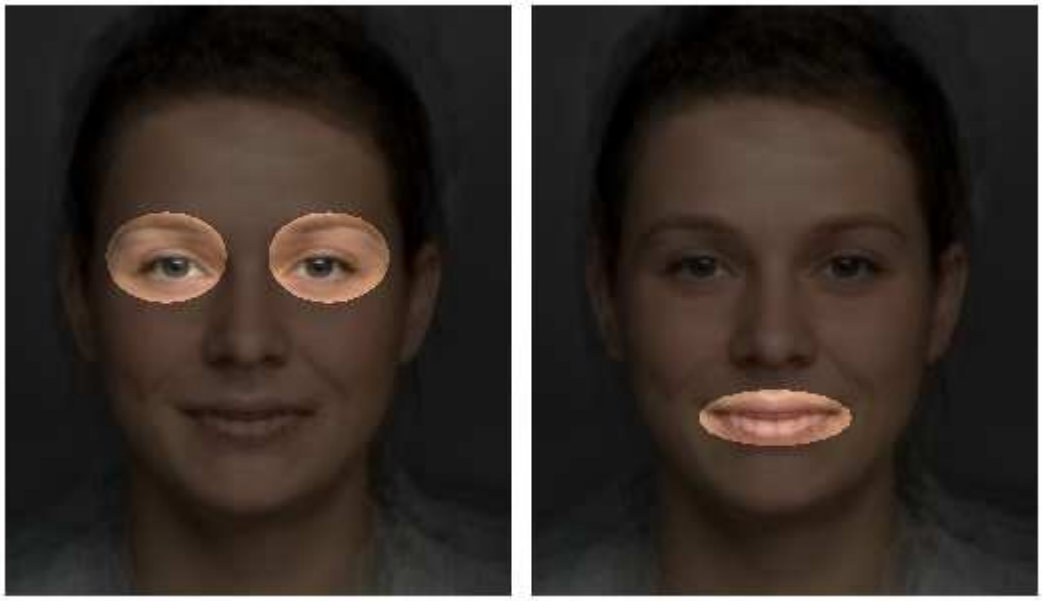


Figure 30: Demonstrated Feature of interest masks We created three feature of interest masks demonstrated above by the ellipses surround the 1) left eye, 2) right eye and 3) mouth region. Eye ellipses including the area of the eyebrow.

Using the feature of interest masks we calculated on a trial-by-trial basis the visibility of each feature (left eye, right eye, mouth), obtained as a scalar value of the sum of pixel visibility within the ellipse of each feature of interest mask. We then split these visibility values into ten equally populated bins ranging from the lowest (bin 1) to the highest (bin 10) visibility values. We used these feature visibility bins to compare how the degree of feature visibility modulated behavioural and ERP responses.

To quantify the effect visibility of each feature had on behavioural judgments, for each participant and feature of interest we calculated for each bin the percentage of correct responses (between 0-1) and Harrell-Davis estimate of median reaction times. We then calculated the *difference* between the 10^{th} bin (most visible) minus the 1^{st} bin (least visible) for each behavioural comparison.

For each behavioural comparison we also calculated a measure of the significance of the difference between high visibility and low visibility of each feature of interest. We did this by shuffling the pixel visibility values of each feature of interest 500 times and calculating MI between pixel visibility and behavioural response. To obtain values associated with a difference between high and low visibility at the significance level of 0.05, we compared the original

MI scores against the 95th percentile of the permutation distribution. For each feature of interest we consider there to be a significant difference in MI between high and low pixel visibility of each feature if the original MI value is above the 95th percentile of the permutation distribution.

To quantify the effect that visibility of each feature had on brain responses, we selected a single electrode of interest from the LE, CE and RE clusters. We optimised electrode selection for each individual in each task, by selecting the electrode from each cluster that had the maximum sum MI across the entire image, between 120-220ms. This broad window was selected to encompass the earliest sensitivity to facial features. Selected LE electrodes were P7/PO7 or their surrounding neighbours. Selected RE electrodes were P8/PO8 or their surrounding neighbours.

We calculated MI between feature visibility and brain activity over the time period of -300 ms pre-stimulus onset and 600 ms post-stimulus onset. For each electrode of interest, we calculated as a scalar value the sum of the pixel visibility within each elliptical feature of interest mask on each trial. We then calculated the time course of the MI about the feature visibility in the bivariate brain response combining EEG voltage with its temporal gradient: MI(feature, [**ERP**, **grad**]).

To compare how feature visibility modulated average ERP waveforms, for each electrode of interest we sorted single trial ERP's into 10 equally populated bins based on a scalar value of the sum of the pixel visibility within the ellipse of each feature of interest mask, ranging from the lowest (bin 1) to the highest (bin 10) visibility values. For each bin, we averaged ERPs across all trials.

Mutual Information Onset Analysis

In order to determine if sensitivity to the eyes began before sensitivity to the mouth, we quantified MI onsets to the eyes and mouth using the causal-filtered datasets. We calculated the maximum MI across all left and right poster lateral electrode clusters between the left and right eye and ERPs and the mouth and ERPs. For each participant, we calculated the median baseline (between -50 ms and 50 ms around stimulus onset). We localised the first peak after 80 ms whose

minimum height was 2.25 times higher than the median baseline. Then, using the ARESLab toolbox (Jekabsons, 2016) we built a piecewise linear model using the Multivariate Adaptive Regression Splines (MARS) method (Friedman, 1991).

Event-Related Potential Analyses

Average ERPs

We computed average ERPs by optimising electrode selection for each participant separately. For each participant, in each task, we calculated the mean ERP waveform across bubble trials for each electrode, and then we selected one electrode from the left and right posterior lateral hemisphere electrode clusters (see *Electrode Selection*). We selected the left and right lateral electrodes that displayed the minimum N170 amplitude in the time window around ~150-250 ms (Jaworska, 2017). Electrodes were PO7/PO8 or their neighbours. These electrode locations have been associated with N170 activity in previous experiments (Jaworska, 2017; Rousselet et al., 2014). We optimised electrodes in the same fashion for non-bubble i.e. practice trials within the time window of ~130-230 ms. This is because N170s peaked earlier in the practice compared to bubble trials. We calculated 95 % percentile bootstrap confidence intervals around the mean ERP waveforms by randomly sampling (N = number of trials) single trial ERPs with replacement 1000 times and calculating the average ERP on each iteration. Using our percentile bootstrap distribution we estimated a 95 % percentile bootstrap confidence interval around the mean ERP.

N170 amplitude and Latency calculations

We calculated the amplitude and latency of the N170 at the left and right hemisphere (using the electrodes as described above) for each participant, in each task. We computed the N170 latency for individual participants as the latency of the minimum ERP in the time window ~150-250 ms (Jaworska, 2017) and its corresponding amplitude. 95 % confidence intervals were computed using the percentile bootstrap technique, with 1000 bootstrap samples.

Between-group Comparisons

In Chapter 5 we calculated the group difference between younger and older adult participants by estimating the Harrell-Davis median of all pairwise differences of the distribution of younger minus older participants for each behavioural and brain comparison of interest. We computed a 95 % confidence interval around this estimate using a percentile bootstrap technique with 500 iterations.

We also calculated the effect size of between-groups comparisons of younger and older participants using a robust statistic: *Cliff's delta* (Cliff, 1996). Cliff's delta is a robust, non-parametric and normalised effect size based upon all pairwise comparisons between observations in two groups. It is not affected by outliers or difference of skewness between two groups. Cliff's delta is computed by calculating the probability that a randomly selected observation from one group (X) is larger than a randomly selected observation from another group (Y), minus the verse probability: *delta* = (sum(X>Y)-sum(X<Y))/NxNy. Cliff's *delta* ranges from 1 (where all values from one group are higher than the values from the other group) to -1 (when all values from one group are lower than the values from the other group). Completely overlapping distributions have a Cliff's *delta* of 0.

We also calculated a 95 % confidence interval of the Cliff's *delta* estimate using a percentile bootstrap with 500 samples. On each iteration, for each group, we sampled the original number of observations from each group (N = 24) with replacement and recalculated Cliff's *delta*. This creates a sampling distribution of bootstrap estimates of delta, from which we calculated the 95 % confidence interval.

We calculated Cliff's *delta* to compare between group differences in behavioural responses (**RT** & **CORRECT**) for younger and older participants. We calculated Cliff's delta independently for each response (**RT** / **CORRECT**), task (**EXNEX**/**GENDER**) and feature of interest (left eye / right eye / mouth). We also calculated Cliff's *delta* to compare between group differences in N170 amplitude and latency differences.

50 % Integration Times

To determine differences in information processing speed of the eyes and mouth in the absence of peaks in MI in some participants, we calculated 50 % integration times of the MI time-course towards the eyes and 50 % integration times of MI time-courses towards the mouth. This analysis takes into account the entire MI waveform (not just peaks) and normalises MI timecourses for comparison between groups. For each participant we computed the cumulated sum of the maximum MI across all electrodes in the time window of 0 - 500 ms. We then normalised the cumulated sum between 0 and 1, such that for each participant their cumulated sum had a value of 0 at stimulus onset and a value of 1 at 500 ms after stimulus onset. Finally, we computed the time necessary to reach 50 % of that function using a linear interpolation and calculated summary statistics based on a 95 % percentile bootstrap confidence interval for the Harrell-Davis median based on 1000 samples.

Group differences in 50 % integration time were calculated using a percentile bootstrap on median integration time to compare independent groups under the experimental hypothesis of there being an effect (rather than the null hypothesis). We used 1000 samples and an alpha of 5 %. Here, if the confidence interval does not contain 0 the difference can be considered statistically significant in the frequentist sense.

We calculated median 50 % integration times and CIs for each group (Younger and Older) for each feature (eyes and mouth) and task separately. We then calculated the difference in integration time between the eyes and mouth between younger and older participants and interaction.

Chapter 4 - Timing of contralateral eye sensitivity compared to other task relevant feature modulation is idiosyncratic

Introduction

In Chapter 2 we have shown that the N170 sensitivity to the contralateral eye area is scale tolerant, and reflects an eye-specific coding mechanism, rather than a generic non-feature specific attentional bias. Using a range of face sizes, we have been able to sufficiently track sensitivity to the iris/pupil area using a reverse correlation method, demonstrating that the N170 is a critical time-window in human face processing mechanisms that reflects predominantly the encoding of the contralateral eye in face detection tasks.

The next stage is to assess whether the N170 sensitivity to the contralateral eye is a task-specific response. If processing of the contralateral eye is a necessary first-stage of face processing, regardless of task, then we would expect there to be evidence of contralateral eye sensitivity in a range of more complex face-processing tasks, such as face identity, facial emotion discrimination, age judgements, judgements of trustworthiness etc. However, if contralateral eye sensitivity is task specific, and the eyes are only prioritised in face processing during particular tasks, such as in face versus noise detection tasks, we would expect contralateral eye sensitivity to be apparent only in a subset of face processing tasks were the eyes carry task-relevant diagnostic information. This could, for example, lead to differentiation where the N170 eye sensitivity is only apparent in tasks where the eyes are task-relevant, for example in face versus noise discrimination tasks (Rousselet et al., 2014), but not in tasks where the eyes are less task relevant, for example in an expression discrimination task (Gosselin & Schyns, 2001).

To this end, the aim in this experiment was to present participants the same facial stimuli under two different task conditions where the relative task diagnosticity of facial features varies between tasks. Participants completed a gender categorisation task (**GENDER**) and an expressive versus non-expressive categorisation task (**EXNEX**). By utilising the Bubbles technique (see Chapter 3: *Stimuli*) we manipulated on a trial by trial basis the amount of information

available about different facial features for the task. During the **GENDER** task, for which the eyes are purported as highly diagnostic (Gosselin & Schyns, 2001) changing the amount of information available from the eyes should affect task performance, whilst changing the amount of information available about the mouth, a feature that is less diagnostic for the task, should affect task performance minimally. For the **EXNEX** task, where the mouth is highly diagnostic (Gosselin & Schyns, 2001), changing the available information about the mouth should affect task performance more than changing information about the eyes, a region that is less task-diagnostic. If the N170 eye sensitivity is task-independent, we would anticipate that during both tasks there would be evidence of contralateral eye sensitivity. If processing the eyes is the first step in both of these tasks, then we would anticipate that contralateral eye sensitivity would begin prior to sensitivity to other facial features, regardless of the extent to which the eyes are task-relevant.

Previous work has suggested based on data from 2 participants that the N170 is modulated by contralateral eye sensitivity in both a **GENDER** and **EXNEX** task (M. L. Smith et al., 2004), despite the mouth being diagnostic in the **EXNEX** task for accuracy and reaction time. Smith et al. (2004) suggest that the N170 reflects a task-independent response to the eyes, whilst the P300 reflects task-dependent encoding of the eyes (in the GENDER task) and mouth (in the EXNEX task). This suggests that a pattern of contralateral eye sensitivity preceding other task relevant feature sensitivity should be observable in future experiments, though Smith et al. (2004) did not quantify the relative timing of the onset of mouth and eye sensitivity in the **EXNEX** task. Smith et al. (2004) also relied upon using grey-scale images with normalised hairstyles and a very long stimulus presentation time (1500 ms). The use of grey-scale face stimuli that are so severely normalised as to be homogeneous is problematic in face perception research (see Chapter 3 for discussion), whilst a long stimulus presentation time that gives viewers the time to scrutinise the stimulus may not result in the same pattern of sensitivity as a shorter presentation time. Real faces are heterogeneous, with variable inter-feature distances, differences in hairline and differences in skin tone. An additional, potentially hugely confounding problem with Smith et al. (2004) stimulus set is the presence of eye makeup worn by the

women models in the stimulus set. This may have changed the contrast of the eye region making the eyes a more salient feature across both tasks.

In an attempt to compare how the N170 sensitivity to facial information varies across 7 'universal' expressions, Schyns, Petro, & Smith (2007) analysed ERPs for 3 participants in a forced choice expression categorisation task using bubbles. The authors suggested that the N170 integrates facial features from the top of the face (e.g. the eyes) downwards, with the N170 peaking when expression-specific diagnostic information is reached, signalling the end of feature integration. The latency of the resulting ERP hence is earlier for 'eye-dependent' expressions (such as 'fear'), and later for 'mouth-dependent' expressions (such as 'happy'). In contrast to Smith et al. (2004) in which the eye region was not diagnostic for the EXNEX task, in Schyns et al. (2007) the eye region was diagnostic for correct expression identification across all emotions (though with idiosyncratic variation between participants). It is worth noting that, unlike the stimuli used by Smith et al. (2004), the stimuli used by Schyns et al. (2007) were FACS coded (Ekman & Friesen, 1978), were not wearing makeup and did not have normalised hairstyles. Similarly, a behavioural-only experiment on the same stimuli used by Schyns et al. (2007) with 14 participants was highly similar, again demonstrating that the eye region was diagnostic in an expression discrimination task (M. L. Smith et al., 2005). Schyns et al. (2007) do however support the notion that eye coding precedes coding of other facial features.

In our experiment, we will treat faces as heterogeneous, and utilising the Bubble-Warp technique (see Chapter 3; *Warped Bubble Masks*) present naturalistic coloured images of faces without makeup and with uncontrolled variation in facial feature inter-distances and hair line. Smith et al. (2004) also ignored individual differences in brain anatomy, with EEG analysis being conducted on the same spatial electrodes for every participant, rather than optimising electrodes to reflect common information processing and functionally equivalent signals across participants. In our experiment we will use mutual information (see Chapter 3: *Mutual Information*) to understand functionally equivalent signals across participants. Lastly, Smith et al. (2004) focussed on differences in N170 amplitude, whilst more recent work has suggested that single-trial N170 latencies code more about the presence of the contralateral eye (Rousselet et al., 2014). In our experiment we will consider both single-trial

N170 latencies and amplitudes to ensure a fuller understanding of how contralateral eye sensitivity modulates the N170 in different tasks.

In summary, we used two tasks (**EXNEX**/**GENDER**) to manipulate the relative task diagnosticity of different facial features, and a heterogeneous face data set to increase the realism and social relevance of our results. We hypothesised that in both tasks, visibility of the contralateral eye would modulate single-trial N170 responses, whilst the processing of other task-relevant facial features, specifically the mouth, would be delayed in comparison.

The primary contribution of this chapter will be demonstrating that visibility of the eyes modulated behaviour (accuracy and reaction times) and brain responses more in the **GENDER** than the **EXNEX** task, whilst the converse is true of mouth visibility. Both the visibility of the eye *and* the mouth region modulate the N170 in both tasks. We will also demonstrate that the onset of sensitivity to the eyes precedes the onset of sensitivity to the mouth in only 60-70 % of participants, suggesting idiosyncratic differences in the timing of feature sensitivity. These results will demonstrate for the first time that contralateral eye sensitivity may be preferentially encoded before sensitivity to other facial features, but that eye coding is not a prerequisite for encoding other facial features, namely the mouth.

Materials and Methods

Participants

The study comprised 24 participants: 16 women, 23 right handed, median age = 23 years (min 18, max 25). Participants did not report any eye condition, history of mental illness, or were currently taking psychotropic medications, suffered from any neurological condition, had diabetes, or had suffered a stroke or a serious head injury. Volunteers were also excluded from participation if they had not had their eyes tested within the last 3 years, in order to minimise the chances that volunteers did not know of an underlying eye condition. Participants' contrast sensitivity and visual acuity were assessed in the lab (Table 3). Contrast sensitivity was assessed using the Mars Letter Contrast Sensitivity set (Arditi, 2005). Visual acuity at 40 cm and 63 cm were assessed

using the Colenbrander mixed contrast card set (Colenbrander & Fletcher, 2004), and 6 m assessed using the Bailey-Lovie Chart (Bailey & Lovie, 1980). All participants had normal or corrected-to-normal visual acuity and normal contrast sensitivity (equal to or above the lower limit of normal Mars letter contrast sensitivity for a person aged 25 years of 1.56 log units (Haymes et al., 2006). Participants were compensated £6/hr for their participation. The experiment was approved by the local ethics committee of the College of Science and Engineering (approval no. 300150158).

LC 40	HC 40	LC 63	HC 63	LC 600	HC 600	CS
96 [89, 103]	105 [94, 110]	99 [84, 105]	107 [99, 110]	97 [79, 104]	104 [90, 109]	1.8 [1.64, 1.88]
0.08 *[0.22, -0.06]*	*-0.10* *[0.12, -0.20]*	*0.02* *[0.32, -0.10]*	*-0.14* *[0.02, 0.20]*	*0.06* *[0.42, -0.08]*	*-0.08* *[0.20, -0.18]*	

Table 3: Younger adult Visual Acuity and Contrast Sensitivity Scores Median Visual acuity and Contrast sensitivity (CS) sores for younger participants. Visual acuity scores are reported for low contrast (LC) and high contrast (HC) charts presented at 40 cm, 63 cm, and 6 m viewing distance, expressed as raw visual acuity scores (VAS). The corresponding logMAR scores are presented below in italics, where higher values indicate poorer vision and negative values represent normal vision (logMAR score of 0 corresponds to 20/20 vision). Square brackets indicate the minimum and maximum scores across participants.

Stimuli, Procedure, EEG Recording and Pre-processing

Chapter 3 outlines the methodology for this chapter, and for the subsequent chapter (Chapter 5). Stimuli, procedures, EEG recording and pre-processing are the same as those outlined in the Methods chapter (see Chapter 3 for details).

Results

Behavioural Results

Behavioural results are given here for trials with bubble masks and practice trials. Practice trials presented without bubble masks revealed the whole face image and were used to familiarise participants with the task. As bubble trials reveal partial face information, we compared behavioural performance between practice and bubble trials.

Reaction Times and Percentage Correct

On practice trials i.e. trials without bubbles where the full face image can be seen, participants were on average faster on the gender than expression task (**GENDER** = 434 ms [396, 472]; **EXNEX** = 444 ms [406, 482]; **Difference** [GENDER minus EXNEX] = -22 ms [11, -56]). Participants were as accurate on the gender and expression tasks (**GENDER** = 97 % [95, 98]; **EXNEX** = 97 % [96, 98]; **Difference** [GENDER minus EXNEX] = 1 % [-2, 1]).

On trials with bubble masks, participants on the whole were slower on the gender than expression task (**GENDER** = 544 ms [490, 599]; **EXNEX** = 489 ms [456, 522]; **Difference** [GENDER minus EXNEX] = 24 ms [-2, 49]). This is in reverse of practice trials on which participants were faster on the gender task. Participants were also less accurate on the gender than expression task for bubble trials (**GENDER** = 82 % [79, 84]; **EXNEX** = 89 % [87, 90]; **Difference** [GENDER minus EXNEX] = -7 % [-4, -10]). Again, this was in contrast to practice trials on which participants performed similarly in the two tasks.

Compared to practice trials, on bubble trials participants were slower (**GENDER** [Practice minus Bubble] = -92 ms [-54, -129]; **EXNEX** [Practice minus Bubble] = -41 [-16, -67]; **Difference** [GENDER minus EXNEX] = -51 ms [-12, -91]) and less accurate (**GENDER** [Practice minus Bubble] = 15 % [13, 17]; **EXNEX** [Practice minus Bubble] = 8 % [7, 10]; **Difference** [GENDER minus EXNEX] = 6 % [3, 10]) on both tasks. Compared to practice trials, bubbling the image had a more negative impact on both reaction times and accuracy for the gender than the expression task, suggesting that the impact of bubbling affected performance on the gender task differentially. This could be due to differentially increasing the difficulty of the gender compared to the expression task when information is limited or fragmented.

This pattern of results is in-keeping with previous research using an adaptive version of bubbles where the number of bubbles is varied to achieve a 75 % correct performance criterion (Schyns et al., 2002). Under these conditions, participants on average required more bubbles i.e. more of the image to be revealed on the **GENDER** task to achieve the same accuracy levels as on the **EXNEX** task.

Behavioural MI Classification Images

To determine what image features are associated with behavioural responses, we looked at mutual information between pixels and reaction times, MI(PIX, **RT**), and pixels and correct responses, MI(PIX, **CORRECT**). Frequencies of significant effects were calculated by a permutation test (see Chapter 3: *Mutual information: Frequency of significant effects*). The classification images shown in Figure 31 are for MI computed using the warped bubble masks (see Chapter 3: *Bubble Warp and Warped Bubble Masks*).

In the **EXNEX** task (Figure 31 panels A-B, top row) all participants RT's and correct responses were modulated by the presence of the mouth. In contrast, in the **GENDER** task (panel A-B, bottom row), there was a very weak relationship between the presence of the eyes and modulation of behavioural responses, which was only significant for **CORRECT** for a maximum of three participants at any one pixel (individual classification images are available in Supplementary 3).

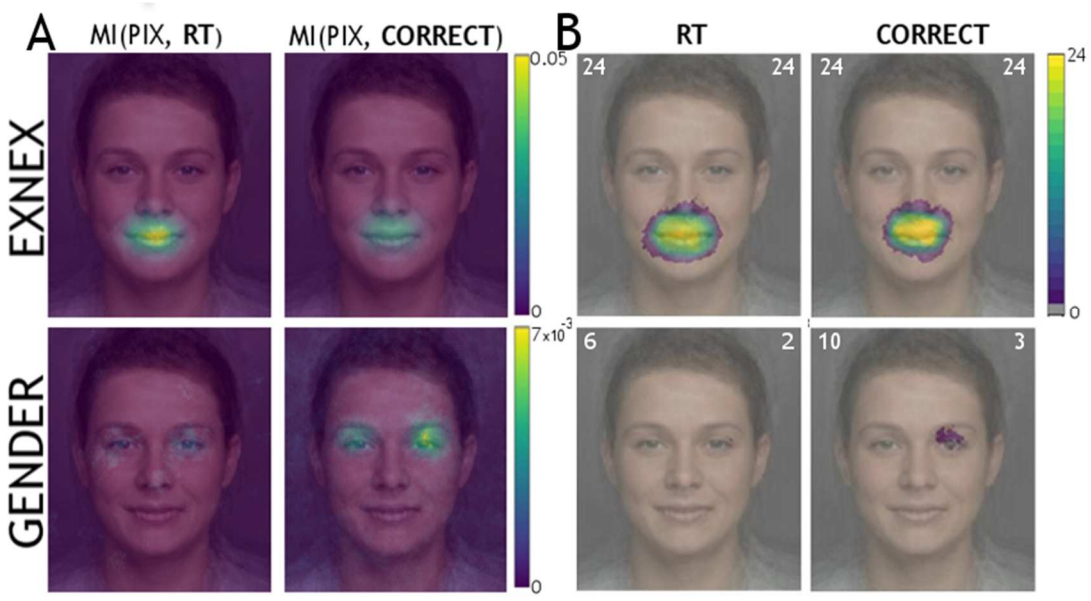


Figure 31: Younger adult behavioural Classification Images and Frequency of Significant Effect (A) Group-average MI maps. Each row corresponds to one task condition, each column corresponds to a different analysis condition (**RT** or **CORRECT**). The average group MI was stronger in **EXNEX** than in **GENDER** and is therefore presented on a different scale. (B) Number of participants showing significant effects based on a permutation test. Small white numbers indicate for each condition the number of participants with significant MI at any pixel (left) and the maximum number of participants with significant MI at the same pixel (right).

As expected, we found evidence of strong and consistent behavioural modulation by mouth visibility in the **EXNEX** task across participants, in keeping with the

mouth being the most diagnostic feature for this task (Gosselin & Schyns, 2001). However, behavioural modulation by eye visibility in the **GENDER** task was weak, despite previous claims of strong diagnosticity in this task (Gosselin & Schyns, 2001). This may be due to problems with the Gosselin & Schyns (2001) stimulus set outlined in Chapter 3.

Whilst we did not find a strong eye modulation for behavioural performance, there may be a disparity between those features driving behavioural responses, and those modulating ERPs.

Reverse Analysis: Behavioural Results

The presence of the mouth was associated with correct responses in the **EXNEX** task, whilst the right eye was weakly associated with correct responses in the **GENDER** task for some participants. However, MI is directionless, in that higher mutual information values can reflect either the presence *or* the absence of a feature modulating behavioural responses. To this end, we used a reverse analysis to check the direction and magnitude of MI effects in their original values. Using a reverse analysis we quantified by how much changing the amount of information about the left eye, right eye or mouth influenced participants behavioural responses.

Using the feature of interest masks (see Chapter 3: *Feature of interest analysis*) we calculated on a trial-by-trial basis the visibility of each feature (left eye, right eye, mouth), and split these visibility values into ten equally populated bins ranging from the lowest (bin 1) to the highest (bin 10) visibility values (see Chapter 3: *Feature of interest analysis* for details). For each participant separately, we calculated the difference in **RT** (measured in milliseconds) and difference in **CORRECT** (measured in percentage points (PP) between 0-1) between the 10^{th} bin (highest visibility) and the 1^{st} bin (least visibility) for each feature mask. For each comparison we also estimated the significance of these differences through the use of a permutation test (see Chapter 3: *Feature of interest analysis* for details).

In the following figures (Figure 32 - Figure 34) we will show that in the **EXNEX** task increased visibility of the mouth resulted in quicker and more accurate

responses, confirming MI results in Figure 31. In addition, our reverse analysis results highlight that, for some participants, visibility of the left and/or right eye also led to quicker and more accurate responses in the **EXNEX** task, which was not highlighted by the MI classification images alone. In the **GENDER** task, increased visibility of the left eye was associated with quicker responses for most (19/24) participants, and more accurate responses for nearly all (23/24) participants, which was not highlighted by the MI classification image results. Additionally, increased visibility of the right eye was also associated with quicker responses for most (21/24) participants, and increased accuracy for all participants, which was not highlighted by the MI results. Increased visibility of the mouth results in quicker reaction times and increased accuracy for half of participants.

MOUTH VISIBILITY

In the **EXNEX** task, increased visibility of the mouth (Figure 32, top panel) was associated with quicker responses (median difference in bin 10 minus bin 1 = -105 ms [-137, -73]). All participants (24/24) were on average faster on bin 10 than bin 1 trials, and MI was significant for 24/24 participants. Increased visibility of the mouth (Figure 32, second panel) was also associated with increased accuracy than when the mouth was not visible (median difference in bin 10 minus bin 1 = 0.25 PP [0.20, 0.30]). All participants (24/24) were on average more accurate on trials where there was increased mouth visibility (Figure 32, Panel B right), and there was significant MI for 17/24 participants. However, even in trials where there was no visibility of the mouth region, participants were still able to discriminate between expressive and non-expressive faces at above chance level, suggesting even in the absence of the mouth as the most diagnostic feature, other features and their combination can be sufficient for categorisation.

In the **GENDER** task (Figure 32, bottom two panels) increased visibility of the mouth was associated with variable behavioural outcomes. Half (12/24) of participants demonstrated faster responses with increased visibility of the mouth (median difference in bin 10 minus bin 1 = -4 ms [-21, 12]). However, MI was significant for 1 participant, whose results were in the opposite direction (slower reaction times for increased mouth visibility). Task order effects could be

associated with directionality of speed of responses with increased mouth visibility. Of those participants demonstrating faster responses with increased mouth visibility, 8/12 participants completed the **EXNEX** task first, where the mouth was the most diagnostic feature for the task. This could potentially represent a cross over effect between tasks, with participants primed to use the mouth region as diagnostic feature in subsequent tasks. However, no clear task order effects are apparent, with patterns of behaviour similar across the two groups (see Supplementary 4).

Similar variability in the **GENDER** task was evident in the association between mouth visibility and correct responses. Whilst nearly no evidence of an effect was evident when comparing median differences across the group (median difference in bin 10 minus bin 1 = 0 PP [-0.03, 0.03]), at an individual level half (12/24) of participants were more accurate with increased visibility of the mouth. There was significant MI evident for 1 participant with a difference in this direction.

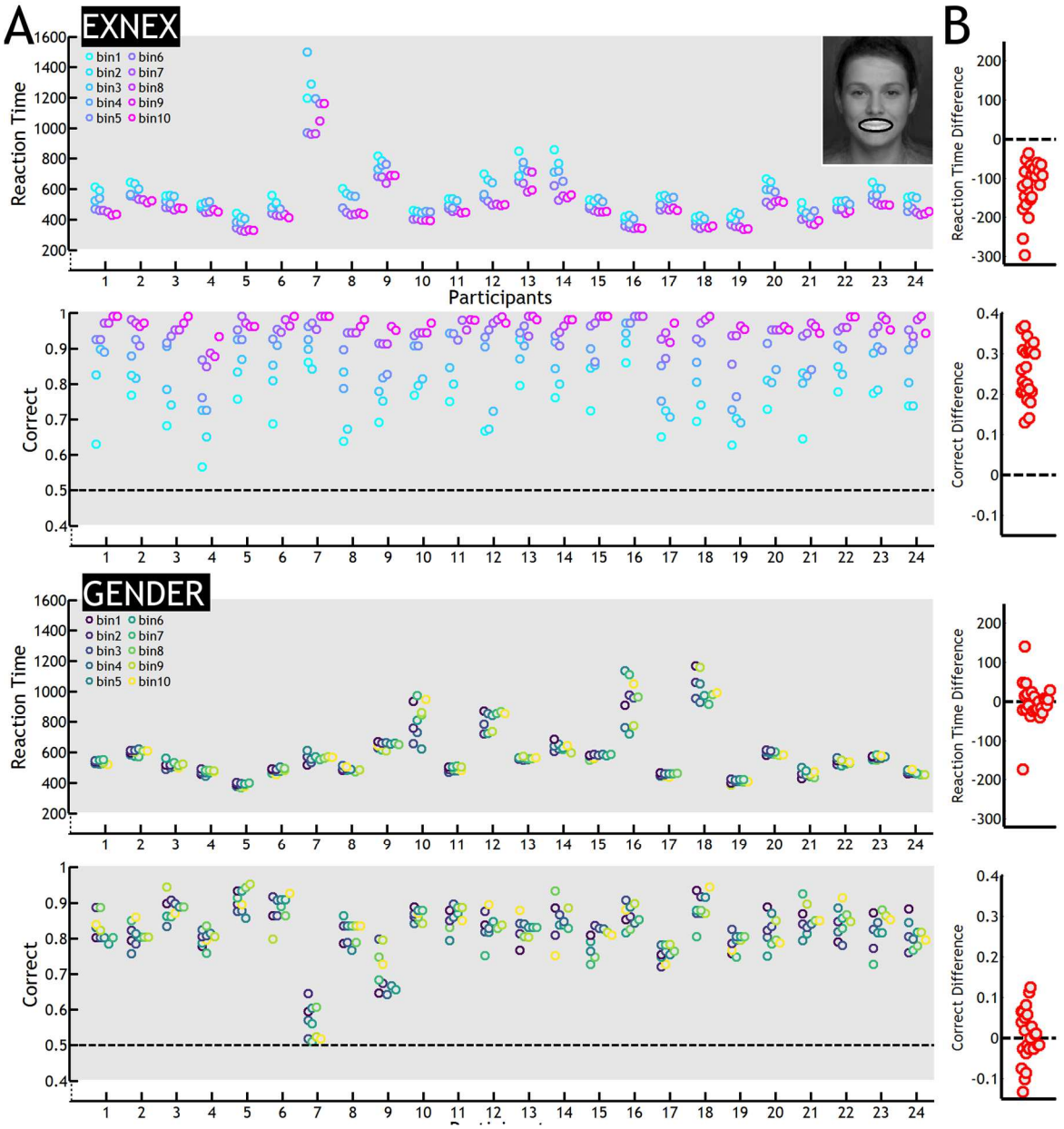


Figure 32: Younger adult reverse analysis of behavioural responses by mouth visibility
Panel A: Individual participant results for each of 10 visibility bins. Top two panels show EXNEX results. Bottom two panels show GENDER results. Panel B: Each circle represents one participant's difference between bin 10 minus bin 1.

In summary, increased visibility of the mouth was associated with increased accuracy and quicker responses in the **EXNEX** task for all participants, though even when there was no visibility of the mouth participants were still able to perform at above chance accuracy levels, suggesting that whilst the mouth is the most diagnostic feature, in its absence participants can use other features well for categorisation. In the **GENDER** task, there was an even split of participants in both directions on speed and accuracy modulations with increased mouth visibility. Whilst some participants were able to use the mouth as a diagnostic

feature for gender categorisation, this could not be explained be a general task-order cross-over effect, and instead seems to reflect individual variation.

LEFT EYE VISIBILITY

Left eye visibility has been shown to modulate reaction times in a face versus noise detection task (Rousselet et al., 2014). Visibility of the left eye has also been implicated in quicker and more accurate responses in gender discrimination tasks (Joyce, Schyns, Gosselin, Cottrell, & Rossion, 2006). In our results, visibility of the left eye (Figure 33) had a very variable effect on response times and accuracy in the two tasks.

Increased visibility of the left eye in the **EXNEX** task (Figure 33, top panel) was associated faster responses in 14/24 participants (median difference in bin 10 minus bin 1 = 7 ms, [-19, 6]). MI was significant for 6 participants, 4/6 participants showing significant MI were faster with increased left eye visibility, and 2/6 participants showing significant MI were slower with increased left eye visibility. Concurrently, increased visibility of the left eye in the **EXNEX** task (Figure 33, panel 2) was associated with an increase in accuracy for 17/24 participants (median difference in bin 10 minus bin 1 = 0.03 PP, [0.01, 0.05]). MI was significant for 1/24 participant who showed an increase in accuracy with increased left eye visibility.

In the **GENDER** task, increased visibility of the left eye (Figure 33, panel 3) was associated with faster responses in 19/24 participants (median difference in bin 10 minus bin 1 = -28 ms, [-41, -16]). MI was significant for 5/25 participants, 4/5 participants showing significant MI were faster with increased left eye visibility. Increased visibility of the left eye (Figure 33, bottom panel) was also associated with increased accuracy in the **GENDER** task (median difference in bin 10 minus bin 1 = 0.07 PP [0.05, 0.08]), for 23/24 participants. MI was significant for 3/24 participants, all of whom showed an increase in correct responses with increased left eye visibility. As MI was calculated based on single trial variability, this discrepancy between the majority of participants showing a difference in medians whilst few participants showing a significant MI, can be explained by median differences being less sensitive than MI calculation to single-trial variability. There was no apparent effects of task order (see Supplementary 5).

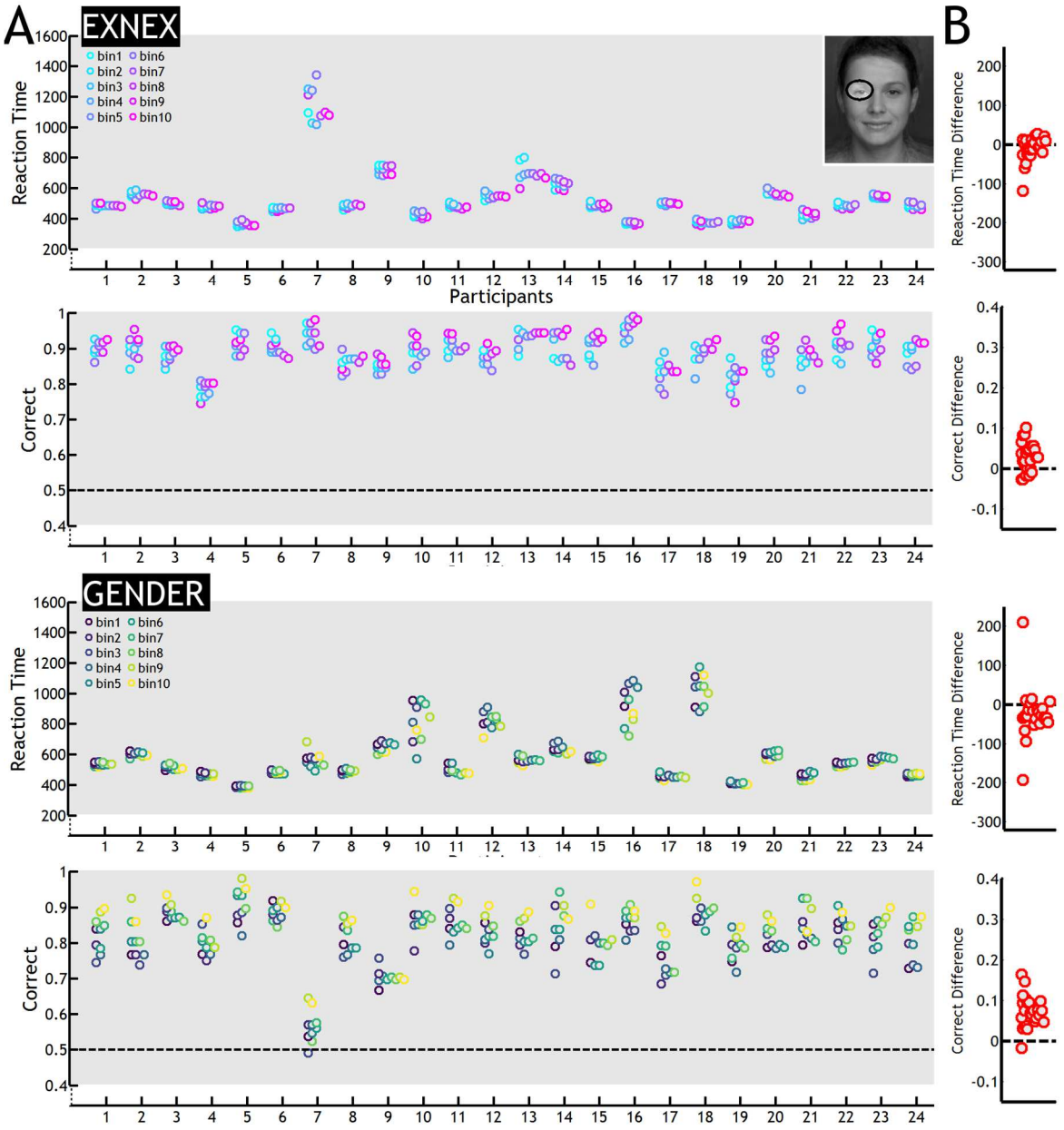


Figure 33: Younger adult reverse analysis of behavioural responses by left eye visibility
Panel A: Individual participant results for each of 10 visibility bins. Top two panels show EXNEX results. Bottom two panels show GENDER results. Panel B: Each circle represents one participant's difference between bin 10 minus bin 1.

In summary, increased visibility of the left eye in the **EXNEX** task was only associated with quicker reaction times and increased accuracy for 14/24 and 17/24 participants respectively. Results for increased visibility of the left eye in the **GENDER** task were also variable, with 19/24 participants showing faster reaction times. Though most (23/24 participants) showed an increase in accuracy with increased left eye visibility in the **GENDER** task, MI was only significant for a fraction of participants (3/24) suggesting large single-trial variability.

Previous work has suggested that the left, but not right eye region drives accurate judgements in male versus female categorisation tasks (Schyns et al., 2002). In our results, nearly all (23/24) participants showed an increase in accuracy with increased left eye visibility in the gender task. However, our results are in contrast to previous work which has suggested that correct categorisation of expressive/non-expressive faces does not involve either of the eye regions (Schyns et al., 2002), as in our results increased visibility of the left eye was related to increased accuracy for over half (14/24) of our participants.

RIGHT EYE VISIBILITY

Whilst left eye visibility has been shown to predominantly modulate behavioural responses in gender categorisation tasks (Schyns et al., 2002), research has demonstrated that a minority of participants demonstrate a converse right eye bias in face versus noise detection tasks (Rousselet et al., 2014). In our results we found a much more consistent bias towards right eye visibility modulation of behaviour in both the **EXNEX** and **GENDER** task than has previously been described (Schyns et al., 2002).

In the **EXNEX** task, visibility of the right eye (Figure 34, top panel) was associated with quicker responses for 14/24 participants (median difference in bin 10 minus bin 1 = -8 ms, [-17, 2]). MI was significant for 1 participant who was faster with increased right eye visibility. Visibility of the right eye (Figure 34, panel 2) was also associated with a small increase in accuracy for 19/24 participants (median difference in bin 10 minus bin 1 = 0.04 PP, [0.03, 0.05]). MI was significant for 1/24 participants, who showed an increase in accuracy with increased right eye visibility.

In the **GENDER** task, increased visibility of the right eye (Figure 34, panel 3) was associated with quicker responses in 21/24 participants (median difference in bin 10 minus bin 1 = -29 ms, [-47, -12]). MI was significant for 5/24 participants showing faster responses with increased right eye visibility. All participants (24/24) showed increased accuracy (median difference in bin 10 minus bin 1 = 0.10 PP [0.07, 0.12]) with increased visibility of the right eye (Figure 34, bottom panel). MI was significant for 10/24 of these participants. There was no apparent effects of task order (see Supplementary 6).

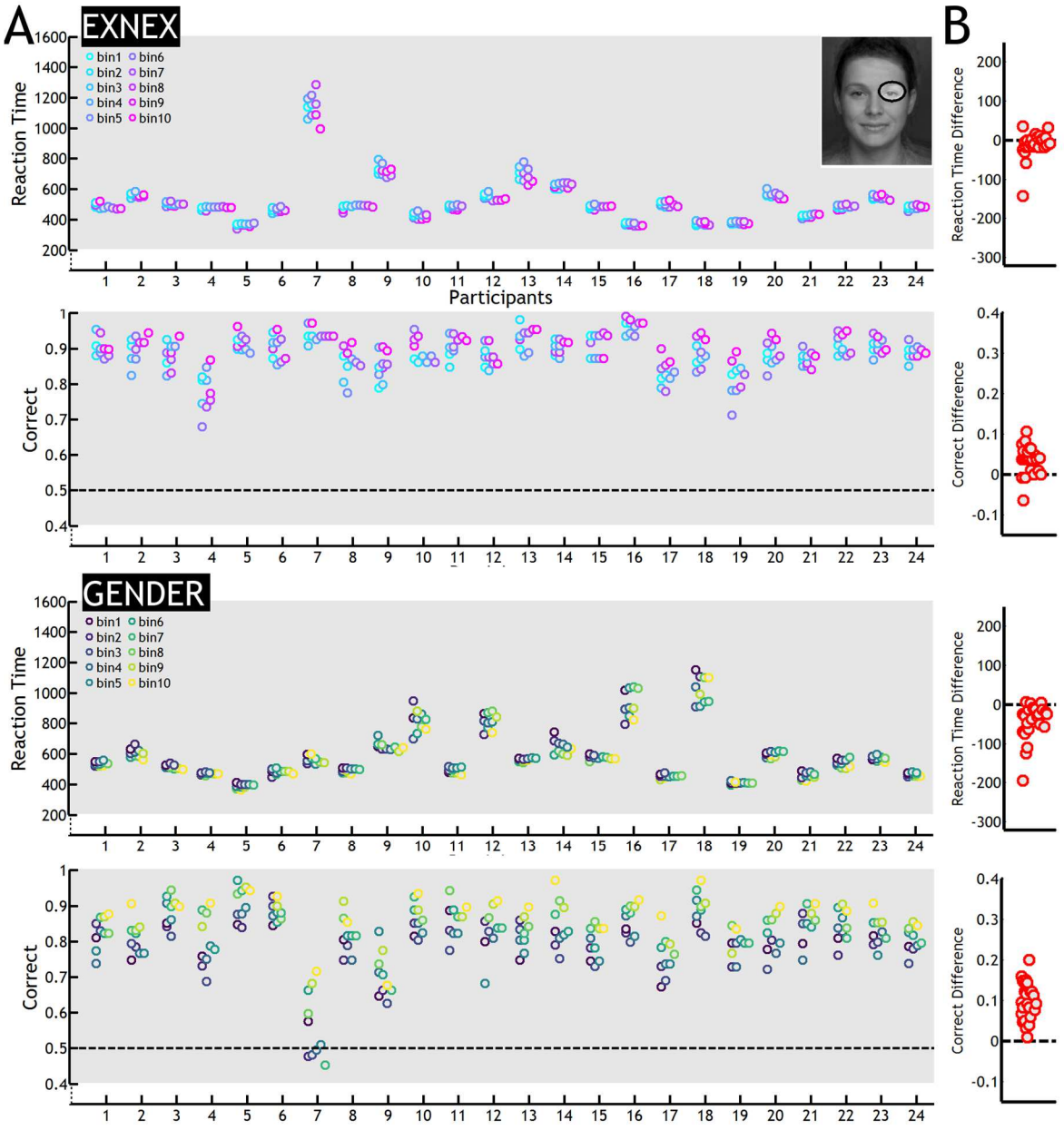


Figure 34: Younger adult reverse analysis of behavioural responses by right eye visibility
Panel A: Individual participant results for each of 10 visibility bins. Top two panels show EXNEX results. Bottom two panels show GENDER results. Panel B: Each circle represents one participant's difference between bin 10 minus bin 1.

In summary, in the **EXNEX** task, all participants were quicker and more accurate with increased visibility of the mouth. Increased visibility of the left eye decreased reaction times for 14/24 participants and increased visibility of the right eye decreased reaction time for 14/24 participants. In the **GENDER** task there was an even split of participants in both directions on speed and accuracy modulations with increased mouth visibility. Whilst some participants were able to use the mouth as a diagnostic feature for gender categorisation, this did not reflect a task-order cross-over effect, but rather individual variation. Increased visibility of the left eye was associated with quicker responses (19/24

participants) and increased accuracy (23/24 participants). Increased visibility of the right eye was associated with quicker responses (21/24 participants) and more accurate responses (24/24 participants). As with previous work (Rousselet et al., 2014), participants with significant MI showed a significant effect for either the left or right eye, but no participants showed significant effects for both eyes in either of the tasks.

In Figure 32 - Figure 34, participant 7 may be considered an outlier. Whilst participants 7's accuracy was comparable to other participants in the EXNEX task, this participant's reaction times were slower. On the GENDER task where reaction times were similar to other participants, this participant's accuracy was reduced. However, this participants has not been removed as intra-individual differences in accuracy and reaction with different visibility of facial features was still evident.

FEATURE VISIBLITY BETWEEN-TASK DIFFERENCES

To compare these tasks differences more clearly, we next calculated for each feature and behavioural comparison, the difference between the **EXNEX** and **GENDER** task difference of bin 10 minus 1.

Comparing differences in reaction time (Figure 35) the majority of participants (19/24) were faster with increased right eye visibility in the **GENDER** than **EXNEX** task. This suggests that increased visibility of the eye area is more task diagnostic for the **GENDER** than **EXNEX** task. Similarly, 16/24 participants were faster with increased left eye visibility in the **GENDER** than **EXNEX** task, again suggesting that increased visibility of the eye area is more task diagnostic for the **GENDER** than **EXNEX** task. Conversely, all but 1 participant was faster with increased mouth visibility in the **EXNEX** than **GENDER** task, highlight the task diagnosticity of the mouth region for the **EXNEX** but not **GENDER** task.

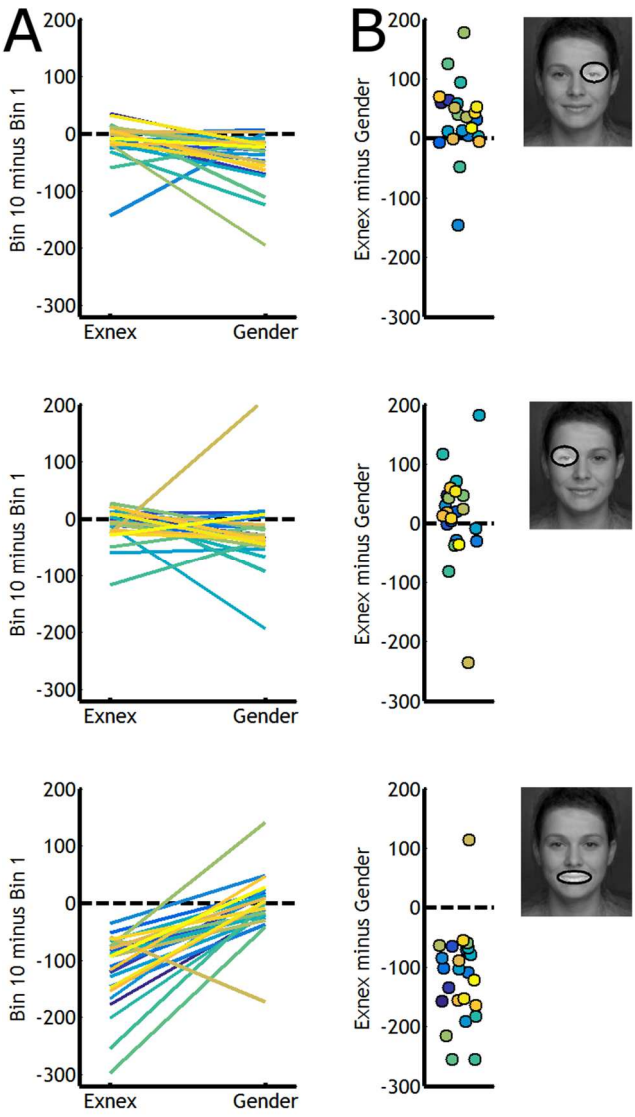


Figure 35: Younger adult reaction time task differences by feature visibility Panel A: The difference in reaction time (ms) between Bin 10 minus Bin 1 in the EXNEX and GENDER task for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in reaction time between Bin 10 minus bin 1 for the EXNEX minus GENDER task for each facial feature of interest.

Comparing differences in accuracy (Figure 36), increased visibility of the right eye resulted in increased gains in accuracy for the **GENDER** than **EXNEX** task. The majority (19/24 participants) of participants demonstrated a larger increase in accuracy with increased right eye visibility in the **GENDER** than **EXNEX** task, suggesting that visibility of the right eye had more of an effect on accuracy in the **GENDER** than **EXNEX** task, and hence more diagnostic value. Similarly, increased visibility on the left eye resulted in increased gains in accuracy for the **GENDER** than **EXNEX** task. The majority (20/24 participants) of participants demonstrated a larger increase in accuracy with increased left eye visibility in the **GENDER** than **EXNEX** task, again demonstrating that visibility of the left eye had more of an effect on accuracy in the **GENDER** than **EXNEX** task, and hence

more diagnostic value. Conversely, all (24/24) participants demonstrated a larger increase in accuracy with increased mouth visibility in the **EXNEX** than **GENDER** task. This suggests that visibility of the mouth had more of an effect on accuracy in the **EXNEX** than **GENDER** task and more diagnostic value.

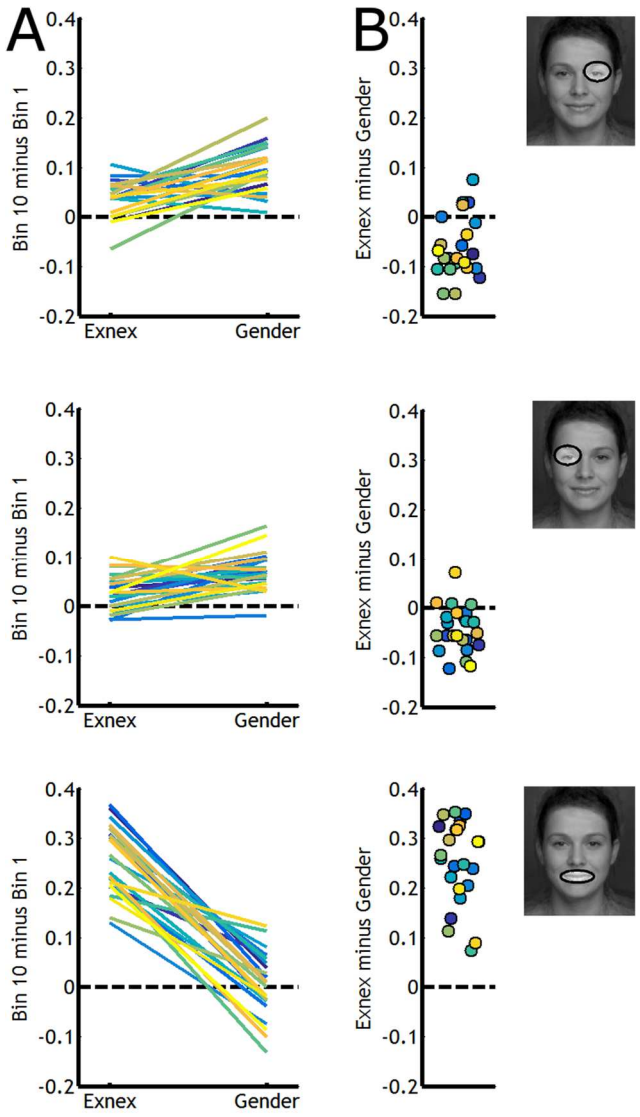


Figure 36: Younger adult accuracy task differences by feature visibility Panel A: The difference in accuracy between Bin 10 minus Bin 1 in the EXNEX and GENDER task for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in accuracy between Bin 10 minus bin 1 for the EXNEX minus GENDER task for each facial feature of interest.

In summary, directly comparing task differences in effects on reaction time of modulating visibility of the left eye, right eye and mouth has demonstrated that increased visibility of the left and right eye led to faster response times and increased accuracy in the **GENDER** than **EXNEX** task. As the eye region is more diagnostic for the **GENDER** than **EXNEX** task, increased visibility of either eye region has more of a reaction time and accuracy advantage for the **GENDER** than

EXNEX task. Conversely, the mouth region is more diagnostic in the **EXNEX** than **GENDER** task. Increasing visibility of the mouth region has more of a reaction time and accuracy advantage for the **EXNEX** than **GENDER** task.

To ensure these differences between tasks were not due to differences in the variability of reaction time and accuracy scores, we conducted a control analysis by normalising our data. For each participant, behavioural comparison and feature we normalised our data between 1 and -1 by dividing the difference between bin 10 minus bin 1, by the difference between bin 1 and bin 10. Normalising our data did not change the pattern of differences (Supplementary 7 - Supplementary 8).

ERP Results

We have seen that, behaviourally, in the **EXNEX** task increased visibility of the mouth resulted in quicker and more accurate responses and that visibility of the left and right eye also lead to quicker and more accurate responses for some participants. In the **GENDER** task, increased visibility of the left eye was associated with quicker responses for most (19/24) participants, and more accurate responses for nearly all (23/24) participants. Additionally, increased visibility of the right eye was also associated with quicker responses for most (21/24) participants, and increased accuracy for all participants. Increased visibility of the mouth results in quicker reaction times and increased accuracy for half of participants.

Next, we will consider ERP results. We predicted that if the N170 eye sensitivity is task-independent, that during both tasks there would be evidence of contralateral eye sensitivity, prior to sensitivity to other facial features, regardless of the extent to which the eyes are task-relevant.

Average ERPs for Practice and Bubble trials

We started by comparing average ERP time courses in the two tasks for all non-bubble i.e. practice trials and bubble trials. Trials with bubbles can be interpreted as similar to completing an object occlusion task where only part of the face is visible. Previous work (Jaworska, 2017) has demonstrated in a face

detection task that on bubble trials younger adults N170 is slightly delayed and larger than on full-face non-bubble trials.

We computed average ERPs for each participant, in each task, for the left and right hemisphere (see Chapter 3: *Average ERPs* for details). We also calculated the median N170 latency and corresponding amplitude for each participant in each task, for the left (LE) and right (RE) hemisphere separately (see Chapter 3: *N170 amplitude and Latency calculations*). We present, for each task the group average mean ERP's for LE and RE (Figure 37) with 95 % confidence intervals, and corresponding median N170 latency and amplitude for practice and bubble trials (Table 4).

		EXNEX		GENDER	
		LE	RE	LE	RE
Practice	LAT	150.1 [149.0, 151.2]	150.0 [149.9, 150.1]	150.0 [149.1, 150.9]	150.0 [149.7, 150.3]
	AMP	-3.7 [-6.3, -1.1]	-3.0 [-5.6, -0.4]	-3.5 [-6.4, -0.7]	-2.6 [-5.4, -0.2]
Bubble	LAT	181.5 [171.4, 191.7]	177.7 [165.0, 190.5]	180.1 [170.3, 189.8]	180.9 [172.4, 189.4]
	AMP	-6.9 [-9.1,-4.8]	-6.0 [-8.6, -3.4]	-7.1 [-9.5, -4.7]	-6.9 [-9.7, -4.1]
Difference	LAT	-27.2 [-36.6, -16.8]	-27.8 [-37.9, -17.6]	-25.5 [-36.0, -15.0]	-29.9 [-40.0, -19.8]
	AMP	3.3 [1.7, 4.9]	4.6 [1.6, 7.6]	3.6 [1.1, 6.1]	5.1 [1.4, 8.8]

Table 4: Younger adults N170 amplitude and latency Median N170 amplitude (AMP) **in ms** and latency (LAT) **in µV** for practice and bubble trials for the left and right hemisphere in the **EXNEX** and **GENDER** task. Difference was calculated as practice minus bubble trials. Square brackets indicate 95 % confidence interval.

Compared to practice (i.e. non-bubble trials), the latency of the N170 in bubble trials was delayed to a similar degree across both tasks and hemispheres (**EXNEX** LE Difference [Practice – Bubble trials] = -27.2 ms [-16.8, -36.6]; **EXNEX** RE Difference [Practice – Bubble trials] = -27.8 ms [-17.6, -37.9]; **GENDER** LE Difference [Practice – Bubble trials] = -25.5 ms [-15.0, -36.0]; **GENDER** RE Difference [Practice – Bubble trials] = -29.9 ms [-19.8, -40.0]). This suggests a general delay in the N170 to bubble stimuli that is not task or hemisphere specific.

There was an increase in the minimum amplitude of the N170 between practice and bubble trials (**EXNEX** LE Difference [Practice – Bubble trials] = 3.3 µV [1.7, 4.9];

EXNEX RE Difference [Practice - Bubble trials] = 4.6 μV [1.6, 7.6]; **GENDER** LE Difference [Practice - Bubble trials] = 3.6 μV [1.1, 6.1]; **GENDER** RE Difference [Practice - Bubble trials] = 5.1 μV [1.4, 8.8]). This may be an effect of occlusion of the image. Individual ERP plots are provided in Supplementary 9 - Supplementary 12.

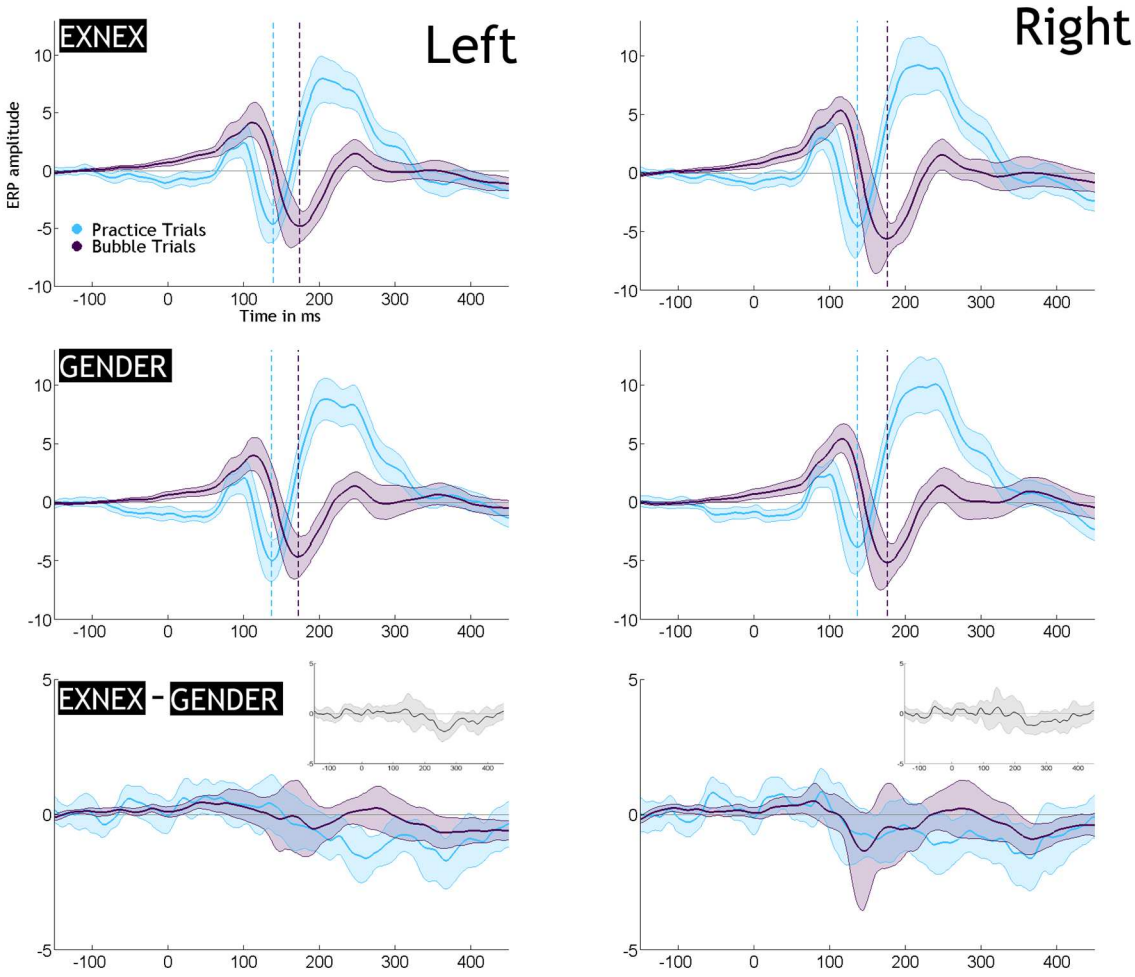


Figure 37: Younger adult average group ERP in Bubble and Non-Bubble Trials Mean bubble and non-bubble trial ERPs for the left and right hemisphere with 95 % confidence intervals around the mean. Vertical lines represent the minimum amplitude peak of the N170 for each task. Bottom panel **EXNEX** minus **GENDER** for bubble and practice trials. Small grey plot shows the pairwise difference of practice minus bubbles trials.

For comparison we also calculated group averaged ERP waveforms using a 20 % trimmed mean. Trimmed means are less affected than outliers and provide a better estimation of the location of the bulk of observations. This may improve the signal to noise ratio of averaged ERP waveforms. We found no difference in the pattern of results when using 20 % trimmed mean (Supplementary 13).

Mutual Information Timecourse

Average ERP's based upon a single electrode of interest as described above are limited, as the above analyses cannot clearly account for single-trial fluctuations around the mean. Mutual information at a given time point is a measure of the relationship between single-trial fluctuations in the signal at that time, and the variation in information from the image.

We began by calculating MI for each participant at each time point in the two tasks between pixels and brain responses for all electrodes. Next, we computed the maximum MI across all pixels at each time point for each electrode. Finally, we computed the maximum MI at each time point across all electrodes in the left and right electrode clusters of interest (see Chapter 3: *Mutual Information Timecourses*). We compared group median MI timecourses in the **EXNEX** and **GENDER** task for the left and right poster electrode clusters, and the difference between tasks (Figure 38). We also calculated the strength of the MI peak for individual participants and the corresponding MI latency, and the group median of the MI peak and latency (Table 5).

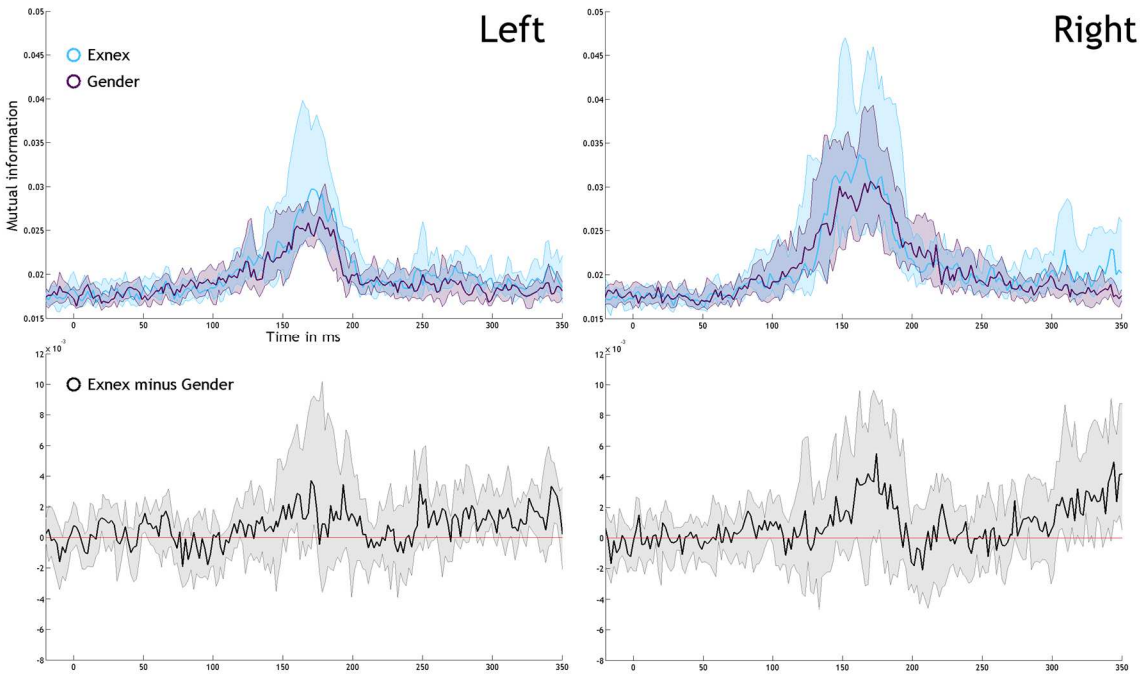


Figure 38: Younger adult MI Timecourses Top row: Group median of individual maximum MI timecourses in the **EXNEX** and **GENDER** task for the LE and RE cluster. Bottom row: Difference in MI timecourses (**EXNEX** minus **GENDER**).

Figure 38 suggests that maximum MI peaked within the time window of the N170 during both tasks. However, due to large degrees of individual variability in MI timecourses (Figure 39), MI peak latencies varied widely between participants, with some individual participants maximum MI peaks occurring outside of the time window of the N170. This could reflect a rebound effect, where the latter part of the N170 codes the presence of the contralateral eye (Ince et al., 2016). Peaks in mutual information for some participants around the time window of the P300, particularly in the **EXNEX** task, could reflect continued representation of diagnostic features over this time window (Van Rijsbergen & Schyns, 2009).

	EXNEX		GENDER	
	LE	**RE**	**LE**	**RE**
LAT	185 [141.9, 228.1]	242 [168.3, 320.7]	164 [132.9, 195.1]	145.6 [123.9, 167.2]
MI	0.04 [0.04, 0.05]	0.04 [0.03, 0.05]	0.03 [0.03, 0.04]	0.03 [0.03, 0.04]

Table 5: Younger adult group average peak MI and latency Peak MI of the group median of individual maximum MI values and corresponding median latency in ms. Square brackets indicate 95 % percentile bootstrap confidence interval around the medians with 1000 samples.

Harrell-Davis estimates suggested that MI peaked later in the **EXNEX** than the **GENDER** task (LE Latency $_{[EXNEX - GENDER]}$ = 33.7 ms [-35.8, 103.2]; RE Latency $_{[EXNEX - GENDER]}$ = 119.5 ms [-5.4, 244.5]). Whilst MI peaked earlier in the LE than RE cluster in the **EXNEX** task, the reverse was evidenced in the **GENDER** task (**EXNEX** difference $_{[LE\ minus\ RE]}$ = -27.8 ms [-118.9, 63.2]; **GENDER** difference $_{[LE\ minus\ RE]}$ = 11.5 ms [-65.9, 88.9]).

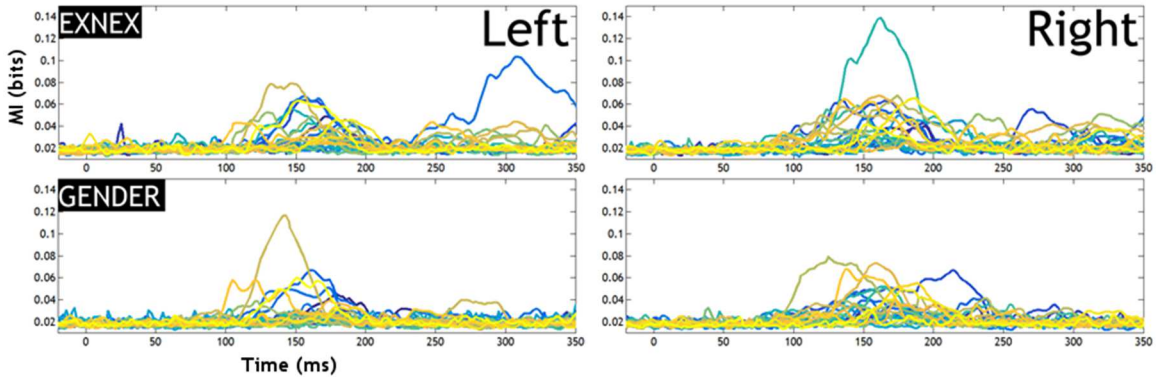


Figure 39: Younger adult individual Maximum MI Timecourses For each task and electrode, individual (N = 24) participants maximum MI time course are shown. Each line represents an individual participant.

Mean MI time courses have demonstrated that predominantly maximum MI peaks within a broad time window around the N170. As we varied the information content available on each trial, averaging MI time courses across all trials is not meaningful as it does not demonstrate how single trial variability in MI time courses are modulated by the information content available on each trial. Next, we took an assumption-free approach to look to understand what facial features specifically modulate this single-trial variability.

Brain Classification Images

To visualise what information is being processed during the entire time course at each electrode cluster, for each participant (Supplementary 15) we calculated for each electrode the max MI value at each pixel across all time points between -300 ms pre-stimulus onset and 600 ms post stimulus onset. We then calculated the maximum MI value for each pixel across all electrodes in the LE and RE clusters separately. We then took the median maximum MI value for each pixel across all participants for each of the two tasks separately, and computed the resulting classification image (Figure 40).

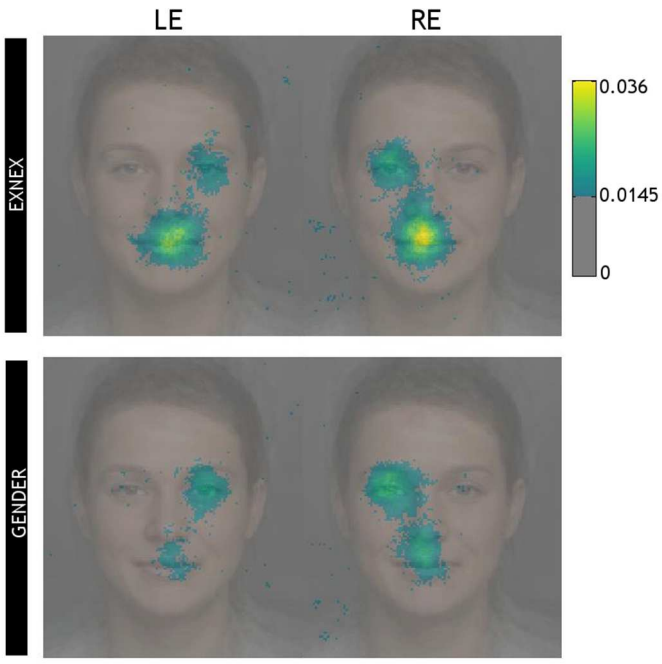


Figure 40: Younger adult group Max Mutual Information Classification Image For each electrode cluster (LE, RE) the median across participants of the maximum MI value for each pixel across the entire time course -300 to 600ms is shown.

In both the **EXNEX** and **GENDER** task, the LE cluster shows a focal hotspot over the contralateral right eye area and the RE cluster over the contralateral left

eye area, consistent with previous results. In the **EXNEX** task, both the LE and RE clusters show strong hotspots over the mouth region, stronger than that in the **GENDER** task where the mouth region is less diagnostic for task performance.

In summary we have shown that group average MI timecourses were broadly similar in the two tasks, though average MI was stronger in the **EXNEX** than the **GENDER** task. When looking at MI values across hemispheres, MI was stronger for the right than left hemisphere in both tasks. Our classification images have shown that, for both tasks, the left and right hemisphere ERPs were modulated by the presence of the contralateral eye and mouth. We have shown that mouth modulation was stronger in the **EXNEX** than **GENDER** task, whilst eye modulation was slightly stronger for the **GENDER** than **EXNEX** task.

Timing of Feature Sensitivity

In both the **EXNEX** and **GENDER** tasks we have highlighted peaks in mutual information between ERPs and pixels around the eye and mouth regions. Next, we will quantify the timing of feature sensitivity to the left eye, right eye and mouth in the two tasks. Given previous research, we would expect to see processing of the contralateral eye on each hemisphere, before processing of other task-relevant facial features, such as the mouth in the **EXNEX** task (M. L. Smith et al., 2004).

First, we examined group MI classification images over time. For each individual, we calculated the maximum MI value across all electrodes within the left and right hemisphere clusters, for every pixel at each time point. We then calculated group median MI value at each time point, presented in Figure 41.

In both tasks, contralateral eye sensitivity and mouth sensitivity are apparent. In both tasks, contralateral eye and mouth sensitivity peaked within the time window of the N170. In the **GENDER** task there was evidence of electrodes on the left hemisphere initially coding the contralateral eye, with later sensitivity to the ipsilateral eye. This is consistent with previous work suggesting that whilst the N170 initially reflects coding of the contralateral eye, this is followed by sensitivity to cross-hemispheric transfer of visual features (Ince et al., 2016).

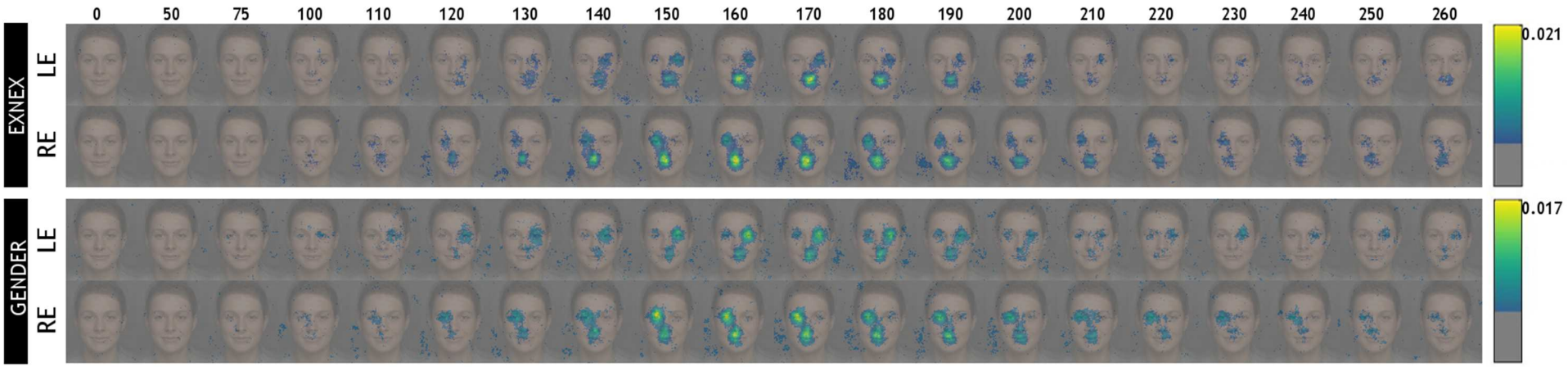


Figure 41: Younger adult mutual information EEG classification images time course Group median of individual maximum MI values at each pixel and time point (in milliseconds, see small numbers on the top row) in each cluster of electrodes of interest in two tasks

Modulation of MI Timecourse by Feature Visibility

Next, we quantified the timing of feature sensitivity seen in our MI classification images by examining the MI time courses between feature of interest visibility and ERPs. We selected 1 left and right lateral posterior electrode and calculated for each feature the time course of the MI about feature visibility – MI(**feature**, [**ERP**, **grad**]) (see Chapter 3: *Feature of Interest Analysis*). We present the mean MI time courses for each electrode, task and feature in Figure 42. We also calculated the group median of individual MI peaks and corresponding latencies.

First, we will address left posterior lateral activity (Figure 42). Left posterior lateral activity was modulated by the contralateral right eye in both the **GENDER** and **EXNEX** task, with MI peaking earlier in the **EXNEX** than **GENDER** task (**EXNEX** = 150.0 ms [103.5, 198.5]; **GENDER** = 175.1 ms [141.3, 208.8]; **Difference** $_{[\text{EXNEX minus GENDER}]}$ = -37.4 ms [-201.9, 127.0]).

MI however was stronger for the **GENDER** than the **EXNEX** task (Difference $_{[\text{EXNEX minus GENDER}]}$ = -0.004 bits [-0.001, -0.007]). This is consistent with our behavioural results showing that the eye region was more diagnostic for the **GENDER** than the **EXNEX** task. There was no evidence of modulation of the left electrode by the ipsilateral left eye.

Left electrode activity was modulated by the mouth in both tasks (**EXNEX** = 0.03 [0.02, 0.04]; **GENDER** = 0.02 [0.01, 0.02]), though MI was stronger in the **EXNEX** than the **GENDER** task (Difference $_{[\text{EXNEX minus GENDER}]}$ = 0.01 bits [0.01, 0.02]). This is consistent with our behavioural results showing that the mouth region was more diagnostic for the **EXNEX** than **GENDER** task. Median MI peaked earlier in the **GENDER** than **EXNEX** task (**EXNEX** = 230.1 ms [119.6, 341.4]; **GENDER** = 173.8 ms [122.9, 224.6]; **Difference** $_{[\text{EXNEX minus GENDER}]}$ = 40.4 ms [-64.4, 145.3]), though this may be due to larger MI values in the later part of the time course in the **EXNEX** task.

Right posterior lateral activity (Figure 42) was modulated by the contralateral left eye in both the **GENDER** and **EXNEX** task (EXNEX = 0.02 [0.01, 0.02]; GENDER = 0.02 [0.02, 0.03]; Difference $_{[\text{EXNEX minus GENDER}]}$ = -0.002 bits [0.0, 0.006]). Peak latency was also similar between tasks (**EXNEX** = 166.8 ms [151.0, 182.5];

GENDER = 169.1 ms [149.8, 188.3]; Difference $_{[EXNEX\ minus\ GENDER]}$ = 3.4 ms [-18.6, 25.5]). There was no evidence of modulation of the left electrode by the ipsilateral left eye.

Right electrode activity was modulated by the mouth in both tasks (**EXNEX** = 0.03 [0.02, 0.04]; **GENDER** = 0.02 [0.02, 0.03]; Difference $_{[EXNEX\ minus\ GENDER]}$ = 0.01 [0.01, 0.02]), though peaked earlier in the **EXNEX** than the **GENDER** task (EXNEX = 180.3 ms [146.5, 214.1]; GENDER = 190.6 ms [123.2, 157.9]; Difference $_{[EXNEX\ minus\ GENDER]}$ = -7.3 ms [-61.2, 46.5]).

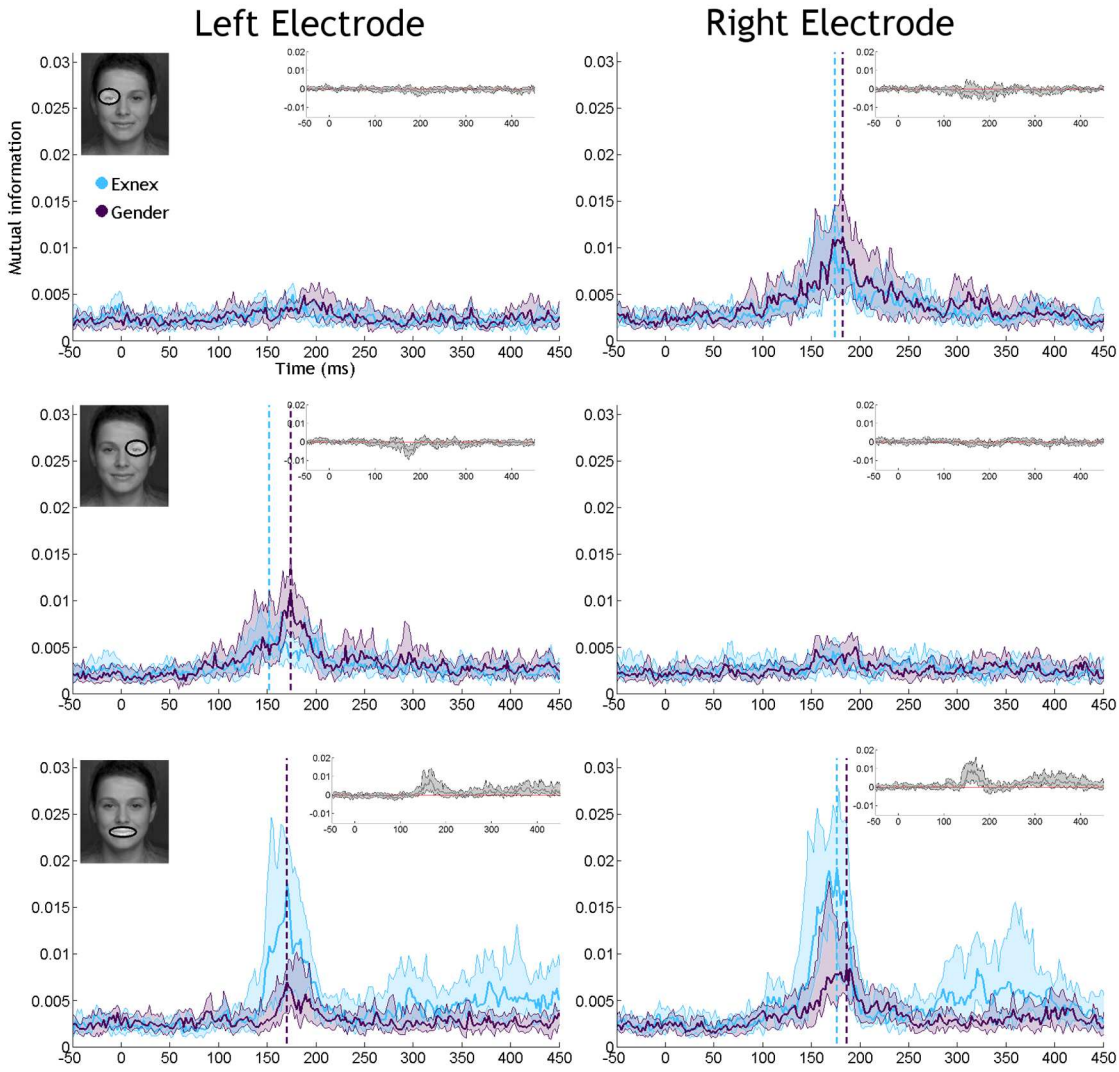


Figure 42: Younger adult mutual Information Timecourse by facial feature Mutual information time course for the right and left posterior lateral electrode towards the left eye (top panel), right eye (middle panel) and mouth (bottom panel) in the **EXNEX** and **GENDER** task. Shaded area corresponds to a 95 % confidence interval calculated by a percentile bootstrap with 1000 samples around the median. Vertical lines indicate the peak of the MI time course.

For each participant and task we also analysed MI time courses to the midline electrode with the maximum sum MI between 120 – 220 ms (see Chapter 3;

Feature of Interest Analysis). Midline electrodes have previously been shown to be sensitive to task-relevant facial features (Schyns, Thut, & Gross, 2011). In particular, in a face versus noise detection task midline electrodes displayed sensitivity to both eyes as well as the nose and mouth areas (Rousselet et al., 2014). The authors suggested that sensitivity to the nose and mouth area peaked at least 20 ms prior to posterior lateral eye sensitivity peaks, and was also present in noise trials, suggesting that midline electrode sensitivity may be a possible low-level effect, rather than an explicit feature integration process. Results from our midline electrode analysis suggested a weaker sensitivity to both the eyes and the mouth, which peaked later than contralateral posterior electrode activity.

Next, we compared MI time courses between features to determine if contralateral eye sensitivity preceded sensitivity to the mouth (Figure 43). For each comparison, we calculated the difference between the MI peak and its latency for eye compared to mouth timecourses within the time window of the N170 (~120 – 220 ms).

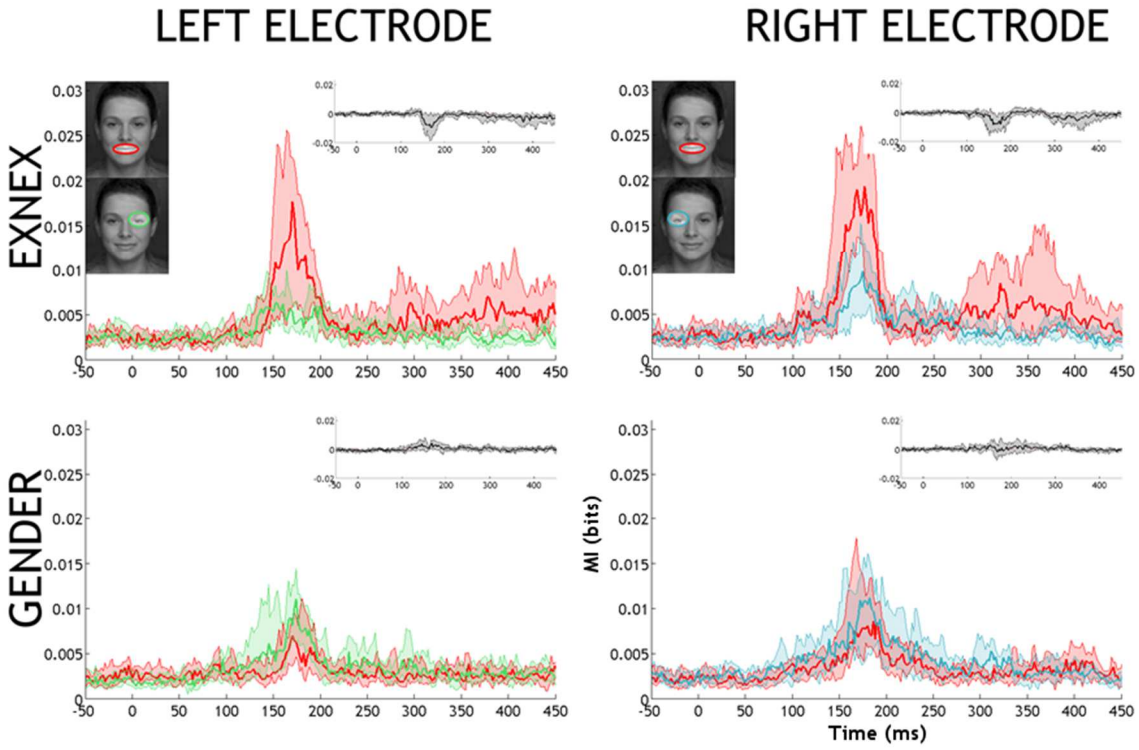


Figure 43: Younger adult Mouth-Eye MI time course differences MI time courses to the contralateral eye and mouth for the left and right electrode in the EXNEX and GENDER task. Grey plots display the pairwise difference in time courses. Shaded area corresponds to the 95 % bootstrap CI around the median.

Right posterior lateral MI for the contralateral left eye peaked earlier than peak MI to the mouth in both the **EXNEX** (Difference $_{[Eye\ minus\ Mouth]}$ = -7.1 ms [-20.3, 6.0]) and **GENDER** (Difference $_{[Eye\ minus\ Mouth]}$ = -5.0 ms [-15.4, 5.4]) task.

Left posterior lateral MI for the contralateral right eye peaked later than peak MI to the mouth in the **EXNEX** (Difference $_{[Eye\ minus\ Mouth]}$ = 9.5 ms [-4.5, 23.6]) and the **GENDER** (Difference $_{[Eye\ minus\ Mouth]}$ = 7.1 ms [-2.3, 16.5]) task.

We found large individual variation in the direction of the difference in MI peaks to the mouth and contralateral eye region that were not explained by task order effects. We present scatter plots of the difference in individual peak MI to the eye and mouth within the time window of the N170 in Supplementary 14.

However, coding of the eye starts well before the peak of the N170 (Rousselet, Ince, van Rijsbergen, & Schyns, 2014), so we should look at MI onset.

MI Onset Analysis

Next, we aimed to calculate the onset of MI to the mouth and eyes. For each participant in each task, we calculated the maximum MI value at each time point combined across both left *and* right posterior lateral electrodes to the mouth and *both* the eyes. We did this as some participants had stronger MI values to the right rather than left eye (Supplementary 15) and using this method avoided assumptions about which eye would be processed first. Using these timecourses, we used a Multivariate Adaptive Regression Splines (MARS) method (Friedman, 1991) to calculate MI onsets for the eyes and mouth (see Chapter 3; *Mutual Information Onset Analysis*).

Individual results for MI onsets in the **EXNEX** task are displayed in Supplementary 16 - Supplementary 17. Four participants' demonstrated timecourses where no clear peak was discernible to the eyes and/or mouth and were removed from analysis. For each participant we calculated the difference in the estimate of MI onset to the eyes minus the mouth. Results are presented in Figure 44.

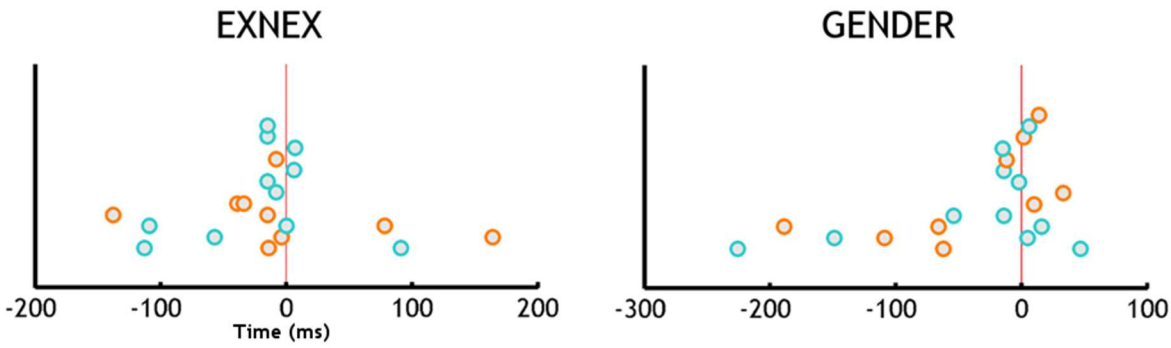


Figure 44: Younger adult difference in MI Onset For each participant the difference in estimated onset times (eyes minus mouth) in the **EXNEX** (left) and **GENDER** (right) task. Orange circles completed the **EXNEX** task first. Blue circles completed the **GENDER** task first.

There was a negatively skewed distribution in differences in MI onset of the eyes compared to the mouth. 14/20 participants has an earlier estimates onset to the eyes compared to the mouth. Group estimates of the median difference were -13 ms [-28, 3]. Whilst some participants demonstrated a preference to processing the eyes first, this was not evident for all participants.

Individual results for MI onsets in the **GENDER** task are displayed in Supplementary 18 - Supplementary 19. Four participants' demonstrated timecourses where no clear peak was discernible to the eyes and/or mouth and were removed from analysis. For each participant we calculated the difference in the estimate of MI onset to the eyes minus the mouth.

There was a negatively skewed distribution in differences in MI onset of the eyes compared to the mouth. 12/20 participants has an earlier estimates onset to the eyes compared to the mouth. Group estimates of the median difference were -12 ms [-46, 21]. Whilst some participants demonstrated a preference to processing the eyes first, this was not evident for all participants.

Reverse Analysis: EEG Results

We have examined MI between pixels and ERPs. We have seen that visibility of the contralateral eye is associated with modulation of left and right hemisphere activity in both the **EXNEX** and **GENDER** task. Similarly, visibility of the ipsilateral eye was associated with weak modulation of left and right hemisphere activity in both tasks. Whilst visibility of the mouth region was associated with modulation of left and right hemisphere activity in the **GENDER** task, this was stronger for the **EXNEX** task.

We have used MI to examine the association between facial features and ERP modulation. However, MI does not reveal how the shapes of the ERPs are affected by our features of interest. To establish how visibility of the left eye, right eye and mouth affect the shape of our ERPs, we next focus on reverse analysis – isolating the specific EEG modulations associated with the presence of specific facial features.

Using the feature of interest masks (see Chapter 3: *Feature of Interest Analysis*) we calculated on a trial-by-trial basis the visibility of each feature (left eye, right eye, mouth), obtained as a scalar value of the sum of pixel visibility within the ellipse of each feature mask. We then split these visibility values into ten equally populated bins ranging from the lowest (bin 1) to the highest (bin 10) visibility values. We then sorted single trial ERP's into 10 bins, based on the feature visibility in each trial. We present group mean ERP's for each level of feature visibility for the left (Figure 45) and right (Figure 46) electrode.

Left hemisphere ERPs (Figure 45) were modulated by the presence of the contralateral (right) eye in both tasks. Modulation of the N170 component was stronger in the **GENDER** than **EXNEX** task, with increased visibility of the contralateral (right) eye increasing amplitude and decreasing latency on the N170 (Figure 45). Increasing visibility of the ipsilateral (left) eye similarly modulated the N170 component of the ERP in both tasks. There was stronger modulation of the N170 in the **GENDER** than the **EXNEX** task, with increased visibility of the ipsilateral (left) eye increasing amplitude and increasing latency on the N170. This increased prominence in modulation of the N170 in the **GENDER** than **EXNEX** task with increased visibility of the eye region indicates that as well as the eye region being diagnostic for behavioural modulation, these features also modulated brain activity. Increasing visibility of the mouth region modulates the N170 component of the ERP in both the **GENDER** and **EXNEX** task. In the EXNEX task, where the mouth is behaviourally task relevant, increased mouth visibility increases N170 amplitude, whilst decreasing N170 latency (Figure 45). In the **GENDER** task, increased mouth visibility increases N170 amplitude, though to a lesser extent than in the **EXNEX** task (Figure 45). This reflects a disparity between behaviour, where the mouth is not task relevant and not related to increased accuracy or speed of performance, and brain activity, where mouth visibility modulates brain activity.

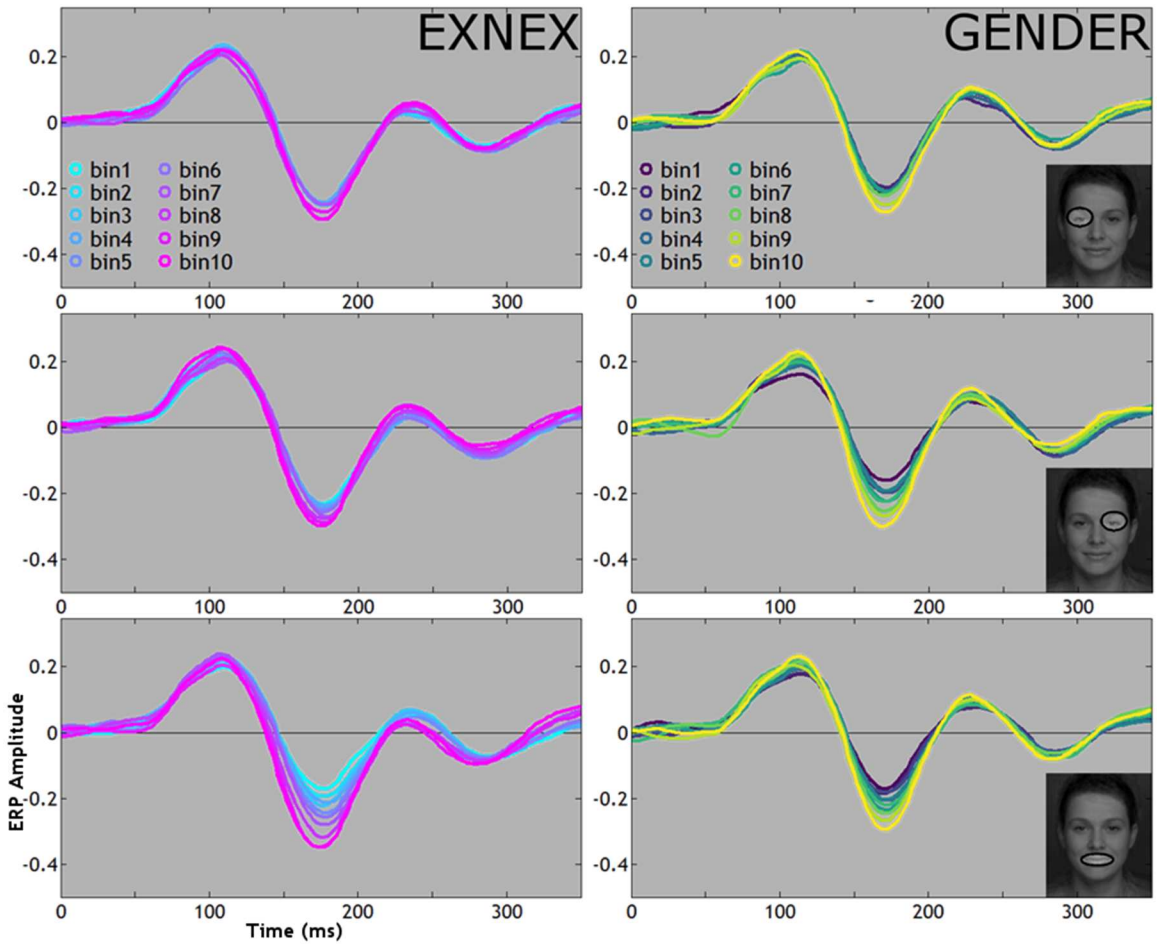


Figure 45: Younger adult left Electrode Binned ERPs by feature visibility Each column represents one task. Column A (left) shows results for the **EXNEX** task, column B (right) shows results for the **GENDER** task. Group mean ERP's are shown for the posterior left electrode. ERPs are binned into 10 levels of visibility ranging from bin 1 (least visible) to bin 10 (most visible) for the left ipsilateral eye (top row), the right contralateral eye (middle row) and the mouth (bottom row).

Right hemisphere ERPs (Figure 46) were modulated by the presence of the contralateral (left) eye in both tasks. Increased visibility of the contralateral (left) eye increased N170 amplitude and decreased N170 latency in both the **EXNEX** and **GENDER** task (Figure 46). Increasing visibility of the ipsilateral (right) eye similarly modulated the N170 component of the ERP in both tasks, though modulation was weaker. Increasingly visibility of the ipsilateral (right) eye increased N170 amplitudes in both tasks (Figure 46). Increasing visibility of the mouth region modulated N170 ERPs in both tasks. As mouth visibility increased, amplitude of the N170 increased in both tasks (Figure 46). Modulation of the N170 by mouth visibility was greater for activity on the right than left hemisphere (Figure 46).

We have demonstrated modulation of single-trial ERPs due to visibility of specific facial features. Left and right hemisphere N170s are modulated more by

the contralateral than ipsilateral eye in both tasks. Contralateral eye modulation of the N170 was stronger for the **GENDER** than **EXNEX** task. Similarly both left and right hemisphere N170s were modulated by the mouth region, though this modulation was stronger in the **EXNEX** than **GENDER** task. Modulation of N170s in both task were strongest when manipulating behavioural diagnostic facial features i.e. the eyes in the **GENDER** task, and the mouth region in the **EXNEX** task.

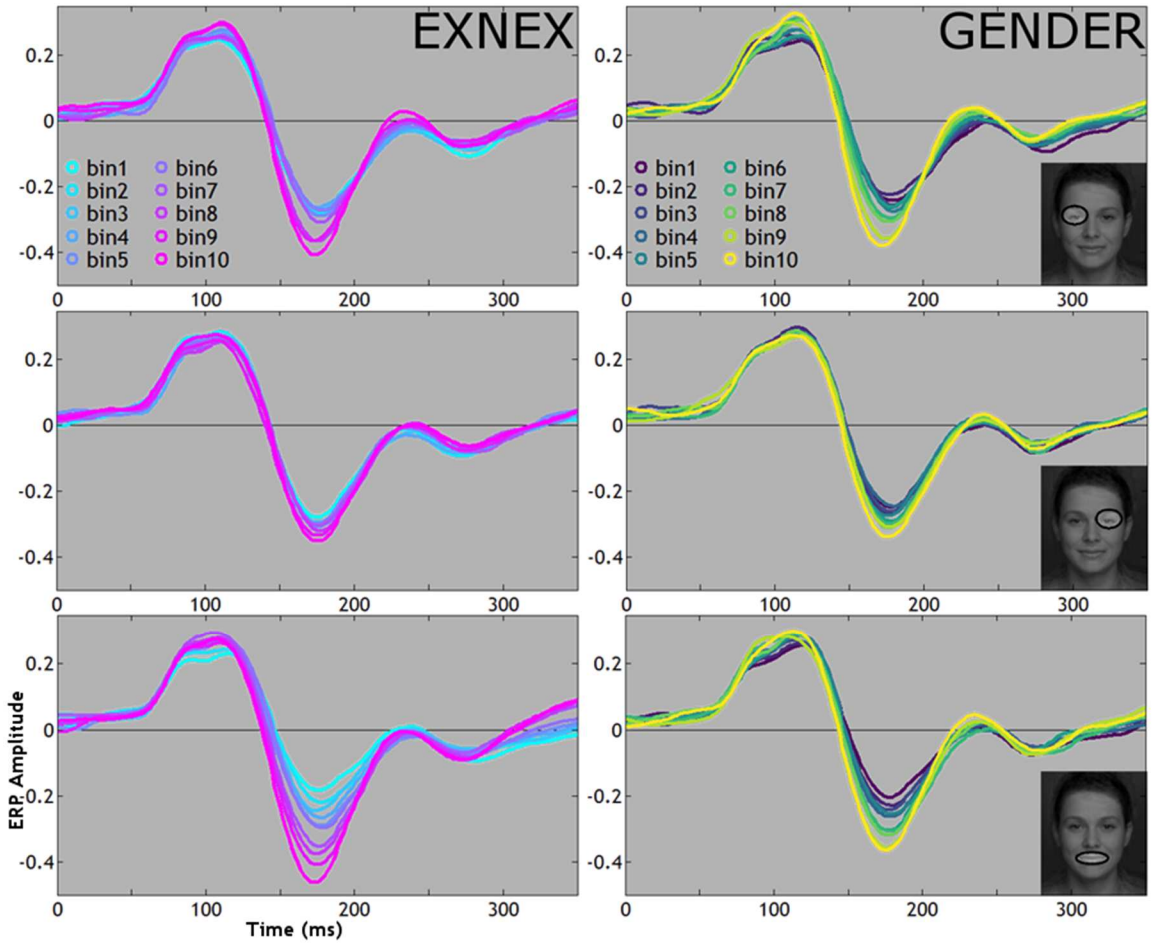


Figure 46: Younger adult right electrode ERPs binned by feature visibility Each column represents one task. Column A (left) shows results for the **EXNEX** task, column B (right) shows results for the **GENDER** task. Group mean ERPs are shown for the posterior right electrode. ERPs are binned into 10 levels of visibility ranging from bin 1 (least visible) to bin 10 (most visible) for the left ipsilateral eye (top row), the right contralateral eye (middle row) and the mouth (bottom row).

Our N170 modulations also demonstrate differences in the relative linearity of ERP responses. With a linear response, we would expect a linear relationship between increasing visibility of a feature and modulation of the N170. This linearity of response can be seen for the right hemisphere modulation of the N170 to increasing mouth visibility. However, right hemisphere N170 responses to the contralateral (left) eye appear more categorical. In the **EXNEX** task, N170

responses among the lowest eye visibility ERPs have minimal variance, with the 3 highest visibility bands displaying a qualitatively difference N170 response.

N170 Latency and Amplitude Differences

We have examined whole ERP modulations in Figure 45 - Figure 46. We have demonstrated that the N170 component of our ERPs is modulated by the relative visibility of specific facial features. Next, we wanted to quantify this modulation of the N170, specifically to compare how the amplitude and latency of the N170 is modulated by the presence of specific facial features between tasks.

We looked to quantify the difference in N170 amplitude and latency between ERPs when facial features were most visible (bin 10) compared to least visible (bin 1). First we calculated for each participant the N170 amplitude and latency at each bin, for each feature in each task, for both posterior lateral electrodes. We then calculated the difference between bin 10 (most visible) minus bin 1 (least visible) for N170 latency in the time window ~150–240 ms post stimulus onset for each task, feature and lateral posterior electrode. Amplitude differences were calculated as a proportion of bin1 (least visible) as a percentage, such that an amplitude difference of 50 % means that the amplitude of bin 10 (most visible) amplitudes were 150 % the size of the amplitudes in bin 1. For example, a point at the x, y coordinates 0, 100 would reflect no latency difference of the N170 between bin 1 (least visible) and bin 10 (most visible) trials but would reflect a amplitude difference of 100 % meaning that the amplitudes of bin 10 (most visible) amplitudes were 200 % the size of the amplitudes in bin 1. In comparison, a point at the x, y coordinates -15, 0 would reflect a latency difference of -15 ms between bin 10 and bin 1 trials (i.e. the peak of the N170 in bin 10 trials was earlier) but would reflect no difference in amplitude.

We present median difference across participants in Figure 47 and Table 6, and individual differences in Figure 48.

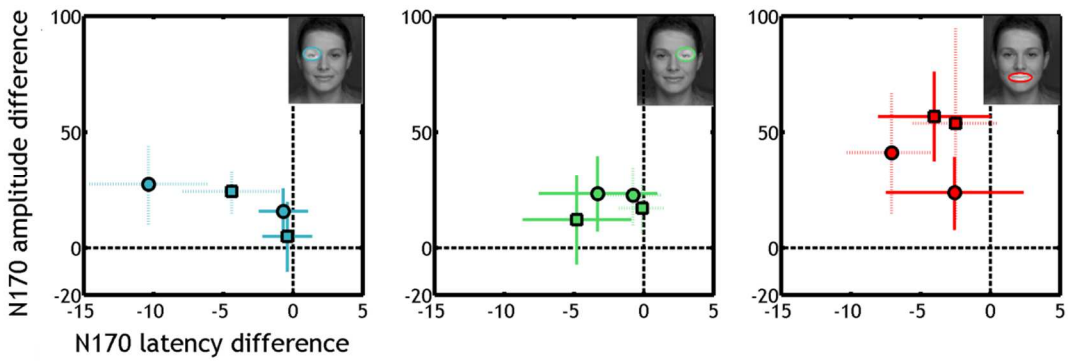


Figure 47: Younger adult N170 Amplitude and Latency Differences by Feature Results are presented for three features of interest: the left eye (blue), right eye (green) and mouth (red). Median N170 latency differences were calculated between bin 10 (most visible) minus bin 1(least visible) and are presented in milliseconds. Median N170 amplitude differences are calculated as a percentage of bin 1, such that an amplitude difference of 50% means that the amplitude of bin 10 was 150% the size of amplitudes in bin1. Vertical and horizontal lines correspond to 95% confidence intervals. **EXNEX** results are plotted with squares, **GENDER** results are plotted with circles. Solid lines are the left electrode; dashed lines are the right electrode.

In both the **GENDER** and **EXNEX** task, increased visibility of the mouth increased N170 amplitudes at both the left and right posterior lateral electrodes. Increased mouth visibility was associated with shorter N170 latencies in the **EXNEX** task at left posterior lateral electrodes, and in the **GENDER** task at right posterior lateral electrodes. This highlights a disparity, with the visibility of the mouth modulating N170 latencies and amplitudes in the **GENDER** task, despite the mouth not being task-relevant in terms of modulating behavioural responses.

Increased visibility of the ipsilateral eye region was associated with increased amplitudes in the **GENDER** task for both the left and right posterior lateral electrodes. In the **EXNEX** task, this was only the case for the right electrode. In both the **EXNEX** and **GENDER** task there was minimal differences in latency modulation. Increased visibility of the contralateral eye region was associated with decreased latencies on the left posterior lateral electrodes in both the **GENDER** and **EXNEX** tasks, but increased amplitude in the **GENDER** task only. At the right posterior electrode, increased visibility of the contralateral (left) eye was associated with increased latency and amplitude of the N170 in both the **EXNEX** and **GENDER** task.

		EXNEX		GENDER	
		LE	**RE**	**LE**	**RE**
LEFT EYE	**LAT**	-0.4 [-2.2, 1.4]	-4.4 [-7.8, -1.0]	-0.7 [-2.4, 1.0]	-10.4 [-14.8, -5.9]
	AMP	4.9 [-10.6, 20.4]	24.3 [15.5, 33.2]	15.8 [5.7, 25.9]	27.3 [9.8, 44.8]
RIGHT EYE	**LAT**	-4.8 [-8.9, -0.8]	-0.1 [-1.8, 1.5]	-3.3 [-7.8, 1.2]	-0.8 [-2.6, 1.1]
	AMP	12.3 [-6.2, 30.7]	17.1 [6.9, 27.3	23.5 [8.3, 38.8]	22.7 [9.9, 35.5]
MOUTH	**LAT**	-4.0 [-8.0, -0.1]	-2.5 [-5.6, 0.6]	-2.6 [-7.3, 2.2]	-7.1 [-10.4, -3.8]
	AMP	56.9 [37.4, 76.5]	53.8 [11.2, 96.5]	23.8 [8.0, 39.6]	41.4 [14.5, 68.2]

Table 6: Younger adult N170 amplitude and latency differences by facial feature Median N170 latency and amplitude differences for the right eye, left eye and mouth for the left and right hemisphere in each task. N170 was measured in the time window between ~150-240ms post stimulus onset. Median N170 latency differences were calculated between bin 10 (most visible) minus bin 1(least visible) and are presented in milliseconds. Median N170 amplitude differences are calculated as a percentage of bin 1, such that an amplitude difference of 50% means that the amplitude of bin 10 was 150% the size of amplitudes in bin1.

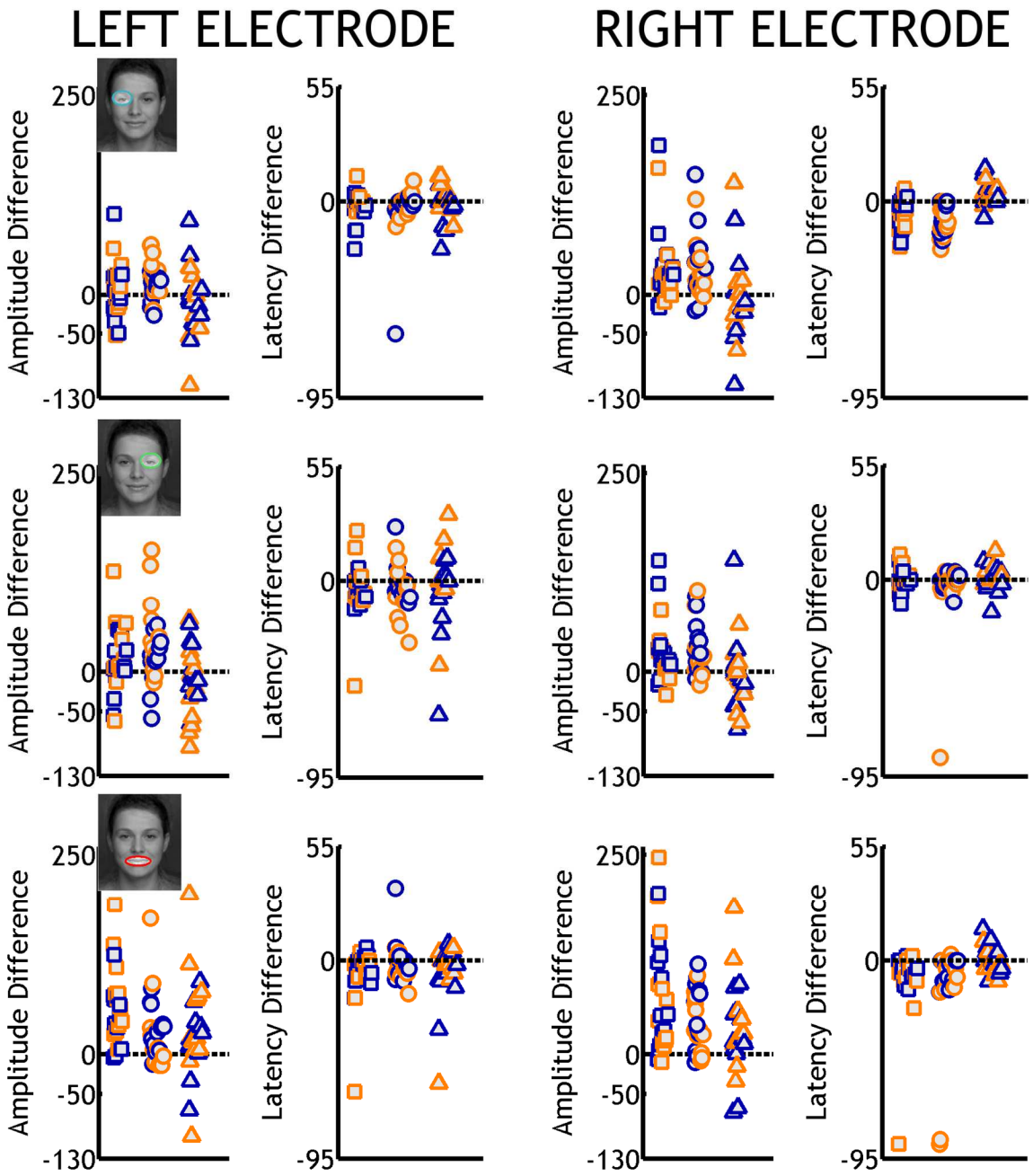


Figure 48: Younger adult N170 individual amplitude and latency differences Individual differences in amplitude and latency between bin 10 minus bin 1. Median N170 latency differences were calculated between bin 10 (most visible) minus bin 1(least visible) and are presented in milliseconds. Median N170 amplitude differences are calculated as a percentage of bin 1, such that an amplitude difference of 50% means that the amplitude of bin 10 was 150% the size of amplitudes in bin1. Squares represent **EXNEX** whilst circles represent **GENDER**. Triangles represent the difference between **EXNEX** minus **GENDER**. Blue symbols represent participants completing the **EXNEX** task first. Orange triangles represent participants completing the **GENDER** task first.

Discussion

To understand whether face processing relies upon a bottom-up, data driven model in which the presence of the contralateral eye is the first processing step succeeding all other task-relevant feature integration, we need to formally quantify contralateral eye sensitivity in face processing tasks during which the eyes are highly or lowly diagnostic.

In the current study using a parametric design we examined what facial information was processed under varying task demands, and linked facial stimulus space to behaviour and brain responses. We found that increased visibility of the mouth led to quicker and more accurate responses in the **EXNEX** task for all participants, though even when there was no visibility of the mouth participants were still able to perform at above chance accuracy levels, suggesting that whilst the mouth is the most diagnostic feature, in its absence participants can use other features well for categorisation. Whilst some of our participants were able to use the mouth as a diagnostic feature for **GENDER** categorisation, predominantly participants' behavioural responses were modulated by visibility of the eye regions.

Increased visibility of the left eye (from the viewers' perspective) increased reaction times and accuracy for some participants in the **EXNEX** task. In the **GENDER** task, increased left eye visibility led to quicker responses for most participants, and more accurate responses for nearly all participants, though MI was only significant for a fraction of participants (3/24) suggesting large single-trial variability. Previous studies have suggested that the presence of the left eye is crucial for the correct detection of faces (Jaworska, 2017; Rousselet et al., 2014). Here we have replicated these results in a more complex socially relevant gender and expressiveness task. Our results suggest that processing of the left eye is not limited to face detection tasks but is apparent for face processing more generally.

Whilst our results were consistent to previous results indicating that the mouth was the most highly diagnostic feature for resolving the **EXNEX** task we did not replicate the strong diagnosticity to the eye region in the **GENDER** task (Schyns et al., 2002). Instead, our results' suggested an idiosyncratic preference for the

eyes that varied in strength across participants. Our results did not indicate a strong dependence on any single feature for resolving the **GENDER** task. Our results may have differed due to using naturalistic coloured images, where there was more information from e.g. the hair and skin pigmentation to resolve the task. However, no participants' classification images highlighted strong sensitivity to the hair region, suggesting that participants predominantly relied on facial feature information. Unlike previous studies which kept performance at 75 % correct responses (Schyns et al., 2002), we did not set a performance threshold for our tasks, and our participants achieved on average 82 % accuracy in the **GENDER** task. Given this higher level of accuracy in our results, we may not have evidenced modulations of accuracy by face features in the **GENDER** task. It may be that if we reduced the accuracy of participants that we may have observed a stronger association between presence of the eye (or other features) and correct responses.

Another potential explanation is that previous studies (Schyns et al., 2002) may have indicated a stronger sensitivity to the eye in the **GENDER** task due to the models in the stimulus set wearing eye makeup. This may have changed the local contrast in the eye region by darkening the eye line and lashes making the eye region more salient in differentiating women and men in the stimulus set (Schyns & Oliva, 1999). This could be readily tested by controlling for the presence or absence of eye makeup in an otherwise identical stimulus set and observing whether the presence of eye makeup results in increased diagnosticity of the eye region.

Having established that participants use different information from a face to perform an **EXNEX** and **GENDER** categorisation task we quantified what information was coded in the brain. We found that single-trial ERPs are modulated by the presence of the contralateral eye region and mouth in both the **EXNEX** and **GENDER** task. We have shown that mouth modulation was stronger in the **EXNEX** than **GENDER** task, whilst eye modulation was slightly stronger for the **GENDER** than **EXNEX** task. This highlights a disparity between the task-relevant information for behaviour and that information coded in the brain.

During the time window of the N170, we highlighted the encoding of both the contralateral eye and mouth regions in both the **EXNEX** and **GENDER** task. This is in stark contrast to Smith et al., (2004) results suggesting that the N170 was modulated by the contralateral eye only. We have shown that in both the **GENDER** and **EXNEX** task, increased visibility of the mouth increased N170 amplitudes at both the left and right posterior lateral electrodes. Increased mouth visibility was also associated with shorter N170 latencies in the **EXNEX** task at left posterior lateral electrodes, and in the **GENDER** task at right posterior lateral electrodes. This suggests that the N170 is not only sensitive to the contralateral eye as suggested by Smith et al. (2004), but is also modulated by other facial features, namely the mouth. We also found evidence of sensitivity to the ipsilateral eye modulating N170 amplitudes during both tasks, particularly for the left ipsilateral electrodes. This is in-keeping with a recent study suggesting that the N170 reflects the coding of the contralateral eye followed by the transfer of communication about the ipsilateral eye (Ince et al., 2016).

We attempted to quantify the timing of feature sensitivity to the eyes and mouth to elucidate whether eye coding began before coding of the mouth. By estimating the onset of MI to the eyes and mouth regions, we found sensitivity to either eye preceded sensitivity to the mouth for 14/20 participants in the **EXNEX** task and 12/20 participants in the **GENDER** task. We evidence negatively skewed distributions indicating a large amount of idiosyncrasy in relative timing of MI onset. As a result these results should be treated with caution, as estimates of onsets on such a small scale may be misleading. Similarly, the large individual differences in onsets should be addressed in future studies.

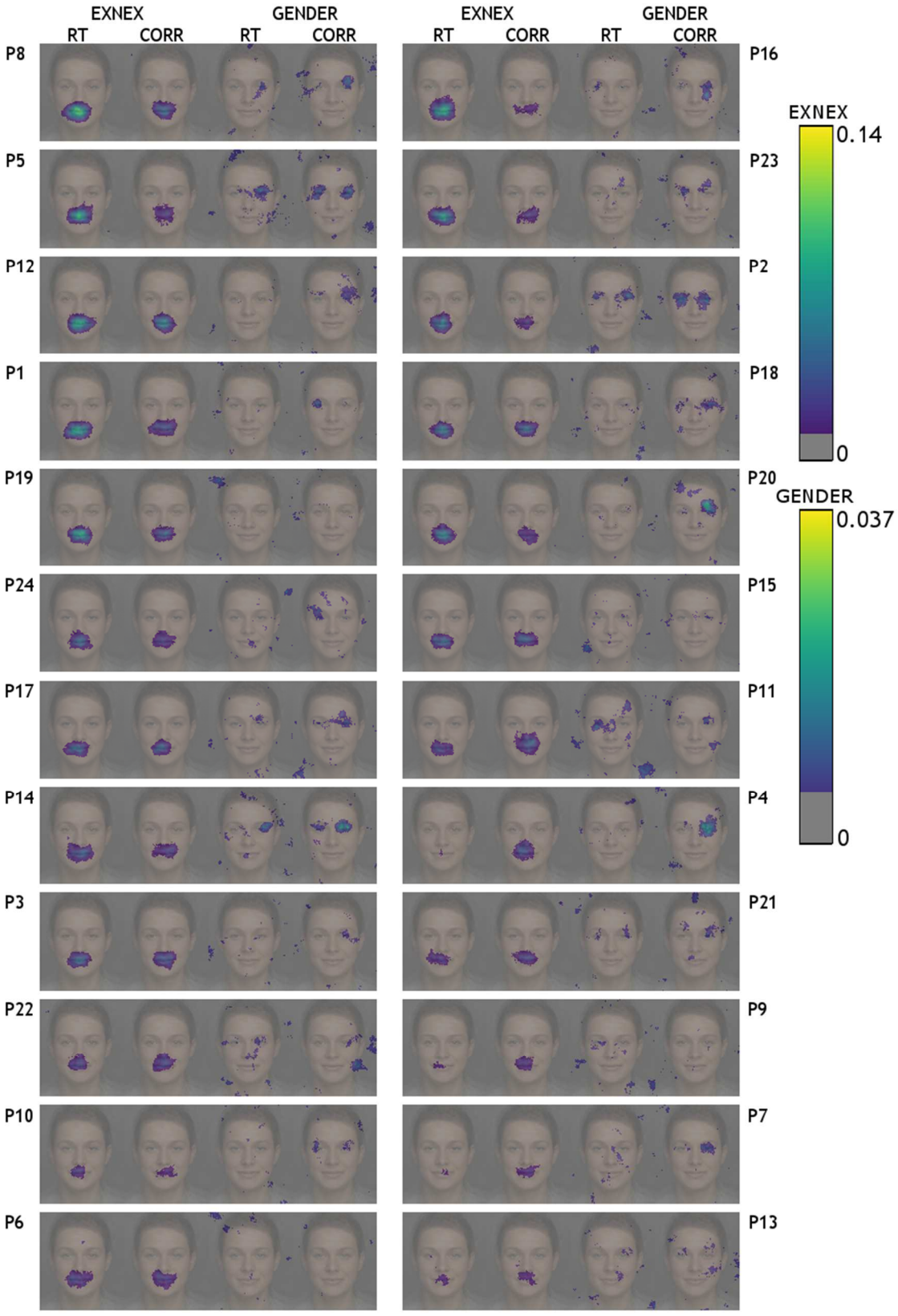
Our results only indicate the use of facial feature information when resolving a very narrow and specific task. Task demands may drive facial feature information processing in a top-down manner. For example, diagnostic information used to correctly categorise a neutral expression (compared another expression/s) differs depending on the task demand and comparison expression/s seen (Smith & Merlusca, 2014). The relative timing of facial feature integration should be considered under multiple and diverse task demands for a more comprehensive understanding of the extent to which eye sensitivity

precedes (or otherwise) sensitivity not just to the mouth but other features and their combination across a wider range of tasks.

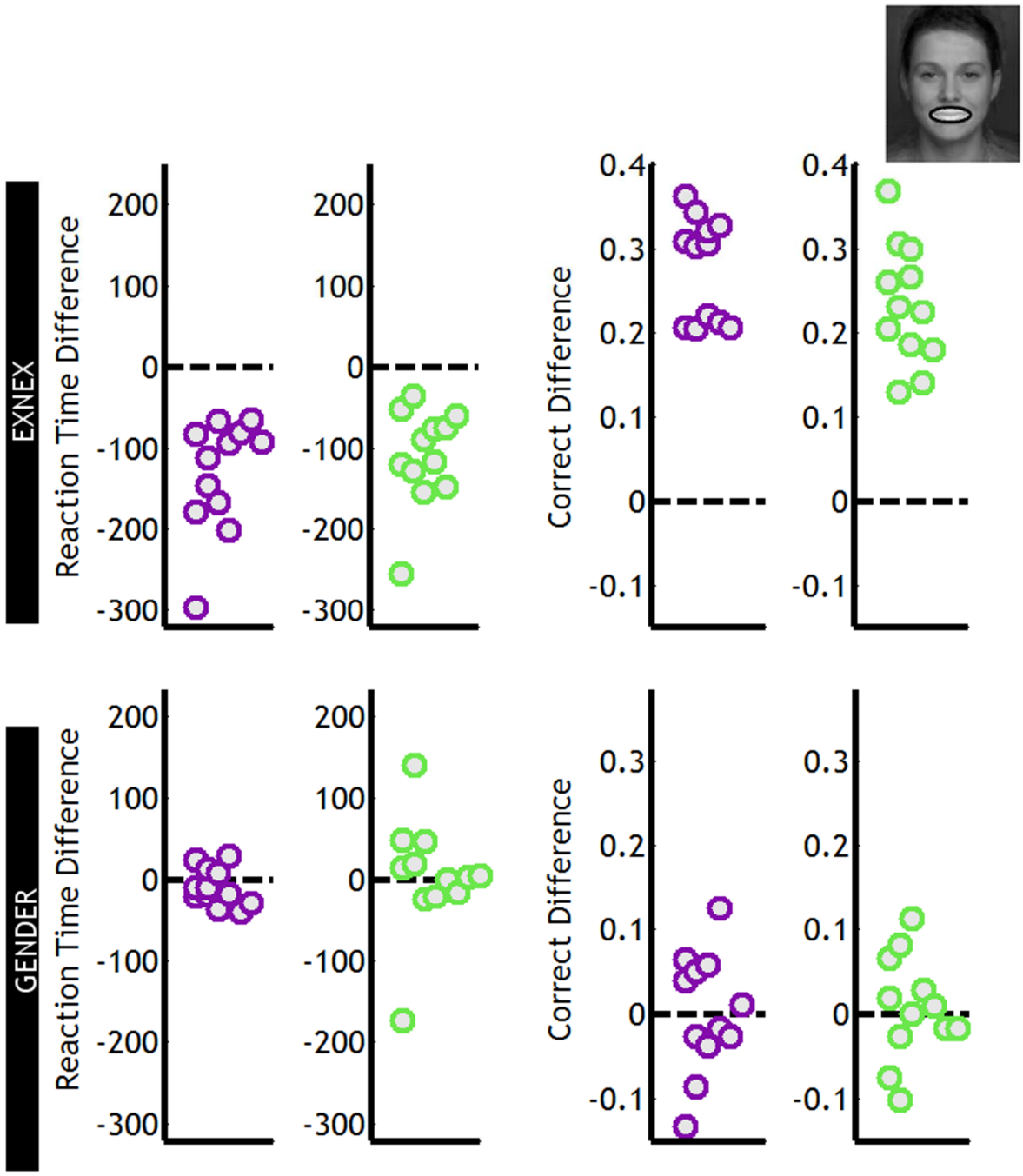
In the current experiment we considered the left eye, right eye and mouth as separate features. Binning based on visibility of a single feature meant that idiosyncratic amounts of visibility of other features were present. As such, we did not consider what Smith et al. (2004) refer to as 'second-order' effects on brain signals (i.e. the combinations of features such as the left eye and mouth). However Smith et al. (2004) found no evidence that configurations of facial features modulated the N170 differentially than considering facial features in isolation. Future studies however may wish to address explicitly if combinations of facial features modulate brain activity.

In conclusion, we have shown that the N170 can be modulated by the mouth as well as the eye region in both a gender and expressive/non-expressive categorisation task. Whilst there is a trend suggesting that sensitivity to the eye regions precedes sensitivity to the mouth, the timing of contralateral eye sensitivity compared to other task relevant feature modulation is idiosyncratic.

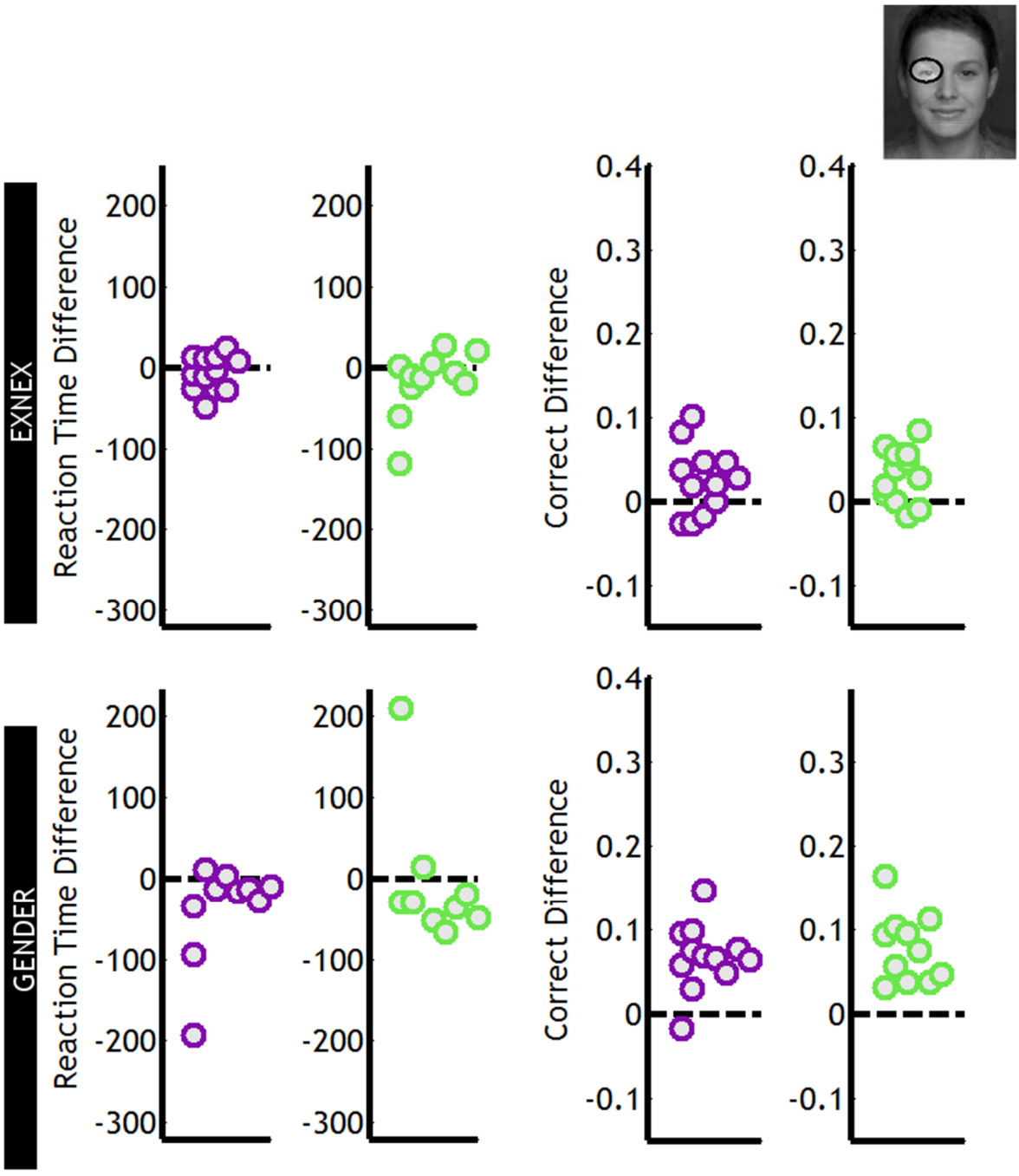
Chapter 4 Supplementary Figures



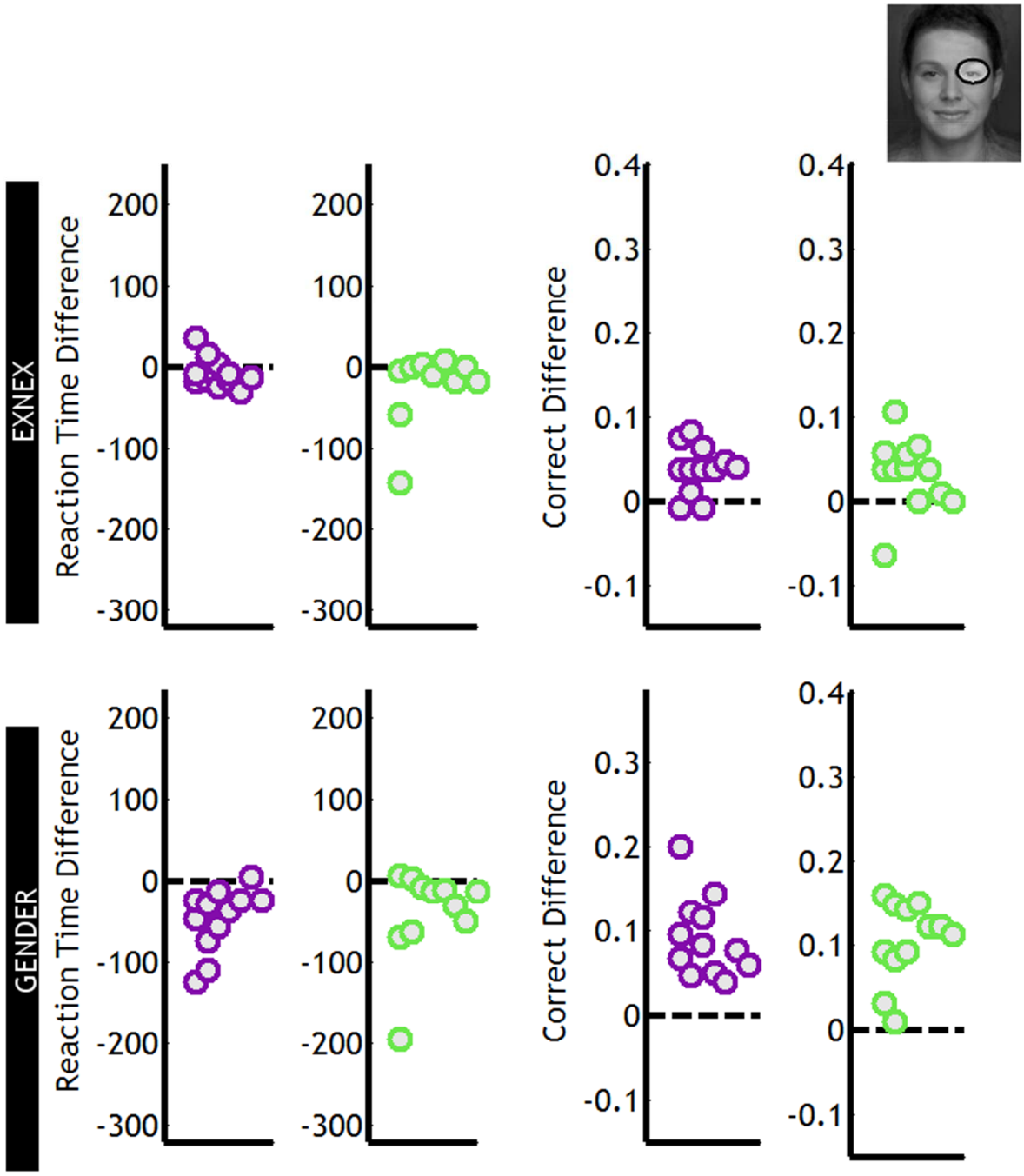
Supplementary 3: Younger adult individual behavioural classification images for all participants. MI(pix, **RT**) and MI(pix,**CORR**) for the **EXNEX** and **GENDER** task. Left column: Participants who completed the **EXNEX** task first. Right column: Participants who completed the **GENDER** task first. Participants are ordered so that the participants with the highest MI values appear at the top.



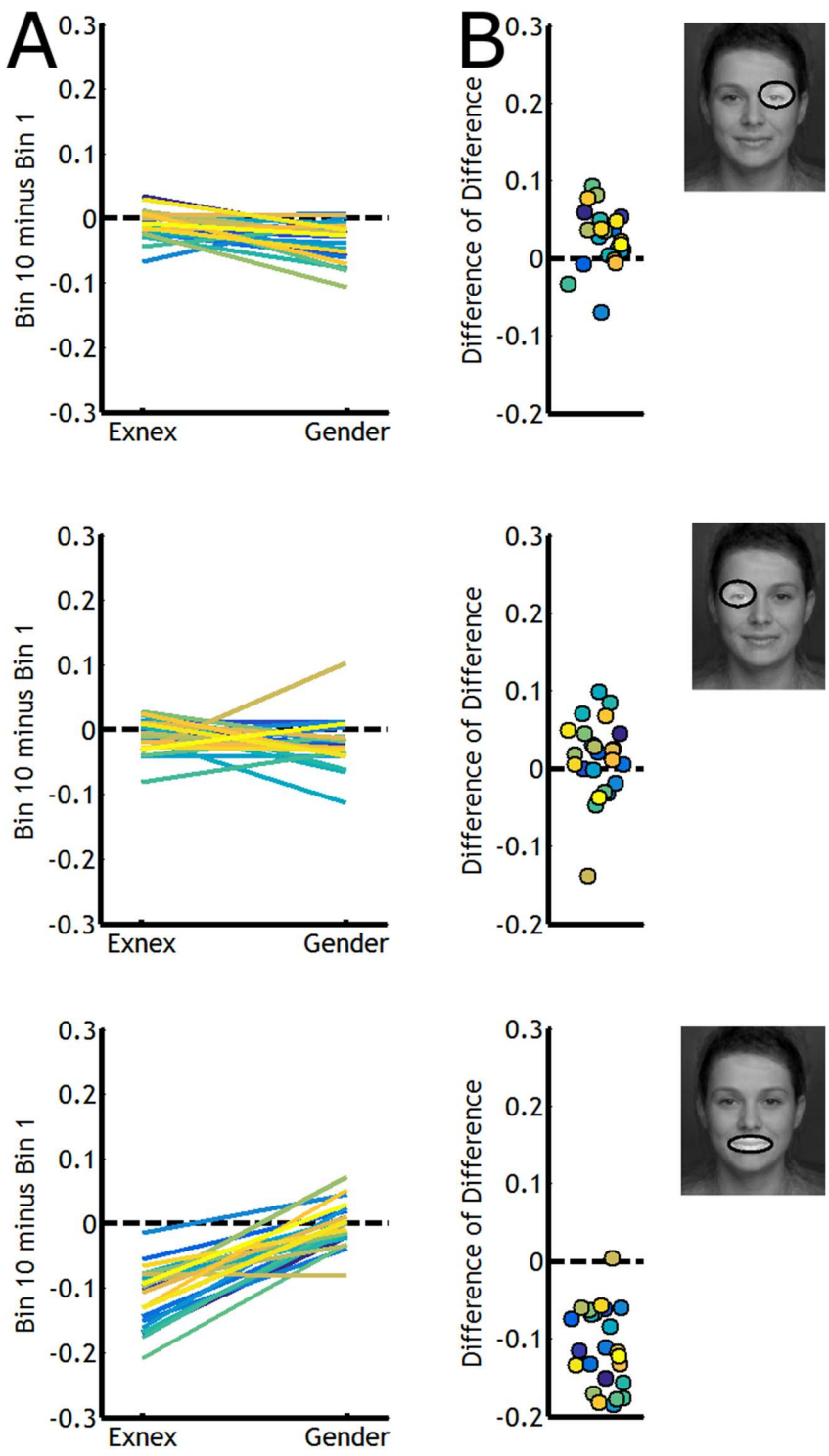
Supplementary 4: Younger adult mouth visibility task order effects Differences between high (bin 10) and low (bin 1) visibility of the mouth in the **EXNEX** (top row) and **GENDER** (bottom row) task. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first. Behavioural differences are broadly the same regardless of task order.



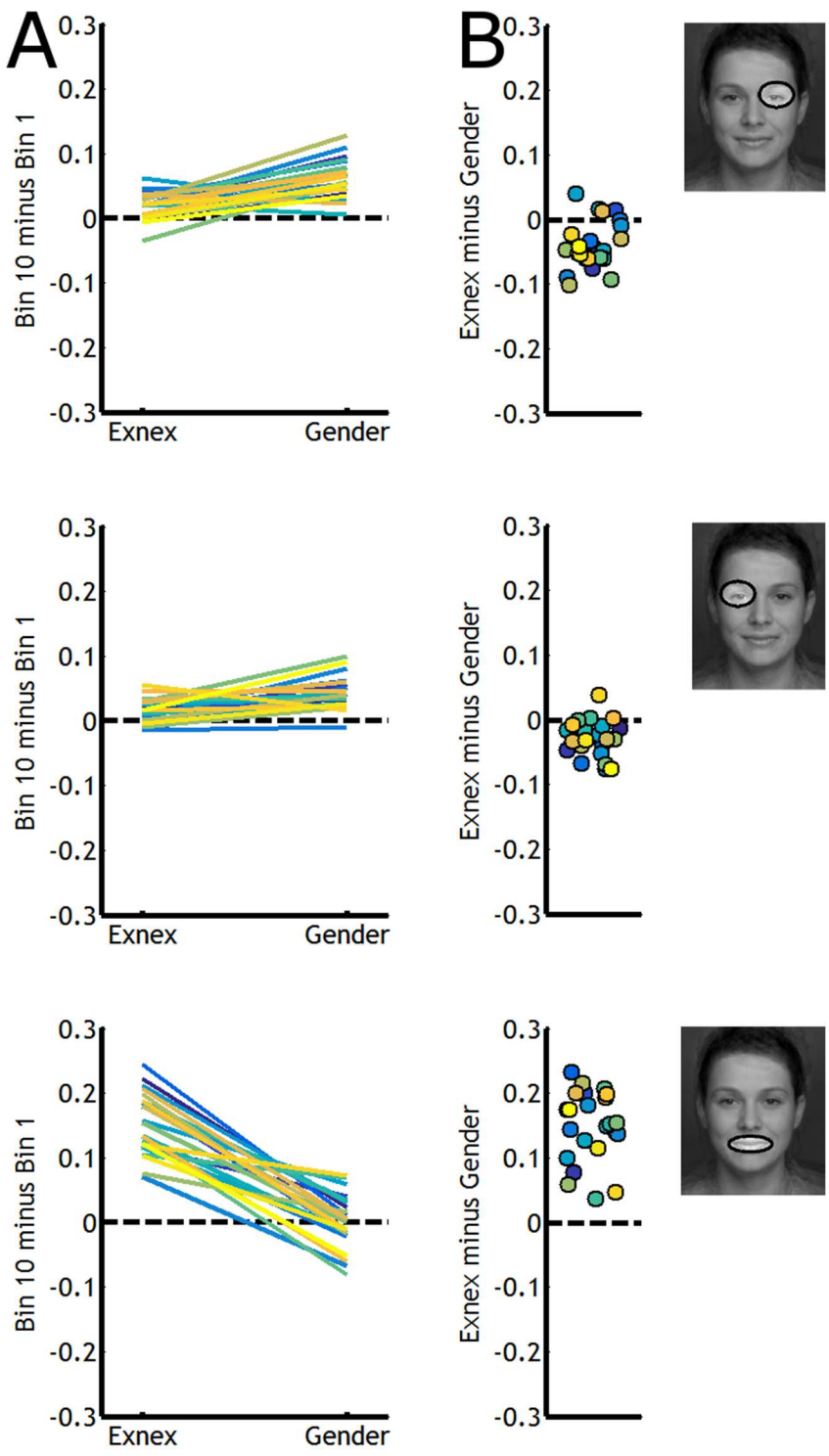
Supplementary 5: Younger adult left eye visibility task order effects Differences between high (bin 10) and low (bin 1) visibility of the mouth in the **EXNEX** (top row) and **GENDER** (bottom row) task. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first. Behavioural differences are broadly the same regardless of task order.



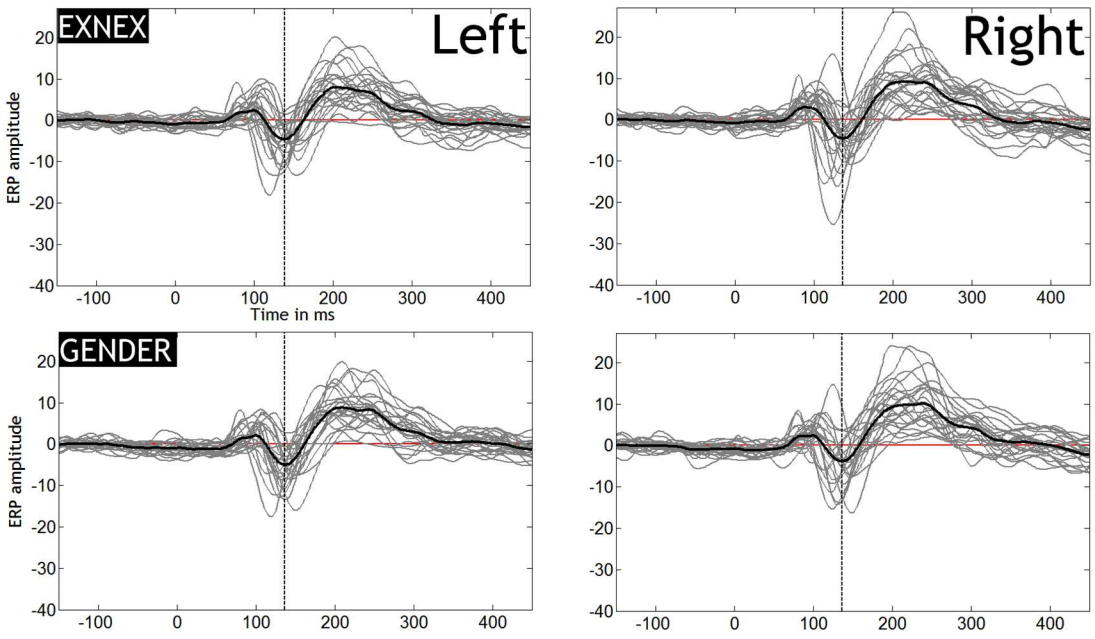
Supplementary 6: Younger adult right eye visibility task order effects Differences between high (bin 10) and low (bin 1) visibility of the mouth in the **EXNEX** (top row) and **GENDER** (bottom row) task. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first. Behavioural differences are broadly the same regardless of task order.



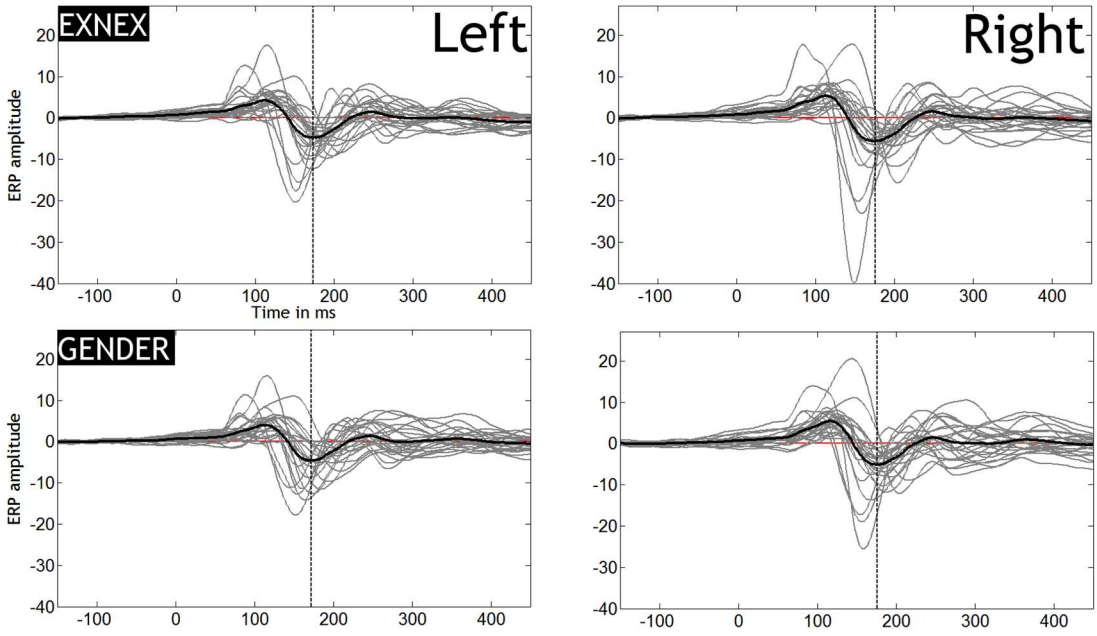
Supplementary 7: Younger adult normalised reaction time task differences by feature visibility Behavioural data normalised by dividing the difference of bin 10 minus bin 1 by the difference of bin 1 + bin 10 for each participant and facial feature. Panel A: The difference in normalised reaction time for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in normalised reaction time for the **EXNEX** minus **GENDER** task for each facial feature of interest.



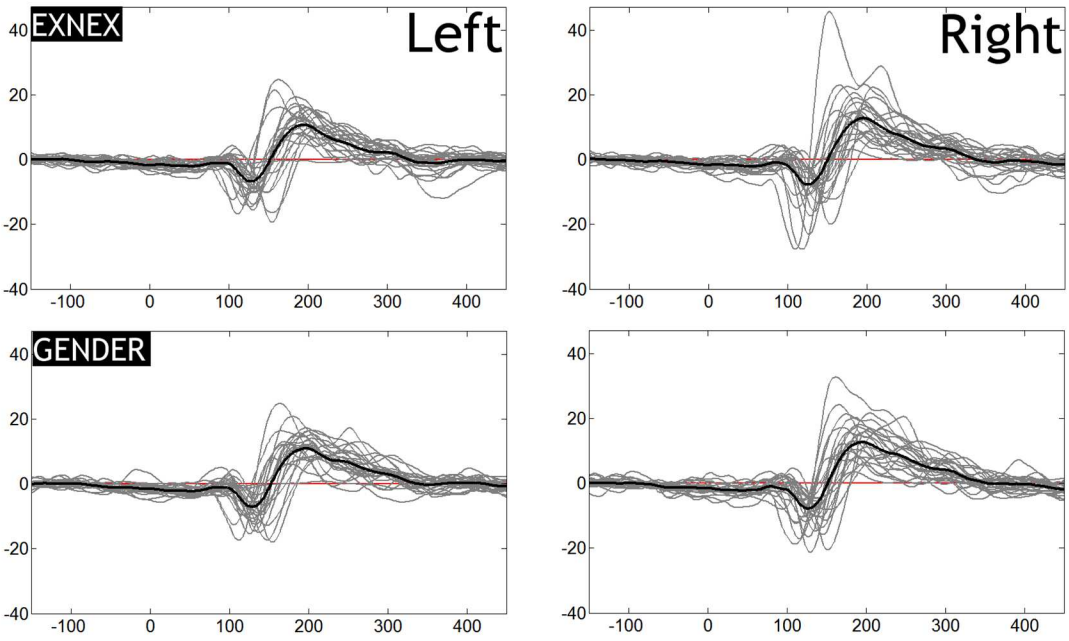
Supplementary 8: Younger adult normalised accuracy time task differences by feature visibility Behavioural data normalised by dividing the difference of bin 10 minus bin 1 by the difference of bin 1 + bin 10 for each participant and facial feature. Panel A: The difference in normalised reaction time for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in normalised reaction time (**EXNEX** minus **GENDER)** for each facial feature of interest.



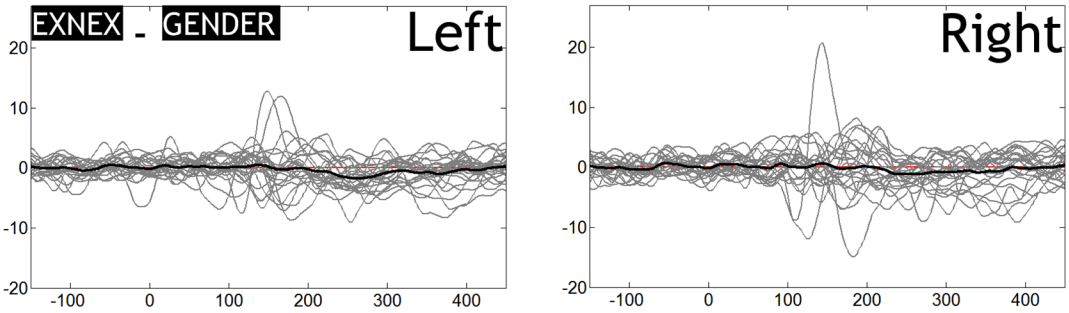
Supplementary 9: Younger adult individual mean ERPs (Practice Trials) Mean ERPs for each participant (N = 24) are superimposed in grey for the left and right hemisphere in each task. Solid black line represents the group mean. Dashed black line represents the latency of the group average minimum N170 amplitude.



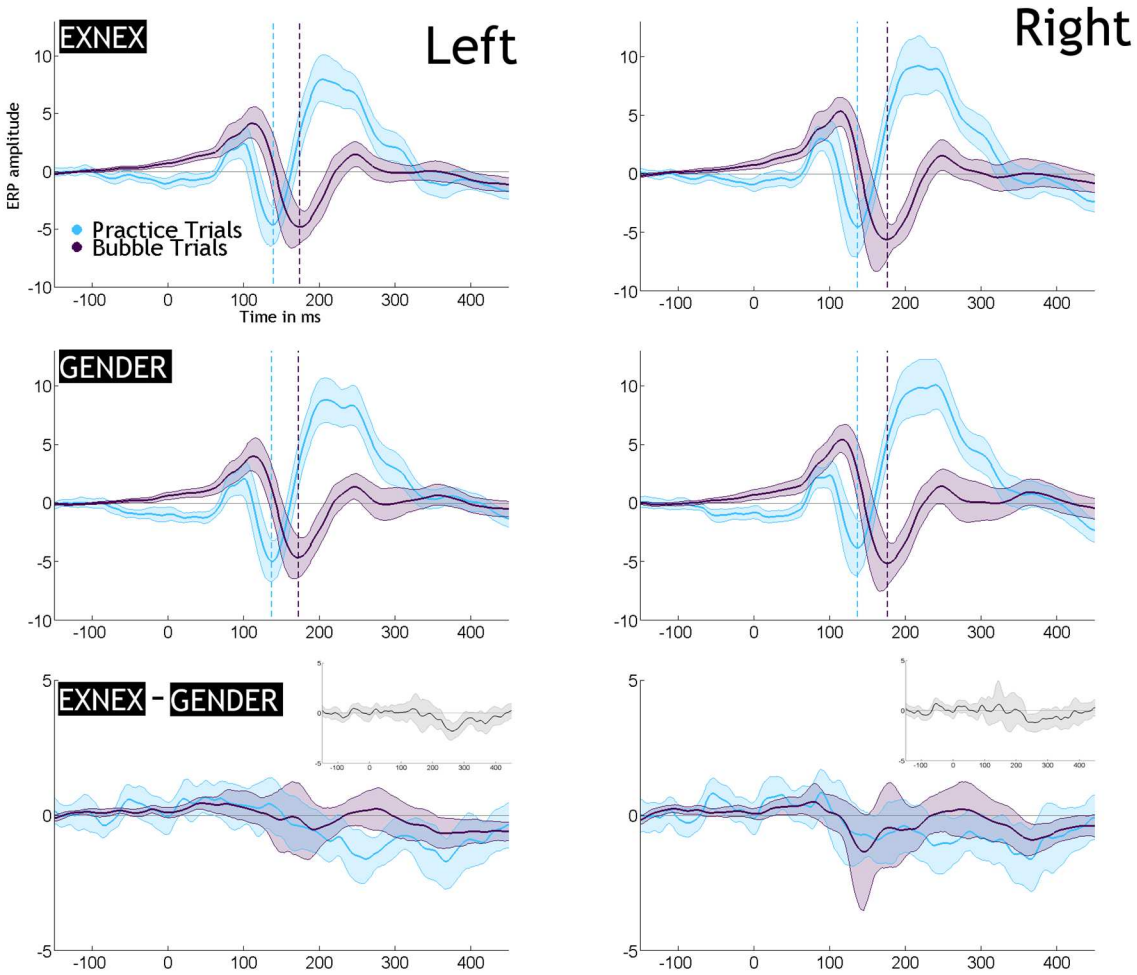
Supplementary 10: Younger adult individual mean ERPs (Bubble Trials) Mean ERPs for each participant (N = 24) are superimposed in grey for the left and right hemisphere in each task. Solid black line represents the group mean. Dashed black line represents the latency of the group average minimum N170 amplitude.



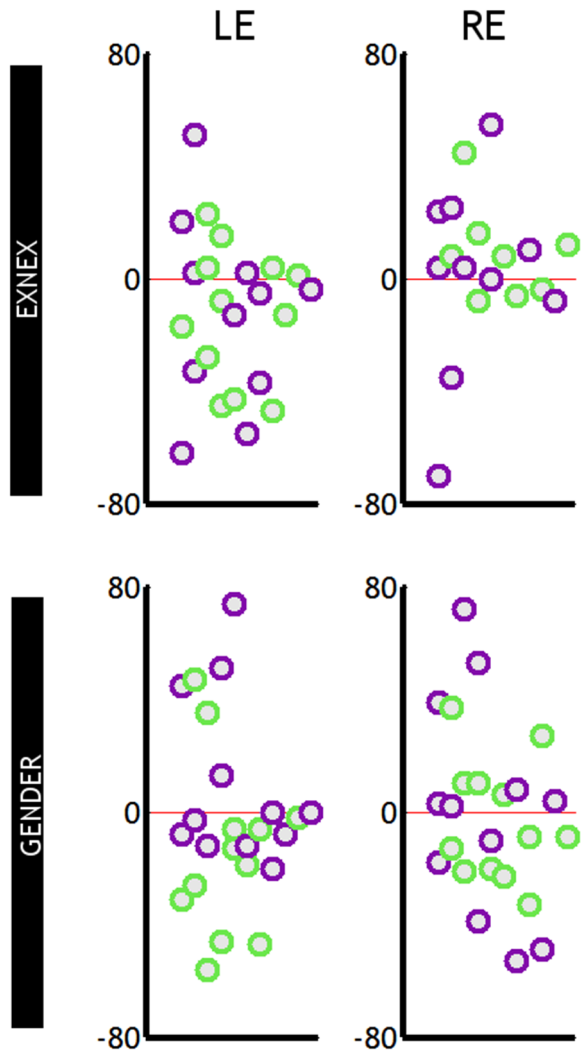
Supplementary 11: Younger adult individual mean ERPs (Bubble minus Practice Trials) Difference in Bubble minus Practice trials mean ERPs for each participant (N = 24) are superimposed in grey for the left and right hemisphere in each task. Solid black line represents the group mean.



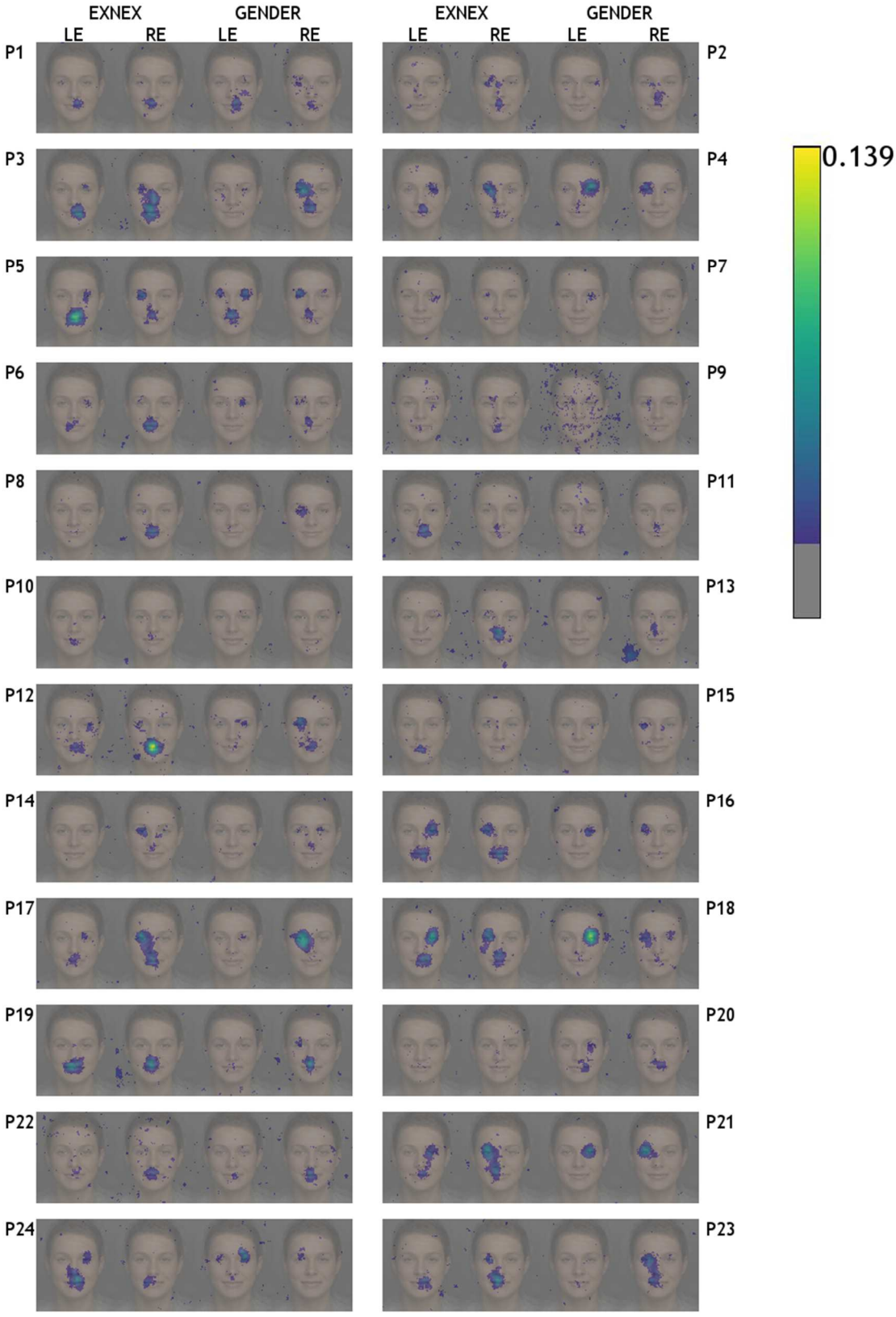
Supplementary 12: Younger adult individual mean ERPs (EXNEX – GENDER) Difference in Bubble minus Practice trials for **EXNEX** minus **GENDER** mean ERPs for each participant (N = 24) are superimposed in grey for the left and right hemisphere in each task. Solid black line represents the group mean.



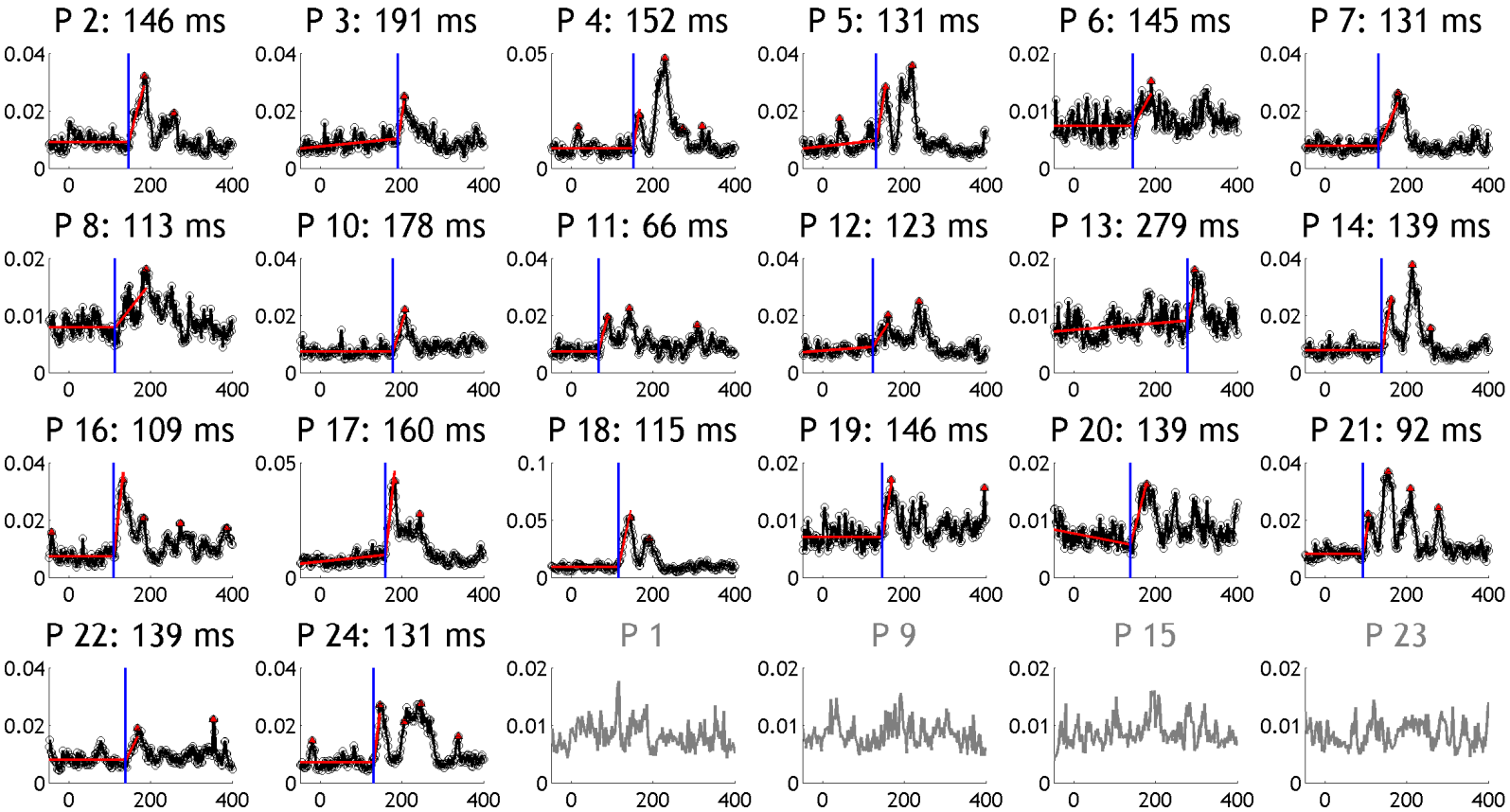
Supplementary 13: Younger adult 20 % trimmed mean average group ERP in Bubble and Non-Bubble Trials 20 % trimmed mean bubble and non-bubble trial ERPs for the left and right hemisphere with 95 % confidence intervals around the 20 % trimmed mean. Vertical lines represent the minimum amplitude peak of the N170 for each task. Bottom panel **EXNEX** minus **GENDER** for bubble and practice trials. Small grey plot shows the pairwise difference of practice minus bubbles trials.



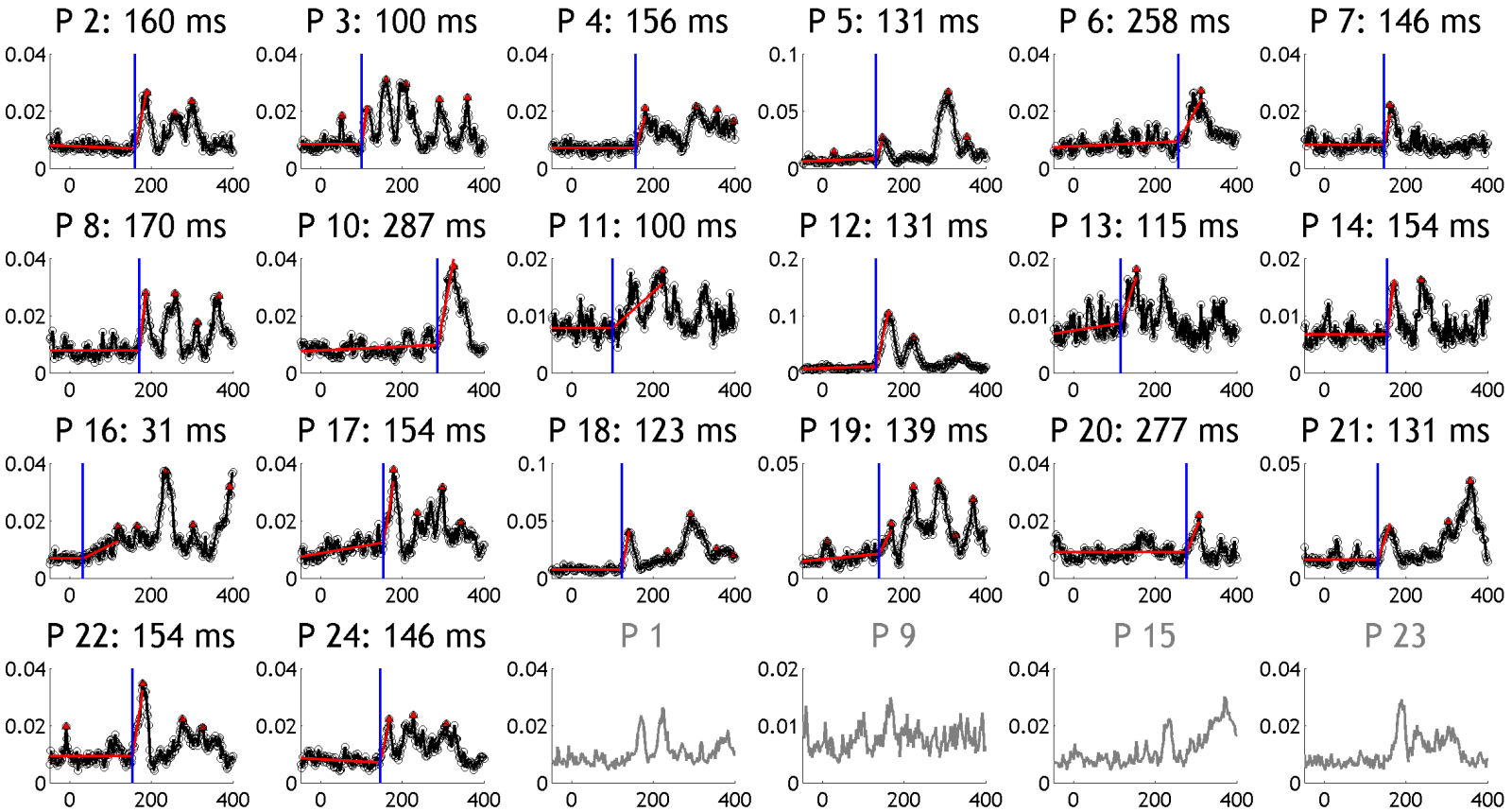
Supplementary 14: Younger adult individual differences in MI peak latency For each participant we calculated the difference in the latency of the peak MI between 120-220 ms for the contralateral eye minus the mouth. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first



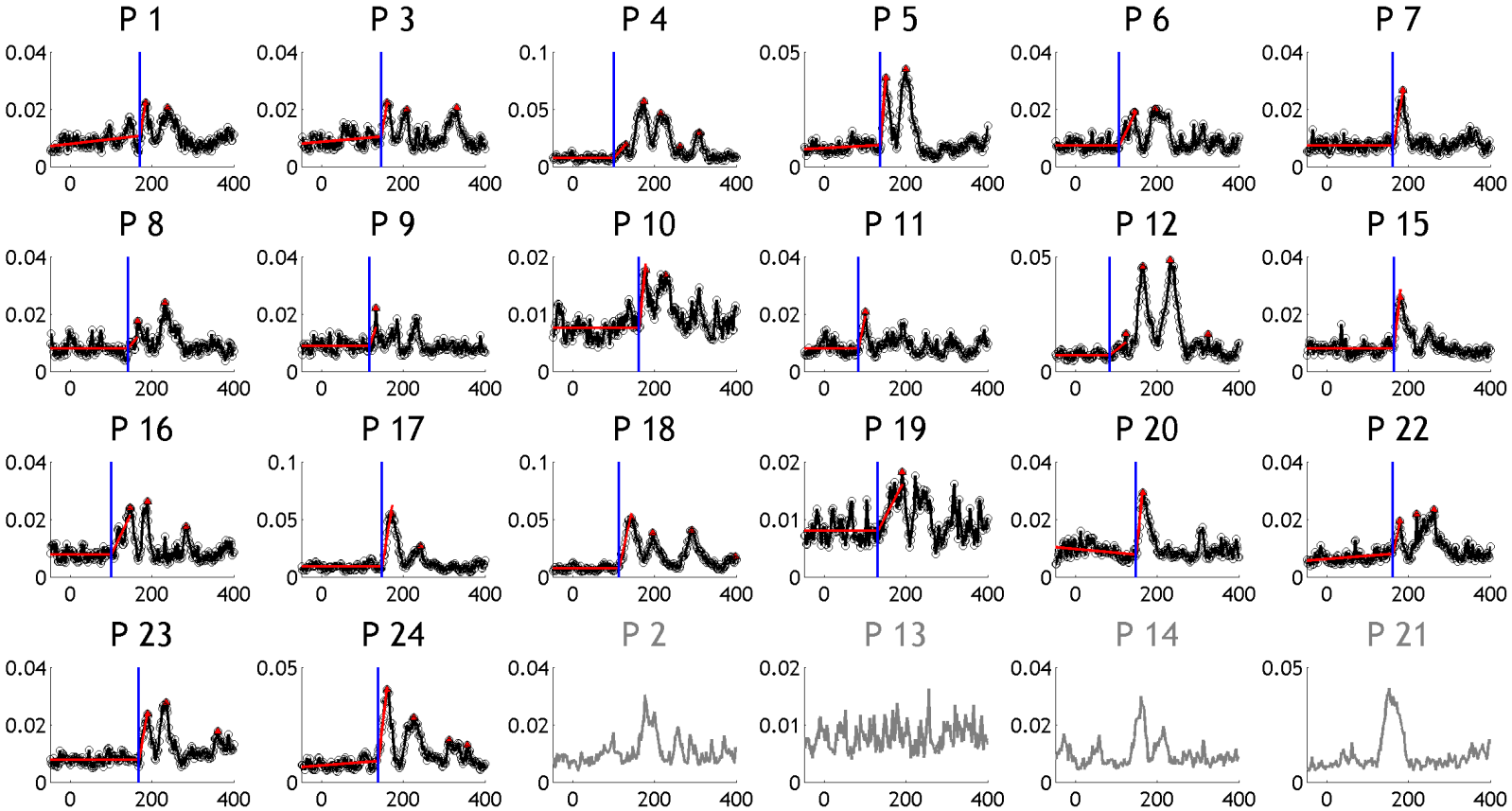
Supplementary 15: Younger adult individual brain classification images for all participants. MI(pix, [ERP,grad]) for the **EXNEX** and **GENDER** task. Left column: Participants who completed the **EXNEX** task first. Right column: Participants who completed the **GENDER** task first.



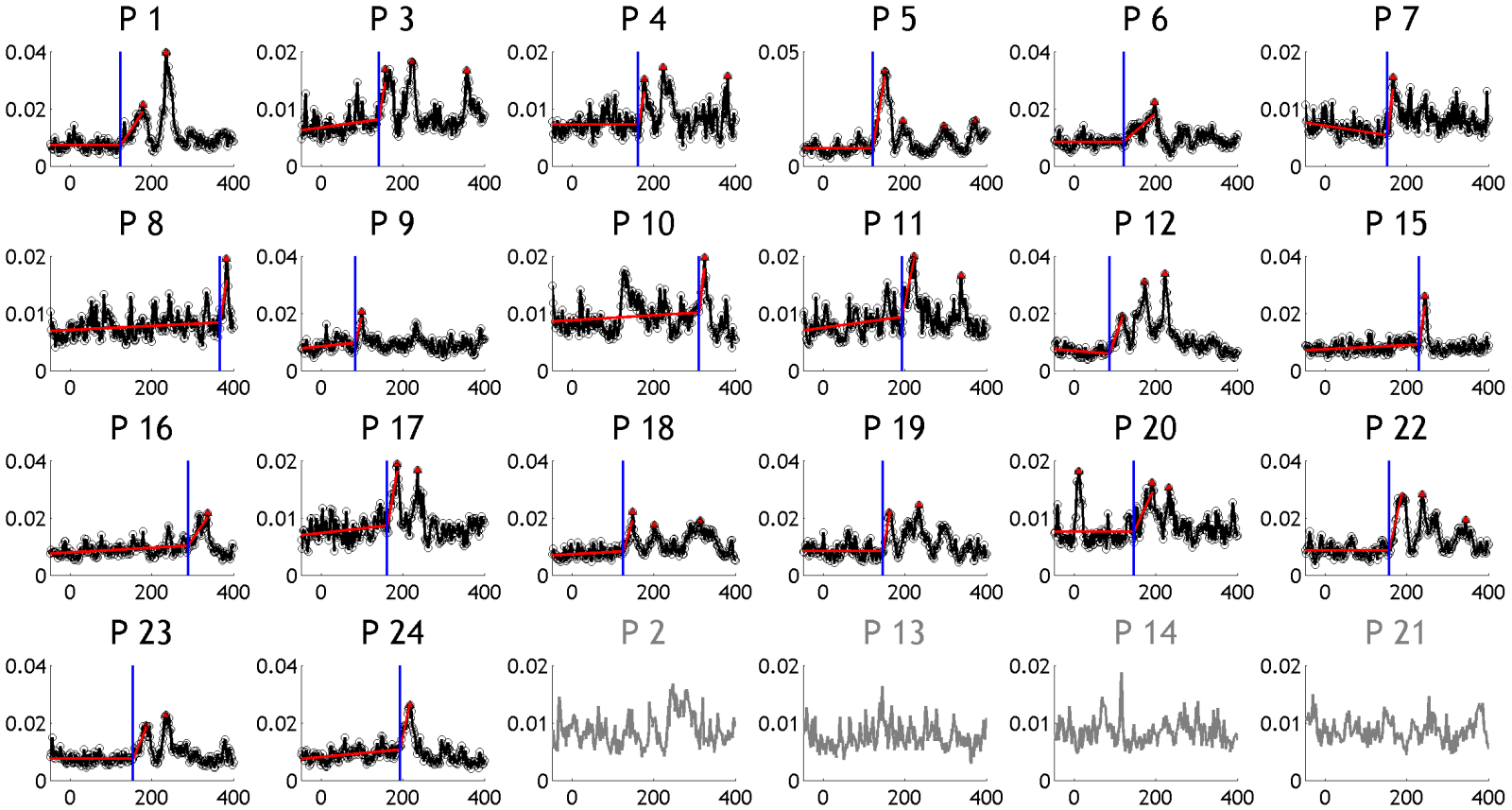
Supplementary 16: Younger adult MI eye onset (EXNEX) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.



Supplementary 17: Younger adult MI mouth onset (EXNEX) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.



Supplementary 18: Younger adult MI eye onset (GENDER) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.



Supplementary 19: Younger adult MI mouth onset (GENDER) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.

Chapter 5 - Eye but not mouth sensitivity is delayed in healthy ageing

Introduction

In Chapter 2, we have seen that the N170 sensitivity to the contralateral eye area was scale tolerant and reflected an eye-specific coding mechanism. In Chapter 4, we have shown that contralateral eye sensitivity is apparent in tasks where the eye is purported to be more task relevant (in a gender discrimination task) and more task irrelevant (in an expressive versus non expressive discrimination task) for behaviour. Unlike previous studies (M. L. Smith et al., 2004) however, we have also shown that the N170 is also sensitive to the mouth. This suggests that the N170 is modulated by multiple facial features, and is not just an eye-specific coding mechanism as previously suggested.

In Chapter 4, we attempted to quantify the timing of feature sensitivity to elucidate whether eye coding is a necessary first-stage of face processing regardless of the task. By estimating the onset of MI to the eyes and mouth regions, we uncovered idiosyncrasies in the onset of feature processing. Whilst there was a trend towards a general preference for encoding the eye first, this was not systematic across all participants, with only 70 % and 60 % of younger participants encoding the eye before the mouth in the **EXNEX** and **GENDER** task respectively. This suggests that whilst eye and mouth coding appears consistently across tasks (sensitivity to the contralateral eye and mouth area was found in both tasks) there are individual idiosyncrasies in the relative timing of onset of sensitivity to each of these features, and these may interact with task.

In this chapter, we investigate how task effects on the N170 sensitivity differ in healthy ageing. Recently, research has shown that during a face versus noise detection task, both younger and older adults responded faster when the contralateral eye region was visible, and older adults were more reliant on the visibility of the eye region for correct responses (Jaworska, 2017). Furthermore, the presence of the contralateral eye modulated ERP's for both younger and older adults, though in older adults this modulation was delayed and weaker, and only the N170 amplitude, not latency, was modulated in older adults (Jaworska, 2017). Together, these results point towards a dissociation of

behavioural and brain responses in older adults, where an increased reliance on presence of the eyes is coupled with a delayed and weaker processing of this very feature during a face detection task.

In this chapter we aimed to explore whether older adults rely on processing the same facial features as younger adults in the same **GENDER** and **EXNEX** categorisation task as in Chapter 4, and quantify the relative timing of feature sensitivity in older compared to younger adults. We predicted that if the N170 eye sensitivity is task-independent, then during both tasks there would be evidence of contralateral eye sensitivity prior to sensitivity to other facial features (such as the mouth), regardless of the extent to which the eyes are task-relevant. We also predicted that whilst older adults would process the same facial features as younger adults (i.e. the mouth in the **EXNEX** task), feature encoding would be delayed and weaker in older compared to younger adults. We aimed to quantify these differences.

The primary contribution of this chapter will be demonstrating that older adults behavioural and brain responses were modulated by the same facial features as younger adults behavioural and brain responses. However, compared to younger adults, older adults relied more on the presence of the mouth for resolving the **EXNEX** task. We will demonstrate that whilst older adults integrated information about both the eyes and mouth more slowly than younger adults this delay in processing speed was not uniform across features and tasks. More specifically, whilst older adults were ~20-23 ms delayed in processing the eyes compared to younger adults across tasks, the delay in mouth integration was task dependent – there was a delay in processing the mouth in the **GENDER** but not **EXNEX** task. Mouth processing was less delayed than eye processing in the **GENDER** task, suggesting that older adults did not demonstrate a general delay that was consistent across all features.

Materials and Methods

Participants

The study comprised of 24 'younger' participants (described in Chapter 4) and 24 'older' participants: 12 women, 22 right handed, median age = 67 years (min 60, max 85). All adults were living in the community at the time of testing. Participants were recruited from a subject pool, which had previously been established through advertising at the University of Glasgow, active age gym classes and a newspaper article. Volunteers were excluded from participation if they reported any current eye condition (i.e. lazy eye, glaucoma, macular degeneration or a significant cataract), had a history of mental illness or were currently taking psychotropic medications, suffered from any neurological condition, had diabetes, or had suffered a stroke or a serious head injury. Volunteers were also excluded from participation if they had not had their eyes tested within the last year, in order to minimise the chances that volunteers did not know of an underlying eye condition. One older participant reported having cataracts removed from both eyes, and one older participant reported having undergone eye surgery to rectify a squint as a child. These participants were included because their corrected vision did not differ from that of the others.

Participants' visual acuity and contrast sensitivity was assessed in the lab (Table 3). Contrast sensitivity was assessed using the Mars Letter Contrast Sensitivity set (Arditi, 2005). Visual acuity at 40 cm and 63 cm were assessed using the Colenbrander mixed contrast card set (Colenbrander & Fletcher, 2004), and visual acuity at 6 m assessed using the Bailey-Lovie Chart (Bailey & Lovie, 1980). All participants had normal or corrected to normal visual acuity and normal contrast sensitivity (equal to or above the lower limit of the normal Mars letter contrast sensitivity for a person aged 60 years of 1.52 log units (Haymes et al., 2006)). During the experimental session, participants wore their habitual eye correction as needed. As a note of caution however, age-related presbyopia causes blurred near vision and, whilst participants wearing bi-focal or vari-foca glasses may have measured within normal visual acuity parameters on the Colenbrander card set, it is possible that participants may still have experienced some optical blur when sitting in front of a computer screen.

Older adults also completed the Montreal Cognitive Assessment (MoCA) to screen for age-related cognitive impairment (Nasreddine et al., 2005). All participants were above the cut-off threshold of ≥ 26 out of 30 (median score = 28 [min = 26, max = 30]) (Nasreddine et al., 2005).

Participants were compensated £6/hr for their participation. The experiment was approved by the local ethics committee of the College of Science and Engineering (approval no. 300150158).

LC 40	HC 40	LC 63	HC 63	LC 600	HC 600	CS
85 [69, 93]	95 [85,103]	91 [80, 105]	100 [90, 110]	88 [77, 96]	100 [88, 105]	1.76 [1.52, 1.84]
0.30 *[0.62, 0.14]*	*0.10* *[0.30, -0.06]*	*0.18* *[0.40, -0.10]*	*0.00* *[0.20, -0.20]*	*0.24* *[0.46, 0.08]*	*0.00* *[0.24, -0.10]*	

Table 7: Older adults Visual Acuity and Contrast Sensitivity Scores Median Visual acuity and Contrast sensitivity (CS) sores for older participants. Visual acuity scores are reported for low contrast (LC) and high contrast (HC) charts presented at 40 cm, 63 cm, and 6 m viewing distance, expressed as raw visual acuity scores (VAS). The corresponding logMAR scores are presented below in italics, where higher values indicate poorer vision and negative values represent normal vision (logMAR score of 0 corresponds to 20/20 vision). Square brackets indicate the minimum and maximum scores across participants. For younger adults visual acuity scores refer to Table 3 (page 103).

Stimuli, Procedure, EEG Recording and Pre-processing

Stimuli, procedures, EEG recording and pre-processing are the same as those outlined in Chapter 3.

Results

Behavioural Results

Behavioural results are given here for trials with bubble masks and practice trials. Practice trials presented without bubble masks revealed the whole face image and were used to familiarise participants with the task. As bubble trials reveal partial face information, we compared behavioural performance between practice and bubble trials.

Reaction Times and Percentage Correct

On practice trials i.e. trials without bubbles where the full face image can be seen, participants were on average faster in the gender than expression task

(**GENDER** = 569 ms [521, 617]; **EXNEX** = 570 ms [505, 634]; **Difference** [**GENDER** minus **EXNEX**] = -17 ms [-84, 49]). Participants were also less accurate in the gender than expression task during practice trials, though performed near-ceiling in both (**GENDER** = 95 % [90, 98]; **EXNEX** = 98 % [97, 99]; **Difference** [**GENDER** minus **EXNEX**] = -3 % [-7, -1]).

On trials with bubble masks, participants were slower on the gender than the expression task (**GENDER** = 781 ms [709, 854]; **EXNEX** = 718 [652, 783]; **Difference** [**GENDER** minus **EXNEX**] = 51 ms [-8, 110]). This is in reverse of practice trials on which participants were faster on the gender task. This pattern was also seen in our younger adult group (see Chapter 4; *Reaction Times and Percentage Correct*). Participants were also less accurate on the gender than expression task for bubble trials (**GENDER** = 65 % [62, 68]; **EXNEX** = 76 % [73, 79]; **Difference** [**GENDER** minus **EXNEX**] = - 11 % [-14, -9]). This is in contrast to the practice trials on which participants performed similarly in the two tasks, but in-keeping with the pattern of results from our younger adults (see Chapter 4: *Reaction Times and Percentage Correct*).

Adding bubbles to the image negatively affected older adults' reaction times and accuracy performance in both tasks. Compared to practice trials, on bubble trials participants were slower (**GENDER** [**Practice** minus **Bubble**] = -217 ms [-301, -134]; **EXNEX** [**Practice** minus **Bubble**] = -112 [-181, -43]; **Difference** [**GENDER** minus **EXNEX**] = -74 [-146, -1]) and less accurate (**GENDER** [**Practice** minus **Bubble**] = 30 % [26, 33]; **EXNEX** [**Practice** minus **Bubble**] = 22 % [19, 25]; **Difference** [**GENDER** minus **EXNEX**] = 8 % [3, 12]) on both tasks. Compared to practice trials, bubbling the image had a more negative impact on both reaction times and accuracy for the gender than the expression task, with an increased delay for the gender compared to the expression task and larger reduction in accuracy. This suggests that the impact of bubbling affected performance on the gender task differentially. This could be due to differential increase of the difficulty of the gender compared to the expression task when information is limited or fragmented. This pattern of results is in-keeping with previous research in younger participants using an adaptive version of bubbles where the number of bubbles is varied to achieve a 75 % correct performance criterion (Schyns et al., 2002). Under these conditions, participants on average required more bubbles i.e. more of the image to be revealed on the

GENDER task to achieve the same accuracy levels as on the **EXNEX** task. This was a similar pattern of effect as demonstrated in our younger adult participants (see Chapter 4: *Reaction Times and Percentage Correct*).

We compared differences in reaction time and accuracy between younger and older participants by computing the Harrell-Davis median of all pairwise differences of the distribution of younger minus older participants for each behavioural comparison of interest. We computed a 95 % confidence interval around this estimate using a percentile bootstrap technique with 500 iterations. We also calculated as a measure of effect size Cliff's delta (Table 8). Cliff's delta is a robust measure of effect size that calculates the probability that a randomly selected observation from one group is larger than a randomly selected observation from another group i.e. P(X<Y) – P(X>Y) (see Chapter 3: *Between-group Comparisons*).

Bubbling the image negatively affected accuracy and reaction times for both age groups, but bubbling was more detrimental for our older adult group, particularly in the **GENDER** task. Whilst older adults performed similarly in terms of accuracy and reaction times on practice trials, with reduced visual information in bubble trials older adults were slower and less accurate, suggesting older adults required more information to perform as accurately as younger adults, as well as more time to integrate information than younger participants, particularly on the **GENDER** task.

	Reaction Times		Accuracy	
	EXNEX	**GENDER**	**EXNEX**	**GENDER**
Practice	-0.6 [-0.3, -0.8]	-0.6 [-0.4, -0.9]	-0.2 [-0.5, 0.1]	0.1 [-0.3, 0.4]
Bubble	-0.8 [-0.6, -1]	-0.7 [-0.5, -0.9]	0.9 [0.7, 1]	0.9 [0.8, 1]

Table 8: Cliff's delta behavioural effect size estimates We estimated Cliff's delta effect size of the difference between younger and older behavioural results. Cliff's delta ranges from 1 (where all values from one group are higher than the values from the other group) to -1 (when all values from one group are lower than the values from the other group). Completely overlapping distributions have a Cliff's delta of 0.

On practice trials, younger adults were faster than older adults on both tasks (**GENDER** = -128 ms [-75, -180]; **EXNEX** = -120 ms [-66, 190]; **Difference** [**GENDER** minus **EXNEX**] = 1 ms [-67, 67]) though both groups demonstrated similar levels of accuracy (**GENDER** = 0 % [-1, 2]; **EXNEX** = -1 % [-2, 0]; **Difference** [**GENDER** minus

EXNEX] = 1 % [0, 3]) responding close to ceiling. In comparison, whilst younger participants were also faster than older participants on bubble trials (**GENDER** = -230 ms [-314, 168]; **EXNEX** = -217 ms [-291, -154]; **Difference** [**GENDER** minus **EXNEX**] = -21 ms [-78, 48]), older adults had an increase delay in bubble compared to practice trials (**GENDER** = 120 ms [58, 120]; **EXNEX** = 71 ms [17, 135]; **Difference** [**GENDER** minus **EXNEX**] = 22 ms [-43, 99]). Younger adults also performed more accurately than older adults on bubble trials (**GENDER** = 17 % [14, 21]; **EXNEX** = 12 % [7, 15]; **Difference** [**GENDER** minus **EXNEX**] = 5 % [1, 9]). Compared to younger adults, older adults had a larger reduction in accuracy in bubble compared to practice trials (**GENDER** = - 16 % [-13, -20]; **EXNEX** = -13 % [-10, -16]; **Difference** [**GENDER** minus **EXNEX**] = -4 % [-8, 0]).

In summary, during practice trials older adults were as accurate in both tasks as younger participants. However, in practice trials it should be noted both age groups performed at near-ceiling accuracy to comparisons between age groups here should be treated with caution. Whilst accuracy declined for both age groups in bubble trials, older adults' accuracy was noticeably more diminished. Whilst older adults were slower than younger adults in the practice trials, reaction times were comparatively more delayed in bubble trials. This is in contrast to previous work indicating that older adults were comparably delayed in practice and bubble trials in a face detection task (Jaworska, 2017). It is not clear to what extent the complexity of our stimuli (coloured images compared to grey scale images in a common oval frame) or the complexity of the task (more stimulus information may be required for **GENDER** discrimination than face detection) contributed to this disparity. This pattern however may be consistent with the indication that older adults' perception is affected by partial occlusion / stimulus fragmentation, and that the extent to which perception is affected may be related to the amount of stimulus information required to successfully complete the task. It may be that age-related inefficiency in utilising fragmented information is more pronounced in a more demanding **GENDER** task requiring greater integration of sparsely sampled features than in a less complex face detection task.

Behavioural MI Classification Images

To determine what image features are associated with behavioural responses, we looked at mutual information between pixels and reaction times MI(PIX, **RT**) and pixels and correct responses MI(PIX, **CORRECT**). For each comparison we calculated the group median MI value for each pixel. We also calculated frequencies of significant effects (see Chapter 3*; Behavioural Permutation*). The classification images shown in Figure 49 are for mutual information computed using the warped bubble masks (See Chapter 3; *Bubble Warp and Warped Bubble Masks)*. In the **EXNEX** task (Figure 49 panels A-B, top row) participants' reaction times and correct responses were modulated by the presence of the mouth. In contrast, in the **GENDER** task (Figure 49 panels A-B, bottom row) there was a very weak relationship between the presence of the eyes and modulation of behavioural responses, that was only significant for **CORRECT** for a maximum of four participants at any one pixel, and for **RT** significant for a maximum of two participants at any one pixel (see Supplementary 20 for individual classification images).

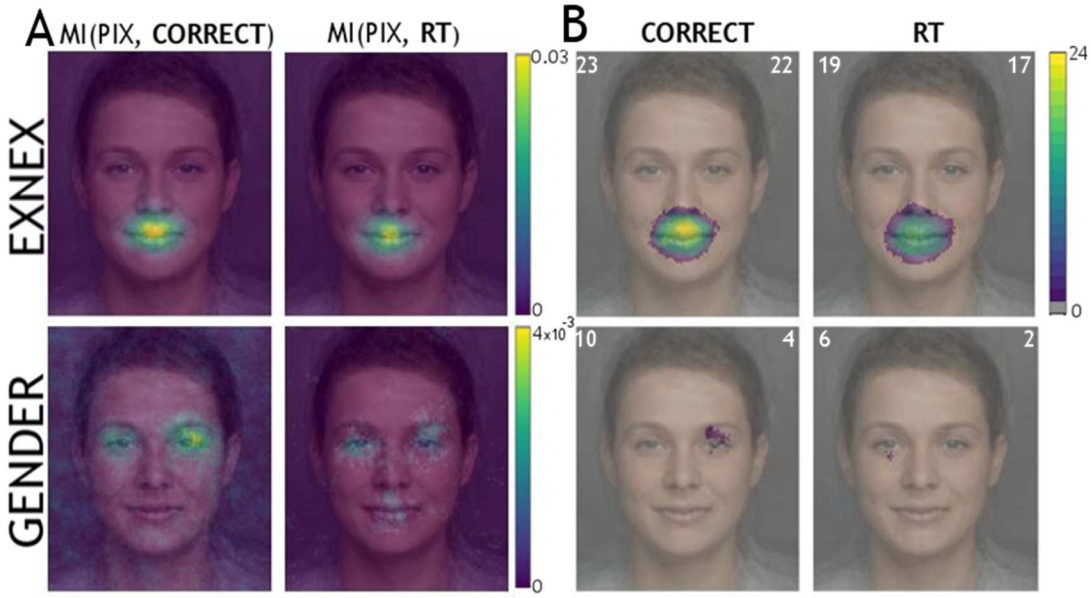


Figure 49: Older adult behavioural Classification Images and Frequency of Significant Effects (A) Group median of mutual information at each pixel. Each row corresponds to one task condition, each column corresponds to a different analysis condition (**RT** or **CORRECT**). Median MI values were stronger in **EXNEX** than in **GENDER**. (B) Number of participants showing significant effects based on a permutation test. Small white numbers indicate for each condition the number of participants with significant MI at any pixel (left) and the maximum number of participants with significant MI at the same pixel (right). For younger adults behavioural classification images see Figure 31 (page 105).

We found evidence of strong and consistent behavioural modulation by mouth visibility in the EXNEX task across participants, in keeping with the mouth being the most diagnostic feature for this task (Gosselin & Schyns, 2001). However, behavioural modulation by eye visibility in the GENDER task was weak, despite previous reports of strong diagnosticity in this task (Gosselin & Schyns, 2001). This was the same pattern of results we observed in our younger participants' (see Chapter 4: Behavioural MI Classification Images). In both cases, this may have been due to other diagnostic cues such as pigmentation or inter-eye distance being available.

To compare how older adults' behavioural responses were modulated compared to younger adults (see Chapter 4: *Behavioural MI Classification Images*), we directly compared the median classification images between the two groups in the two tasks. We computed the difference (**Younger** minus **Older**) in group-median MI maps for reaction times and accuracy in each task. Relative to younger participants, older participants displayed weaker MI values around the mouth for MI(PIX, **RT**) in the **EXNEX** task (Figure 50), suggesting that younger adults' reaction times were more modulated by mouth visibility than in older adults. Only weak differences between younger and older adults were seen in any other behavioural comparison. This suggests that, except for MI(PIX, **RT**) in the **EXNEX** task, older adults behavioural MI estimates were not weaker than younger adults at the group level.

Figure 50: Younger and Older Average MI Differences Group median MI differences between Younger minus Older adults.

Comparing group medians can mask the large inter-subject variability in responses. To explore differences between groups more thoroughly, we calculated for each individual in each age group separately the sum of MI values within 3 ROI feature masks – the left eye, right eye and mouth (see Chapter 3:

Feature of Interest Analysis). We also calculated Cliff's delta as a measure of effect size (see Chapter 3 *Between-group comparisons*).

As can be seen in Figure 51, whilst the younger participants had a comparatively stronger relationship between mouth visibility and reaction times than older adults when comparing group medians (as denoted by the red rectangle), there was a large overlap between younger and older participants at the individual subject level, with the upper estimate of Cliff's delta approaching 0. We found weak estimates of effect size when comparing age-related differences in medians and a large overlap in the spread of results between younger and older participants for all other behavioural comparisons. This suggests that older adults did not have weaker sensitivity than younger adults to the eyes in terms of behavioural responses, though reaction times tended to be less modulated by the mouth in the **EXNEX** task.

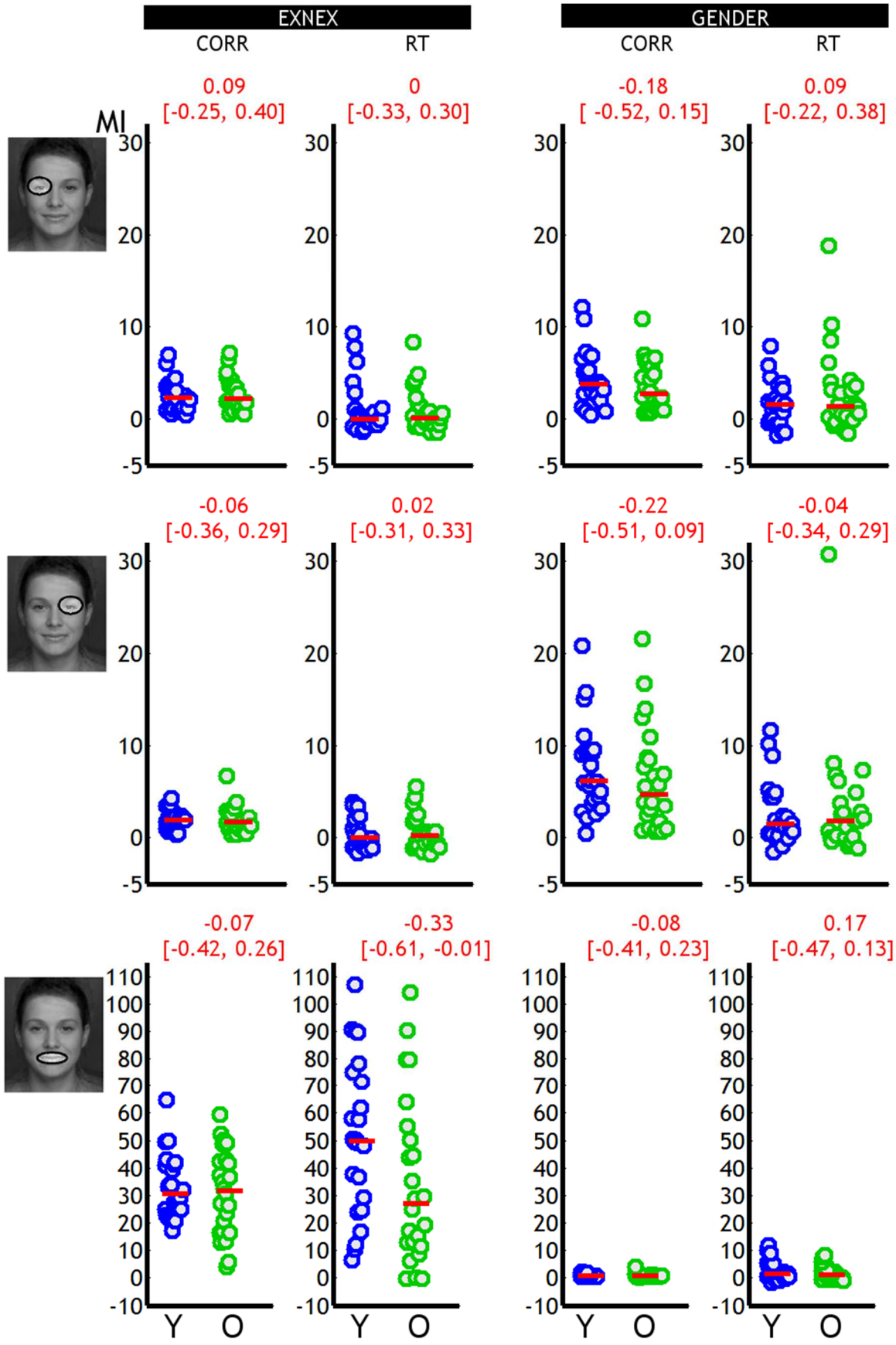


Figure 51: Younger and Older adult individual sum MI within regions of interest For each task and each behavioural comparison, we calculated the sum of all MI values within each feature mask for younger (blue) and older (green) participants. Red bars indicate group median. Red values are Cliff's delta measure of effect size with 95 % percentile bootstrapped confidence interval.

Reverse Analysis: Behavioural Results

The presence of the mouth was associated with behavioural responses in the **EXNEX** task, whilst the right eye was weakly associated with correct responses in the **GENDER** task for some participants. However, MI is directionless, in that higher MI values can reflect either the presence or the absence of a feature modulating behavioural responses. Using a reverse analysis we quantified by how much changing the amount of information about the left eye, right eye or mouth influenced participants' behavioural responses. This reverse analysis also helps appreciate the effects on more natural scales.

Using the feature of interest masks we calculated on a trial-by-trial basis how the visibility of each feature modulated behavioural responses for each comparison of interest. To quantify the effect of visibility of each feature on behavioural judgments, for each participant separately, we calculated the **RT** and **CORRECT** difference between the 10th bin (highest visibility) and the 1st bin (least visibility) for each feature mask. For each comparison and individual participant we also performed a permutation test (see Chapter 3: *Feature of Interest Analyses*).

MOUTH VISIBILITY

In the **EXNEX** task, increased mouth visibility (Figure 52, top panel) was associated with quicker responses (median difference in bin 10 minus bin 1 = -158 ms [-216, -100]). Nearly all (23/24) participants were on average faster on bin 10 than bin 1 trials, and these differences were statistically significant in 21/24 participants. Increased mouth visibility was also associated with increased accuracy (median difference in bin 10 minus bin 1 = 0.30 PP [0.26, 0.34]). All participants (24/24) were on average more accurate on trials where there was increased mouth visibility, which was statistically significant in 20/24 participants. On trials with little to no mouth visibility, some participants performed below chance level, and all participants showed a marked decline in response accuracy. These results suggest that in the absence of the mouth as the most diagnostic feature, other features and their combination, for at least some participants, were insufficient for categorisation. This is in contrast to our younger participants' who were all able to use other features and their combinations in the absence of the mouth. This suggests that the absence of

mouth visibility has a greater effect on accuracy in our older than younger participants, and that older adults may rely more upon the presence of the mouth to make correct responses. In contrast younger participants could use other features and their combination to correctly discriminate 'happy' from 'neutral' faces.

In the **GENDER** task (Figure 52, bottom two panels) increased mouth visibility had weak and variable effects on behaviour. Only 13/24 participants demonstrated faster responses with increased mouth visibility (median difference in bin 10 minus bin 1 = -14 ms, [-46, 18]). MI was significant for 5/24 participants, 4 of whom were faster with increased mouth visibility, and 1 slower with increased mouth visibility.

It is possible that task order might explain the directionality of effect that increased mouth visibility had on reaction times. For example, completing the **EXNEX** task where visibility of the mouth strongly modulated reactions times (see Figure 49) first, may have resulted in carry-over effects to the **GENDER** task, even when the feature was no longer task relevant. Older adults have been shown to have a deficit in suppressing task-irrelevant information, even with prior knowledge of stimulus relevance (Zanto, Hennigan, Östberg, Clapp, & Gazzaley, 2010) thus a failure to suppress attention to previously task-relevant information is plausible. In our data set however this does not seem be the case. Of those participants demonstrating faster responses in the **GENDER** task with increased mouth visibility only 7/13 participants completed the **EXNEX** task first, where the mouth was the most diagnostic feature for the task. Of those showing significant MI, 3/4 participants whom were faster with increased mouth visibility actually completed the **GENDER** task first. There was no discernible pattern of cross over effect between tasks in our older adult sample (see Supplementary 21).

Variability in the **GENDER** task was also evident in the association between mouth visibility and correct responses. The group was evenly split between participants demonstrating an increase/decrease in accuracy with increased mouth visibility (median difference in bin 10 minus bin 1 = -0.01 PP [-0.06, 0.04]). MI was not significant for 11/24 participants who were more accurate with increased mouth visibility. MI however was significant for 2 participants,

with a difference in the opposite direction (i.e. less accurate responses with increased mouth visibility).

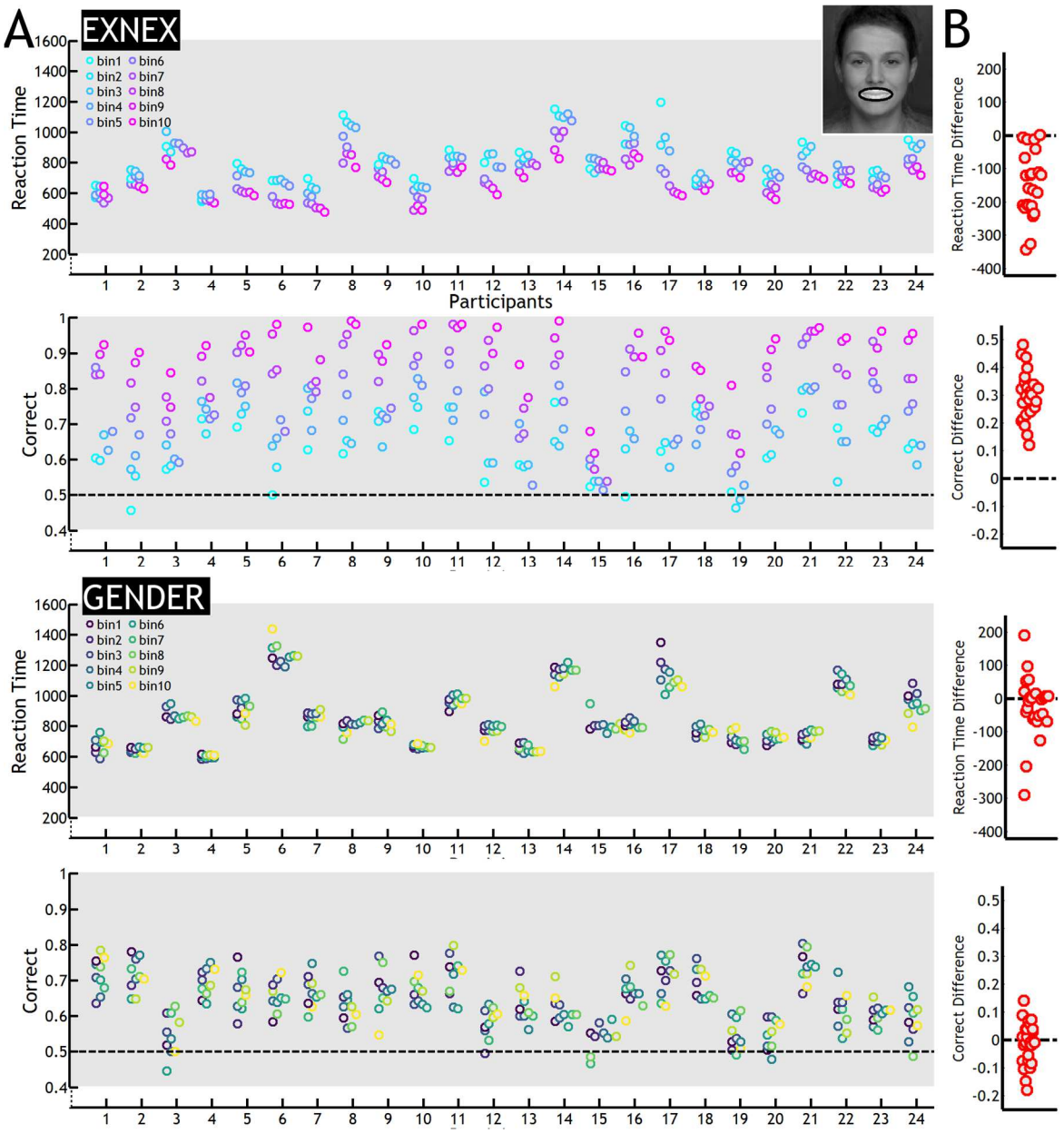


Figure 52: Older adult reverse analysis of behavioural responses by mouth visibility Panel A: Individual participant results for each of 10 visibility bins. Top two panels show **EXNEX** results. Bottom two panels show **GENDER** results. Panel B: Each circle represents one participant's difference between bin 10 minus bin 1.For younger adult results see Figure 32 (page 109).

In summary, increasing mouth visibility was associated with increased accuracy (24/24 participants) and faster responses (23/24 participants) in the **EXNEX** task. When there was little to no mouth visibility, some participants performed below chance level, confirming that the mouth is a diagnostic feature for this task, and that older adults rely more upon the presence of the mouth for making correct judgments than younger participants. In the **GENDER** task, there was a near

even split of participants in both directions on speed and accuracy modulations with increased mouth visibility. Whilst some participants were able to use the mouth as a diagnostic feature for gender categorisation, there did not appear to be a task-order cross-over effect, but rather individual variation.

LEFT EYE VISIBILITY

Left eye visibility (from the viewers perspective) has been shown to modulate reaction times in a face versus noise detection task (Rousselet et al., 2014). Visibility of the left eye has also been implicated in quicker and more accurate responses in gender discrimination tasks (Joyce et al., 2006). In our results, visibility of the left eye (Figure 53) had a very variable effect on responses times and accuracy in the two tasks.

Increased visibility of the left eye in the **EXNEX** task (Figure 53, top panel) was very weakly associated with faster responses for 8/24 participants (median difference in bin 10 minus bin 1 = -8 ms [-21, 6]). MI was significant for 3/24 participants, all of whom were faster with increased left eye visibility. Increased visibility of the left eye increased accuracy for 18/24 participants (median difference in bin 10 minus bin 1 = 0.04 [0.01, 0.06]). MI was significant for 2/24 participants both of whom showed increased accuracy with increased left eye visibility.

In the **GENDER** task, increased visibility of the left eye (Figure 53, panel 3) was associated with faster reaction times (median difference in bin 10 minus bin 1 = -25 ms [-45, -5]) for 18/24 participants. MI was significant for 6/24 participants, 5/6 were faster with increased left eye visibility, 1/6 was slower with increased left eye visibility. Increased left eye visibility was also associated with slightly increased accuracy (median difference in bin 10 minus bin 1 = 0.09 [0.06, 0.12]) for 22/24 participants. MI was significant for 5/24 participants, all of whom were more accurate with increased visibility of the left eye.

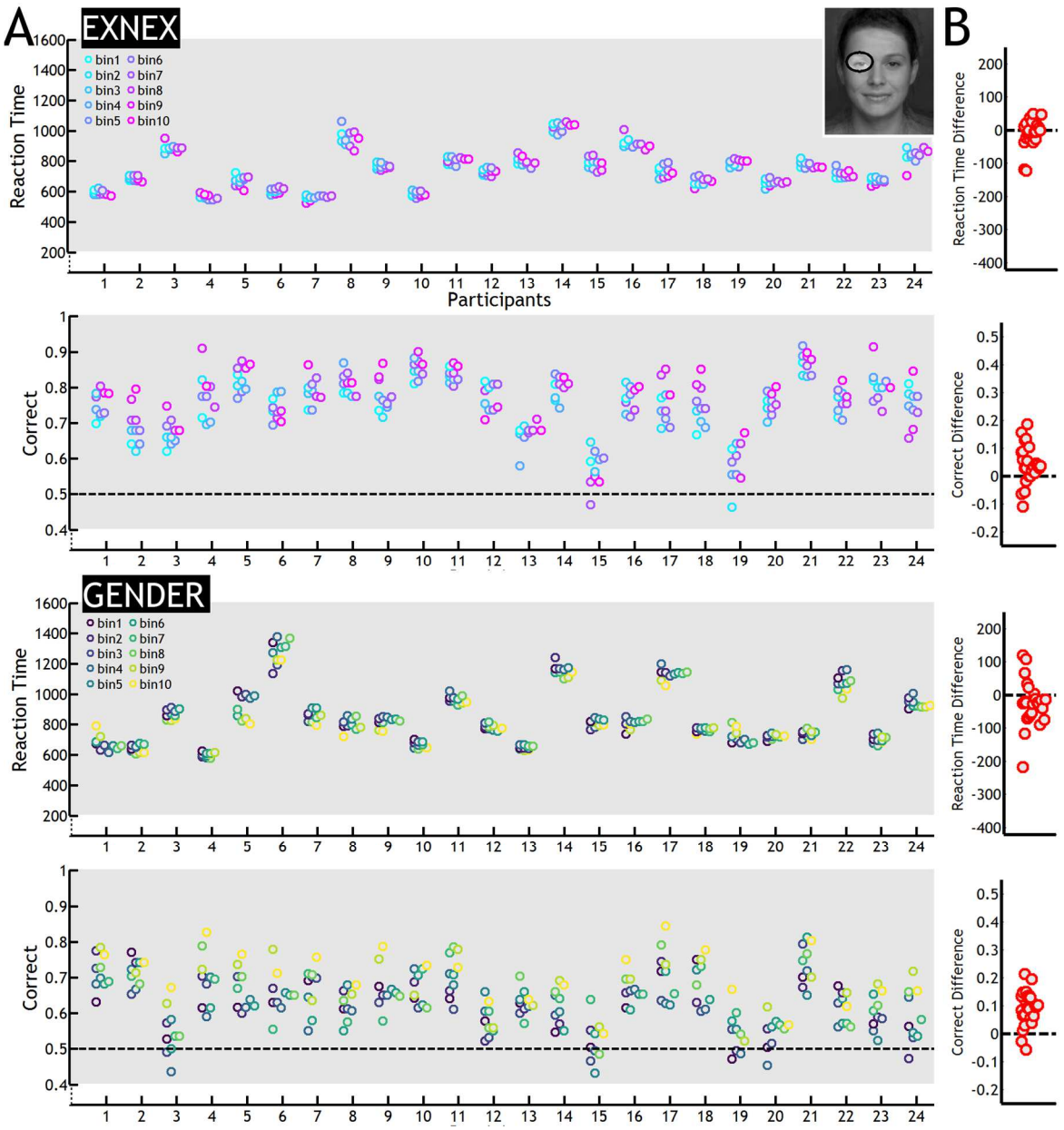


Figure 53: Older adult reverse analysis of behavioural responses by left eye visibility Panel A: Individual participant results for each of 10 visibility bins. Top two panels show **EXNEX** results. Bottom two panels show **GENDER** results. Panel B: Each circle represents one participant's difference between bin 10 minus bin 1. For younger adult results see Figure 33 (page 111).

This is inconsistent with previous reports suggesting that the most diagnostic feature in a gender discrimination task is the left eye (Schyns et al., 2002). Whilst most of our participants were on average more accurate on bin 10 than bin 1 trials, there was large inter-trial variability leading to a few participants having significant MI differences. Whilst we controlled for the degree of visibility of the left eye, trials binned on this basis reflect a mixture of visibility of other facial features and their combinations – for example on bin 1 trials with limited to no visibility of the left eye, there may have been visibility of the right eye, mouth, hair etc. or combination of these features. As most of our older adults

managed to maintain an above chance accuracy even with no visibility of the left eye, it may be that for the majority of our participants some combination of visibility of other features (such as the right eye and/or mouth) was enough to resolve the task in the absence of the left eye being visible.

In summary, increased visibility of the left eye in the **EXNEX** task was weakly associated with quicker reaction times and increased accuracy. In the **GENDER** task, increased visibility of the left eye was associated with quicker reaction times (18/24 participants) and more accurate responses (22/24 participants). There was no apparent effects of task order (see Supplementary 3).

Previous work has suggested that the left, but not right eye region drives accurate judgements in male versus female categorisation tasks (Schyns et al., 2002). In our results, nearly all (22/24) participants showed an increase in accuracy with increased left eye visibility in the gender task. However, our results are in contrast to previous work which has suggested that correct categorisation of expressive/non-expressive faces does not involve either of the eye regions (Schyns et al., 2002), as in our results increased visibility of the left eye was related to increased accuracy for over half (18/24) of our participants.

RIGHT EYE VISIBILITY

Whilst left eye visibility has been shown to predominantly modulate behavioural responses in **GENDER** categorisation tasks (Schyns et al., 2002), a minority of participants demonstrate a converse right eye bias in face versus noise detection tasks (Rousselet et al., 2014). In our results, we found mixed effects of increasing right eye visibility in the **EXNEX** and **GENDER** tasks (Figure 54).

In the **EXNEX** task, increased visibility of the right eye (Figure 54, top panel) was associated with faster reaction times for 11/24 participants (median difference in bin 10 minus bin 1 = -5 ms [-25, 35]). MI was significant for 5/24 participants, 4 of whom were faster with increased visibility of the right eye. Increased visibility of the right eye was also associated with increased accuracy for 17/24 participants (median difference in bin 10 minus bin 1 = 0.04 [<0.01, 0.08]). MI was significant for 2/24 participants. 1 participant was more accurate with increased right eye visibility, whilst the other participant was less accurate.

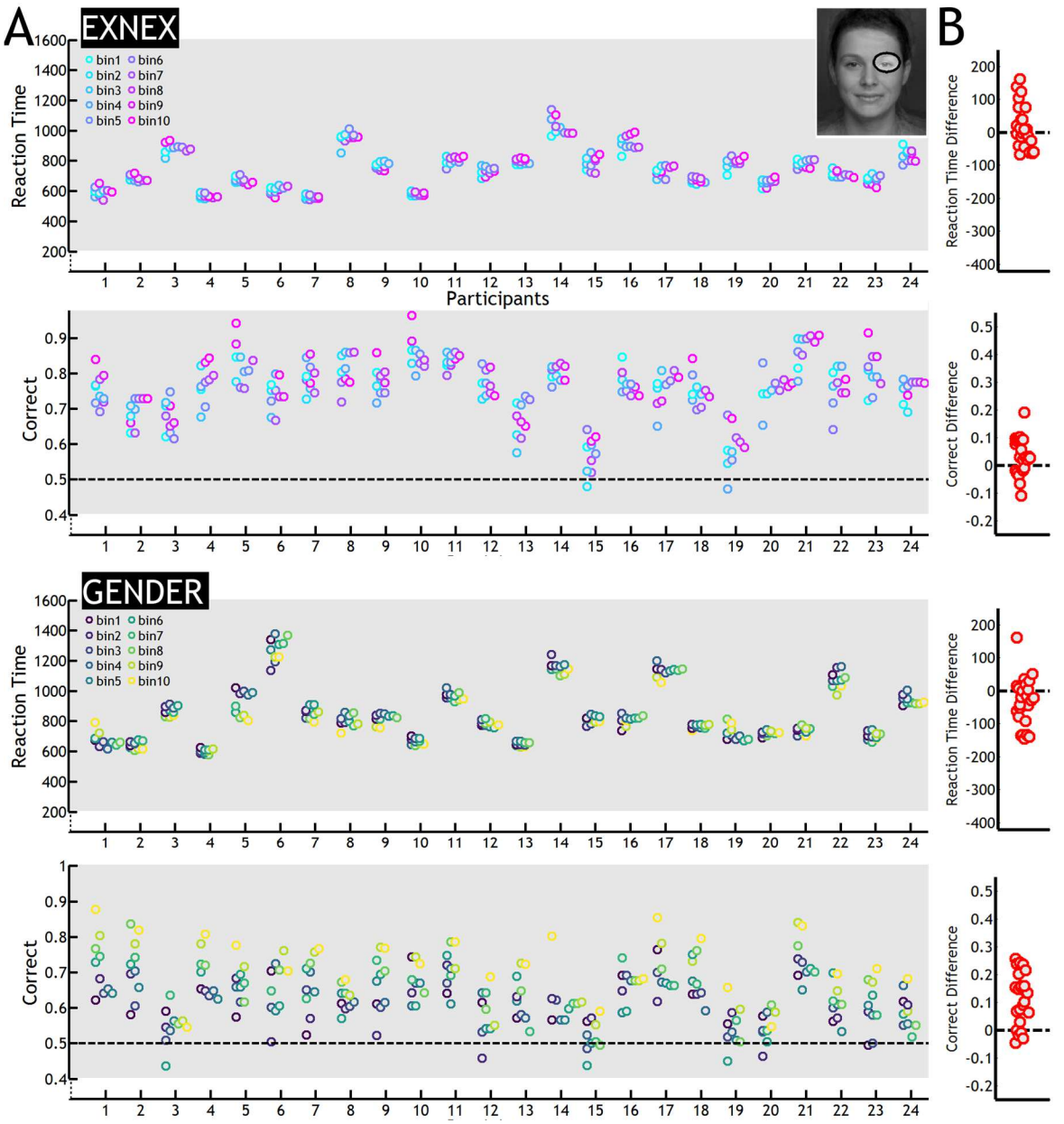


Figure 54: Older adult reverse analysis of behavioural responses by right eye visibility Panel A: Individual participant results for each of 10 visibility bins. Top two panels show **EXNEX** results. Bottom two panels show **GENDER** results. Panel B: Each circle represents one participant's difference between bin 10 minus bin 1. For younger adult results see Figure 34 (page 113).

In the **GENDER** task, increased visibility of the right eye (Figure 54, panel 3) was associated with faster responses for 15/24 participants (median difference in bin 10 minus bin 1 = -28 ms [-65, 8]). MI was significant for 5/24 participants, 3/5 participants were faster with increased right eye visibility, whilst 2/5 participants were slower with increased right eye visibility. Increased visibility of the right eye also increased accuracy for most (19/24) participants (median difference in bin 10 minus bin 1 = 0.13 [0.07, 0.18]). MI was significant for 6/24 participants, all of whom were more accurate with increased visibility of the right eye.

In summary, in the **EXNEX** task increased visibility of the right eye was mixed. Whilst some participants were faster and more accurate with increased visibility of the right eye, some participants showed differences in the opposite direction. Increased visibility of the right eye during the **GENDER** task was similarly mixed, with some participants demonstrating faster responses, whilst other participants were slower. Nearly all (19/24) participants were more accurate with increased visibility of the right eye. There was no apparent effect of task order (see Supplementary 23).

We have demonstrated that in the **EXNEX** task, increased mouth visibility was associated with quicker and more accurate responses. Increased visibility of the left eye was only weakly associated with quicker reaction times and increased accuracy in the **EXNEX** task. Increased visibility of the right eye increased accuracy and decreased reaction times in some participants in the **EXNEX** task, though some participants showed differences in the opposite direction. In the **GENDER** task, modulation of behavioural responses showed approximately an even split between increases/decreases in accuracy and reaction time with increased mouth visibility. Increased visibility of the left eye was associated with quicker and more accurate responses for most participants. Increased visibility of the right eye had a mixed effect on reaction times, with some participants demonstrating faster responses. Nearly all participants though were more accurate with increased visibility of the right eye.

FEATURE VISIBLITY BETWEEN-TASK DIFFERENCES

Having calculated for each individual the difference in accuracy and reaction time for bin 10 minus bin 1 trials for each feature of interest (Figure 52 - Figure 54), we now moved to compare these differences between tasks. For each participant we compare how feature visibility affected reaction times and accuracy differentially for the two tasks (Figure 55 - Figure 56).

Comparing differences in reaction times (Figure 55), the majority of participants (20/24 participants) were faster with increased right eye visibility in the **GENDER** than **EXNEX** task, though 1 participant appears to be an outlier showing a stronger effects in the opposite direction. This suggests that increased visibility of the right eye area was more diagnostic for the **GENDER** than **EXNEX**

task. Similarly, 17/24 participants were faster with increased left eye visibility in the **GENDER** than **EXNEX** task, again suggesting that increased visibility of the eye area is more task diagnostic for the **GENDER** than **EXNEX** task. Whilst 1 participant may be an outlier due to a larger comparative difference in reaction times between bin 10 and bin 1 trials, their pattern of behaviour is in line with other participants (i.e. faster with increased left eye visibility in the GENDER than EXNEX task).

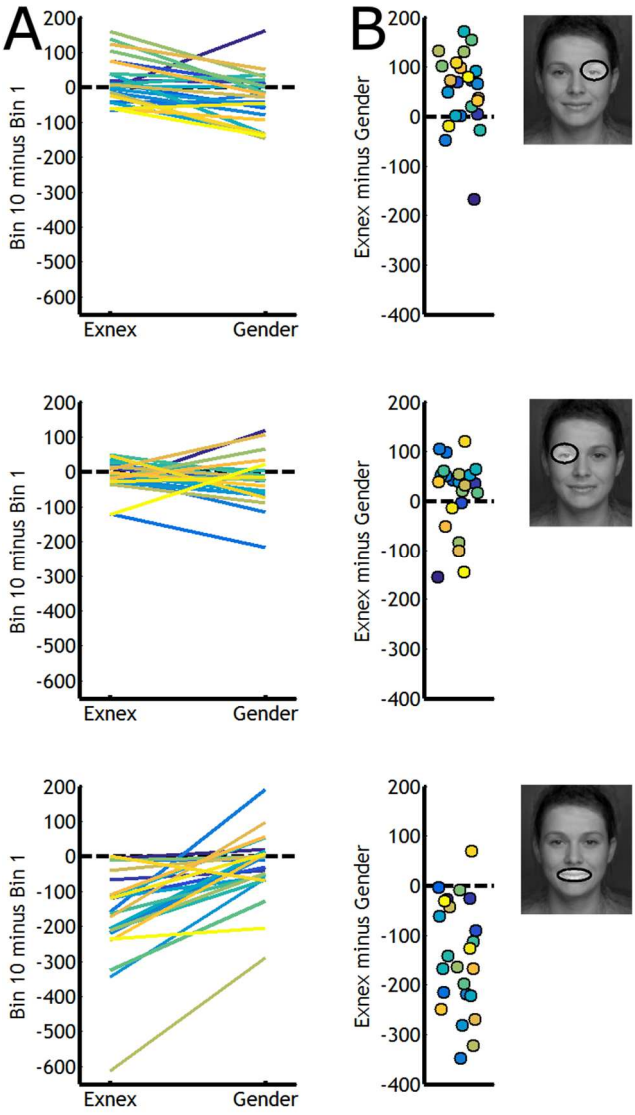


Figure 55: Older adult reaction time task differences by feature visibility Panel A: The difference in reaction time (ms) between Bin 10 minus Bin 1 in the **EXNEX** and **GENDER** task for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in reaction time (Bin 10 minus bin 1) in the **EXNEX** minus **GENDER** task for each facial feature of interest. For younger adult results see Figure 35 (page 115).

Conversely, all but 1 participant was faster with increased mouth visibility in the EXNEX than GENDER task, highlighting the task diagnosticity of the mouth region

for the EXNEX but not GENDER task. Whilst 1 participant may be an outlier due to a larger comparative difference in reaction times between bin 10 and bin 1 trials, their pattern of behaviour is in line with other participants (i.e. faster with increased mouth visibility in the EXNEX than GENDER task).

Comparing differences in accuracy (Figure 56), increased visibility of the right eye resulted in increased gains in accuracy for the **GENDER** than the **EXNEX** task. The majority (19/24) of participants demonstrated a larger increase in accuracy with increased right eye visibility in the **GENDER** than **EXNEX** task, suggesting that visibility of the right eye had more of an effect on accuracy in the **GENDER** than **EXNEX** task, and hence more diagnostic value. Similarly, increased visibility of the left eye had more of an effect on accuracy in the **GENDER** than **EXNEX** task. The majority (19/24) of participants demonstrated a larger increase in accuracy with increased left eye visibility in the **GENDER** than **EXNEX** task, again demonstrating that visibility of the left eye had more of an effect on accuracy in the **GENDER** than **EXNEX** task, and hence more diagnostic value. Conversely, all (24/24) participants demonstrated a larger increase in accuracy with increased mouth visibility in the **EXNEX** than **GENDER** task. This suggests that mouth visibility had more of an effect on accuracy in the **EXNEX** than **GENDER** task and more diagnostic value.

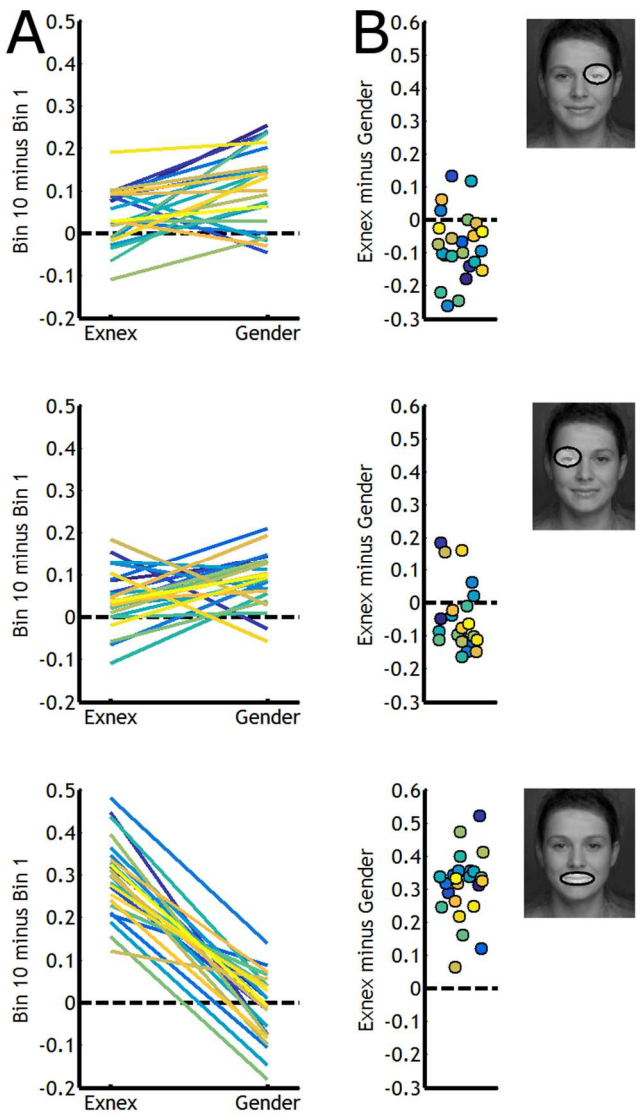


Figure 56: Older adult accuracy task differences by feature visibility Panel A: The difference in accuracy between Bin 10 minus Bin 1 in the **EXNEX** and **GENDER** task for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in accuracy between Bin 10 minus bin 1 for the **EXNEX** minus **GENDER** task for each facial feature of interest. For younger adult results see Figure 36 (page 116).

In summary, increased visibility of the left and right eye led to faster response times and increased accuracy in the **GENDER** than **EXNEX** task, whilst the opposite was true with increased mouth visibility. As the eye region is more diagnostic for the **GENDER** than **EXNEX** task, increased visibility of either eye region has more of a reaction time and accuracy advantage for the **GENDER** than **EXNEX** task. Conversely, the mouth region is more diagnostic in the **EXNEX** than **GENDER** task. Increasing mouth visibility region has more of a reaction time and accuracy advantage for the **EXNEX** than **GENDER** task.

To ensure these differences between tasks were not due differences in the variability of reaction time and accuracy scores between the two tasks, we normalised our data. For each participant, behavioural comparison and feature we normalised our data between 1 and -1 by dividing the difference between bin 10 minus bin 1, by the sum of bin 1 and bin 10. Normalising our data did not change the pattern of differences (Supplementary 24 - Supplementary 25).

Young minus Older Behavioural Differences

We have seen that behaviourally our group of older adults are faster and more accurate in the **GENDER** task when there is increased visibility of the left or right eye region, and faster and more accurate in the **EXNEX** task when there is increased mouth visibility region.

Next, we wanted to quantify how the modulation of behaviour by facial feature visibility compares between our older and younger (see Chapter 4) participants. We begin here by comparing differences in reaction time and accuracy for the two groups. We consider comparisons of MI later.

For each facial feature (left eye, right eye, mouth) we compared distributions of the difference between bin 10 minus bin 1 visibility of each facial feature. We then calculated the median of these differences between age groups (**Younger** minus **Older**) and the corresponding confidence interval (Figure 57). For each behavioural comparison, we also calculated a corresponding Cliff's *delta* with 95 % confidence interval as a measure of effect size (Table 9) of the between-group differences (see Chapter 3: *Between-group comparisons*).

	Reaction Times		Accuracy	
	EXNEX	**GENDER**	**EXNEX**	**GENDER**
Left Eye	-0.01 [-0.34, 0.33]	-0.01 [-0.38, 0.31]	-0.14 [-0.46, 0.20]	-0.23 [-0.55, 0.13]
Right Eye	-0.22 [-0.52, 0.14]	-0.08 [-0.40, 0.25]	-0.03 [-0.37, 0.28]	-0.15 [-0.48, 0.18]
Mouth	0.24 [-0.09, 0.55]	0.18 [-0.15, 0.50]	-0.28 [-0.58, 0.03]	0.07 [-0.28, 0.39]

Table 9: Older and Younger adult behavioural differences effect sizes Cliff's delta between-group comparisons of reaction times and accuracy. Square brackets indicate 95 % percentile bootstrap confidence interval with 500 samples.

In the **EXNEX** task (Figure 57 left panel) increased visibility of the left eye (bin 10 minus bin 1) leads to negligible group differences in reaction time

modulations (younger minus older difference of medians = -0.65 ms [-16, 15]) and accuracy modulations (-0.01 PP [-0.04, 0.02]). Increased visibility of the right eye had more of an effect, with increased visibility leading to a stronger decrease in reaction times for younger than older participants (-13.47 ms [-4, 8]), though accuracy differences were still negligible (-0.01 PP [-0.04, 0.02]) and estimated effect sizes for both comparisons were weak. Increased mouth visibility had more of an effect on reaction times for older than younger participants, with older adults having comparatively quicker responses for bin 10 than bin 1 trials than younger adults (41.05 ms [-23.53, 99.71]), though in absolute terms younger adults were still faster than older adults. This may suggest that older adults in particular rely on visibility of the mouth region for making quicker responses, though differences in accuracy modulations were negligible (-0.04 PP [-0.09, 0.01]) suggesting that older adults could still resolve the task without visibility of the mouth, albeit slower. In both cases, estimated effect sizes for both comparisons crossed 0.

In the **GENDER** task (right panel) increased visibility of the left eye (bin 10 minus bin 1) lead to negligible differences in reaction time modulations (younger minus older difference of medians = -0.81 ms [-23.16, 24.23]) and accuracy modulations (-0.02 PP [-0.05, 0.01]). Increased visibility of the right eye had more of an effect, with increased right eye visibility being associated with more of a decrease in reaction times for younger than older participants (-8.19 ms [-38.37, 23.17]), though accuracy differences were still negligible (-0.02 PP [-0.06, 0.03]) and for both comparisons estimated effect sizes crossed 0. It may be that younger adults were faster to use the right eye as an additional cue for resolving that task than older participants. Increased mouth visibility had more of an effect on reaction times for older than younger participants, with older adults having comparatively quicker responses for bin 10 than bin 1 trials than younger adults (15.98 ms [-13.63, 43.50]), though in absolute terms younger adults were still faster than older adults. This may suggest that older adults could rely on visibility of the mouth region for making quicker responses, though differences in accuracy were negligible (0.01 PP [-0.03, 0.06]), suggesting that older adults could still resolve the task without visibility of the mouth, albeit slower. In both cases, estimated effect sizes for both comparisons crossed 0.

We found an interaction between age group and task differences for correct performance with increased levels of mouth visibility. Older adults' correct performance relied more heavily upon visibility of the mouth than younger adults in the **EXNEX** but not **GENDER** task. Older adults accuracy in the **EXNEX** task was very high on trials with high visibility on the mouth region (refer to Figure 52) achieving levels of performance similar to younger adults. When the mouth was barely visible, older adults' performance dropped substantially, where as in younger adults they continued to be able to resolve the task from other features and their combination.

In summary, both younger and older adults' behavioural responses were modulated heavily by the mouth region in the **EXNEX** task, with both younger and older adults showing a marked increase in accuracy and quicker responses with increased mouth visibility region. This pattern was stronger for the older adults, who showed a larger increase in accuracy and larger decrease in reaction times with increased mouth visibility in the **EXNEX** task than younger adults. All other behavioural differences were comparatively weak. However, for all behavioural comparisons our estimates of effect size crossed 0.

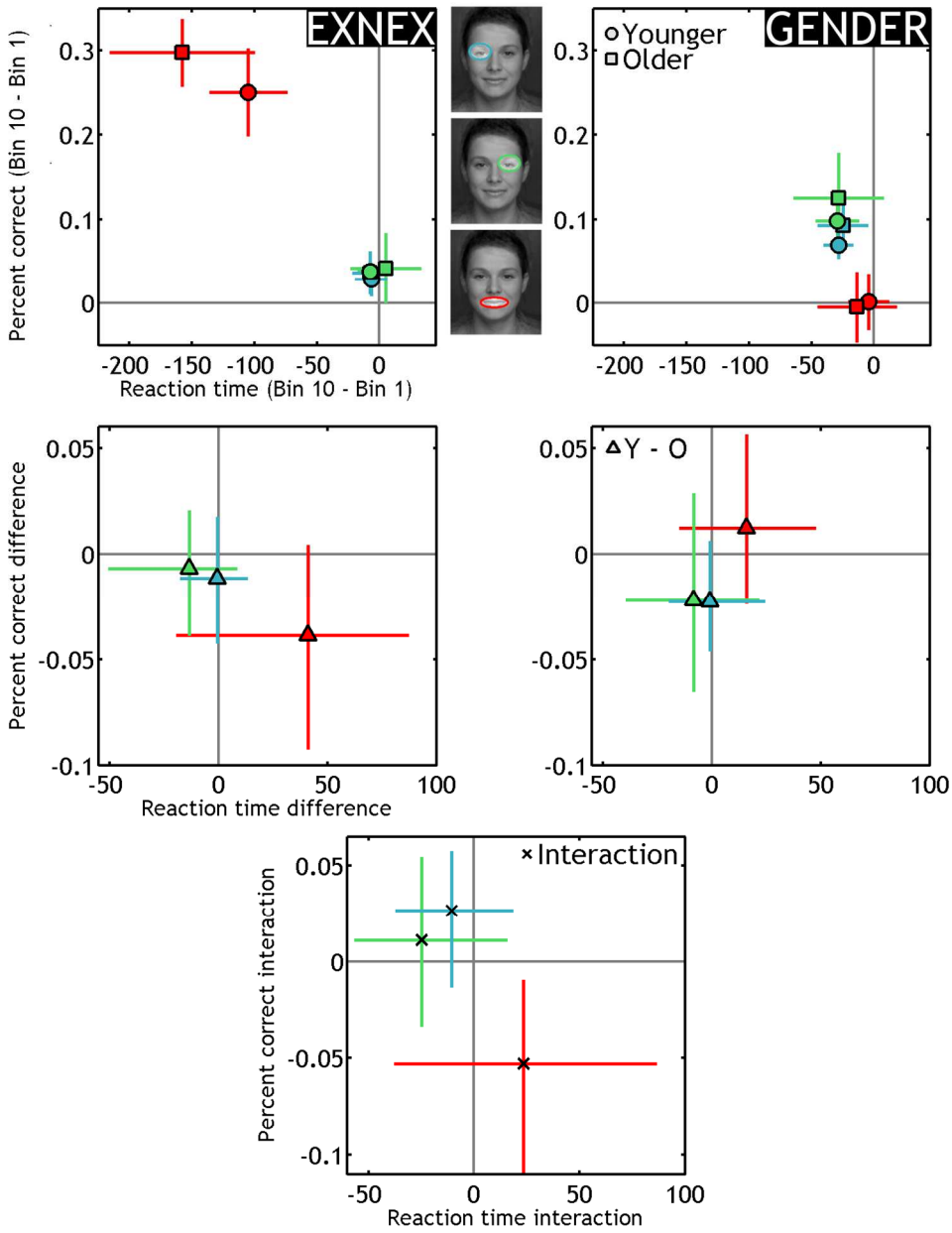


Figure 57: Task by age group interaction of behavioural differences. Top panel: Differences in reaction time (ms) and accuracy between bin 10 minus bin 1, for younger (circles) and older (square) participants. For each facial feature (left eye = blue, right eye = green, mouth = red) the median difference between bin 10 minus bin 1 for each age group is shown. Bottom panel: Differences between younger minus older medians. Vertical and horizontal lines represent the 95 % confidence interval.

ERP Results

We have seen that, behaviourally, in the **EXNEX** task increased mouth visibility resulted in quicker and more accurate responses for both age groups, an effect that was stronger in older adults. Increased visibility of the left and right eye led to quicker and more accurate responses for most, but not all, of our older and younger participants, with negligible differences between age groups. In the **GENDER** task, increased mouth visibility had negligible effects on accuracy and reaction time for both age groups. Increased visibility of the left and right eye led to quicker and more accurate responses for most older adults and nearly all younger adults, but group differences were not significant.

Next, we will consider ERP results. We predicted that if the N170 eye sensitivity is task-independent, then during both tasks there would be evidence of contralateral eye sensitivity prior to sensitivity to other facial features (such as the mouth), regardless of the extent to which the eyes are task-relevant. We also predicted that whilst older adults would process the same facial features as younger adults (i.e. the mouth in the **EXNEX** task), feature encoding would be delayed and weaker in older compared to younger adults. We aimed to quantify these differences.

Average ERP Timecourses

The bubbles sampling technique could affect the shape of ERPs differently in older compared to younger adults. As such, we started by comparing ERP time courses in the two tasks for all non-bubble i.e. practice trials and bubble trials in older compared to younger adults to establish if there was an age-related differences in the effect of the bubble manipulation on ERP timecourses. We found that whilst older adults N170 was no more delayed than younger participants on bubble compared to non-bubble trials, the amplitude of the N170 was affected differentially for older compared to younger adults.

As we have already demonstrated, older adults experienced reduced accuracy and slower reaction times to bubble compared to practice trials. Diminished behavioural performance in older adults exceeded the corresponding reductions in speed and accuracy in younger adults. As well as affecting behavioural

performance, bubbles can also affect the timing of ERPs. Previous work (Jaworska, 2017) has demonstrated in a face detection task that both older and younger adults ERPs are delayed in bubble compared to non-bubble face trials. Whilst older adults N170 latencies to faces were delayed compared to younger adults in both bubble and non-bubble trials, this delay was similar between practice and non-practice face trials suggesting that the N170 is no more delayed by bubbles in older compared to younger participants.

We demonstrated a delay in bubble compared to non-bubble trials in our younger participants (see Chapter 4). To understand if older adults experienced a comparable delay, and to ensure older adults were not more adversely affected by the bubbles manipulation we computed average ERPs for each older adult participant, in each task, for the left and right hemisphere (see Chapter 3: *Average ERPs* for details). We also calculated the median N170 latency and corresponding amplitude for each participant in each task, for the left (LE) and right (RE) hemisphere separately (see Chapter 3: *N170 amplitude and Latency calculations*).

We present, for each task, the group average mean ERP's for LE and RE (Figure 58) with 95 % confidence intervals, and corresponding median N170 latency and amplitude for practice and bubble trials (Table 10).

		EXNEX		GENDER	
		LE	**RE**	**LE**	**RE**
Practice	**LAT**	151.4 [147.3, 155.4]	151.7 [147.3, 156.2]	152.6 [148.5, 156.7]	152.3 [146.4, 158.1]
	AMP	-9.4 [-11.8, -7.0]	-10.0 [-12.0, -8.0]	-9.5 [-10.9, -8.2]	-10.0 [-12.0, -8.0]
Bubble	**LAT**	194.0 [184.3, 203.7]	200.9 [184.8, 217.0]	193.4 [178.1, 208.6]	195.1 [181.7, 208.5]
	AMP	-5.6 [-7.3, -4.0]	-7.0 [-8.6, -5.3]	-6.8 [-8.7, -4.9]	-7.5 [-9.2, -5.8]
Difference	**LAT**	-40.5 [-46.9, -34.0]	-42.4 [-55.1, -29.6]	-36.0 [-45.4, -26.6]	-39.2 [-48.9, -29.5]
	AMP	-3.3 [-4.9, -1.8]	-2.9 [-5.2, -0.6]	-3.3 [-5.9, -0.7]	-2.2 [-4.3, 0]

Table 10: Older adults N170 amplitude and latency Median N170 amplitude (AMP) and latency (LAT) for practice and bubble trials for the left and right hemisphere in the EXNEX and GENDER task. Square brackets indicate 95 % confidence interval.

Compared to practice trials (i.e. non-bubble trials), the latency of the N170 in bubble trials was delayed to a similar degree across both tasks and hemispheres (**EXNEX** LE Difference [Practice – Bubble trials] = -40.5 ms [-46.9, -34.0]; **EXNEX** RE Difference [Practice – Bubble trials] = -42.4 ms [-55.1, -29.6]; **GENDER** LE Difference [Practice – Bubble trials] = -36.0 ms [-45.4, -26.6]; **GENDER** RE Difference [Practice – Bubble trials] = -39.2 ms [-48.9, -29.5]). This suggests a general delay in the N170 to bubble stimuli that is not task or hemisphere specific.

There was a decrease in the minimum amplitude of the N170 between practice and bubble trials (**EXNEX** LE Difference [Practice – Bubble trials] = -3.3 μV [-4.9, -1.8]; **EXNEX** RE Difference [Practice – Bubble trials] = -2.9 μV [-5.2, -0.6]; **GENDER** LE Difference [Practice – Bubble trials] = -3.3 μV [-5.9, -0.7]; **GENDER** RE Difference [Practice – Bubble trials] = -2.2 μV [-4.3, 0]). This is the opposite effect seen in younger participants whose N170 amplitude increased on bubble compared to practice trials (see Chapter 4: *Average ERPs for Practice and Bubble trials*). This may be an effect of occlusion of the image. Individual ERP plots are provided in Supplementary 26 - Supplementary 28.

For comparison we also calculated group averaged ERP waveforms using a 20 % trimmed mean. Trimmed means are less affected by outliers and provide a better estimation of the location of the bulk of observations. This may improve the signal to noise ratio of averaged ERP waveforms. We found no difference in the pattern of results when using a 20 % trimmed mean (Supplementary 29).

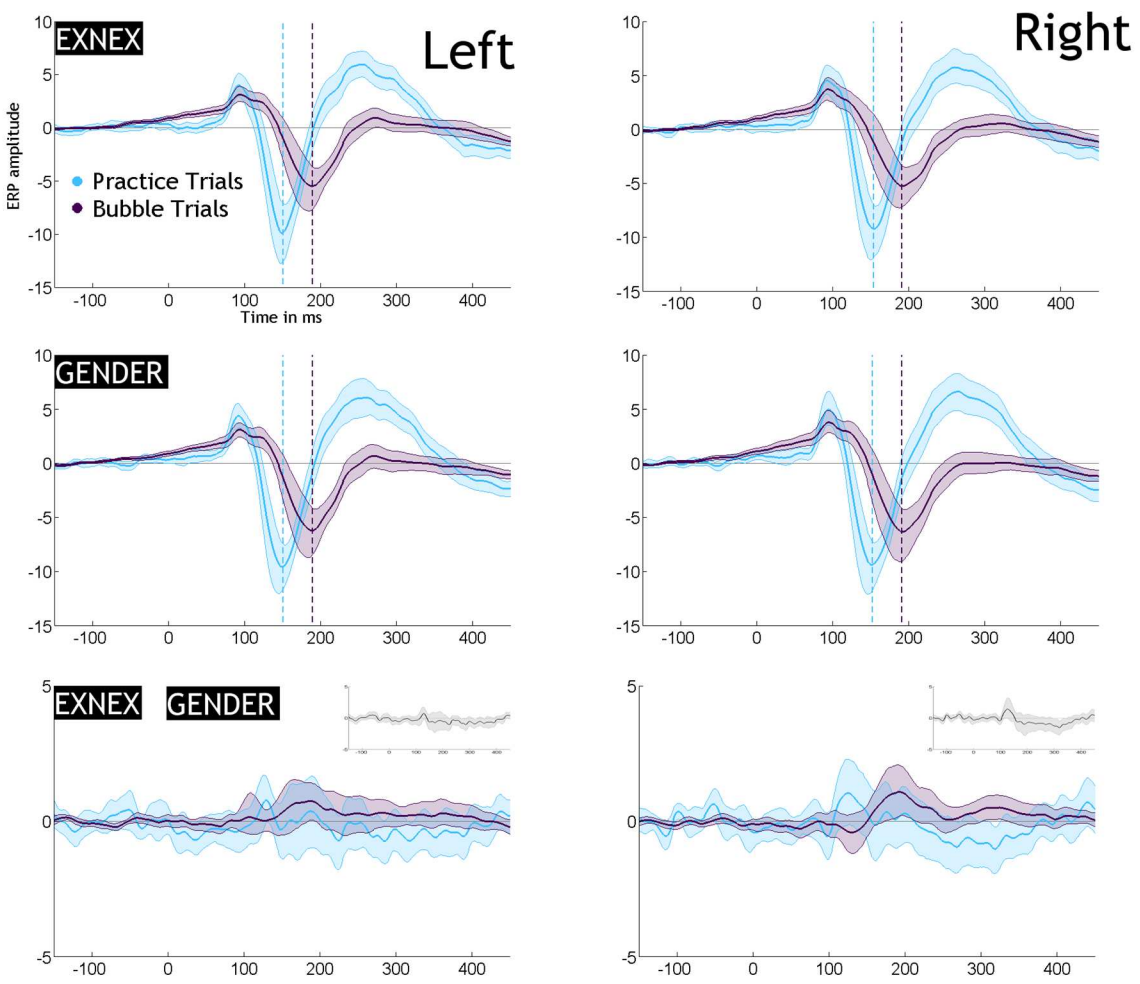


Figure 58: Older adult average group ERP in Bubble and Non-Bubble Trials Mean bubble and non-bubble trial ERPs for the left and right hemisphere with 95 % confidence intervals around the mean. Vertical lines represent the minimum amplitude peak of the N170 for each task. Bottom panel **EXNEX** minus **GENDER** for bubble and practice trials. Small grey plot shows the pairwise difference of practice minus bubbles trials.

Next we examined how younger and older adults' differences in mean ERPs in practice (Figure 59) and bubble (Figure 60) trials compared. For each comparison of interest, we calculated the group difference between younger and older adult participants N170 latency and amplitude by estimating the Harrell-David median of all pairwise differences of the distribution of younger minus older participants. We computed a 95 % confidence interval around this estimate using a percentile bootstrap technique with 500 iterations. We also calculated the effect size (Table 11) of between-groups comparisons using Cliff's *delta* (See Chapter 3: *Between-group Comparisons*).

		EXNEX		GENDER	
		LE	**RE**	**LE**	**RE**
Practice	**LAT**	-0.6 [-0.8, -0.4]	-0.5 [-0.8, -0.3]	-0.6 [-0.8, -0.3]	-0.6 [-0.8, -0.3]
	AMP	0.5 [0.2, 0.7]	0.5 [0.2, 0.7]	0.6 [0.4, 0.8]	0.6 [0.4, 0.8]
Bubble	**LAT**	-0.4 [-0.7, -0.1]	-0.4 [-0.7, -0.2]	-0.4 [-0.6, 0]	-0.4 [-0.6, 0]
	AMP	-0.1 [-0.4, 0.2]	-0.1 [-0.4, 0.2]	0 [-0.3, 0.3]	0 [-0.2, 0.4]
Difference	**LAT**	-0.0 [-0.4, 0.3]	-0.1 [-0.4, 0.2]	-0.1 [-0.4, 0.2]	-0.1 [-0.4, 0.3]
	AMP	-0.7 [-0.9, -0.5]	-0.6 [-0.8, -0.4]	-0.6 [-0.8, -0.4]	-0.6 [-0.9, -0.3]

Table 11: Cliff's delta N170 effect size estimates We estimated Cliff's delta effect size of the difference between younger and older N170 amplitude and latency differences. Cliff's delta ranges from 1 (where all values from one group are higher than the values from the other group) to -1 (when all values from one group are lower than the values from the other group). Completely overlapping distributions have a Cliff's delta of 0.

Previous results have demonstrated that, compared to younger adults, older adults N170 latencies for practice (i.e. non-bubble) trials and non-practice (i.e. bubble trials) are delayed in a face versus noise detection task (Jaworska, 2017). We also observed this pattern of results, with a delay in the peak of the N170 in older compared to younger adults in both practice (**EXNEX LE** = -12 ms [-19.0, -6.3]; **GENDER LE** = -11.7 ms [-19.6, -5.1]; **EXNEX RE** = -10 ms [-17.4, -5.8]; **GENDER RE** = -12.2 ms [-21.2, -7.2]) and experimental (**EXNEX LE** = -13.3 ms [-22.7, -1.2]; **GENDER LE** = -13.2 [-27.4, -0.9]; **EXNEX RE** = -17.5 ms [-30.0, -4.9]; **GENDER RE** = -14.0ms [-27.4, -1.8]) trials. Compared to younger adults, older adults had an increased N170 amplitude i.e. more negative in practice (**EXNEX LE** = 4.6 μV [1.4, 6.8]; **GENDER LE** = 5.0 μV [2.0, 7.6]; **EXNEX RE** = 4.3 μV [1.2, 7.2]; **GENDER RE** = 6.4 μV [3.1, 8.9]) but not experimental (**EXNEX LE** = -0.9 μV [-3.1, 1.3]; **GENDER LE** = 0.1 μV [-2.1, 2.6]; **EXNEX RE** = -0.7 μV [-3.6, 1.4]; **GENDER RE** = 0.5 μV [-1.7, 2.7]) trials.

Next we calculated the difference between practice minus bubble trials, for Young minus Older participants. There was no group difference between the differences in N170 peak latency to practice and bubble trials in either the expression (**EXNEX LE** = -1.7 [-9.8, 6.4]; **EXNEX RE** = -4.6 [-14.3, 5.2]) or gender task (**GENDER LE** = -2.4 [-12.4, 6.5]; **GENDER RE** = -1.5 [-10.7, 6.0]). This suggests that the N170 is no more delayed by bubbles in older compared to younger participants and that any age-related delays in the processing of facial

features cannot be attributed to the presence of bubbles. There was however a group difference between the differences in N170 peak amplitude to practice and bubble trials in both expressive (**EXNEX LE** = -5.1 [-7.6, -2.7]; **EXNEX RE** = -5.7 [-8.4, -2.7]) and gender (**GENDER LE** = -4.9 [-7.1, -2.5]; **GENDER RE** = -5.3 [-8.1, -2.7]) tasks. This suggests that the N170 peak amplitude is significantly reduced in bubble compared to non-bubble trials for older compared to younger participants. This difference is due to two phenomena. Firstly, older participants demonstrated an enhanced N170 (Figure 59) to practice compared to bubble trials. Secondly, younger participants demonstrated a converse slightly enhanced N170 to bubble compared to practice trials (see Chapter 4).

Larger N170 amplitudes to full images of faces in older than younger participants were also observed by Daniel & Bentin (2012). They compared the N170 amplitude of full faces to 'inner components' of faces (where the eyes, nose and mouth remained visible, but the hair, ears and face contour was removed) and inner component scrambled faces (where the inner components appeared in a random configuration). They found that compared to full images of faces, younger participants had an enhanced N170 to inner component only faces and inner component scrambled faces, whilst older adults demonstrated a slight (though not significant) reduction in amplitude for inner component only faces, and a significant reduction in amplitude for scrambled inner component faces.

On bubble trials, we varied what information was available on each trial. Like the inner component faces described above, on bubble trials we removed varying degrees of visibility of the hair, ears and face contour. Whilst we did not scramble the configuration of facial features, our bubble images may have resulted in increased uncertainty as to the configuration of the face underlying the image when minimal inner components were visible. Hence our manipulation may have led to the same pattern of results as seen by Daniel & Bentin (2012).

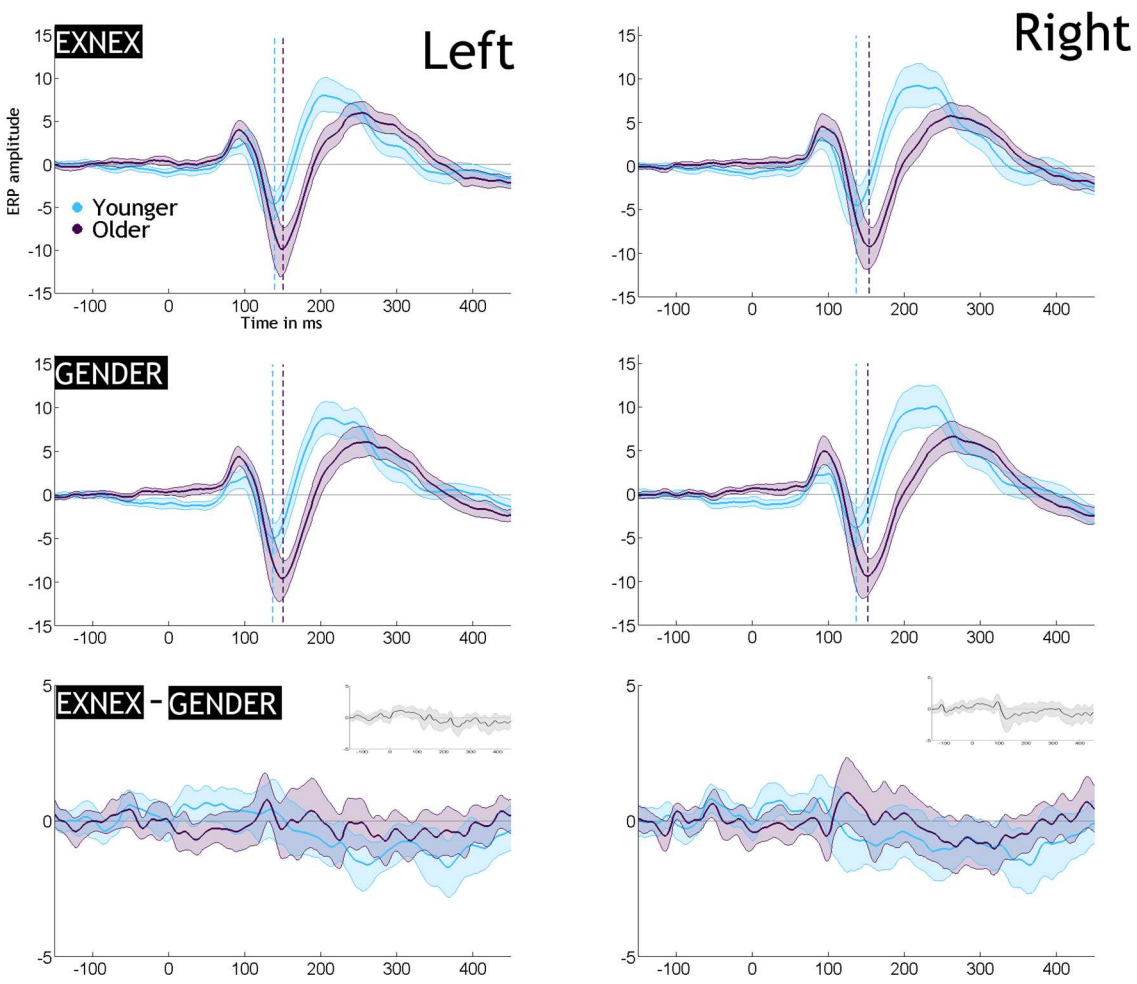


Figure 59: Non-Bubble ERP task differences by age group Mean practice trial ERPs for the Older and Younger by left and right hemisphere with 95 % confidence intervals around the mean. Vertical lines represent the minimum amplitude peak of the N170 for each task. Bottom panel EXNEX minus GENDER for bubble and practice trials. Small grey plot shows the pairwise difference of practice minus bubbles trials.

In summary, we have seen that younger and older adults N170 timecourses are delayed in bubble compared to practice trials and that whilst older adults' N170 is delayed compared to younger adults in both bubble and practice trials, older adults' N170 is no *more* delayed by bubbles compared to practice trials. This suggest that, whilst older adults accuracy and reaction times are more affected than younger adults' in bubble compared to practice trials, the processing time of bubbles compared to full face images was not different in the two age groups. However, we did observe differences in the effect of bubbles on the amplitude on the N170. Older adults have an enhanced N170 to full face images compared to younger subjects. Whilst bubbling the image significantly decreased the amplitude of the N170 for older participants, it slightly increased the amplitude of the N170 for younger participants.

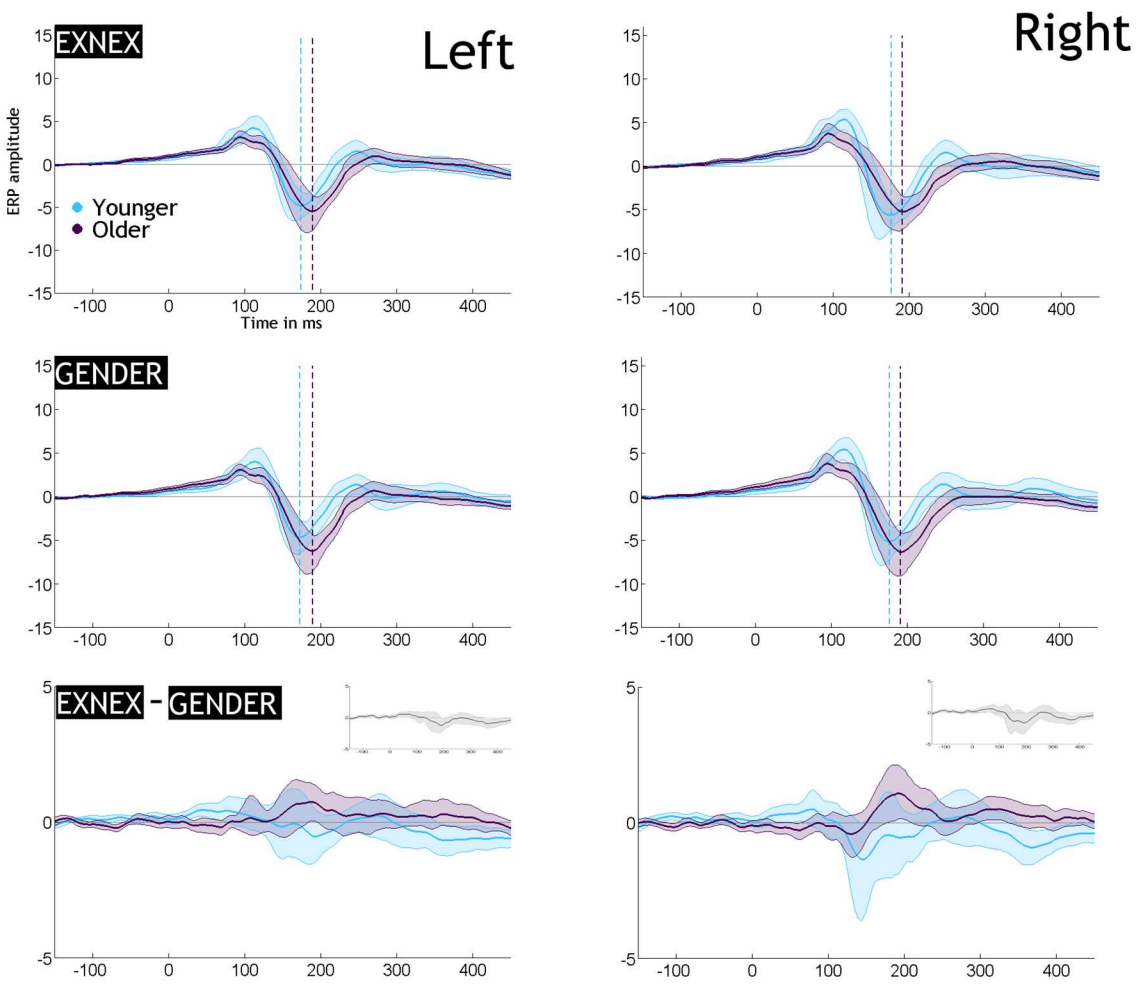


Figure 60: Bubble ERP task differences by age group Mean practice trial ERPs for the Older and Younger by left and right hemisphere with 95 % confidence intervals around the mean. Vertical lines represent the minimum amplitude peak of the N170 for each task. Bottom panel EXNEX minus GENDER for bubble and practice trials. Small grey plot shows the pairwise difference of practice minus bubbles trials.

Mutual Information Timecourse

We have seen that on average ERPs are delayed in older compared to younger participants. However, average ERPs based upon a single electrode of interest as described above are limited, as the above analyses cannot clearly account for single-trial fluctuations around the mean. Mutual information at a given time point is a measure of the relationship between single-trial fluctuations in the signal at that time, and the variation in information from the image. We expected to also observe a delay in the mutual information time course of older compared to younger participants.

We began by calculating MI for each participant at each time point in the two tasks between pixels and brain responses for all electrodes. Next, we computed the maximum MI across all pixels at each time point for each electrode. Finally,

we computed the maximum MI at each time point across all electrodes in the left and right electrode clusters of interest (see Chapter 3*: Mutual Information Timecourses*). We compared group medians of maximum MI timecourses in the **EXNEX** and **GENDER** task for the left and right posterior electrode clusters, and the difference between tasks (Figure 61). We also calculated the amplitude of the MI peak for individual participants and the corresponding MI latency, and the group median of the MI peak and latency (Table 12).

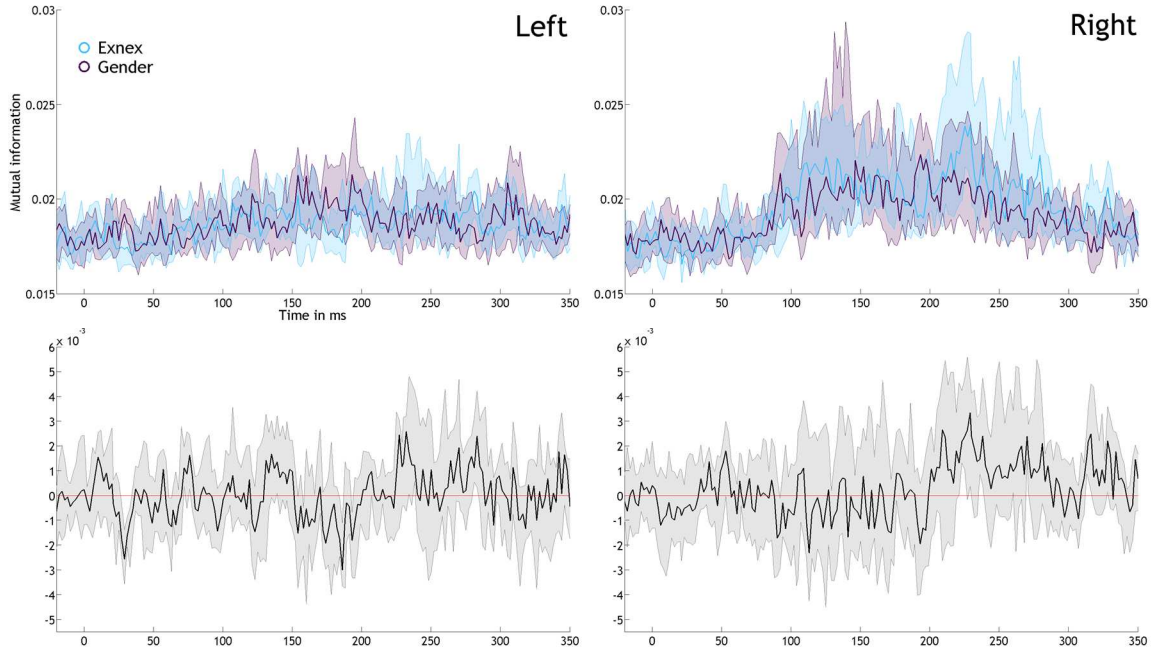


Figure 61: Older adult MI Timecourses Top row: Group median of individual maximum MI timecourses in the **EXNEX** and **GENDER** task for the LE and RE cluster. Bottom row: Difference in MI timecourses (**EXNEX** minus **GENDER**).

As can be seen in Figure 61, there was little evidence of a clear group peak when comparing median time courses across participants. Comparing group medians masks the large amount of individual variability in MI timecourses (Figure 62). MI peak latencies varied widely between participants, with some participants maximum MI peaks occurring outside of the time window of the N170. This could reflect a rebound effect – MI first peaks due to encoding of the contralateral eye, and then peaks a second time where the latter part of the N170 codes the presence of the contralateral eye (Ince et al., 2016). Peaks in mutual information for some participants around the time window of the P300, particularly in the **EXNEX** task, could reflect continued representation of diagnostic features over this time window (Van Rijsbergen & Schyns, 2009). Other participants had a flat MI time course.

	EXNEX		GENDER	
	LE	**RE**	**LE**	**RE**
LAT	218.8 [90.8, 346.8]	180 [129.6, 231.0]	189.7 [112.8, 266.7]	198.6 [137.7, 259.5]
MI	0.03 [0.03, 0.03]	0.03 [0.03, 0.03]	0.03 [0.03, 0.03]	0.03 [0.02, 0.03]

Table 12: Group average peak MI and latency Peak MI of the group median of individual maximum MI values and corresponding median latency in ms. Square brackets indicate 95 % percentile bootstrap confidence interval around the medians with 1000 samples

When comparing the time courses of MI peaks between **EXNEX** and **GENDER**, we found only a weak group difference in peak latency (LE Latency $_{[EXNEX - GENDER]}$ = -4.0 ms [-84.3, 76.2]; RE Latency $_{[EXNEX - GENDER]}$ = -23.0 ms [-71.2, 35.2]).

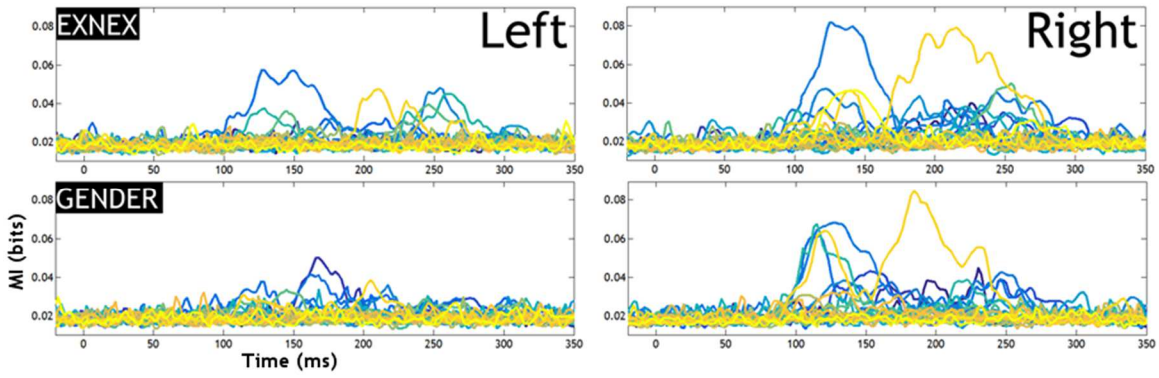


Figure 62: Older adult individual Maximum MI Timecourses For each task and electrode, individual (N = 24) participants maximum MI time course are shown. Each line represents an individual participant.

To visualise what information is being processed during the entire time course at each electrode cluster, for each participant (Supplementary 31) we took the calculated for each electrode their max MI value for each pixel across all time points between -300 ms pre-stimulus onset and 600 ms post stimulus onset. We then calculated the maximum MI value across all electrodes in the LE and RE clusters separately. We then took the median maximum MI value for each pixel across all participants for each of the two tasks separately, and computed the resulting classification image (Figure 63).

In both **EXNEX** and **GENDER**, the LE cluster shows a weak focal hotspot over the contralateral right eye area and the RE cluster over the contralateral left eye area, consistent with previous results. In the **EXNEX** task, both the LE and RE clusters show hotspots over the mouth region, stronger than that in the **GENDER** task where the mouth region is less diagnostic for task performance.

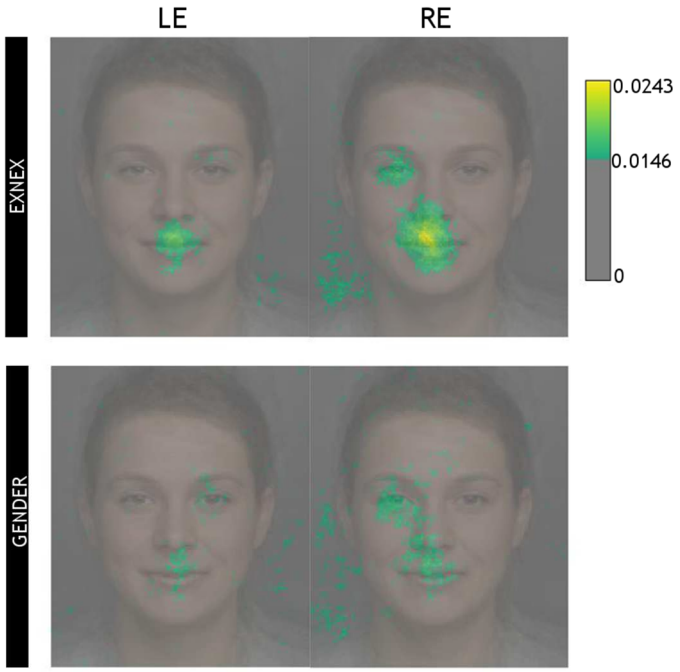


Figure 63: Older adult group Max Mutual Information Classification Images For each electrode cluster (LE, CE, RE) the median across participants of the maximum MI value for each pixel across the entire time course -300 to 600 ms is shown.

In summary we have shown that group average MI timecourses were broadly similar in the two tasks, though large inter-subject variability in the strength and timing of MI was evidenced. Our classification images have shown that, for both tasks, the left and right hemisphere ERPs were modulated by the presence of the contralateral eye and mouth. We have shown that mouth modulation was stronger in the **EXNEX** than **GENDER** task, though again there was large inter-subject variability.

Timing of Feature Sensitivity

In the **EXNEX** task we have highlighted peaks in mutual information between ERPs and pixels around the mouth and left eye region. We found a similar, though weaker pattern of results in the **GENDER** task. Next, we quantify the timing of feature sensitivity to the left eye, right eye and mouth in the two tasks. Given previous research, we expected to see processing of the contralateral eye, before processing of other task-relevant facial features, such as the mouth in the **EXNEX** task (M. L. Smith et al., 2004).

First, we examined group MI classification images over time. For each individual, we calculated the maximum MI value across all electrodes within the left and right hemisphere clusters, for every pixel at each time point. We then calculated

the group median MI value at each time point, presented in Figure 64. In both tasks, contralateral eye sensitivity and mouth sensitivity are apparent. Sensitivity is stronger and earlier on the right than left hemisphere in both tasks.

As it is unclear when sensitivity to the eye and mouth begin, next, we quantified the timing of feature sensitivity seen in our MI classification images by examining the MI time courses between feature of interest visibility and ERPs. We selected 1 left and right lateral posterior electrode and calculated for each feature the time course of the MI about feature visibility – MI(**feature**, [**ERP**, **grad**]) (see Chapter 3: *Feature of Interest Analysis*). We present the median MI time courses for each electrode, task and feature in Figure 66. We present Younger adults MI time courses as presented originally in Figure 41 again in Figure 65 for ease of comparison

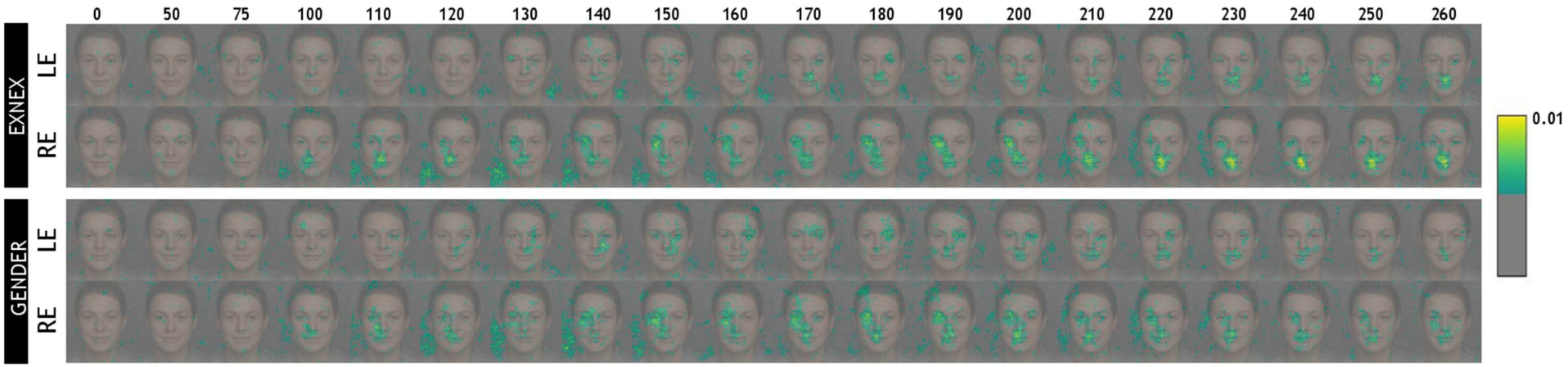


Figure 64: Older Adult Mutual information EEG classification images time course Group median of individual maximum MI values at each pixel and time point(in milliseconds, see small numbers on the top row) in each cluster of electrodes of interest in two tasks

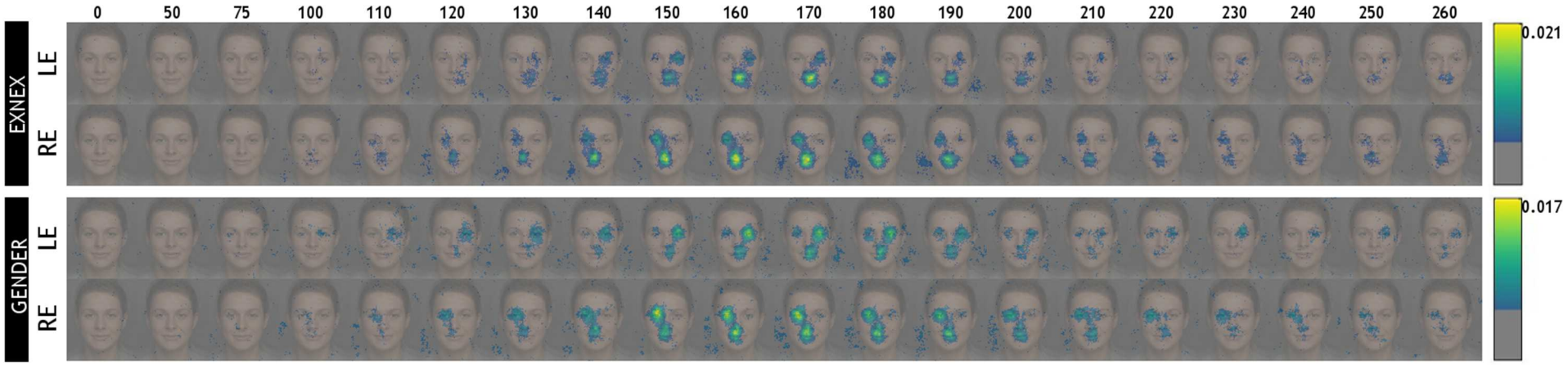


Figure 65: Younger Adult Mutual information EEG classification images time course Group median of individual maximum MI values at each pixel and time point (in milliseconds, see small numbers on the top row) in each cluster of electrodes of interest in two tasks

Left and right posterior lateral activity was modulated by the contralateral eye in both the **GENDER** and **EXNEX** task, though modulation was weaker in the **EXNEX** task and the modulation for both tasks was stronger on the right than left hemisphere. This is consistent with our behavioural results showing that the eye region was more diagnostic for the **GENDER** than the **EXNEX** task. Unsurprisingly, there was no apparent modulation of the ERPs at either the left or right electrode by the ipsilateral left and right eye in either task. Whilst left and right electrode activity was modulated by the mouth in both tasks, MI was stronger in the **EXNEX** than **GENDER** task, though delayed. This is consistent with our behavioural results showing that the mouth region was more diagnostic for the **EXNEX** than **GENDER** task.

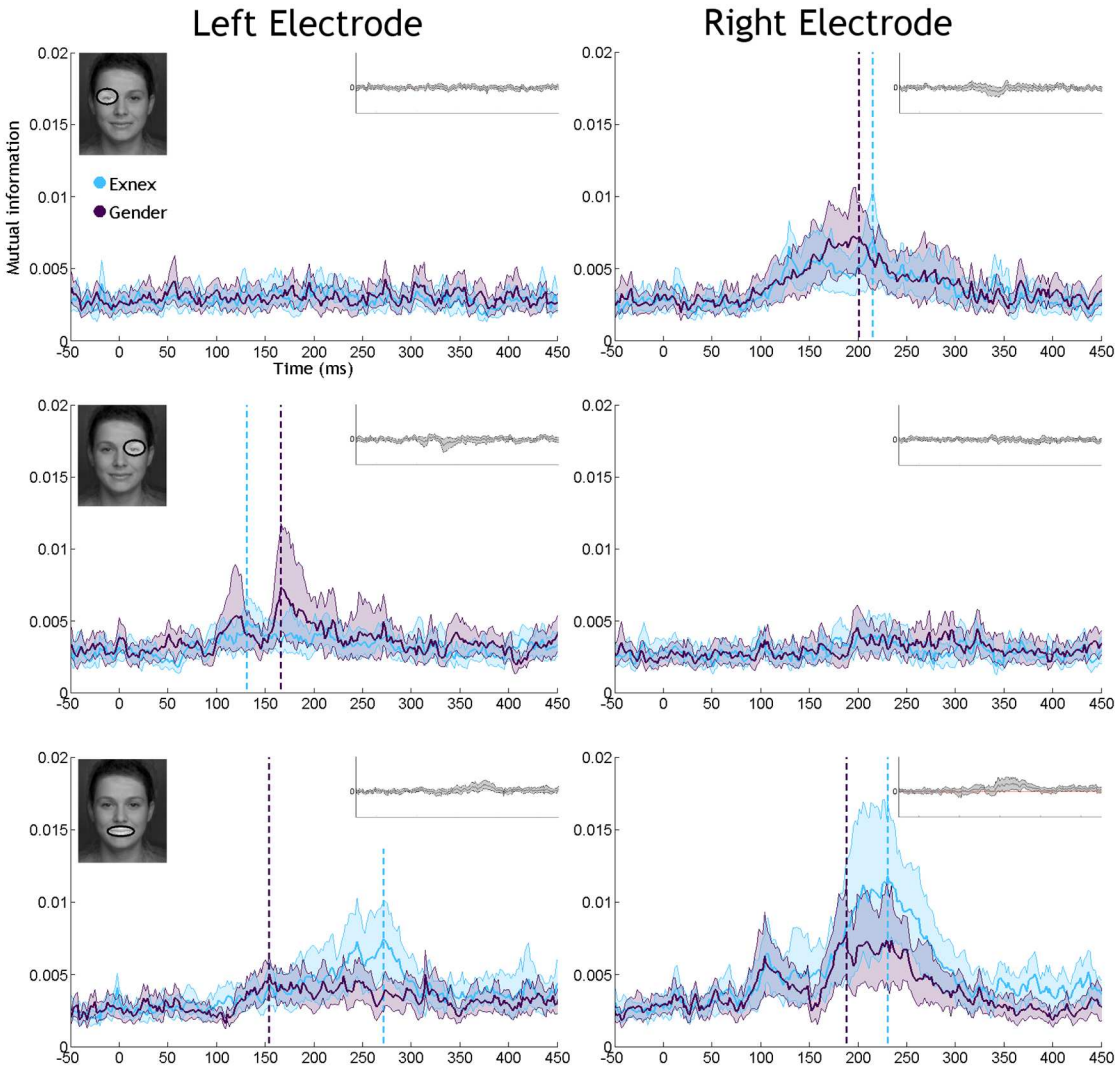


Figure 66: Older adult mutual Information Timecourse by facial feature Mutual information time course for the left and right posterior lateral electrode towards the left eye (top panel), right eye (middle panel) and mouth (bottom panel) in the **EXNEX** and **GENDER** task. Shaded area corresponds to a 95 % confidence interval calculated by a percentile bootstrap with 1000 samples around the mean. Vertical lines indicate the peak of the MI time-course.

For each participant and task we also analysed MI time courses to the midline electrode with the maximum sum MI between 120 – 220 ms (see Chapter 3; *Feature of Interest Analysis)*. Midline electrodes have previously been shown to be sensitive to task-relevant facial features (Schyns et al., 2011). In particular, in a face versus noise detection task midline electrodes displayed sensitivity to both eyes as well as the nose and mouth areas (Rousselet et al., 2014). The authors suggested that sensitivity to the nose and mouth area peaked at least 20 ms prior to posterior lateral eye sensitivity peaks, and was also present in noise trials, suggesting that midline electrode sensitivity may be a possible low-level effect, rather than an explicit feature integration process. Results from our midline electrode analysis suggested a weaker sensitivity to both the eyes and the mouth, which peaked later than contralateral posterior electrode activity.

Next, we compared MI time courses between features to determine if contralateral eye sensitivity preceded sensitivity to the mouth (Figure 67). For each comparison, we calculated the difference between the MI peak and its latency for eye compared to mouth timecourses within the time window of the N170 (~120 – 220 ms).

Right posterior lateral MI for the contralateral left eye peaked earlier than peak MI to the mouth in the EXNEX (Difference [Eye minus Mouth] = -32.2 ms [-60.5, -3.8]) but not in the GENDER (Difference [Eye minus Mouth] = = 4.6 ms [-20.1, 27.2]) task.

Left posterior lateral MI for the contralateral right eye peaked earlier than peak MI to the mouth in the EXNEX (Difference [Eye minus Mouth] = - 15.7 ms [-34.8, 3.4]) but not in the GENDER (Difference [Eye minus Mouth] = 8.8 ms [-6.1, 23.8]) task.

We found large individual variation in the direction of the difference in MI peaks to the mouth and contralateral eye region that were not explained by task order effects. We present scatter plots of the difference in individual peak MI to the eye and mouth within the time window of the N170 in Supplementary 30. However, coding of the eye starts well before the peak of the N170 (Rousselet, Ince, van Rijsbergen, & Schyns, 2014), so we should also look at MI onset.

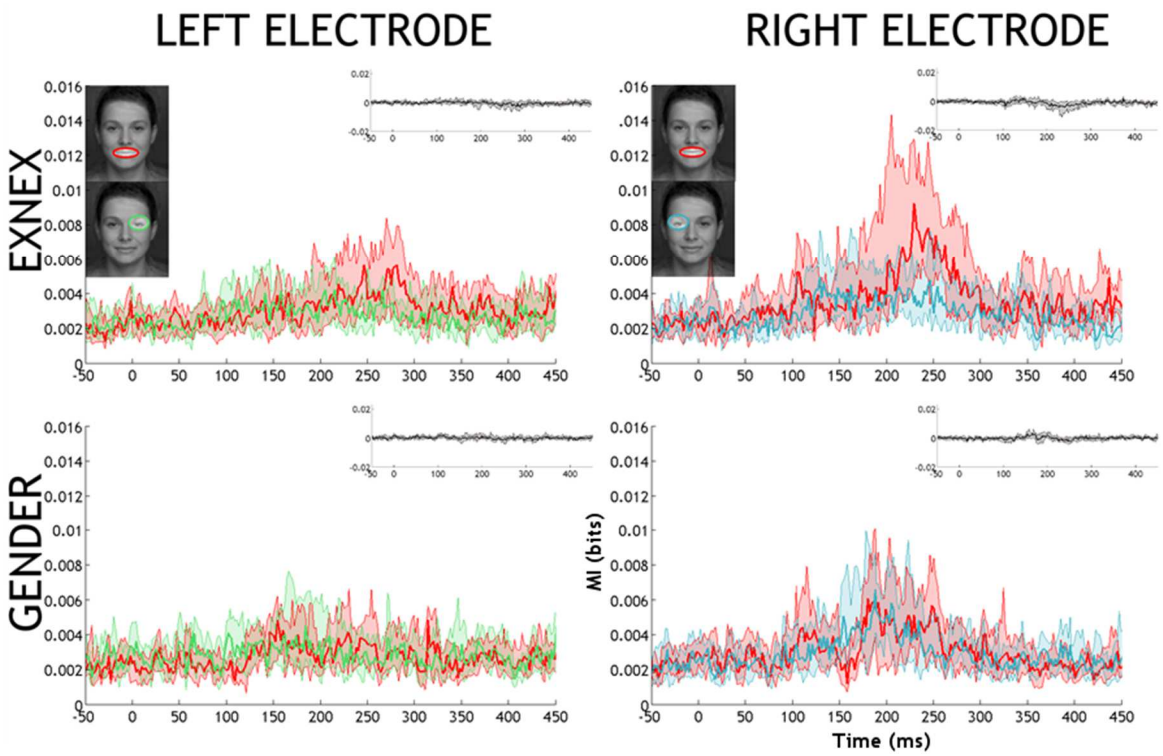


Figure 67: Older adult Mouth-Eye MI time course differences MI time courses to the contralateral eye and mouth for the left and right electrode in the EXNEX and GENDER task. Grey plots display the pairwise difference in time courses. Shaded area corresponds to the 95 % bootstrap CI around the median.

MI Onset Analysis

Next, we aimed to calculate the onset of MI to the mouth and eyes. For each participant in each task, we calculated the maximum MI value at each time point combined across both left *and* right posterior lateral electrodes to the mouth and *both* the eyes. We did this as some participants had stronger MI values to the right rather than left eye (Supplementary 31) and using this method avoided assumptions about which eye would be processed first. Using these timecourses, we used a Multivariate Adaptive Regression Splines (MARS) method (Friedman, 1991) to calculate MI onsets for the eyes and mouth (see Chapter 3; *Mutual Information Onset Analysis*).

Individual results for MI onsets in the **EXNEX** task are displayed in Supplementary 32 - Supplementary 33, and **GENDER** in Supplementary 34 - Supplementary 35.

Only 8/24 participants' in the EXNEX task, and 9/24 in the GENDER task had a detectable peak in MI towards both the eyes and mouth. All other participants' timecourses were flat for either the eyes and/or mouth. For those participants

where a peak was identifiable, we calculated the difference in the estimate of MI onset to the eyes minus the mouth. Results are presented in Figure 68.

In the **ENXEX** task, 6/8 participants had an earlier estimated onset to the eyes compared to the mouth, with a group median difference of -13 ms. In the **GENDER** task 5/9 participants had an earlier estimated onset to the eyes compared to the mouth, with a group median difference was -26 ms. However, the very small sample size does now allow the drawing of conclusions.

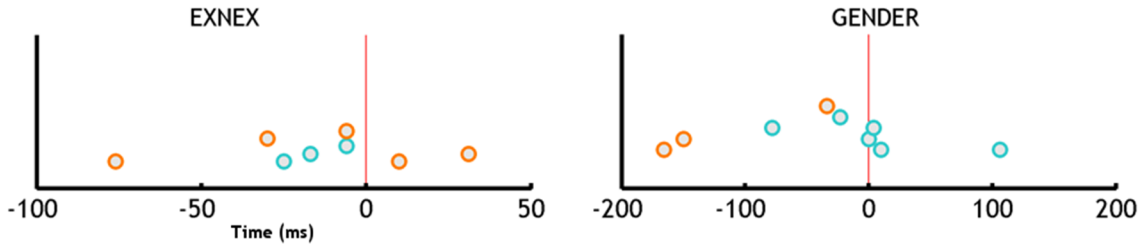


Figure 68: Older adult difference in MI Onset For each participant the difference in estimated onset times (eyes minus mouth) in the **EXNEX** (left) and **GENDER** (right) task. Orange circles completed the **EXNEX** task first. Blue circles completed the **GENDER** task first.

50 % Integration Time

Due to the lack of clear MI peaks in our older adult participants drawing conclusions on onset of MI to the eyes and mouth was not possible. As an alternative approach, we instead moved to calculating 50 % integration times of the MI timecourse towards the eyes and mouth (See Chapter 3; *50 % Integration times*). 50 % integration times can be used as a measure of processing speed which takes into account the entire MI waveform (not just peaks) and so offered benefits over our MI onset analysis where clear peaks needed to be identified. Calculating 50 % integration times also offers the advantage of normalising MI timecourses for comparison between groups. We calculated median 50 % integration times and CIs for each group (Younger and Older) for each feature (eyes and mouth) and task separately (Table 13). We then calculated the difference in 50 % integration time for the eyes and the 50 % in integration time for the mouth between younger and older participants.

We found that older adults integrated information about both the eyes and mouth more slowly than younger adults. This delay in processing speed was not uniform across features and tasks. More specifically, whilst older adults were

~20-23 ms delayed in processing the eyes compared to younger adults across tasks, the delay in mouth integration was task dependent – there was a delay in processing the mouth in the **GENDER** but not **EXNEX** task. Mouth processing was less delayed than eye processing in the **GENDER** task, suggesting that older adults did not demonstrate a general delay that was consistent across all features. There were no significant interactions.

	EXNEX		GENDER	
	Younger	**Older**	**Younger**	**Older**
Eyes	231.4 [224.9, 240.2]	250.1 [244.4, 254.3]	227.7 [221.1 235.5]	254.0 [248.9 261.2]
Mouth	274.3 [259.5 287.7]	271.7 [258.9 284.0]	243.0 [237.9 248.4]	253.0 [247.6 262.6]

Table 13: 50 % integration time estimates. Median integration times and 95 % confidence interval for the eyes (top row) and mouth (bottom row) in the EXNEX (left) and GENDER (right) task for younger and older adults.

EXNEX FEATURE INTEGRATION

In the **EXNEX** task (Figure 69; top panel), at the group level, both younger and older adults integrated information about the eyes faster than information about the mouth (Younger feature difference = -38.7 ms [-53.4 -27.9]; Older feature difference = -23.7 ms [-32.4 -12.7]). At the individual level, 22/24 younger and 22/24 older participants' integrated eye information faster than mouth information. However, the distributions of integration times were positively skewed in both age groups, with large idiosyncrasies in timing of integration of the two features.

Comparing the speed of feature integration between the two groups, younger adults integrated information about the eyes faster than older adults (Group difference = -20 ms [-26.1, -7.6]). However, there was a large overlap of timing at the individual level (Figure 70). There was no group difference in time taken to integrate the mouth (Group difference = 3.5 ms [-17.2, 24.7]). Older adults slower integration of eye but not mouth information suggests that feature integration is qualitatively different in older than younger subjects – older adults did not demonstrate a general delay that was consistent across all features.

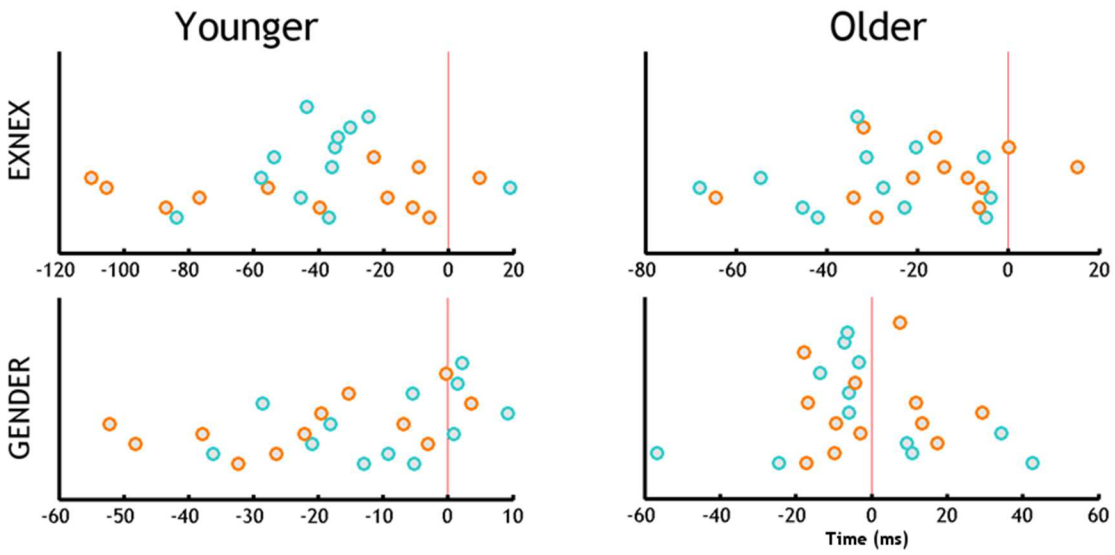


Figure 69: Difference in integration time between the eyes and mouth Difference in integration time (eyes minus mouth) for each age group and task. Blue circles indicate participants who completed the **EXNEX** task first. Orange circles indicate participants who completed the **GENDER** tsk first.

GENDER FEATURE INTEGRATION

In the **GENDER** task (Figure 69; bottom panel), at the group level, younger but not older adults integrated information about the eyes faster than information about the mouth (Younger feature difference = -13.7 ms [-22.5 -4.9]; Older feature difference = -3.8 ms [-8.6 7.0]). At the individual level, 19/24 younger and 15/24 older participants' integrated eye information faster than mouth information. The distributions of integration times show large idiosyncrasies in timing of integration of the two features.

Comparing the speed of feature integration between the two groups, younger adults integrated information about the eyes faster than older adults (Group difference = -23.2 ms [-39.3 -16.8]. Whilst there was some overlap of timing at the individual level (Figure 70), there was a clearer pattern of group differences than in the **EXNEX** task. Younger adults also took less time to integrate the mouth than Older adults (Group difference = -9.3 ms [-21.3 -1.5]), though there was less of a delay for mouth than eye integration. Older adults were more delayed in integration of eye than the mouth compared to younger adults – older adults did not demonstrate a general delay that was consistent across all features.

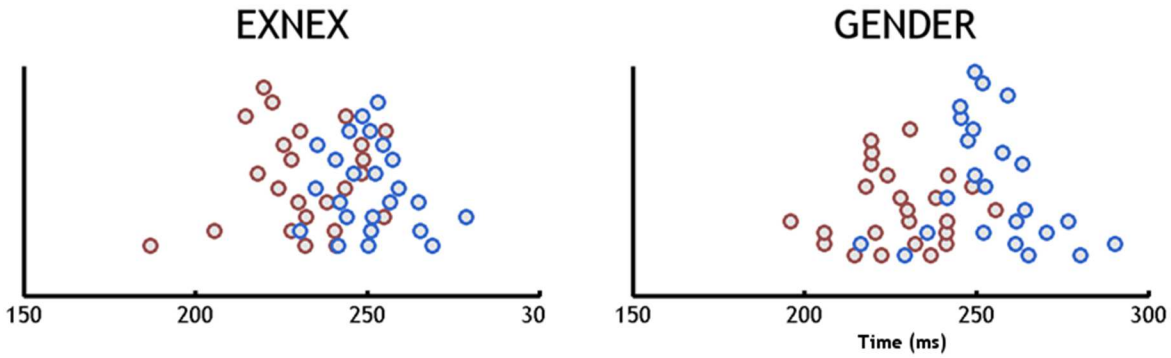


Figure 70: Eye integration time by age 50 % integration time in the EXNEX (left) and GENDER (right) tasks for the eyes between younger (brown) and older (blue) age groups

Reverse Analysis: EEG Results

We have examined mutual information between pixels and brain signals and shown that older adults' brain signals are sensitive to the contralateral eye and mouth, but are these face features coded by the N170 amplitude and latency in older adults?

Using the feature of interest masks (see Chapter 3: *Feature of Interest Analysis*) we calculated on a trial-by-trial basis the visibility of each feature (left eye, right eye, mouth), obtained as a scalar value of the sum of pixel visibility within the ellipse of each feature mask. We then split these visibility values into ten equally populated bins ranging from the lowest (bin 1) to the highest (bin 10) visibility values. We then sorted single trial ERP's into 10 bins, based on the feature visibility in each trial. We present group mean ERP's for each level of feature visibility for the left (Figure 71) and right (Figure 72) electrode.

Weak modulations of the N170 amplitude and latency were apparent in both tasks. Modulations were strongest for mouth sensitivity on the right electrode in the two tasks (Figure 72, bottom row).

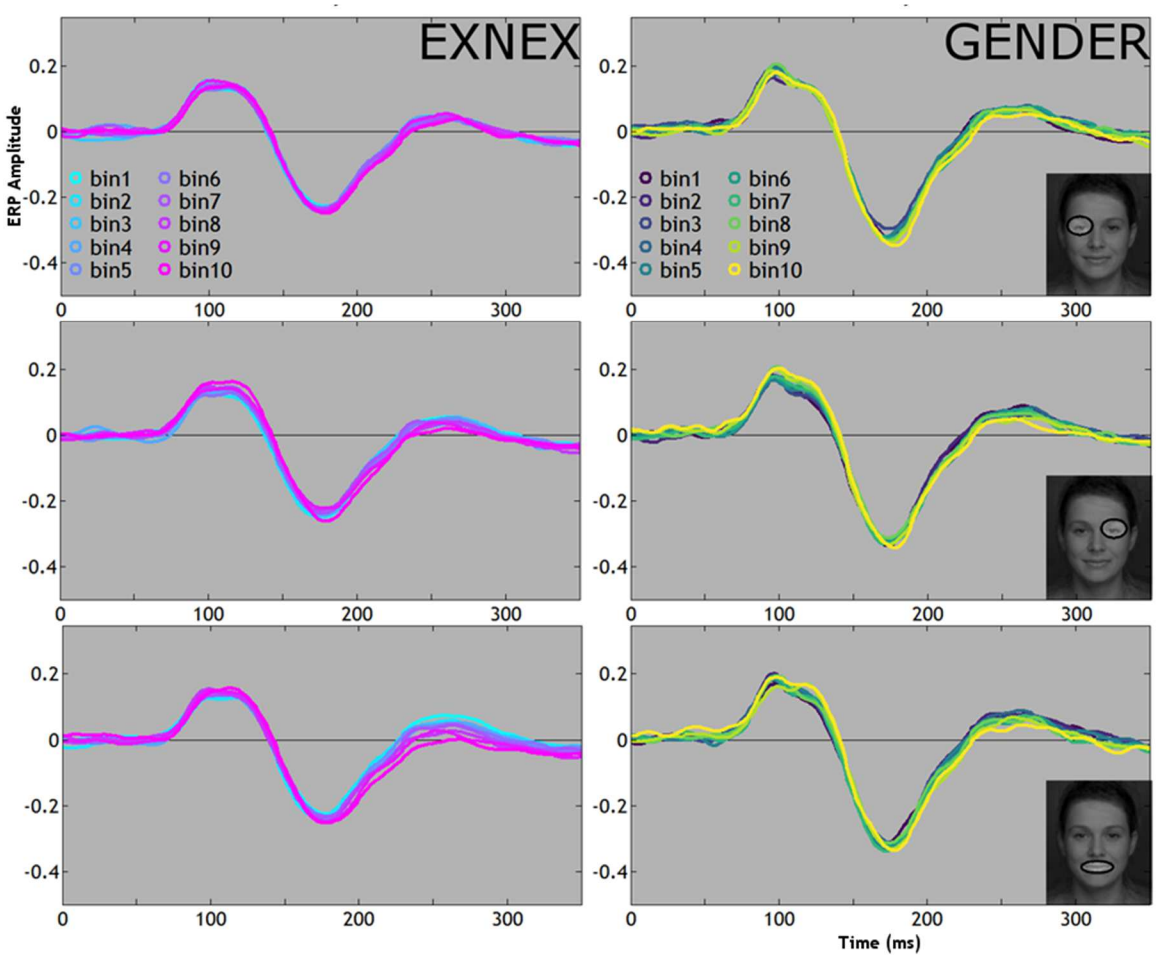


Figure 71: Older adult Left Electrode Binned ERPs by feature visibility Each column represents one task. Column A (left) shows results for the **EXNEX** task, column B (right) shows results for the **GENDER** task. Group mean ERP's are shown for the posterior left electrode. ERPs are binned into 10 levels of visibility ranging from bin 1 (least visible) to bin 10 (most visible) for the left ipsilateral eye (top row), the right contralateral eye (middle row) and the mouth (bottom row).

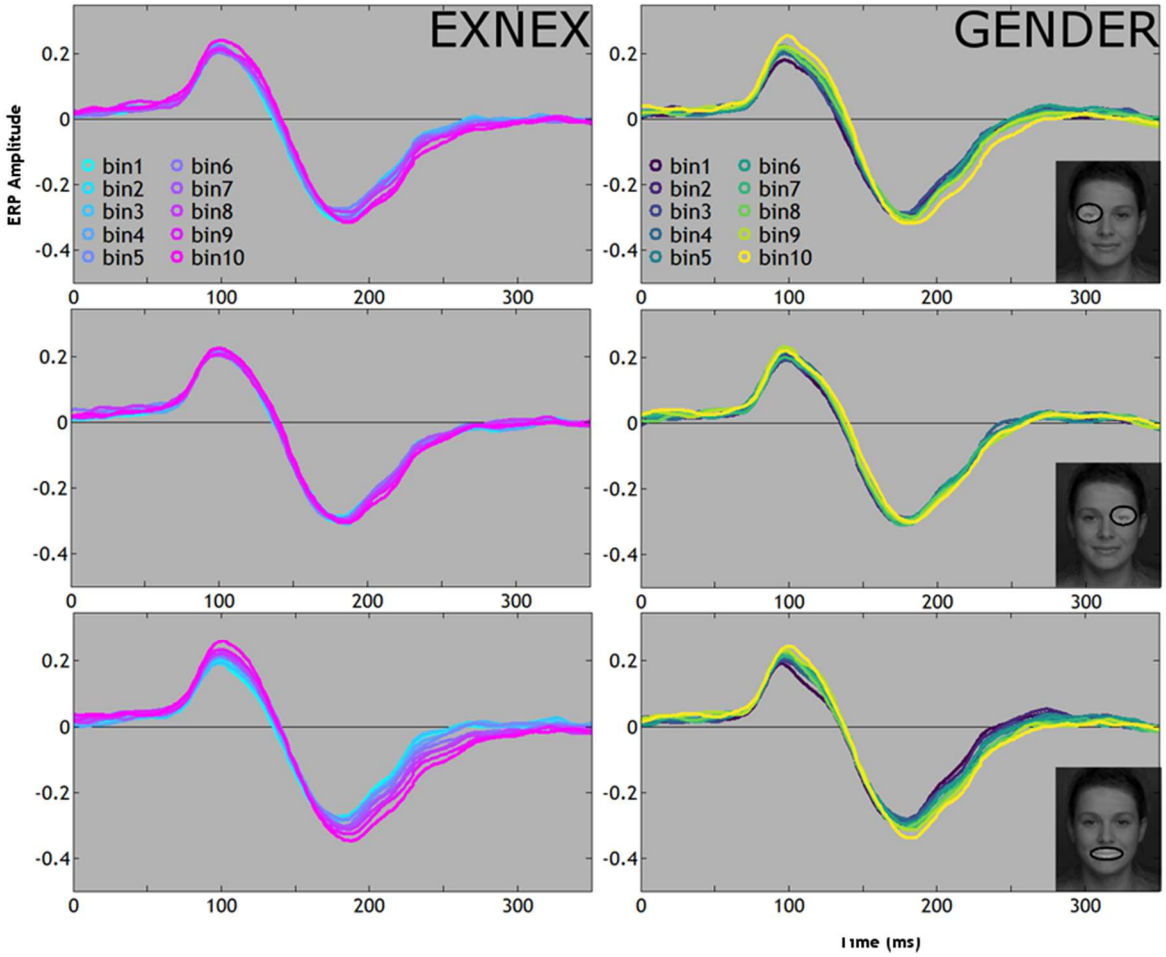


Figure 72: Older adult Right Electrode Binned ERPs by feature visibility Each column represents one task. Column A (left) shows results for the **EXNEX** task, column B (right) shows results for the **GENDER** task. Group mean ERP's are shown for the posterior left electrode. ERPs are binned into 10 levels of visibility ranging from bin 1 (least visible) to bin 10 (most visible) for the left ipsilateral eye (top row), the right contralateral eye (middle row) and the mouth (bottom row)

N170 Latency and Amplitude Differences

Next, we wanted to quantify this modulation of the N170, specifically to compare how the amplitude and latency of the N170 is modulated by the presence of specific facial features between tasks.

We quantified the difference in N170 amplitude and latency between ERPs when facial features were most visible (bin 10) compared to least visible (bin 1). First we calculated for each participant the N170 amplitude and latency at each bin, for each feature in each task, for both posterior lateral electrodes. We then calculated the difference between bin 10 (most visible) minus bin 1 (least visible) for N170 latency in the time window ~150–240 ms post stimulus onset for each task, feature and lateral posterior electrode. Amplitude differences were

calculated as a proportion of bin1 (least visible) as a percentage, such that an amplitude difference of 50 % means that the amplitude of bin 10 (most visible) amplitudes were 150 % the size of the amplitudes in bin 1. We present median difference across participants in Figure 73 and Table 14, and individual differences in Supplementary 36.

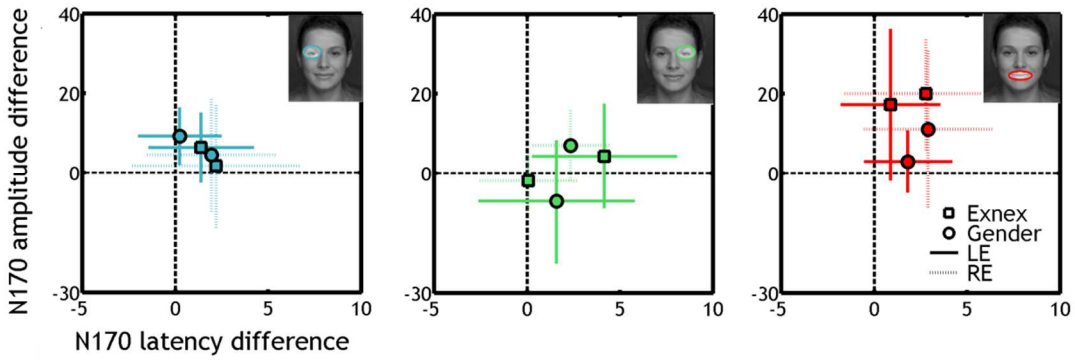


Figure 73: Older adult N170 Amplitude and Latency Differences Results are presented for three features of interest: the left eye (blue), right eye (green) and mouth (red). N170 was measured in the time window between ~150-240ms post stimulus onset. Median N170 latency differences were calculated between bin 10 (most visible) minus bin 1(least visible) and are presented in milliseconds. Median N170 amplitude differences are calculated as a percentage of bin 1, such that an amplitude difference of 50% means that the amplitude of bin 10 was 150% the size of amplitudes in bin1. Vertical and horizontal lines correspond to 95% confidence intervals. **EXNEX** results are plotted with squares, **GENDER** results are plotted with circles. Solid lines are the left electrode; dashed lines are the right electrode.

For the majority of comparisons there were no significant differences as 95 % confidence intervals contained 0. However, in the **EXNEX** task, increased visibility of contralateral right eye modulated the N170 latency at the left electrode (Figure 73; middle box, square with solid lines), whilst the right hemisphere N170 amplitude was modulated by increased visibility of the mouth (Figure 73; right box, square with dashed lines). In the **GENDER** task, the ipsilateral left eye modulated the amplitude of the N170 at the left electrode (Figure 73; left box, circle with solid lines), whilst increased visibility ipsilateral right eye modulated N170 latency at the right electrode (Figure 73; middle box, circles with dashed lines). Increased visibility of the mouth modulated the latency of the N170 also at the left electrode (Figure 73; right box, circle with solid lines).

		EXNEX		GENDER	
		LE	**RE**	**LE**	**RE**
LEFT EYE	**LAT**	1.4 [-1.2, 4.0]	2.2 [-2.1, 6.4]	0.2 [-1.9, 2.4]	1.9 [-1.5, 5.4]
	AMP	6.4 [-2.8, 15.7]	1.8 [-13.9, 17.5]	9.3 [2.2, 16.4]	4.7 [-9.7, 19.0]
RIGHT EYE	**LAT**	4.2 [0.3, 8.0]	0.1 [-2.5, 2.6]	1.6 [-2.4, 5.5]	2.3 [0.2, 4.5]
	AMP	4.4 [-8.5, 17.3]	-1.9 [-9.1, 5.4]	-7.1 [-22.8, 8.7]	7.1 [-2.4, 16.6]
MOUTH	**LAT**	2.7 [-0.6, 6.0]	1.7 [-1.0, 4.4]	2.9 [0.4, 5.4]	2.1 [-2.3, 6.5]
	AMP	17.3 [-1.1, 35.7]	20.1 [6.4, 33.8]	3.0 [-4.7, 10.7]	11.2 [-8.9, 31.2]

Table 14: Older adult N170 amplitude and latency differences by facial feature Median N170 latency and amplitude differences for the right eye, left eye and mouth for the left and right hemisphere in each task. N170 was measured in the time window between ~150-240ms post stimulus onset. Median N170 latency differences were calculated between bin 10 (most visible) minus bin 1(least visible) and are presented in milliseconds. Median N170 amplitude differences are calculated as a percentage of bin 1, such that an amplitude difference of 50% means that the amplitude of bin 10 was 150% the size of amplitudes in bin1

Next, we tested the interaction between age and N170 latency and amplitude modulation of the N170 i.e. whether the difference in latency and amplitude modulation by the presence of a particular feature differs between the two age groups. We computed the effect size estimates for the group difference at each electrode cluster (Table 14). We found significant group effects, with the presence of the contralateral left eye having a stronger effect on the N170 latency and amplitude modulation in younger than older participants at the right hemisphere in both tasks. We also found that the presence of the contralateral right eye had a stronger effect on the N170 latency in younger than older participants at the right hemisphere in both tasks, whilst the group difference in amplitude modulation was only significant in the **GENDER** but not **EXNEX** task.

We also found significant group effects with the presence of the mouth having a stronger effect on the N170 latency and amplitude modulation in younger than older participants at both hemispheres and tasks.

In summary our feature of interest analysis suggests modulation of the N170 latency and amplitude as a mechanism involved in face detection in younger and older participants. This mechanism changes with ageing, where both **EXNEX** and

GENDER categorisation tasks are associated with changes in latency and amplitude of the N170.

		EXNEX		GENDER	
		LE	**RE**	**LE**	**RE**
LEFT EYE	**LAT**	-2.2 [-5.8, 0.2] *-0.27* *[-0.57, 0.04]*	-8.0 [-13.1, -3.7] *-0.55* *[-0.78, -0.27]*	-2.0 [-4.2, 1.1] *-0.21* *[-0.49, 0.11]*	-13.5 [-18.3, -8.9] *-0.81* *[-0.94, -0.62]*
	AMP	- 0.1 [-11.9, 11.6] *-0.0* *[-0.25, 0.32]*	-21.0 [-34.6, -0.85] *0.34* *[0.02, 0.64]*	-5.3 [-16.4, 4.4] *0.18* *[-0.18, 0.54]*	-26.1 [-44.9, -7.7] *0.48* *[0.19, 0.74]*
RIGHT EYE	**LAT**	-8.0 [-12.5, -4.1] *-0.53* *[-0.77, -0.26]*	0.0 [-2.8, 3.9] *0.01* *[-0.34, 0.28]*	-5.8 [-11.1, -0.1] *-0.34* *[-0.63, -0.04]*	-3.9 [-6.2, -1.6] *-0.43* *[-0.72, -0.11]*
	AMP	-13.8 [-33.1, 3.0] *0.25* *[-0.08, 0.56]*	-17.8 [-30.9, -4.3] *0.43* *[0.11, 0.72]*	-30.0 [-49.4, -12.7] *0.46* *[0.18, 0.73]*	-16.4 [-30.1, -3.7] *0.40* *[0.14, 0.66]*
MOUTH	**LAT**	-7.3 [-10.9, -3.6] *-0.58* *[-0.81, -0.33]*	-7.8 [-12.2, -4.0] *-0.55* *[-0.80, -0.26]*	-4.0 [-7.9, -1.5] *-0.43* *[-0.68, -0.13]*	-10.7 [-15.1, -5.6] *-0.64* *[-0.86, -0.41]*
	AMP	-37.1 [-59.9, -16.3] *0.52* *[0.24, 0.76]*	-31.9 [-69.3, -5.8] *0.39* *[0.11, 0.66]*	-21.7 [-35.5, -3.4] *0.42* *[0.12, 0.69]*	-24.6 [-45.7, -3.5] *0.38* *[0.06, 0.64]*

Table 15: Estimates for group differences in N170 modulation Estimate for group differences (Younger minus Older) in N170 latency (LAT) and amplitude (AMP). Latency values correspond to median latencies expressed in milliseconds. Amplitude values are calculated as a percentage points, so that an amplitude difference of - 50 means the amplitude difference in younger adults was 50 percentage points larger than the amplitude difference in older adults. Square brackets indicate 95% confidence intervals. A corresponding Cliff's delta estimate is shown in italics.

Discussion

In the current study we explored whether there are similarities in the facial features processed by younger and older adults in a gender and expression discrimination task, and quantified age-related differences in modulation of ERPs by feature visibility. We found that whilst younger and older adults both relied on the same facial features for accurate categorisation in both tasks, older adults were more dependent than younger adults on visibility of the mouth for making accurate responses in the EXNEX task. We found a delay in information processing in older compared to younger adults, with eye encoding being delayed by ~25 ms. In comparison, we found only a slight ~9 ms delay in mouth processing in the GENDER task and no delay in mouth processing in the EXNEX task. This suggests there is not a uniform age-related delay in processing of all facial features.

BEHAVIOUR

We found that behaviourally, older adults relied on the same facial features as young adults used as diagnostic cues to resolve gender and expressiveness categorisation tasks. However, compared to our younger adults, older adults were more heavily dependent upon the presence of the mouth to correctly and quickly determine expressive and non-expressive faces, whereas younger adults could complete the task using other features in the absence of the mouth.

There was no comparable over-reliance on eye information for resolving the gender task. Recently Jaworska (2017) showed that during face versus noise detection tasks, both younger and older adults responded faster when the contralateral eye region was visible, and older adults were more reliant on the visibility of the eye region for correct responses. Older adults in the current experiment may not have had over-reliance on the eye region in the gender task due to several other cues being available for resolving the task, such as the hair or other facial features, or inter-eye distance. This suggests that the absence of mouth visibility has a greater effect on accuracy in our older than younger participants, and that older adults may rely more upon the presence of the mouth to make correct responses. In contrast younger participants could use

other features and their combination to correctly discriminate ‘happy’ from ‘neutral’ faces.

Despite older adults relying more upon the mouth for correct behavioural responses in the EXNEX task, the strength MI between the mouth and reaction times (calculated be summing MI within each region of interest) was weaker in older compared to younger adults. This suggests that, in the EXNEX task dependency on the mouth region for correct responses didn’t translate to modulating reaction times for older adults. We did not find any difference in the strength of MI towards the eye regions in either task.

EEG INFORMATION

We calculated MI between pixels and brain responses. We found that older adults’ brain signals were modulated by the same facial features as per younger adults, but that MI was weaker.

AGE-RELATED DIFFERENCES IN EYE ENCODING

We attempted to quantify when information is encoded in the brain. We calculated 50 % integration time of MI towards the eyes as a measure of processing speed. We found that older adults integrated information about the eyes consistently slower than younger adults – approximately 20 ms slower in the EXNEX and 23 ms in the GENDER task. We calculated integration times over both left and right posterior lateral electrodes, with results consistent with those of Jaworska (2017) suggesting a 23 – 25 ms delay in processing the contralateral left and right eye on right and left posterior lateral electrodes respectively in a face detection task. This suggests that the delay in feature processing previously observed in a face versus noise detection task is similar to that observed in our experiment, suggesting that the delay in processing of the eyes is similarly delayed in a range of face processing tasks. Whilst in both the face detection task reported by Jaworska (2017) and GENDER task reported here, the eyes were task relevant for behaviour, in the EXNEX task the eyes were minimally task diagnostic. This is consistent with the claim that the eyes are processed automatically regardless of their task-relevancy and goes a step further by suggesting that the delay in eye sensitivity is relatively consistent (at least at

the group level) across tasks. Further research could continue to test the robustness of this finding by diversifying the range of face processing tasks utilised and explore the variability in the timing of eye sensitivity across individuals in different tasks.

AGE-RELATED DIFFERENCES IN MOUTH ENCODING

We compared younger and older adults' speed of processing the mouth in the two tasks. In the EXNEX task, we did not find any difference in the processing speed of mouth information between younger and older adults. This suggests that whilst older adults have delayed processing of the eyes, there is no delay in mouth processing. In the GENDER task older adults experienced a delay in mouth processing of approximately 9 ms. It is unclear why there is a differential delay in mouth processing in the GENDER than EXNEX task. Some explanations could be:

1) Weaker eye coding and increased mouth processing in the EXNEX task may reflect older adults processing the mouth more readily, reducing age related differences in mouth processing speed

2) Older adults spend more time encoding the eyes in the GENDER than EXNEX task which may delay mouth encoding.

EYE-MOUTH ENCODING DELAY

Previously, it has been suggested that contralateral eye processing precedes the processing of other, task relevant features (Rousselet et al., 2014). In the EXNEX task both younger and older adults processed information about the eyes faster than information about the mouth, though there were large idiosyncrasies in relative timing and not all participants showed this effect. In the GENDER task, whilst younger adults processed information about the eyes faster than the mouth, there was not a significant difference in relative processing times for the older age group. The difference between eye and mouth integration was larger in the EXNEX than GENDER task for both age groups. This may reflect greater encoding of the mouth in the EXNEX than the GENDER task in both age groups.

We found that age-related differences in speed of processing is different for different features and tasks – there is not a uniform age-related delay in processing speed of all facial features in all face processing tasks.

AGE-RELATED DIFFERENCES IN N170 ENCODING

We compared how the presence of facial features modulated the amplitude and latency of the N170 in older compared to younger adult participants. We found that the presence of the contralateral left eye resulted in earlier and stronger N170s in younger than older participants in both tasks, whilst the presence of the contralateral right eye modulated N170 latencies only (earlier in younger than older participants). Similarly increased visibility of the mouth also led to earlier and stronger N170s in younger than older participants in both tasks.

BUBBLES MANIPULATION

In the current experiment, older participants' reaction times and accuracy were more negatively affected by bubbling the image than behavioural outcomes in younger participants. This suggests that older adults required more information to complete the task to the same level of accuracy as younger participants. The bubbles manipulation, where only parts of the image is revealed, may be akin to partially occluded or fragmented object perception. It has been suggested previously that older adults experience an age-related decline in perceptual tasks where visual information is missing or fragmented, such as in perceptual closure tasks (Frazier & Hoyer, 1992; Whitfield & Elias, 1992). Previous work using bubbles with older adults in a face versus noise detection task (Jaworska, 2017) has also suggested that older adults experience a larger decline in accuracy than younger adults in bubble compared to non-bubble trials.

However, whilst bubbling affected behavioural responses, our differences in brain responses cannot be understood as a result of the bubbles manipulation. When comparing brain data, both younger and older adults N170 were delayed by the bubbles manipulation. There was however only weak group differences between the differences in N170 peak latency to practice and bubble trials in either the expression or gender task. This suggests that age-related delays in processing of facial features cannot be attributed to the presence of bubbles as an experimental manipulation. However, we did observe differences in the

effect of bubbles on the amplitude on the N170. Whilst bubbling the image significantly decreased the amplitude of the N170 for older participants, it slightly increased the amplitude of the N170 for younger participants. It is not clear why we observed this effect, though a related pattern was also observed by Daniel & Bentin (2012).

COMPARISON WITH SCHYNS ET AL. (2002)

Whilst we found evidence of strong mouth modulation of behaviour in the EXNEX task, we found the increased visibility of the eye had only a weak effect on behavioural modulation. We found eye sensitivity in some participants, but this was weaker than that found previously. Unlike previous work, we found that the mouth did modulate the N170 in both younger and older participants.

Our results may have differed from Schyns et al. (2002) as our naturalistic images may have resulted in less dependence upon the eye region due to

1) Models not wearing eye makeup in the current experiment which may have reduced the contrast between images of men and women and reduced the saliency of attending to the eye region;

2) Coloured images used in the current experiment may have increased the diagnostic information across the whole face for example due to pigmentation; and

3) Other cues were available in the current experiment for resolving the task, for example the hair. For some older adults brain classification images showed modulation of ERPs with increased visibility of an area above the shoulder where either the background or hair would be visible (depending on hair length and style).

GRAND AVERAGES

The widespread problem in neuroscience of averaging data and comparing means is well known. Comparing grand averages between groups, such as younger and older adults, masks the rich patterns and complexities of individual differences. Within this thesis individual differences are presented, which have not been

reported in previous studies. For example, in Chapter 2 we demonstrated that our group results reflected previous findings of contralateral eye sensitivity, but that there were undocumented individual variations in this pattern. In Chapters 4 and 5 we demonstrated that, among other findings, that the onset of feature sensitivity is idiosyncratic. The implications of these results are that previous studies should be interpreted with caution – whist average effects may be reproducible, on an individual level, participants vary widely. In the most extreme cases, average results may not represent any individual patterns in the underlying data. Future research should more faithfully report individual differences, and how many individuals show patterns consistent with average data when used.

LIMITATIONS AND FUTURE DIRECTIONS

Whilst we attempted to ensure our older adults could be considered as ageing healthily through screening for mild cognitive impairment, a healthy older adult sample cannot be guaranteed on the basis of behavioural screening tests alone. Anatomical and functional brain differences occur before the beginnings of behaviourally evident cognitive decline (Beason-Held et al., 2013; Braak & Braak, 1991). However, engaging older adults in more stringent screening procedures is often not financially feasible and may discourage older adults from participating in research if participation is seen as being more demanding (i.e. increase on time or number of visits required) or invasive.

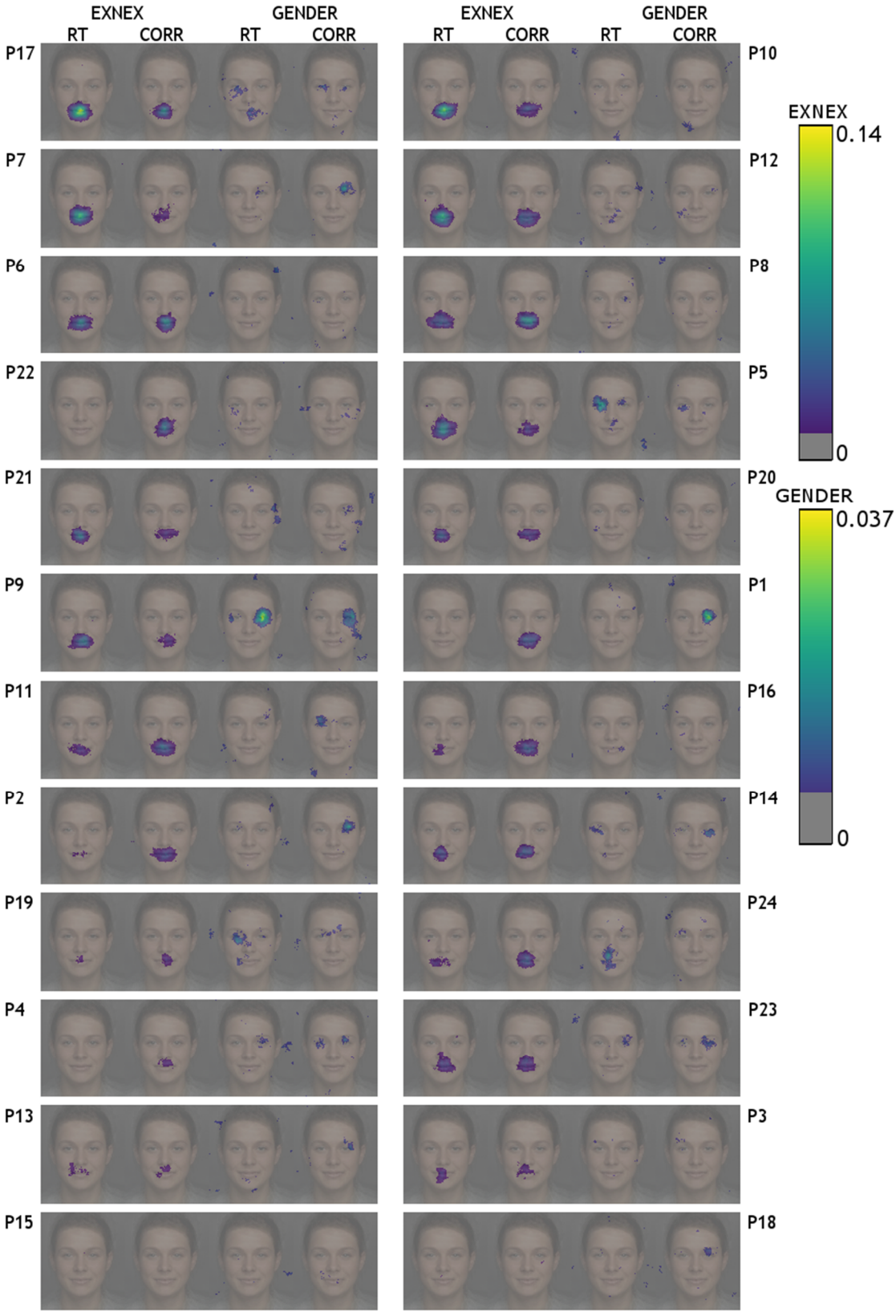
Ageing-related macular degeneration and other optical changes with age may also be a problem. While we attempted to ensure our older adults had normal or correct to normal vision by asking that participants have visited an optician within the last year, Owsley (2011) suggests that more stringent measures of reporting about the ageing eye should be undertaken. The authors criticise using self-report and visual acuity cut-off points as measures of healthy ageing, and warn that ophthalmologists and optometrists can differ in how they consider older adults' eyes as being clinically 'normal'. Given these views it may be more appropriate to consider using more objective and standardized criteria for assessing good retinal health in older adult populations in future studies.

Our results are limited due to our choice of stimuli. In our study, our stimulus set only consisted of images of young adult faces. We did not account for how the perceived age of the face may affect recognition in regard to the facial features that are diagnostic for particular facial processing judgments, nor how face processing strategies may vary depending on the age of the observer relative to the age of the stimuli shown. The age of a face may have consequences for some face processing tasks, such as accurate emotion recognition. For example, the occurrence of wrinkles and folds with increasing face age may lead to increased difficulty in emotion recognition of older faces (Fölster, Hess, & Werheid, 2014). Brain responses will also change with the age of the face displayed. The N170 may be larger for older relative to younger adults faces, regardless of the age of the observer (Komes, Schweinberger, & Wiese, 2015; Wiese et al., 2008), though not all studies have suggested that this difference in amplitude is significant (Ebner, He, Fichtenholtz, McCarthy, & Johnson, 2011).

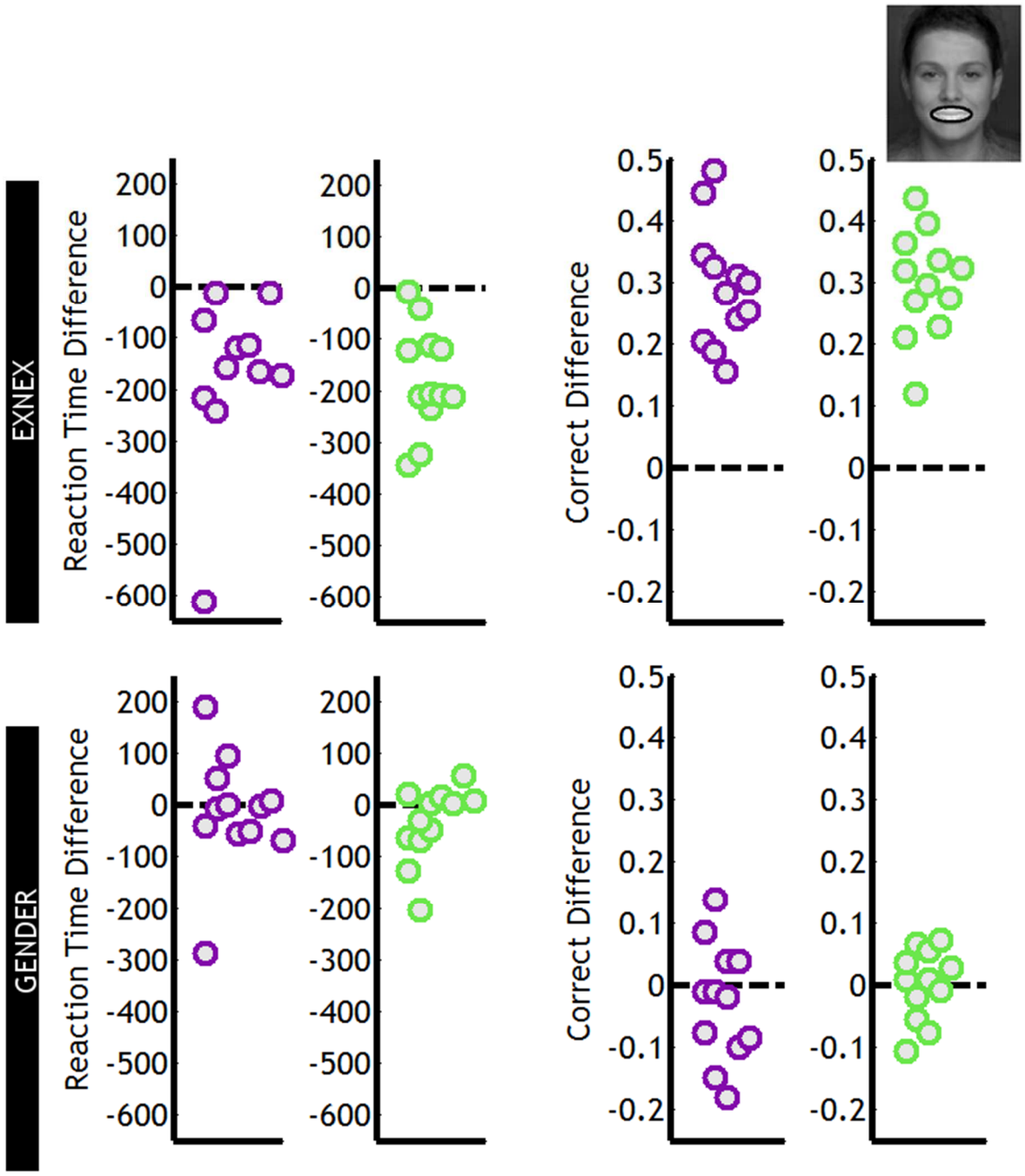
An own-age bias (OAB) i.e. better recognition of faces of a similar age to ourselves (Anastasi & Rhodes, 2005; Wiese et al., 2008) may mean that face processing varies depending on the relative age of the face compared to the observer. OAB cannot be explained by an increase in the allocation of attention to own versus other ages faces however (Neumann, End, Luttmann, Schweinberger, & Wiese, 2015) and expertise or contact based models, in which greater expertise with own age faces may be due to supposed increased contact with our own-age versus other aged faces, is controversial. Thus it is still unclear why OAB is apparent. However, if older adults may perform differently when viewing own versus other aged faces, it follows that comparing older adults face processing mechanisms only to younger, but not older faces, may not reflect the whole picture of age-related differences in face processing. To resolve such an issue, the experiment conducted in this chapter could be extended to include images of older adult faces that are part of the stimulus database used for this experiment with the same participants to quantify if diagnostic features or timing in feature encoding is sensitive to differences in the age of face stimuli shown.

In conclusion, we have shown that the content of early visual ERPs in older adults does not differ from that of younger adults, though the processing of these same features is weaker and delayed in healthy ageing. Specifically, older adults' processing of the eyes is comparatively delayed compared to younger adults across tasks. However, the relative timing of processing other facial features, specifically the mouth, is idiosyncratic, with the relative timing of feature processing not consistent across tasks, and not uniformly delayed with ageing.

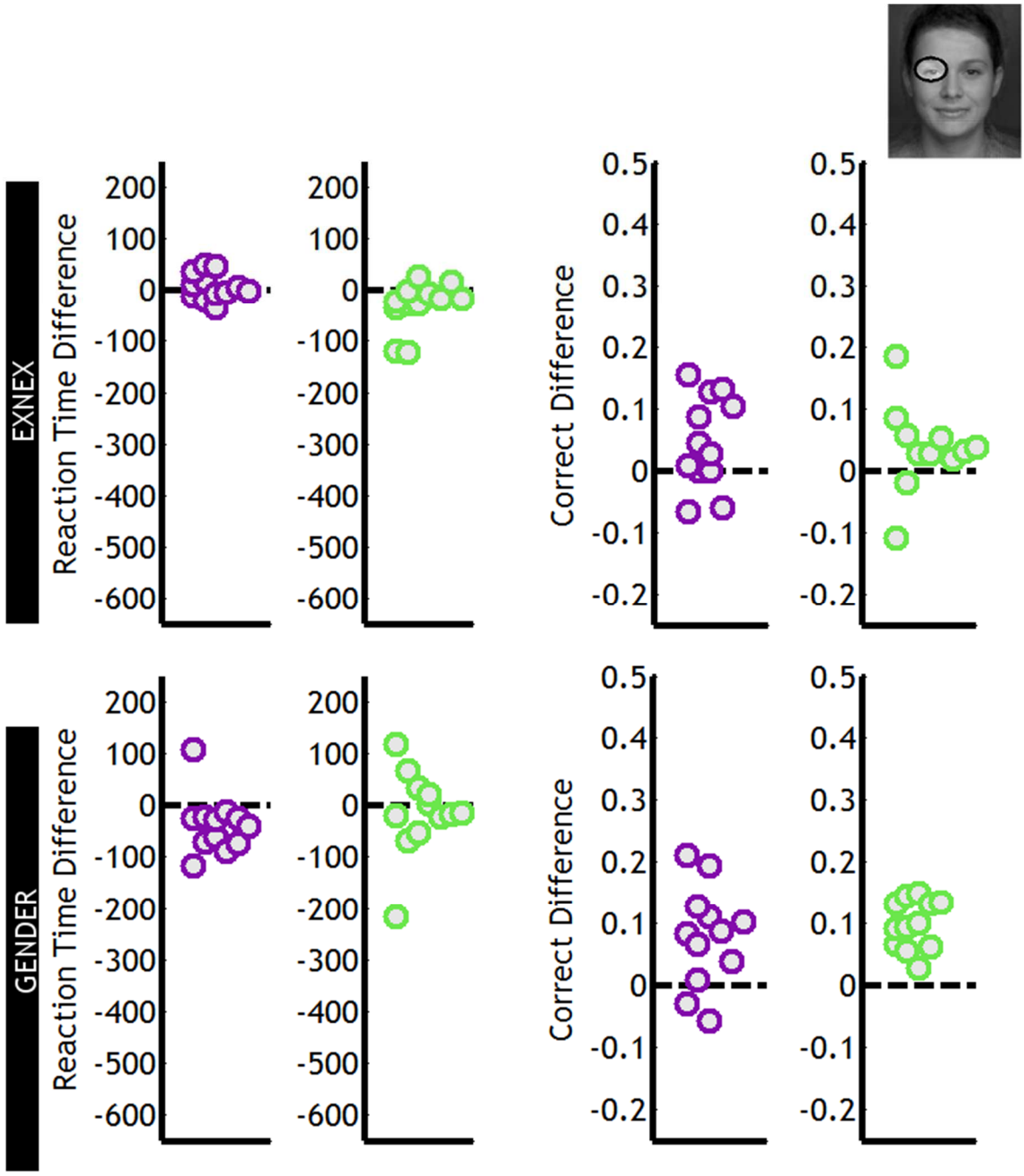
Chapter 5 Supplementary Figures



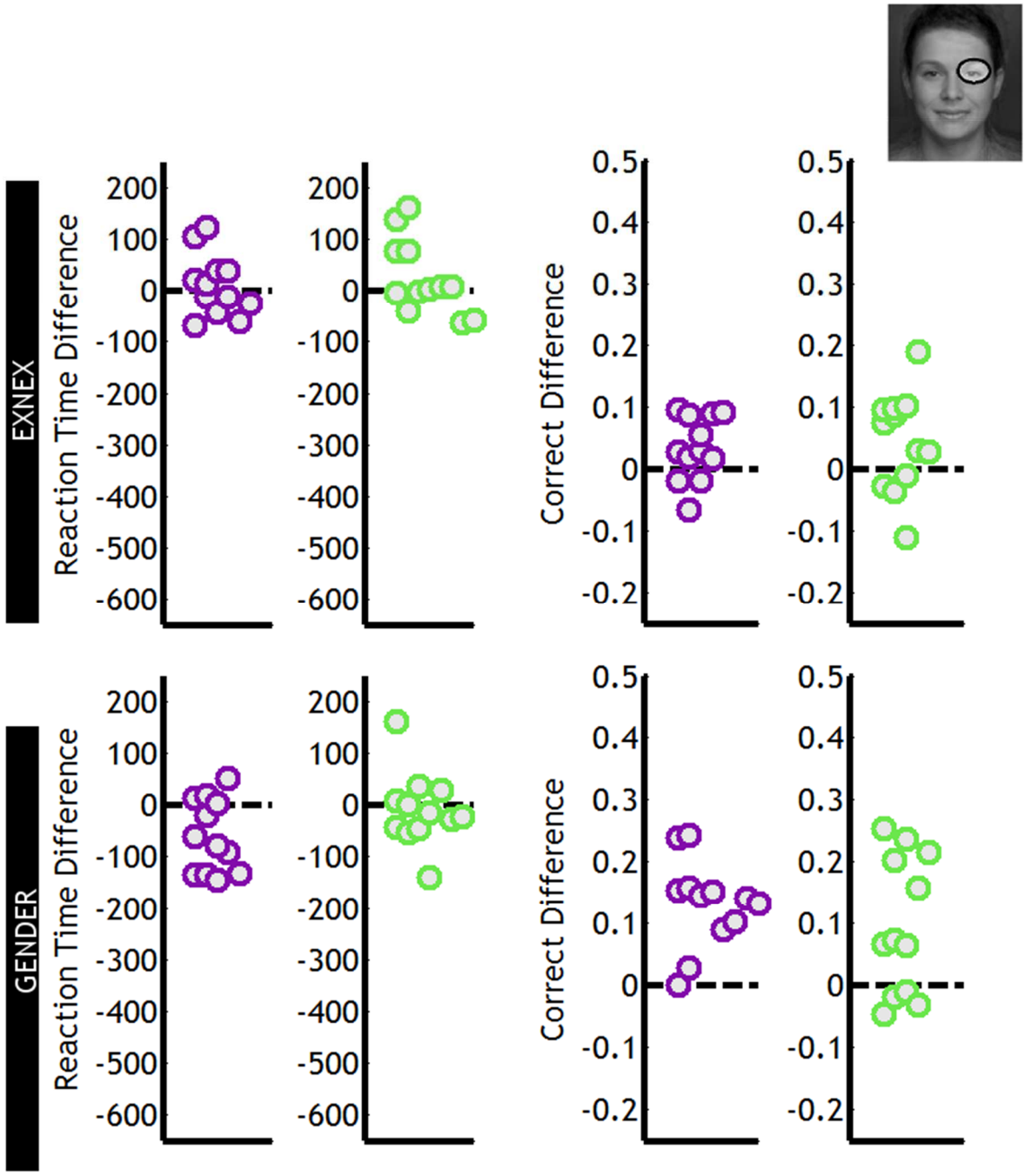
Supplementary 20: Older adult Individual behavioural classification images for all participants. MI(pix, **RT**) and MI(pix,**CORR**) for the **EXNEX** and **GENDER** task. Left column: Participants who completed the **EXNEX** task first. Right column: Participants who completed the **GENDER** task first.



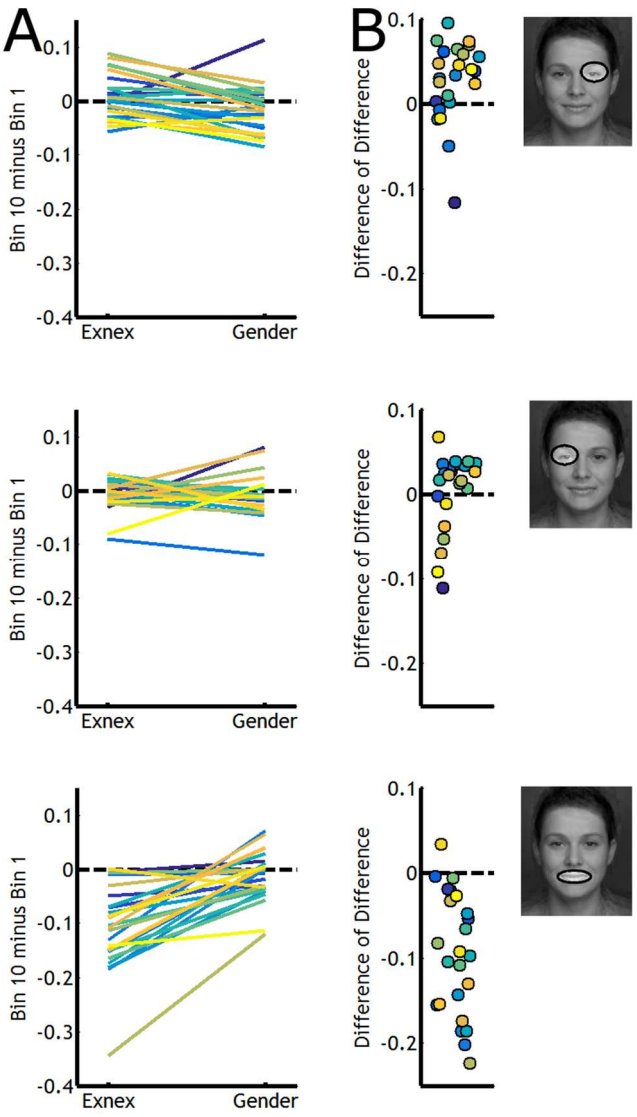
Supplementary 21: Older adult mouth visibility task order effects Differences between high (bin 10) and low (bin 1) visibility of the mouth in the **EXNEX** (top row) and **GENDER** (bottom row) task. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first. Behavioural differences are broadly the same regardless of task order.



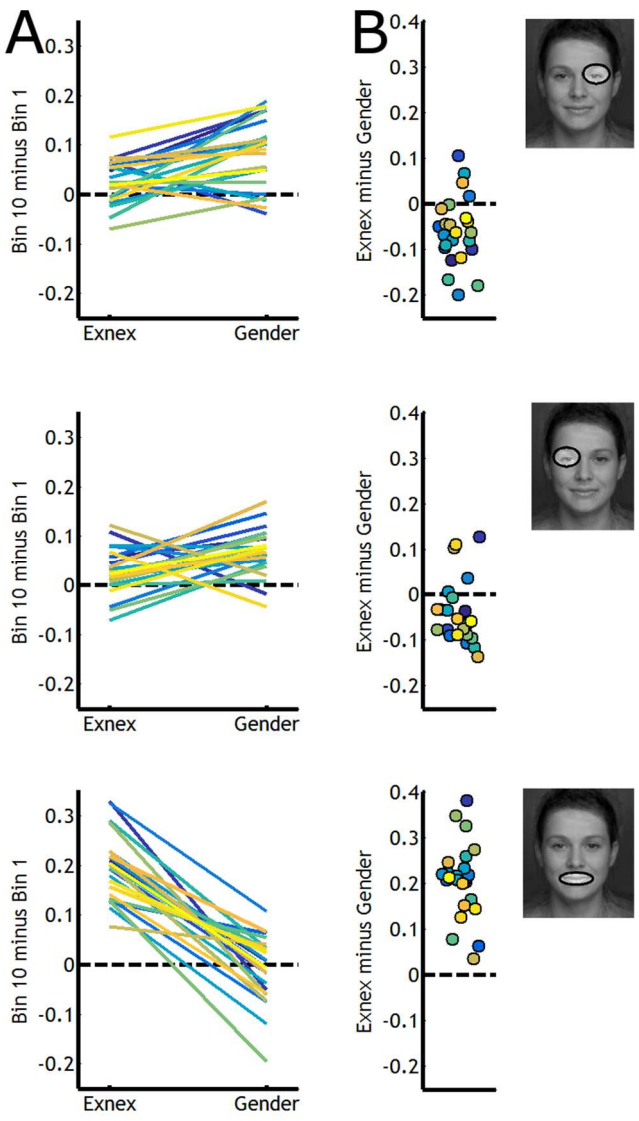
Supplementary 22: Older adult left eye visibility task order effects Differences between high (bin 10) and low (bin 1) visibility of the mouth in the **EXNEX** (top row) and **GENDER** (bottom row) task. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first. Behavioural differences are broadly the same regardless of task order.



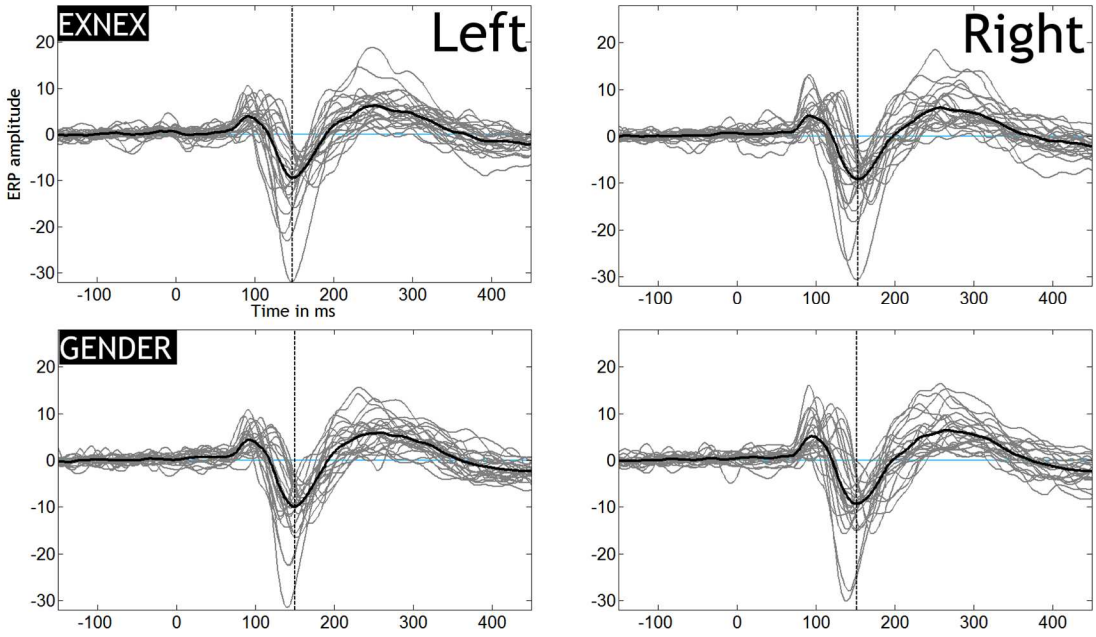
Supplementary 23: Older adult right eye visibility task order effects Differences between high (bin 10) and low (bin 1) visibility of the mouth in the **EXNEX** (top row) and **GENDER** (bottom row) task. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first. Behavioural differences are broadly the same regardless of task order.



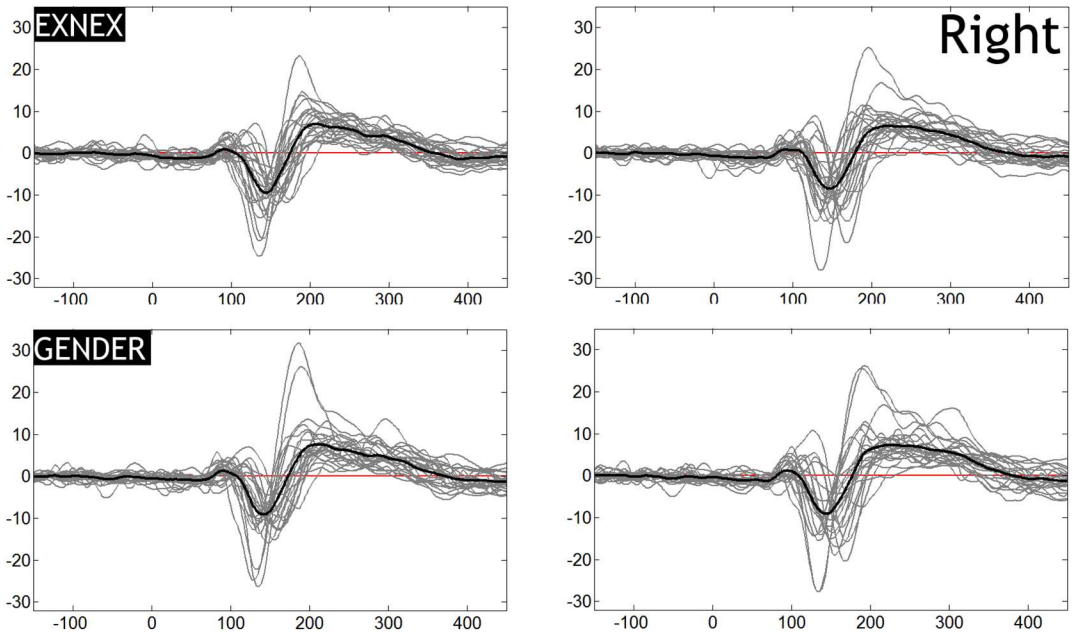
Supplementary 24: Older adult normalised Reaction time task differences by feature visibility Behavioural data normalised by dividing the difference of bin 10 minus bin 1 by the difference of bin 1 + bin 10 for each participant and facial feature. Panel A: The difference in normalised reaction time for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in normalised reaction time for the **EXNEX** minus **GENDER** task for each facial feature of interest.



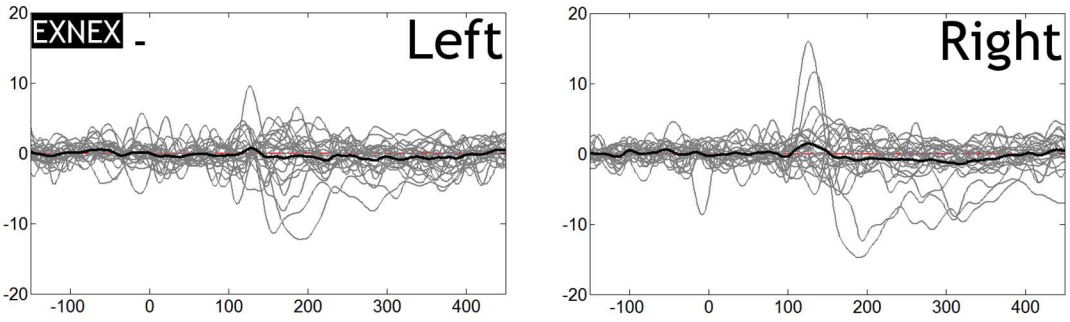
Supplementary 25: Older adult normalised Accuracy task differences by feature visibility
Behavioural data normalised by dividing the difference of bin 10 minus bin 1 by the difference of bin 1 + bin 10 for each participant and facial feature. Panel A: The difference in normalised accuracy for the left eye (top), right eye (middle) and mouth (bottom). Each line represents one participant. Panel B: Difference of differences. The difference in normalised accuracy for the **EXNEX** minus **GENDER** task for each facial feature of interest.



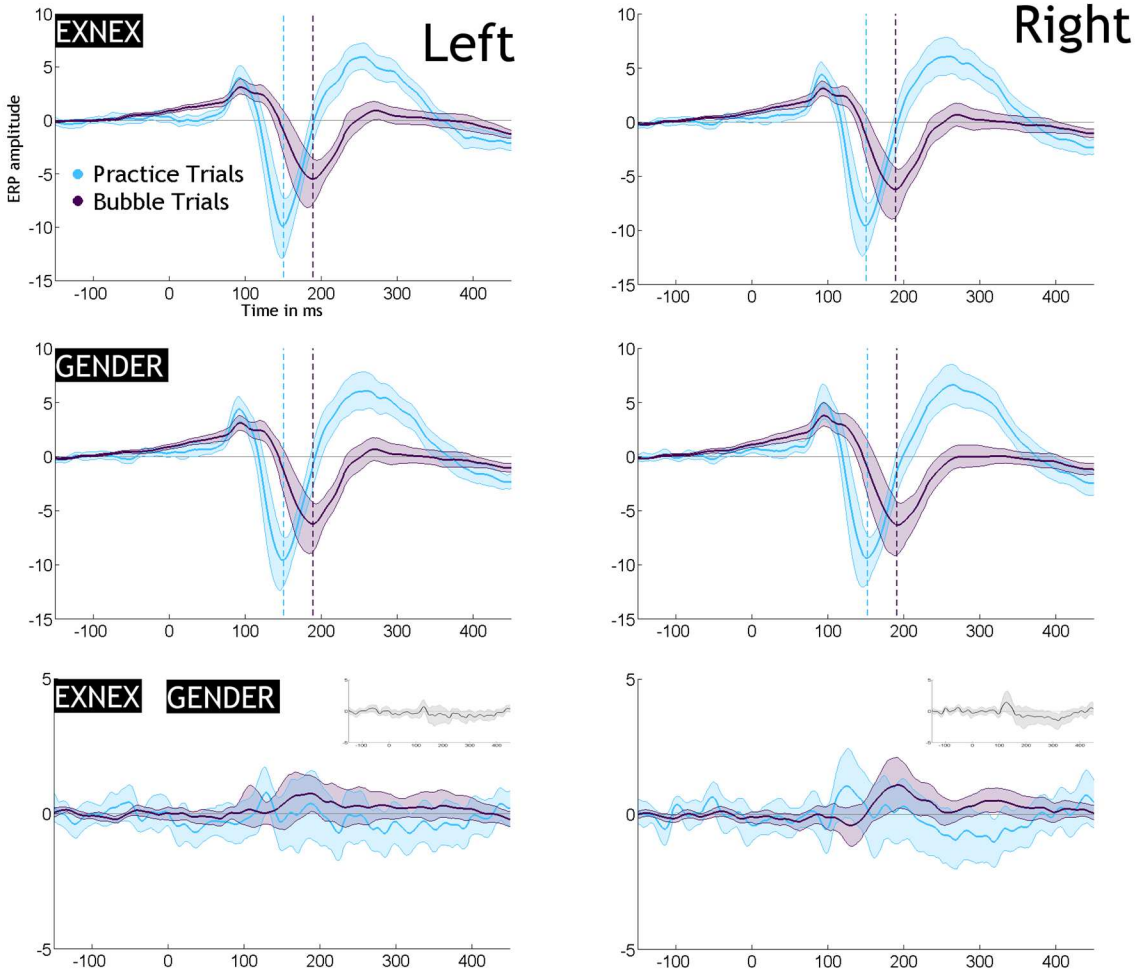
Supplementary 26: Older adult individual Mean ERPs (Practice Trials) Mean ERPs for each participant (N = 24) are superimposed in grey for the left and right hemisphere in each task. Solid black line represents the group mean. Dashed black line represents the latency of the group average minimum N170 amplitude.



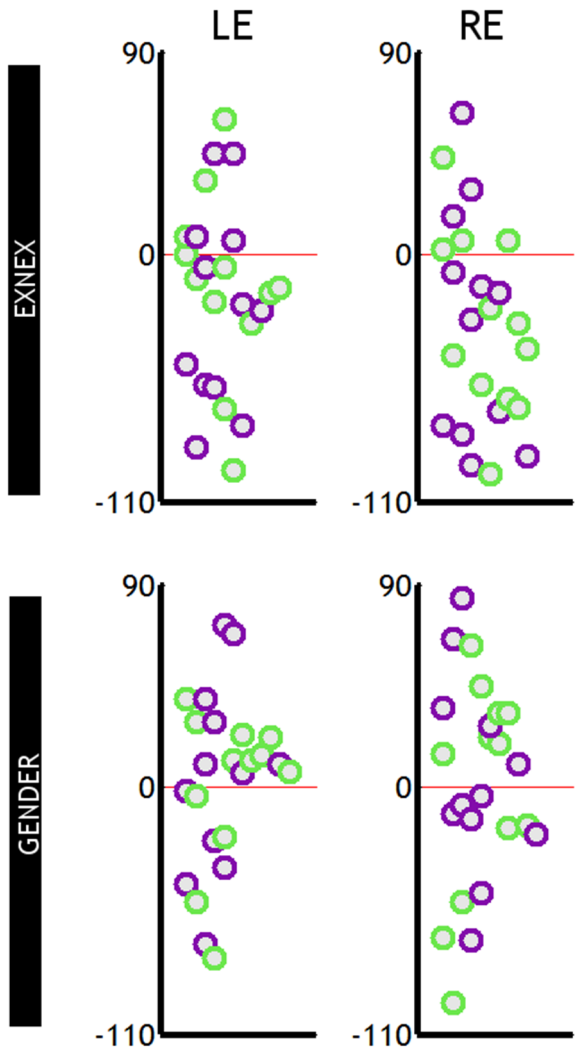
Supplementary 27: Older adult individual Mean ERPs (Bubble minus Practice Trials) Difference in Bubble minus Practice trials mean ERPs for each participant (N = 24) are superimposed in grey for the left and right hemisphere in each task. Solid black line represents the group mean.



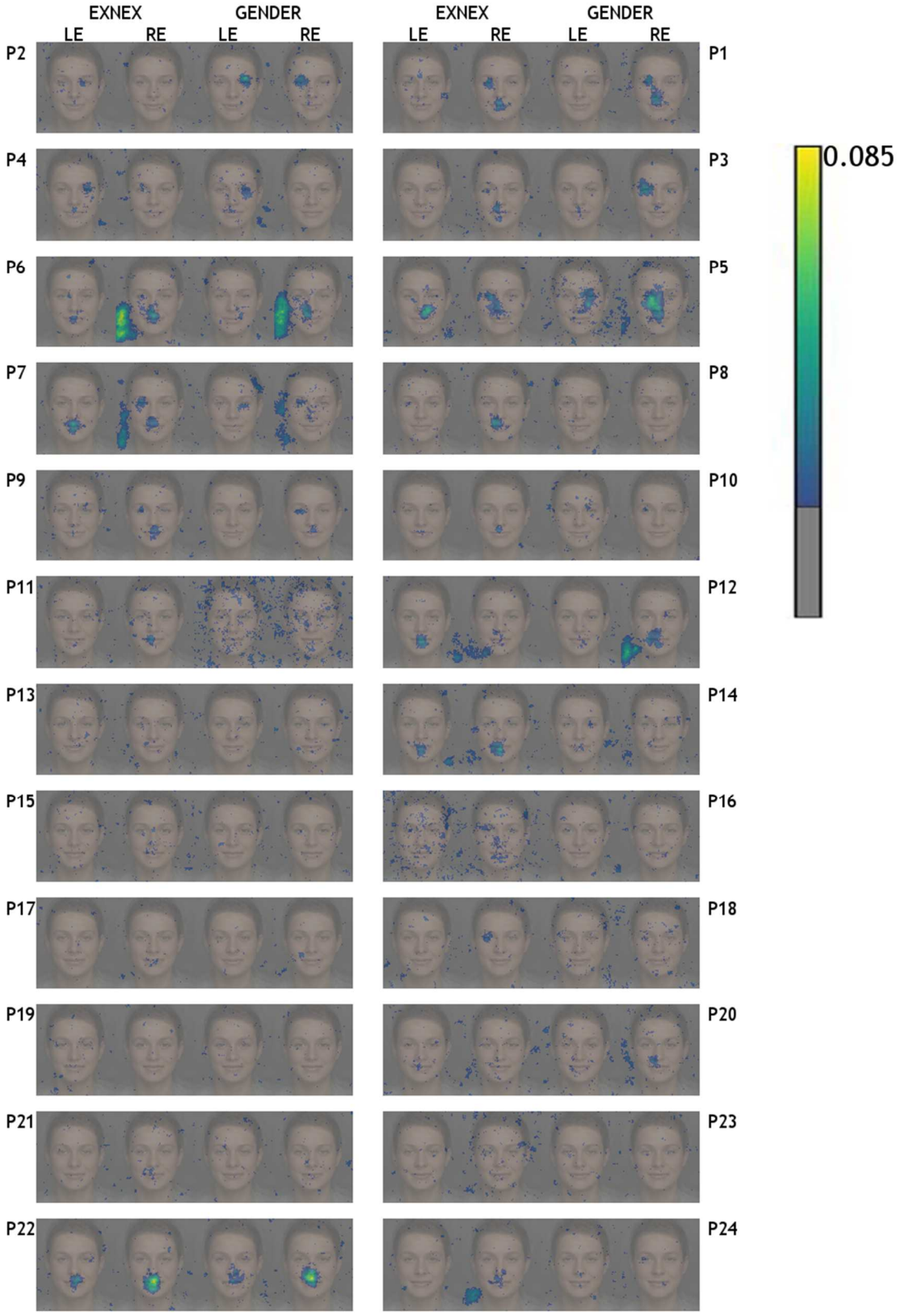
Supplementary 28: Older adult individual Mean ERPs (EXNEX – GENDER) Difference in Bubble minus Practice trials for EXNEX minus GENDER mean ERPs for each participant (N = 24) are superimposed in grey for the left and right hemisphere in each task. Solid black line represents the group mean.



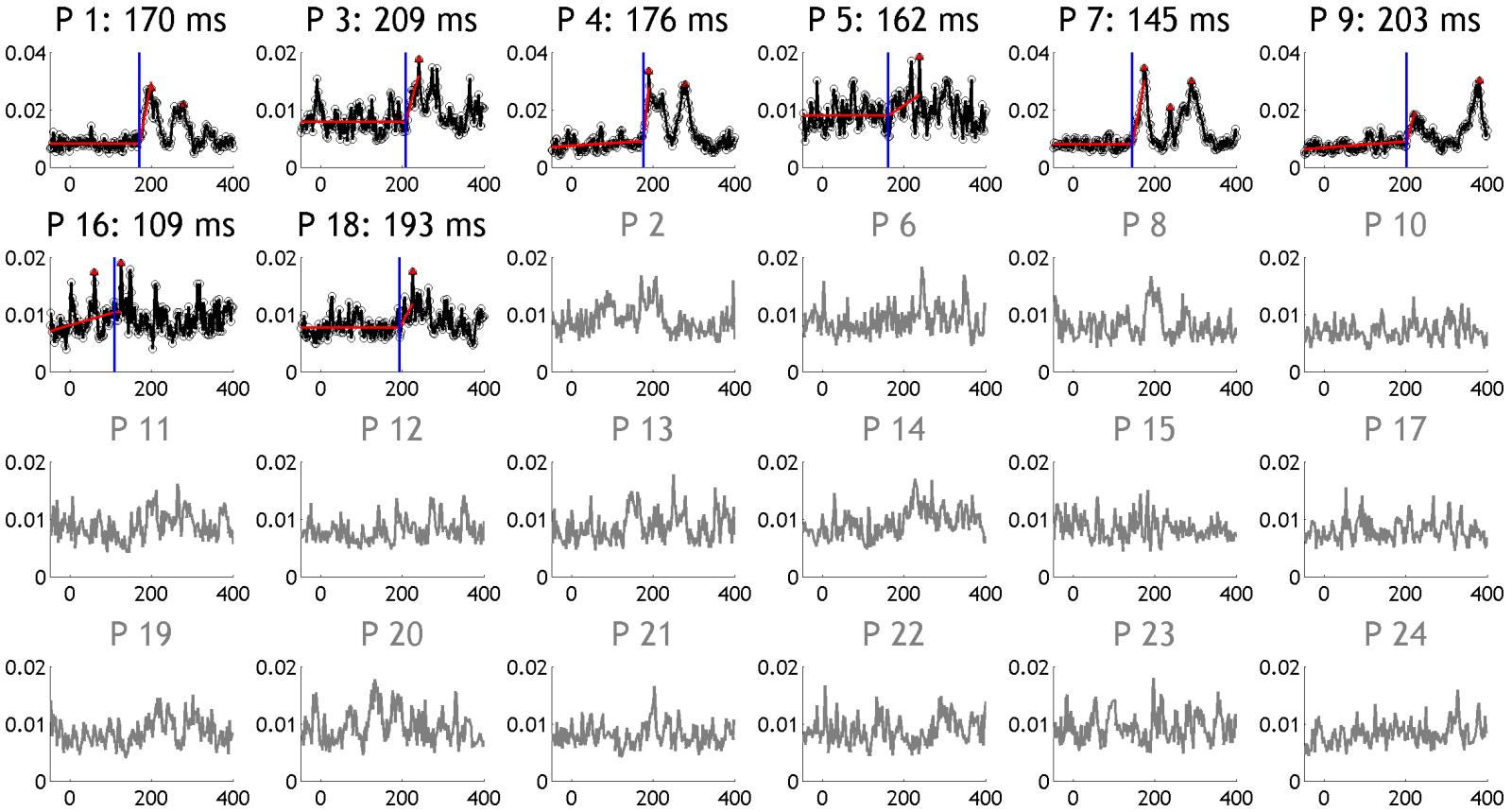
Supplementary 29: Older adult 20 % trimmed mean average group ERP in Bubble and Non-Bubble Trials 20 % trimmed mean bubble and non-bubble trial ERPs for the left and right hemisphere with 95 % confidence intervals around the 20 % trimmed mean. Vertical lines represent the minimum amplitude peak of the N170 for each task. Bottom panel **EXNEX** minus **GENDER** for bubble and practice trials. Small grey plot shows the pairwise difference of practice minus bubbles trials.



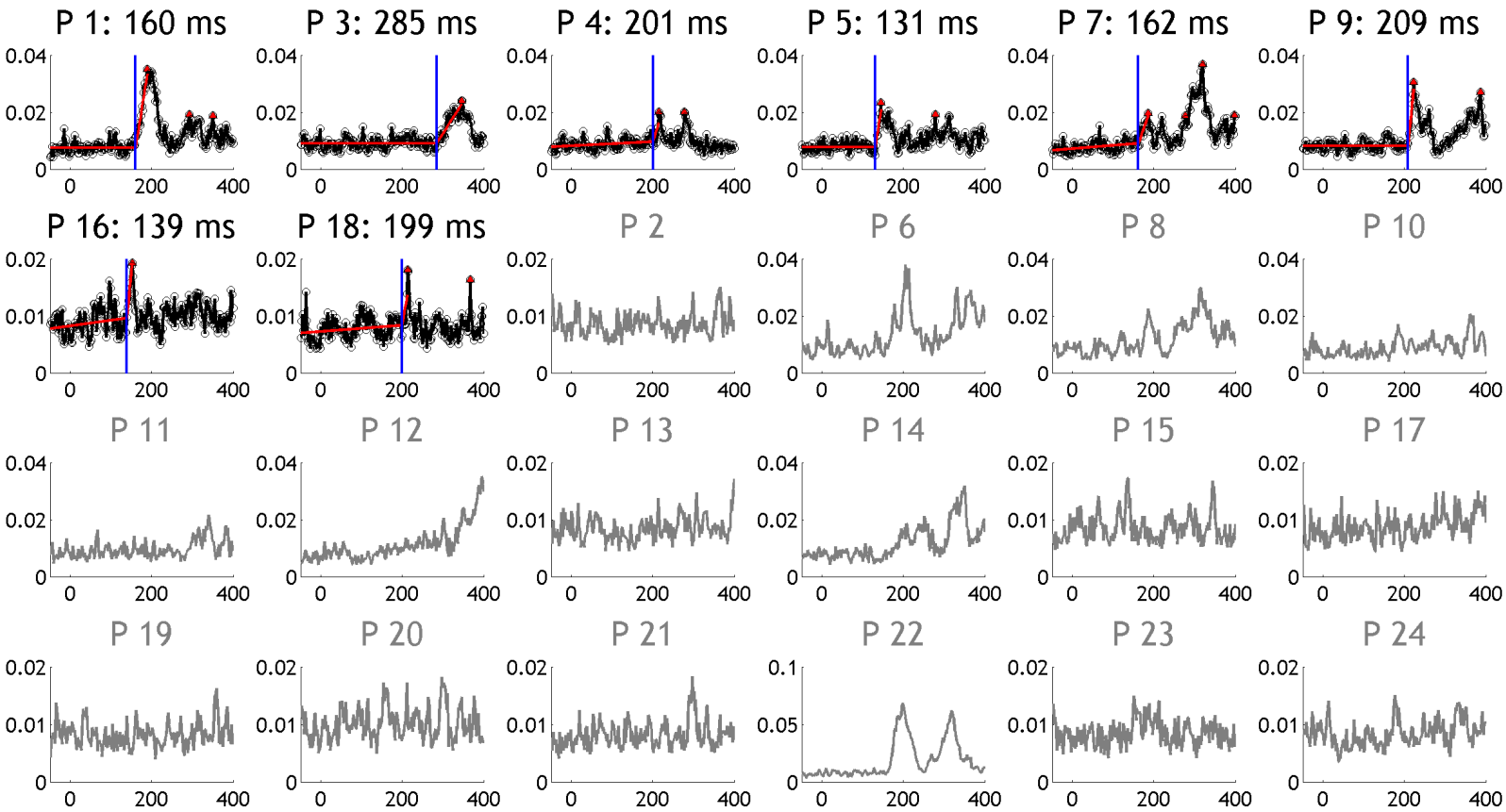
Supplementary 30: Older adult individual Differences in MI peak latency For each participant we calculated the difference in the latency of the peak MI between 120-220 ms for the contralateral eye minus the mouth. Purple circles are participants who completed the **EXNEX** task first. Green circles are participants who completed the **GENDER** task first



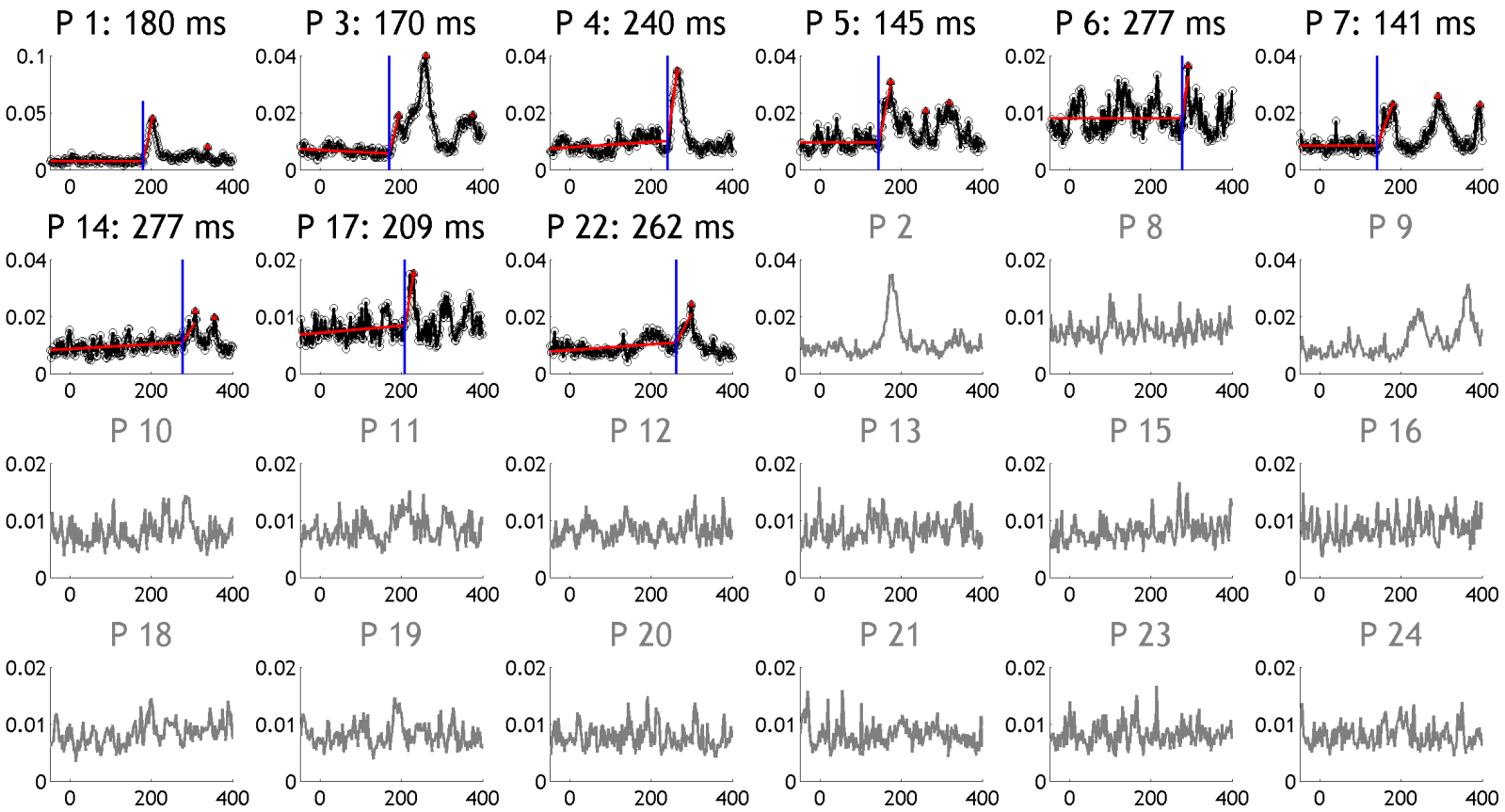
Supplementary 31: Older adult individual brain classification images for all participants. MI(pix, [**ERP,grad**]) for the **EXNEX** and **GENDER** task. Left column: Participants who completed the **EXNEX** task first. Right column: Participants who completed the **GENDER** task first.



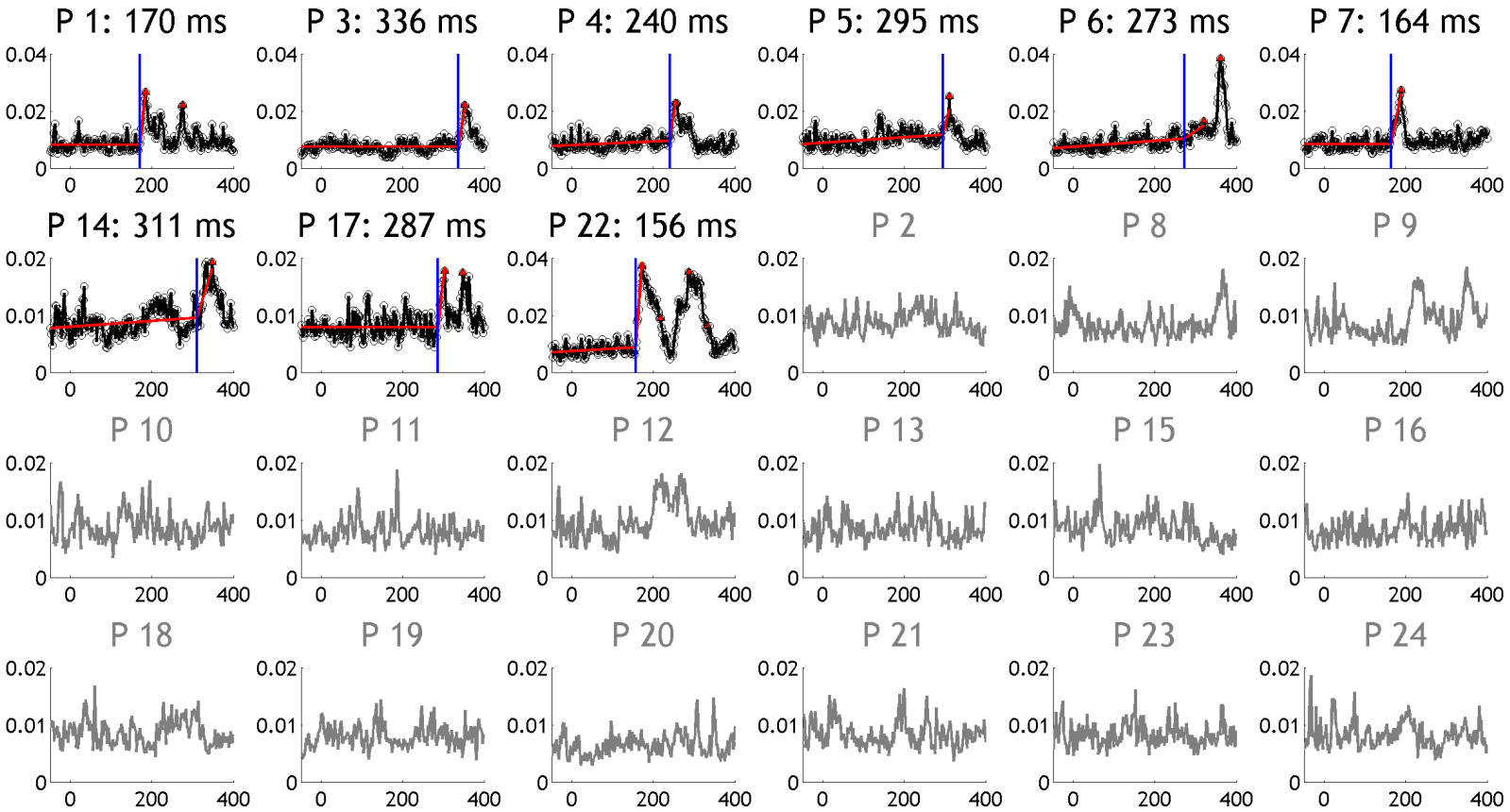
Supplementary 32: Older adult MI eye onset (EXNEX) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.



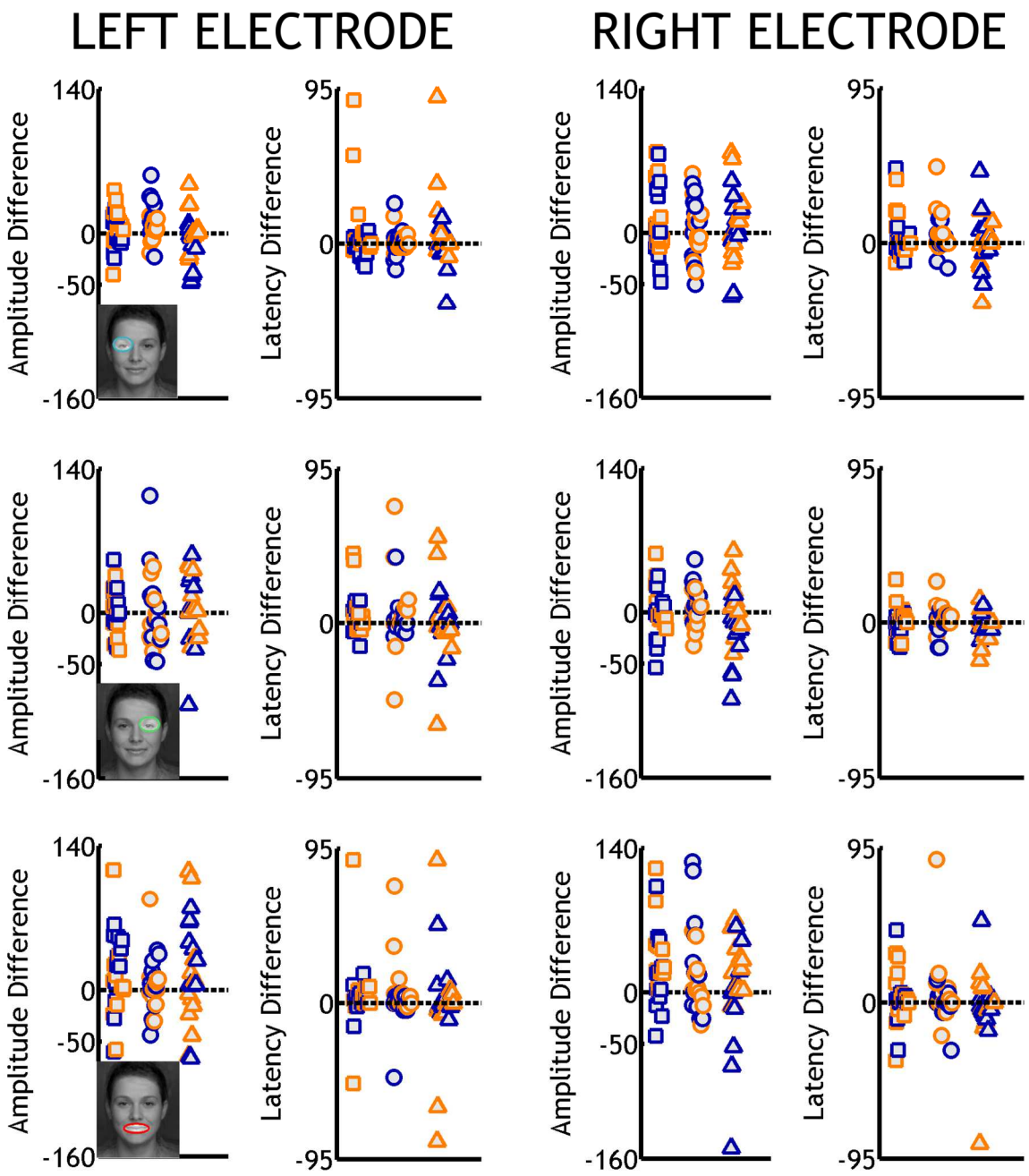
Supplementary 33: Older adult MI mouth onset (EXNEX) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.



Supplementary 34: Older adult MI eye onset (GENDER) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.



Supplementary 35: Older adult MI mouth onset (GENDER) For each participant we plotted the maximum MI across the left and right electrode to the left and right eyes. For each participant the time course is shown in black. Peaks 2.25 times larger than median baseline are highlighted by a red triangle. Red line depicts the model estimation. Blue lines indicate the estimation of the onset of MI, which is stated above each plot. Plots in grey show individuals whose data was not included in the analysis as no peak was identified for the eyes and/or mouth time course.



Supplementary 36: Older adult N170 individual amplitude and latency differences. Individual differences in amplitude and latency between bin 10 minus bin 1. Median N170 latency differences were calculated between bin 10 (most visible) minus bin 1(least visible) and are presented in milliseconds. Median N170 amplitude differences are calculated as a percentage of bin 1, such that an amplitude difference of 50% means that the amplitude of bin 10 was 150% the size of amplitudes in bin1. Squares represent **EXNEX** whilst circles represent **GENDER**. Triangles represent the difference between **EXNEX** minus **GENDER**. Blue symbols represent participants completing the **EXNEX** task first. Orange triangles represent participants completing the **GENDER** task first.

General Discussion

Faces are important social stimuli for achieving mutual understanding and communicating a wide range of signals, such as threat and aggression. We engage in many basic face processing tasks automatically in naturalistic settings, such as judging the age, sex and emotional state from others faces. However, our understanding of the computational stages that visual processing undergoes from initial sensory input to decision making, and how these information processing stages are affected by healthy ageing remains elusive.

Through investigating the information processing steps during the most basic face processing task, face detection, previous work has revealed that increased visibility of the contralateral eye was associated with quicker and more accurate behavioural responses, as well as increased eye visibility modulating the EEG signal (Rousselet et al., 2014). Furthermore, this pattern of results was also evident in older adults, though this association was weaker and delayed (Jaworska, 2017). Coupled together, these results suggest that eye encoding is the first stage of visual processing in a face detection task across the lifespan, and that an age-related slowing of information processing of the same facial feature occurs. This age related slowing is in line with Salthouse's (1996) theory that perceptual and cognitive impairments in ageing can be accounted for by a generalised slowing down of neural information processing with age, as well as extending previous results charting a 1 ms per year slowing down of visual processing speed (Rousselet et al., 2009).

In this thesis we aimed to demonstrate that contralateral eye sensitivity in a face detection task is not an artefact of a left attentional or left gaze bias, nor due to alignment between the size of the Gaussian apertures used to reveal the stimulus and the size of the eye region of the stimulus. In Chapter 2, we used 4 image sizes, maintaining the same Gaussian aperture size across all the stimulus sizes, whilst varying the number of the apertures to ensure an approximately equal percentage of the stimulus space was revealed. We found that reaction times and accuracy were modulated by visibility the eye region for face but not noise trials. With increasing image size, we found a shrinking ‘hotspot’ of sensitivity, revealing sensitivity to the pupil/iris area with increasing specificity. We also found N170 sensitivity to the contralateral eye, also with a hotspot that

shrank with increasing face sizes – however, N170 contralateral eye sensitivity lateralisation was almost lost for very small faces (3 degrees of visual angle). These results suggest that, at least in a face detection task, the feature driving behavioural and brain signals is iris/pupil area and that contralateral eye sensitivity is size tolerant for all but very small faces. Our failure to find lateralisation in the smallest face size was not due to using a single Gaussian aperture, as contralateral eye sensitivity has been shown for larger image sizes using a single large bubble (Rousselet et al., 2015), but rather seems to be a consequence of the very small size of the stimulus being presented foveally. The 'tracking' of the eye across all other stimulus sizes, coupled with no sensitivity to the equivalent area on noise tasks, refutes alternative left attentional or gaze bias explanations of these results.

Whilst face detection tasks give us the opportunity to understand face processing in its most basic elementary stages and is a good starting point for understanding *what* information the brain processes when it detects a face, in Chapter 4 we aimed to shed light on the information processing steps in more complex face processing tasks of gender and expressiveness categorisations. During face detection tasks trial-by-trial uncertainty is introduced as to whether a face will be present or not. In our tasks, participants had top-down knowledge that faces would always be presented. Removing this uncertainty and changing the task demands from a face-detection to a gender or expressiveness detection task may alter the features initially processed. For example, it is possible that if processing the eyes is only the initial processing step in a particular subset of face processing tasks, including face verses noise detection for example, that eye processing may not be evidenced in an expressive versus no expressiveness task if the eyes provide no diagnostic information for the task. Alternatively, it is also possible that eye processing is a universal initial face processing step, in which case eye processing may proceed the processing of other facial features, even when the eyes provide no diagnostic information for the task.

Previously it has been reported that participants rely on the eye and mouth regions to resolve a gender categorisation task, whilst they rely only on the mouth to resolve an expressive/non-expressive categorisation task (Gosselin & Schyns, 2001). Whilst there has been some inconsistency in the supposed

features driving behavioural responses in the gender task using the same stimuli (e.g. Schyns, Bonnar, & Gosselin (2002) suggest that only the left eye from the viewers perspective is used in the gender task, whilst Schyns, Jentzsch, Johnson, Schweinberger, & Gosselin (2003) suggests both eyes but not the mouth are used), a general consensus from these results can be formed that during the EXNEX task participants rely on the mouth, whilst in the GENDER task they rely on at least 1 if not both eyes and possibly also the mouth region. In contrast, in both tasks modulation of the N170 has been suggested to be associated with encoding of the contralateral eye, whilst the later P300 component reflects the processing of diagnostic features (the mouth in the EXNEX task and eyes in the GENDER task) (M. L. Smith et al., 2004).

We attempted to replicate these results in Chapter 4 with a new stimulus set that offered the advantage of coloured images of faces that had not been manipulated to conform to a normalised hairstyle and were not wearing makeup. We also employed a new image sampling technique - *BubbleWarp* – that we outlined in Chapter 3, which allowed us to present stimuli with naturalistic variation in e.g. inter-feature distances. In addition, we used mutual information to quantify the dependence between stimulus features and behavioural and brain responses, which offered the advantage of being non-parametric and able to detect both linear and non-linear associations.

We found the strongest behavioural associations between the mouth and increased accuracy and reaction times in the EXNEX task, and the eyes and increased accuracy and reaction times in the GENDER task. However, diagnosticity to the eye region was weak, and our results suggested an idiosyncratic preference for the eyes that varied in strength across participants for the gender task. We found that single-trial ERPs were modulated by the presence of the contralateral eye region and mouth in both the EXNEX and GENDER task. We showed that mouth modulation was stronger in the EXNEX than GENDER task, whilst eye modulation was slightly stronger for the GENDER than EXNEX task though weaker than mouth modulation. This highlights a disparity between the task-relevant information for behaviour and that information coded in the brain. During the time window of the N170, we highlighted the encoding of both the contralateral eye and mouth regions in both the EXNEX and GENDER

task. This is in stark contrast to Smith et al.'s (2004) results suggesting that the N170 was modulated by the contralateral eye only. We have shown that in both the GENDER and EXNEX task, increased visibility of the mouth increased N170 amplitudes at both the left and right posterior lateral electrodes. Increased mouth visibility was also associated with shorter N170 latencies in the EXNEX task at left posterior lateral electrodes, and in the GENDER task at right posterior lateral electrodes. This suggests that the N170 is not only sensitive to the contralateral eye as suggested by Smith et al. (2004), but is also modulated by other facial features, such as the mouth. We also found evidence of sensitivity to the ipsilateral eye modulating N170 amplitudes during both tasks, particularly for the left ipsilateral electrodes. This is in-keeping with a recent study suggesting that the N170 reflects the coding of the contralateral eye followed by the transfer of communication about the ipsilateral eye (Ince et al., 2016).

Vitally, through estimating the onset on MI towards the eyes and mouth regions, we quantified the timing of sensitivity to the eyes and mouth, suggesting that the onset of MI to the eyes preceded the onset of MI to the mouth region by 13 ms in the EXNEX and 12 ms in the GENDER task. However, these results should be treated with caution as at the individual level we evidenced a heavily skewed distribution and the direction of this effect was only apparent in 70 % of participants in the EXNEX task and 60 % of participants in the GENDER task. Thus we conclude that there is an idiosyncratic preference for contralateral eye sensitivity preceding sensitivity to other facial features.

In Chapter 5, we extended our results to understand how older adults compared to younger adults. We found that behaviourally, older adults relied on the same facial features as young adults used as diagnostic cues to resolve gender and expressiveness categorisation tasks. However, compared to our younger adults, older adults were more heavily dependent upon the presence of the mouth to correctly and quickly determine expressive and non-expressive faces.

Comparing the time taken to integrate 50 % of their MI time course, we found that older adults integrated information about the eyes consistently slower than younger adults – approximately 20 ms slower in the EXNEX and 23 ms in the GENDER task, consistent with the 23 – 25 ms delay in processing the

contralateral left and right eye on right and left posterior lateral electrodes respectively in a previously reported face detection task (Jaworska, 2017). Interestingly, our relatively similar delay in eye coding as in a less complex face detection task may be in disparity with Salthouse's (1996) theory of generalised slowing, as this theory would point towards expectations of an increased delay in processing in more complex tasks. However, our results should be treated with caution, as, whilst bubbles delayed ERPs in both younger and older adults, using bubbled stimuli may have more of an effect on older adults perception, as it has been suggested that eye processing is not delayed or weaker in older compared to younger adults when the face context is preserved (Jaworska, 2017).

In contrast to the age-related delay in eye processing, we found a substantially reduced delay in processing the mouth in the GENDER task (approximately 9 ms) and no delay in mouth processing in the EXNEX task. This suggests that whilst older adults have delayed eye integration compared to younger adults, this does not result in a simple shift of the entire time course (where feature processing is uniformly delayed compared to younger adults) but rather a qualitative shift with a variable delay in feature processing that is inconsistent between features. These results are inconsistent with the idea of a generalised slowing of neural information processing proposed by Salthouse (1996) and instead suggest a more differential slowing. Future studies should be conducted to explore whether these results can be replicated with another group of participants and in other tasks for example. Also, extending analysis to the P300 would be interesting to understand how the later ERP time course compared between younger and older participants, as our results suggest greater activity around the P300 in older adults.

In the EXNEX task, we found that both younger and older adults processed information about the eyes faster than information about the mouth (though with large idiosyncrasies) though only younger adults processed the eyes faster in the GENDER task. However, our results should be treated with caution, particularly in older adults where several participants had flat MI time courses.

We did not record eye movements in either experiment, though differential gaze behaviour may be optimised for peak performance in a given perceptual task. For example, in a gender discrimination task, first saccades were to the left of

the nose (from the viewers perspective) and below the eye, whilst in comparison in an emotion discrimination task fixation points were systematically shifted downward (Peterson & Eckstein, 2012). These fixation locations have functional importance, with perceptual performance degrading when fixations were forced away from preferred fixation locations. Based upon an ideal observer, the eye region was identified as containing the most diagnostic information for the two tasks, whilst the mouth was also informative, particularly for emotion discrimination. However, saccade distributions were not consistent with a strategy of fixating on the most informative feature for each task. Instead, results were consistent with an optimally foveated strategy. For example, in a happy versus neutral discrimination task the foveated ideal observer estimation of fixating on the tip of the nose is reminiscent of the distribution of first saccades in real participants (Peterson & Eckstein, 2012) with the assumption being that visibility of the eyes and mouth are maintained in the periphery. However, future studies should consider the influence of ageing on the degradation of peripheral vision, as the "useful field of view" may shrink with ageing (Ball, Beard, Roenker, Miller, & Griggs, 1988) and result in changes in the most optimal fixation strategy. For example, in Chapter 4 older adults were more heavily dependent on the presence of the mouth to correctly determine happy versus neutral faces, so it may be that older adults' fixations are shifted more downward than in younger adults.

Similarly, eye tracking could also be used to ensure participants were fixating at the centre of the screen. Short stimulus presentation time means that our participants may have fixated off the central fixation cross in a systematic way, for example due to pre-emptive orientating towards their preferred fixation point for a particular task that could have affected results.

Contribution of the results

The results of this thesis have contributed to our understanding of face processing in several important ways.

Firstly, we have expanded on previous research by demonstrating that contralateral eye sensitivity in face processing is tolerant to a range of stimulus and task modulations. Contralateral eye sensitivity persists despite changes to the size of the stimulus, the type of stimulus set used, the age of participants and the type of task being completed (i.e. face detection, gender categorisation, expression categorisation). This suggests that contralateral eye sensitivity seems to be a consistent finding during face processing tasks. However, we have also demonstrated that contralateral eye sensitivity is subject to idiosyncratic differences between subjects – in our experiments not all participants demonstrated contralateral eye sensitivity, and there were differences in the timing and strength of contralateral eye sensitivity in those who demonstrated this result. Thus, our results suggest that contralateral eye sensitivity is not a *necessary* first stage in face processing as suggested by others (Rousselet et al., 2014; Schyns et al., 2007). Instead, we suggest that processing the contralateral eye is idiosyncratically preferential. This has contributed to the discussion within the research field as to the consistency of the contralateral eye sensitivity finding, as well as suggesting that more research is needed to understand the purpose of contralateral eye processing.

Given this result, it is not clear whether contralateral eye sensitivity offers a processing advantage, or why there is inter-subject variability in its presence or strength. Further research could investigate this matter, for example by investigating groups of “contralateral eye processors” versus “non-eye processors” to investigate whether there are behavioural differences between these groups, and whether there is intra-subject variability in contralateral eye sensitivity over different tasks for example to investigate if these differences in contralateral eye sensitivity preferences are stable within participants.

Secondly, our results have contributed to the discussion around age-related differences in processing speed. Consistent with other studies, we suggest that ageing is characterised by a slowing down of processing. However, unlike other studies, we suggest that age-related differences in processing speed may be feature and task dependent – whilst there was delayed processing of the eyes in both tasks, the delay in mouth processing was substantially reduced and potentially task dependent. This is a novel finding, and should be investigated

more thoroughly in the future for example by testing the same individuals on a greater range of face processing tasks, as well as collecting data from middle-aged adults to assess when age-related differences begin. In the current experiments we did not assess how the recruitment of difference brain areas may contribute to age-related differences in processing speed. In the future, studies should consider combined EEG-MEG or EEG-fMRI methods to more fully investigate how any underlying changes in recruitment of brain areas relates to differences in processing speed.

Finally, our results have contributed to the discussion in the literature surrounding age-related differences in the N170. Consistent with other studies, we suggest that the N170 in older compared to younger adults is delayed for both full face and bubbled images, consistent with the idea that there is a delay in information processing with older adults. However, we did not find a clear difference in the amplitude of the N170 between younger and older adults during bubbled images, but did during full-face images. This suggests that the N170 is affected by our stimulus manipulation and that our lack of age-related differences in amplitude may be due to this manipulation.

Limitations and Future Directions

In Chapter 2 faces were interspersed with texture trials. Future research should expand investigating contralateral eye sensitivity in tasks during which faces are always presented (compared to presenting faces intermixed with noise trials) to understand the role of habituation on contralateral eye sensitivity. Repeated presentation of a small number of face stimuli may lead to habituation of brain responses, such as a habituation in the FFA (Gauthier et al., 2000). The N170 may also habituate to repeated face presentations, such as a decline in amplitude and latency for repeated faces (Heisz et al., 2006b; Itier & Taylor, 2002). It is unclear how repetition of small face set might affect contralateral eye sensitivity over successive repetitions. This could be tested directly using the data from Chapter 4 and 5 where faces were always presented, for example by quantifying whether there is habituation of the N170 over repeated trials, and whether habituation of the N170 is different in older adults. This could for example help to elucidate why we found weaker levels of mutual information to

the eyes in Chapter 4 and 5 than in a face detection task (Rousselet et al., 2014). For example, if the function of the N170 is due to a process of face detection, top-down knowledge that faces will always be presented and our restricted number of face stimuli being repeated may reduce the N170 response and strength of MI towards the eye region.

One limitation of Chapter 4 & 5 is the construct of gender. *Gender* is a social and cultural construct that is not synonymous with biological sex. In Chapter 4 & 5 participants were asked only to discriminate between 'male' and 'female' – suggestive of a gender or sex binary. We do not know whether participants interpreted the instructions of the task as instructing them to consider the sex or the gender of the faces. This could have had consequences for our results. For example, in Chapter 3 we outlined several criticisms of the stimuli used by Gosselin & Schyns, (2001). In particular we highlighted that models appeared to be wearing make up in these stimuli which may have altered the contrast and diagnosticity of the eye region. However, in a purely *gender* discrimination task makeup (worn by any sex) may be a diagnostic feature for identifying an individuals' expressed gender (but not necessarily congruent with an individuals' sex). An investigation into whether diagnosticity of the eyes changes if participants are instructed to categorise a faces sex or gender, and the potential confound of eye makeup could be interesting. It is possible that the presence of eye makeup is negligible and that the eyes are still diagnostic even in the absence of eye makeup – but is there biological sexual dimorphism in the region of the eyes that might be diagnostic in a *sex* discrimination task? Evidence suggests that there is sexual dimorphism in the anatomy of the eye orbit (Samal, Subramani, & Marx, 2007). In particular the inter-canthal width (distance between the inner corner of the left and right eye) and outer-canthal width (distance between the outer corners of the left and right eye) is significantly different between adult males and females (Kesterke et al., 2016). However, sexual dimorphism in the anatomy of the eye orbit may be *less* pronounced than the sexual dimorphism of the lips and mouth (Samal et al., 2007). Whilst some participants were more accurate in the gender task with increased visibility of the mouth region, the increased visibility of the eye region had more of an effect on accuracy. It would be interesting in the future to use image

manipulation of the eye region to understand what information from the eyes is useful in categorising sex.

Another limitation of Chapter 5 is ensuring our older adult participants were a representative sample of 'healthy' older adults. Here, there are two potential issues in our study. Firstly, whilst older adult participants passed the MoCA cognitive screening test, it is possible that our participants already have pathological brain changes, as changes may occur before behaviourally evident cognitive decline. Therefore, we cannot guarantee that our 'healthy' older adult sample can be considered 'healthy' on the basis of behavioural screening tests alone. However, the time and financial burden of more detailed testing is often not practical. Conversely, our older adult participants could also be considered non-representative due to potentially being 'super-agers'. Our older adult participants were unusual – older adult participants were recruited within the vicinity of the university and also through university links, such as through the university staff alumni network. As such it is likely that our sample of older adults could be considered more healthy and well-educated compared to their peers. Golomb et al. (2012) suggest that with increasing age, older adult participants becoming increasingly non-representative of the population they are intended to represent, as less vigorous and less healthy older adult volunteers may be unable or less inclined to participate in research. Many of our older adult volunteers have participated in several EEG research studies, suggesting that they are more motivated volunteers. Participation in prior studies may be problematic – increased test familiarity may mean that our older adult participants behaved unusually in the experiments themselves (for example having developed strategies for bubble stimuli due to previous exposure). It is also possible that our older adult volunteers also had increased familiarity with the MoCA screening test, which may have reduced the effectiveness of using the MoCA as a screening tool if test familiarity could compensate enough to mask early indicators of cognitive decline. Future research should attempt to ensure a more representative sample. For example Shafto et al. (2014) have successfully recruited participants using a population-based representative sample through GP registrations. It would also be advisable for future studies to consider carefully the familiarity of older volunteers to any cognitive screening tests, and potentially diversify the cognitive tests utilised within labs as necessary. Whilst

the MMSE is less sensitive than the MoCA (Smith, Gildeh, & Holmes, 2007) and may not make a suitable cognitive screening candidate for research purposes, a new computerised modified version of the Cambridge Brain Sciences Battery (mCBS) may offer a promising solution. Used alongside the MoCA the mCBS may provide useful additional information in determining whether older adults with borderline MoCA scores should be considered as experiencing mild cognitive decline or not (Brenkel, Shulman, Hazan, Herrmann, & Owen, 2017).

A second limitation of Chapter 5 was comparing older and younger adults in a cross-sectional manner in which age was treated as a categorical rather than continuous variable. Our results could be expanded e.g. through the recruitment and testing of middle-aged participants to examine continuous changes across the adult lifespan. For example by recruiting middle-aged participants it may be possible to determine if the age-related decline in eye sensitivity is progressive or non-linear, for example by using a similar approach as Rousselet et al. (2010) who identified a 1 ms/year decline in noise sensitivity. Future studies should consider using more longitudinal designs rather than cross-sectional designs. Discrepancies between the results of longitudinal and cross-sectional designs have been highlighted, such as whether there is a reduction in frontal cortex recruitment with ageing (Nyberg et al., 2010). It may be possible using our current older adult volunteer base to investigate longitudinal differences of some of our older adult volunteers who have participated in research for several years. The feasibility of such an approach is yet to be fully explored.

Our results are also limited due to the low spatial resolution of EEG. It remains unclear from the experiments described in this thesis which brain areas contributed to the reported effects, nor whether the same sources or processes were responsible for the effects reported in the two age groups. This question is particularly interesting given the finding of age-related delays in eye processing not being followed by a uniform delay in mouth processing. Understanding what is happening at the source level could elucidate these results. This issue could be addressed by using MEG for example.

Whilst we have expanded our feature encoding to different stimulus sizes and tasks, there are many face processing tasks we have not explored, for example, categorising facial identity, eye gaze direction or facial age. The extent to which

contralateral eye sensitivity is present in these tasks, and quantification on the timing of eye sensitivity compared to the sensitivity of other facial features could provide greater understanding of the role of contralateral eye sensitivity in face processing.

In this thesis we introduced a new technique to allow for more heterogeneous stimulus sets to be utilised. Our stimulus set however was still relatively uniform – all stimuli were front-view faces for example, and facial features were still generally close to the average face. The strength of this new technique is yet to be fully tested. For example, it would be possible to use this technique to explore face processing within natural scenes. This offers potential for future studies to explore face processing under more ecologically valid conditions.

List of References

Anastasi, J. S., & Rhodes, M. G. (2005). An own-age bias in face recognition for children and older adults. *Psychonomic Bulletin & Review*, *12*(6), 1043–1047. https://doi.org/10.3758/BF03206441

Andersen, G. J. (2012). Aging and vision: Changes in function and performance from optics to perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *3*(3), 403–410. https://doi.org/10.1002/wcs.1167

Arditi, A. (2005). Improving the Design of the Letter Contrast Sensitivity Test. *Investigative Opthalmology & Visual Science*, *46*(6), 2225. https://doi.org/10.1167/iovs.04-1198

Bailey, I. L., & Lovie, J. E. (1980). The design and use of a new near vision chart. *Optometry and Vision Science*, *57*(6), 378–387. https://doi.org/10.1097/00006324-198006000-00011

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: expanding the useful field of view. *Journal of the Optical Society of America A*, *5*(12), 2210–2219. https://doi.org/10.1364/JOSAA.5.002210

Batty, M., & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, *17*(3), 613–620. https://doi.org/10.1016/S0926-6410(03)00174-5

Beason-Held, L. L., Goh, J. O., An, Y., Kraut, M. A., O'Brien, R. J., Ferrucci, L., & Resnick, S. M. (2013). Changes in Brain Function Occur Years before the Onset of Cognitive Impairment. *Journal of Neuroscience*, *33*(46), 18008–18014. https://doi.org/10.1523/JNEUROSCI.1402-13.2013

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological Studies of Face Perception in Humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. https://doi.org/10.1162/jocn.1996.8.6.551

Bentin, S., & Golland, Y. (2002). Meaningful processing of meaningless stimuli: The influence of perceptual experience on early visual processing of faces. *Cognition*, *86*(1), 1–14. https://doi.org/10.1016/S0010-0277(02)00124-5

Bentin, S., Sagiv, N., Mecklinger, A., Friederici, A., & von Cramon, Y. D. (2002). Priming Visual Face-Processing Mechanisms: Electrophysiological Evidence. *Psychological Science*, *13*(2), 190–193. https://doi.org/10.1111/1467-9280.00435

Bieniek, M. M., Frei, L. S., & Rousselet, G. A. (2013). Early ERPs to faces: Aging, luminance, and individual differences. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00268

Bindemann, M., & Burton, A. M. (2009). The role of color in human face detection. *Cognitive Science*, *33*(6), 1144–1156. https://doi.org/10.1111/j.1551-6709.2009.01035.x

Bötzel, K., Schulze, S., & Stodieck, S. R. G. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental Brain Research*, *104*(1), 135–143. https://doi.org/10.1007/BF00229863

Boutet, I., Taler, V., & Collin, C. A. (2015). On the particular vulnerability of face recognition to aging: a review of three hypotheses. *Frontiers in Psychology*, *6*, 1139. https://doi.org/10.3389/fpsyg.2015.01139

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, *82*(4), 239–259. https://doi.org/10.1007/BF00308809

Brenkel, M., Shulman, K., Hazan, E., Herrmann, N., & Owen, A. M. (2017). Assessing Capacity in the Elderly: Comparing the MoCA with a Novel Computerized Battery of Executive Function. *Dementia and Geriatric Cognitive Disorders Extra*, *7*(2), 249–256. https://doi.org/10.1159/000478008

Burianová, H., Lee, Y., Grady, C. L., & Moscovitch, M. (2013). Age-related

dedifferentiation and compensatory changes in the functional network underlying face processing. *Neurobiology of Aging*, *34*(12), 2759–2767. https://doi.org/10.1016/J.NEUROBIOLAGING.2013.06.016

Burra, N., Baker, S., & George, N. (2017). Processing of gaze direction within the N170/M170 time window: A combined EEG/MEG study. *Neuropsychologia*, *100*, 207–219. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2017.04.028

Cabeza, R. (2002). Hemispheric Asymmetry Reduction in Older Adults: The HAROLD Model. *Psychology and Aging*, *17*(1), 85–100. https://doi.org/10.1037//0882-7974.17.1.85

Cabeza, R., Anderson, N. D., Locantore, J. K., & McIntosh, A. R. (2002). Aging gracefully: Compensatory brain activity in high-performing older adults. *NeuroImage*, *17*(3), 1394–1402. https://doi.org/10.1006/nimg.2002.1280

Carbon, C. C., Grüter, M., & Grüter, T. (2013). Age-Dependent Face Detection and Face Categorization Performance. *PLoS ONE*, *8*(10), e79164. https://doi.org/10.1371/journal.pone.0079164

Carmel, D., & Bentin, S. (2002). Domain specificity versus expertise: Factors influencing distinct processing of faces. *Cognition*, *83*(1), 1–29. https://doi.org/10.1016/S0010-0277(01)00162-7

Chen, J., Liu, B., Chen, B., & Fang, F. (2009). Time course of amodal completion in face perception. *Vision Research*, *49*(7), 752–758. https://doi.org/10.1016/J.VISRES.2009.02.005

Cliff, N. (1996). *Ordinal Methods For Behavioral Data Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates. Colenbrander,. https://doi.org/10.4324/9781315806730

Colenbrander, A., & Fletcher, D. C. (2004). Evaluation of a new Mixed Contrast Reading Card. *ARVO Meeting Abstracts*, *45*(5), 4352. Retrieved from http://iovs.arvojournals.org/article.aspx?articleid=2409886

Dailey, M. N., Joyce, C., Lyons, M. J., Kamachi, M., Ishi, H., Gyoba, J., & Cottrell, G. W. (2010). Evidence and a Computational Explanation of Cultural Differences in Facial Expression Recognition. *Emotion*, *10*(6), 874–893. https://doi.org/10.1037/a0020019

Dalrymple, K. A., Fletcher, K., Corrow, S., das Nair, R., Barton, J. J. S., Yonas, A., & Duchaine, B. (2014). "A room full of strangers every day": The psychosocial impact of developmental prosopagnosia on children and their families. *Journal of Psychosomatic Research*, *77*(2), 144–150. https://doi.org/10.1016/J.JPSYCHORES.2014.06.001

Dalrymple, K. A., Oruç, I., Duchaine, B., Pancaroglu, R., Fox, C. J., Iaria, G., … Barton, J. J. S. (2011). The anatomic basis of the right face-selective N170 IN acquired prosopagnosia: A combined ERP/fMRI study. *Neuropsychologia*, *49*(9), 2553–2563. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2011.05.003

Daniel, S., & Bentin, S. (2012). Age-related changes in processing faces from detection to identification: ERP evidence. *Neurobiology of Aging*, *33*(1), 206.e1–206.e28. https://doi.org/10.1016/J.NEUROBIOLAGING.2010.09.001

Davis, S. W., Dennis, N. A., Daselaar, S. M., Fleck, M. S., & Cabeza, R. (2008). Qué PASA? the posterior-anterior shift in aging. *Cerebral Cortex*, *18*(5), 1201–1209. https://doi.org/10.1093/cercor/bhm155

DeBruine, L. M. (2018). debruine/webmorph: Beta release 2 (Version v0.0.0.9001). Zenodo. https://doi.org/10.5281/zenodo.1162670

Deffke, I., Sander, T., Heidenreich, J., Sommer, W., Curio, G., Trahms, L., & Lueschow, A. (2007). MEG/EEG sources of the 170-ms response to faces are co-localized in the fusiform gyrus. *NeuroImage*, *35*(4), 1495–1501. https://doi.org/10.1016/J.NEUROIMAGE.2007.01.034

Dering, B., Martin, C. D., Moro, S., Pegna, A. J., & Thierry, G. (2011). Face-sensitive processes one hundred milliseconds after picture onset. *Frontiers in Human Neuroscience*, *5*(93). https://doi.org/10.3389/fnhum.2011.00093

Earp, B. D., & Everett, J. A. C. (2013). Is the N170 face specific? Controversy, context, and theory. *Neuropsychological Trends*, *13*(1), 7–26. https://doi.org/10.7358/neur-2013-013-earp

Ebaid, D., Crewther, S. G., MacCalman, K., Brown, A., & Crewther, D. P. (2017). Cognitive processing speed across the lifespan: Beyond the influence of motor speed. *Frontiers in Aging Neuroscience*, *9*(62). https://doi.org/10.3389/fnagi.2017.00062

Ebner, N. C., He, Y., Fichtenholtz, H. M., McCarthy, G., & Johnson, M. K. (2011). Electrophysiological correlates of processing faces of younger and older individuals. *Social Cognitive and Affective Neuroscience*, *6*(4), 526–535. https://doi.org/10.1093/scan/nsq074

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES-a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*(1), 351–362. https://doi.org/10.3758/BRM.42.1.351

Eimer, M. (1998). Does the face-specific N170 component reflect the activity of a specialized eye processor? *NeuroReport*, *9*(13), 2945–2948. https://doi.org/10.1097/00001756-199809140-00005

Eimer, M. (2000). The face-specific N170 component reflects late stages in the structural encoding of faces. *NeuroReport*, *11*(10), 2319–2324. https://doi.org/10.1097/00001756-200007140-00050

Ekman, P. Friesen, W. V. (1978). Facial action coding system: A technique for the measurement of facial action. *Manual for the Facial Action Coding System*.

éthier-Majcher, C., Joubert, S., & Gosselin, F. (2013). Reverse correlating trustworthy faces in young and older adults. *Frontiers in Psychology*, *4*, 592. https://doi.org/10.3389/fpsyg.2013.00592

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A

practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Fölster, M., Hess, U., & Werheid, K. (2014). Facial age affects emotional expression decoding. *Frontiers in Psychology*, *5*, 30. https://doi.org/10.3389/fpsyg.2014.00030

Friedman, J. H. (1991). Multivariate Adaptive Regression Splines. *The Annals of Statistics*, *19*(1), 1–67. https://doi.org/10.1214/aos/1176347963

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, *12*(3), 495–504. https://doi.org/10.1162/089892900562165

Gazzaley, A., Clapp, W., Kelley, J., McEvoy, K., Knight, R. T., & D'Esposito, M. (2008). Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(35), 13122–13126. https://doi.org/10.1073/pnas.0806074105

Gill, D., DeBruine, L., Jones, B., & Schyns, P. (2015). Bubble-Warp: a New Approach to the Depiction of High-Level Mental Representation. *Journal of Vision*, *15*(12), 420. https://doi.org/10.1167/15.12.420

Giorgio, A., Santelli, L., Tomassini, V., Bosnell, R., Smith, S., De Stefano, N., & Johansen-Berg, H. (2010). Age-related changes in grey and white matter structure throughout adulthood. *NeuroImage*, *51*(3), 943–51. https://doi.org/10.1016/j.neuroimage.2010.03.004

Gittings, N. S., & Fozard, J. L. (1986). Age related changes in visual acuity. *Experimental Gerontology*, *21*(4–5), 423–433. https://doi.org/10.1016/0531-5565(86)90047-1

Gold, J., Bennett, P. J., & Sekuler, A. B. (1999). Identification of band-pass

filtered letters and faces by human and ideal observers. *Vision Research*, *39*(21), 3537–3560. https://doi.org/10.1016/S0042-6989(99)00080-2

Golomb, B. A., Chan, V. T., Evans, M. A., Koperski, S., White, H. L., & Criqui, M. H. (2012). The older the better: Are elderly study participants more non-representative? A cross-sectional analysis of clinical trial and observational study samples. *BMJ Open*, *2*(6), e000833. https://doi.org/10.1136/bmjopen-2012-000833

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A techniques 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

Grady, C. (2012). The cognitive neuroscience of ageing. *Nature Reviews Neuroscience*, *13*(7), 491–505. https://doi.org/10.1038/nrn3256

Grady, C. L., Maisog, J., Horwitz, B., Ungerleider, L., Mentis, M., Salerno, J., … Haxby, J. (1994). Age-related Changes in Cortical Blood Flow Activation during Visual Processing of Faces and Location. *J Neurosci*, *14*(3), 1450–1462. https://doi.org/doi.org/10.1523/JNEUROSCI.14-03-01450.1994

Guo, K., Meints, K., Hall, C., Hall, S., & Mills, D. (2009). Left gaze bias in humans, rhesus monkeys and domestic dogs. *Animal Cognition*, *12*(3), 409–418. https://doi.org/10.1007/s10071-008-0199-3

Guo, K., Smith, C., Powell, K., & Nicholls, K. (2012). Consistent left gaze bias in processing different facial cues. *Psychological Research*, *76*(3), 263–269. https://doi.org/10.1007/s00426-011-0340-9

Guo, K., Tunnicliffe, D., & Roebuck, H. (2010). Human spontaneous gaze patterns in viewing of faces of different species. *Perception*, *39*(4), 533–542. https://doi.org/10.1068/p6517

Haig, N. D. (1985). How faces differ - A new comparative technique. *Perception*, *14*(5), 601–615. https://doi.org/10.1068/p140601

Harrell, F., & Davis, C. (1982). A new distribution-free quantile estimtor. *Biometrika*, *69*(3), 635–640. https://doi.org/10.1093/BIOMET/69.3.635

Hasselmo, M. E., & Sarter, M. (2011). Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology*, *36*(1), 52–73. https://doi.org/10.1038/npp.2010.104

Hawkley, L. C., & Cacioppo, J. T. (2007). Aging and loneliness: Downhill quickly? *Current Directions in Psychological Science*, *16*(4), 187–191. https://doi.org/10.1111/j.1467-8721.2007.00501.x

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. https://doi.org/10.1016/S1364-6613(00)01482-0

Haymes, S. A., Roberts, K. F., Cruess, A. F., Nicolela, M. T., LeBlanc, R. P., Ramsey, M. S., … Artes, P. H. (2006). The Letter Contrast Sensitivity Test: Clinical Evaluation of a New Design. *Investigative Opthalmology & Visual Science*, *47*(6), 2739. https://doi.org/10.1167/iovs.05-1419

Heath, R. L., Rouhana, A., & Ghanem, D. A. (2005). Asymmetric bias in perception of facial affect among Roman and Arabic script readers. *Laterality*, *10*(1), 51–64. https://doi.org/10.1080/13576500342000293

Heisz, J. J., Watter, S., & Shedden, J. M. (2006a). Automatic face identity encoding at the N170. *Vision Research*, *46*(28), 4604–4614. https://doi.org/10.1016/J.VISRES.2006.09.026

Heisz, J. J., Watter, S., & Shedden, J. M. (2006b). Progressive N170 habituation to unattended repeated faces. *Vision Research*, *46*(1–2), 47–56. https://doi.org/doi.org/10.1016/j.visres.2005.09.028

Henke, K., Schweinberger, S. R., Grigo, A., Klos, T., & Sommer, W. (1998). Specificity of Face Recognition: Recognition of Exemplars of Non-Face Objects In Prosopagnosia. *Cortex*, *34*(2), 289–296. https://doi.org/10.1016/S0010-9452(08)70756-1

Hinojosa, J. A., Mercado, F., & Carretié, L. (2015). N170 sensitivity to facial expression: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *55*, 498–509. https://doi.org/10.1016/j.neubiorev.2015.06.002

Hui-Wen Hsiao, J., & Cottrell, G. (2008). Two Fixations Suffice in Face Recognition. *Psychological Science*, *19*(10), 998–1006. https://doi.org/10.1111/j.1467-9280.2008.02191.x.

Ince, R. A. A., Giordano, B. L., Kayser, C., Rousselet, G. A., Gross, J., & Schyns, P. G. (2016). A statistical framework for neuroimaging data analysis based on mutual information estimated via a Gaussian copula. *bioRxiv*, *38*(3), 1–53. https://doi.org/10.1101/043745

Ince, R. A. A., Giordano, B. L., Kayser, C., Rousselet, G. A., Gross, J., & Schyns, P. G. (2017). A statistical framework for neuroimaging data analysis based on mutual information estimated via a gaussian copula. *Human Brain Mapping*, *38*(3), 1541–1573. https://doi.org/10.1002/hbm.23471

Ince, R. A. A., Jaworska, K., Gross, J., Panzeri, S., Van Rijsbergen, N. J., Rousselet, G. A., & Schyns, P. G. (2016). The Deceptively Simple N170 Reflects Network Information Processing Mechanisms Involving Visual Feature Coding and Transfer Across Hemispheres. *Cerebral Cortex*, *26*(11), 4123–4135. https://doi.org/10.1093/cercor/bhw196

Ince, R. A. A., Petersen, R. S., Swan, D. C., & Panzeri, S. (2009). Python for information theoretic analysis of neural data. *Frontiers in Neuroinformatics*, *3*, 4. https://doi.org/10.3389/neuro.11.004.2009

Issa, E. B., & Dicarlo, J. J. (2012). Precedence of the eye region in neural processing of faces. *Journal of Neuroscience*, *32*(47), 16666–16682. https://doi.org/10.1523/JNEUROSCI.2391-12.2012

Itier, R. J., & Taylor, M. J. (2002). Inversion and contrast polarity reversal affect both encoding and recognition processes of unfamiliar faces: A repetition study using ERPs. *NeuroImage*, *15*(2), 353–372. https://doi.org/10.1006/nimg.2001.0982

Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal Differences between Object and Face Processing Using ERPs. *Cerebral Cortex*, *14*(2), 132–142. https://doi.org/10.1093/cercor/bhg111

Itier, R. J., & Taylor, M. J. (2004). Source analysis of the N170 to faces and objects. *NeuroReport*, *15*(8), 1261–1265. https://doi.org/10.1097/01.wnr.0000127827.73576.d8

Jackson, A. F., & Bolger, D. J. (2014). The neurophysiological bases of EEG and EEG measurement: A review for the rest of us. *Psychophysiology*, *51*(11), 1061–1071. https://doi.org/10.1111/psyp.12283

Jaworska, K. (2017). *Understanding age-related differences in the speed of information processing of complex object categories measured with electroencephalography (EEG)*. The University of Glasgow. Retrieved from http://theses.gla.ac.uk/8112/7/2017JaworskaPhD.pdf

Jeffreys, D. A. (1989). A face-responsive potential recorded from the human scalp. *Experimental Brain Research*, *78*(1), 193–202. https://doi.org/10.1007/BF00230699

Jeffreys, D. a., Tukmachi, E. S. a, & Rockley, G. (1992). Evoked potential evidence for human brain mechanisms that respond to single, fixated faces. *Experimental Brain Research*, *91*(2), 351–362. https://doi.org/10.1007/BF00231669

Jekabsons, G. (2016). ARESLab Adaptive Regression Splines toolbox for Matlab/Octave. Retrieved August 9, 2018, from http://www.cs.rtu.lv/jekabsons/

Jeong, S. K., & Xu, Y. (2016). Behaviorally Relevant Abstract Object Identity Representation in the Human Parietal Cortex. *The Journal of Neuroscience*, *36*(5), 1607–1619. https://doi.org/10.1523/JNEUROSCI.1016-15.2016

Johnston, R. A., & Edmonds, A. J. (2009). Familiar and unfamiliar face recognition: A review. *Memory*, *17*(5), 577–596.

https://doi.org/10.1080/09658210902976969

Joyce, C. A., Schyns, P. G., Gosselin, F., Cottrell, G. W., & Rossion, B. (2006). Early selection of diagnostic facial information in the human visual cortex. *Vision Research*, *46*(6–7), 800–813. https://doi.org/10.1016/j.visres.2005.09.016

Joyce, C., & Rossion, B. (2005). The face-sensitive N170 and VPP components manifest the same brain processes: The effect of reference electrode site. *Clinical Neurophysiology*, *116*(11), 2613–2631. https://doi.org/10.1016/J.CLINPH.2005.07.005

Kalkstein, J., Checksfield, K., Bollinger, J., & Gazzaley, A. (2011). Diminished Top-Down Control Underlies a Visual Imagery Deficit in Normal Aging. *Journal of Neuroscience*, *31*(44), 15768–15774. https://doi.org/10.1523/JNEUROSCI.3209-11.2011

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311. https://doi.org/10.1098/Rstb.2006.1934

Kato, Y., Muramatsu, T., Kato, M., Shintani, M., Yoshino, F., Shimono, M., & Ishikawa, T. (2004). An earlier component of face perception detected by seeing-as-face task. *Neuroreport*, *15*(2), 225–229.

Kerchner, G. A., Racine, C. A., Hale, S., Wilheim, R., Laluz, V., Miller, B. L., & Kramer, J. H. (2012). Cognitive Processing Speed in Older Adults: Relationship with White Matter Integrity. *PLoS ONE*, *7*(11), e50425. https://doi.org/10.1371/journal.pone.0050425

Kirschstein, T., & Köhling, R. (2009). What is the Source of the EEG? *Clinical EEG and Neuroscience*, *40*(3), 146–149. https://doi.org/10.1177/155005940904000305

Komes, J., Schweinberger, S. R., & Wiese, H. (2015). Neural correlates of

cognitive aging during the perception of facial age: the role of relatively distant and local texture information. *Frontiers in Psychology*, *6*, 1420. https://doi.org/10.3389/fpsyg.2015.01420

Konar, Y., Bennett, P. J., & Sekuler, A. B. (2013). Effects of aging on face identification and holistic face processing. *Vision Research*, *88*, 38–46. https://doi.org/10.1016/j.visres.2013.06.003

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

Levine, S. C., Banich, M. T., & Koch-Weser, M. (1984). Variations in patterns of lateral asymmetry among dextrals. *Brain and Cognition*, *3*(3), 317–334. https://doi.org/10.1016/0278-2626(84)90024-1

Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. London: MIT Press.

Luke, C. J., & Pollux, P. M. J. (2016). Lateral presentation of faces alters overall viewing strategy. *PeerJ*, *4*, e2241. https://doi.org/10.7717/peerj.2241

Luo, W., Feng, W., He, W., Wang, N.-Y., & Luo, Y.-J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, *49*(2), 1857–1867. https://doi.org/10.1016/J.NEUROIMAGE.2009.09.018

Malcolm, G. L., Lanyon, L. J., Fugard, A. J. B., & Barton, J. J. S. (2008). Scan patterns during the processing of facial expression versus identity: An exploration of task-driven and stimulus-driven effects. *Journal of Vision*, *8*(8). https://doi.org/10.1167/8.8.2

Mäntyjärvi, M., & Laitinen, T. (2001). Normal values for the Pelli-Robson constrast sensitivity test. *J Cataract Refract Surg*, *27*(2), 261–266. https://doi.org/10.1016/S0886-3350(00)00562-9

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2014). Holistic processing and reliance on global viewing strategies in older adults' face perception. *Acta Psychologica*, *151*, 155–163. https://doi.org/10.1016/J.ACTPSY.2014.06.001

Mercure, E., Dick, F., Halit, H., Kaufman, J., & Johnson, M. H. (2008). Differential lateralization for words and faces: Category or psychophysics? *Journal of Cognitive Neuroscience*, *20*(11), 2070–2087. https://doi.org/10.1162/jocn.2008.20137

Mercure, E., Kadosh, K. C., & Johnson, M. H. (2011). The N170 Shows Differential Repetition Effects for Faces, Objects, and Orthographic Stimuli. *Frontiers in Human Neuroscience*, *5*, 6. https://doi.org/10.3389/fnhum.2011.00006

Nakajima, K., Minami, T., & Nakauchi, S. (2017). Interaction between facial expression and color. *Scientific Reports*, *7*, 41019. https://doi.org/10.1038/srep41019

Nakamura, A., Yamada, T., Abe, Y., Nakamura, K., Sato, N., Horibe, K., … Ito, K. (2001). Age-related changes in brain neuromagnetic responses to face perception in humans. *Neuroscience Letters*, *312*(1), 13–16. https://doi.org/10.1016/S0304-3940(01)02168-1

Nasreddine, Z. S., Phillips, N. A., BÃ©dirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nemrodov, D., Anderson, T., Preston, F. F., & Itier, R. J. (2014). Early sensitivity for eyes within faces: A new neuronal account of holistic and featural processing. *NeuroImage*, *97*, 81–94. https://doi.org/10.1016/j.neuroimage.2014.04.042

Neumann, M., End, A., Luttmann, S., Schweinberger, S. R., & Wiese, H. (2015). The own-age bias in face memory is unrelated to differences in attention—

Evidence from event-related potentials. *Cognitive, Affective, and Behavioural Neuroscience*, *15*, 180–194. https://doi.org/10.3758/s13415-014-0306-7

Nguyen, V. T., & Cunnington, R. (2014). The superior temporal sulcus and the N170 during face processing: Single trial analysis of concurrent EEG–fMRI. *NeuroImage*, *86*, 492–502. https://doi.org/10.1016/j.neuroimage.2013.10.047

Nicholls, M. E. R., Ellis, B. E., Clement, J. G., & Yoshino, M. (2004). Detecting hemifacial asymmetries in emotional expression with three-dimensional computerized image analysis. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1540), 663–668. https://doi.org/10.1098/rspb.2003.2660

Nyberg, L., Salami, A., Andersson, M., Eriksson, J., Kalpouzos, G., Kauppi, K., … Nilsson, L.-G. (2010). Longitudinal evidence for diminished frontal cortex function in aging. *Proceedings of the National Academy of Sciences*, *107*(52), 22682–22686. https://doi.org/10.1073/pnas.1012651108

Okubo, M., Ishikawa, K., & Kobayashi, A. (2018). The cheek of a cheater: Effects of posing the left and right hemiface on the perception of trustworthiness. *Laterality: Asymmetries of Body, Brain and Cognition*, *23*(2), 209–227. https://doi.org/10.1080/1357650X.2017.1351449

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

Owsley, C. (2011). Aging and vision. *Vision Research*, *51*(13), 1610–1622. https://doi.org/10.1016/j.visres.2010.10.020

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research*, *23*(7), 689–699. https://doi.org/10.1016/0042-6989(83)90210-9

Pelli, D. G., & Robson, J. G. (1988). The design of a new letter chart for

measuring contrast sensitivity. *Clinical Vision Science*, *2*(3), 187–199. https://doi.org/10.1016/j.parkreldis.2012.11.013

Pernet, C. R., Chauveau, N., Gaspar, C., & Rousselet, G. A. (2011). LIMO EEG: A toolbox for hierarchical linear modeling of electroencephalographic data. *Computational Intelligence and Neuroscience*, 831409. https://doi.org/10.1155/2011/831409

Pernet, C. R., Sajda, P., & Rousselet, G. A. (2011). Single-Trial Analyses: Why Bother? *Frontiers in Psychology*, *2*, 322. https://doi.org/10.3389/fpsyg.2011.00322

Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences*, *109*(48), 3314–3323. https://doi.org/10.1073/pnas.1214269109

Peterson, M. F., & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science*, *24*(7), 1216–1225. https://doi.org/10.1177/0956797612471684

Piepers, D. W., & Robbins, R. A. (2012). A Review and Clarification of the Terms "holistic," "configural," and "relational" in the Face Perception Literature. *Frontiers in Psychology*, *3*, 559. https://doi.org/10.3389/fpsyg.2012.00559

Piguet, O., Double, K. L., Kril, J. J., Harasty, J., Macdonald, V., McRitchie, D. A., & Halliday, G. M. (2009). White matter loss in healthy ageing: a postmortem analysis. *Neurobiology of Aging*, *30*(8), 1288–95. https://doi.org/10.1016/j.neurobiolaging.2007.10.015

Pitchaimuthu, K., Wu, Q., Carter, O., Nguyen, B. N., Ahn, S., Egan, G. F., & McKendrick, A. M. (2017). Occipital GABA levels in older adults and their relationship to visual perceptual suppression. *Scientific Reports*, *7*(1), 14231. https://doi.org/10.1038/s41598-017-14577-5

Pitcher, D., Walsh, V., & Duchaine, B. (2011). The role of the occipital face area

in the cortical face perception network. *Experimental Brain Research*. https://doi.org/10.1007/s00221-011-2579-1

Proverbio, A. M., & Galli, J. (2016). Women are better at seeing faces where there are none: an ERP study of face pareidolia. *Social Cognitive and Affective Neuroscience*, *11*(9), 1501–1512. https://doi.org/10.1093/scan/nsw064

Racca, A., Guo, K., Meints, K., & Mills, D. S. (2012). Reading faces: Differential lateral gaze bias in processing canine and human facial expressions in dogs and 4-year-old children. *PLoS ONE*, *7*(4), e36076. https://doi.org/10.1371/journal.pone.0036076

Reuter-Lorenz, P. A., & Campbell, K. A. (2008). Neurocognitive Ageing and The Compensation Hypothesis. *Current Directions in Psychological Science*, *17*(3), 177–182. https://doi.org/10.1111/j.1467-8721.2008.00570.x

Rieth, C. A., Lee, K., Lui, J., Tian, J., & Huber, D. E. (2011). Faces in the mist: illusory face and letter detection. *I-Perception*, *2*(5), 458–76. https://doi.org/10.1068/i0421

Rossion, B., & Jacques, C. (2008, February 15). Does physical interstimulus variance account for early electrophysiological face sensitive responses in the human brain? Ten lessons on the N170. *NeuroImage*. Academic Press. https://doi.org/10.1016/j.neuroimage.2007.10.011

Rossion, B., & Jacques, C. (2011). *The N170: Understanding the Time Course of Face Perception in the Human Brain*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0064

Rousselet, G. A., Gaspar, C. M., Pernet, C. R., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2010). Healthy aging delays scalp EEG sensitivity to noise in a face discrimination task. *Frontiers in Psychology*, *1*, 1–14. https://doi.org/10.3389/fpsyg.2010.00019

Rousselet, G. A., Gaspar, C. M., Wieczorek, K. P., & Pernet, C. R. (2011).

Modeling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects). *Frontiers in Psychology*, *2*, 137. https://doi.org/10.3389/fpsyg.2011.00137

Rousselet, G. A., Gilman, H. L., Ince, R. A. A., & Schyns, P. G. (2015). The N170 is mostly sensitive to pixels in the contralateral eye area. In *Vision Sciences Society Annual Meeting Abstract*. https://doi.org/doi:10.1167/15.12.687

Rousselet, G. A., Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2008). Time course and robustness of ERP object and face differences. *Journal of Vision*, *8*(12), 3–3. https://doi.org/10.1167/8.12.3

Rousselet, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., & Sekuler, A. B. (2009). Age-related delay in information accrual for faces: Evidence from a parametric, single-trial EEG approach. *BMC Neuroscience*, *10*, 114. https://doi.org/10.1186/1471-2202-10-114

Rousselet, G. A., Ince, R. A. A., van Rijsbergen, N. J., & Schyns, P. G. (2014). Eye coding mechanisms in early human face event-related potentials. *Journal of Vision*, *14*(13), 1–24. https://doi.org/10.1167/14.13.7.doi

Rousselet, G. A., Ince, R. A. A., van Rijsbergen, N. J., & Schyns, P. G. (2014). Eye coding mechanisms in early human face event-related potentials. *Journal of Vision*, *14*(13), 7–7. https://doi.org/10.1167/14.13.7

Rousselet, G. A., Mace, M. J.-M., & Fabre-Thorpe, M. (2004). Animal and human faces in natural scenes: How specific to human faces is the N170 ERP component? *Journal of Vision*, *4*(1). https://doi.org/10.1167/4.1.2

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, *32*(4), 863–881. https://doi.org/10.1016/J.NEUBIOREV.2008.01.001

Sagiv, N., & Bentin, S. (2001). Structural Encoding of Human and Schematic Faces: Holistic and Part-Based Processes. *Journal of Cognitive Neuroscience*,

13(7), 937–951. https://doi.org/10.1162/089892901753165854

Salthouse, T. A. (1996). The Processing-Speed Theory of Adult Age Differences in Cognition. *Psychological Review*, *103*(3), 403–428. https://doi.org/10.1037/0033-295X.103.3.403

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, *16*, 754–760. https://doi.org/10.1017/S1355617710000706

Samal, A., Subramani, V., & Marx, D. (2007). Analysis of sexual dimorphism in human face. *Journal of Visual Communication and Image Representation*, *18*(6), 453–463. https://doi.org/10.1016/j.jvcir.2007.04.010

Schultz, S. R., Ince, R. A. A., & Panzeri, S. (2015). Applications of Information Theory to Analysis of Neural Data. *Encyclopedia of Computational Neuroscience*, 199–203. https://doi.org/10.1007/978-1-4614-7320-6_280-1

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). SHOW ME THE FEATURES! Understanding Recognition From the Use of Visual Information. *Psychological Science*, *13*(5), 402–409. https://doi.org/10.1111/1467-9280.00472

Schyns, P. G., Jentzsch, I., Johnson, M., Schweinberger, S. R., & Gosselin, F. (2003). A principled method for determining the functionality of brain responses. *Neuroreport*, *14*(13), 1665–9. https://doi.org/10.1097/01.wnr.0000088408.04452.e9

Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: when categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*(3), 243–265. https://doi.org/10.1016/S0010-0277(98)00069-9

Schyns, P. G., Petro, L. S., & Smith, M. L. (2007). Dynamics of Visual Information Integration in the Brain for Categorizing Facial Expressions. *Current Biology*, *17*(18), 1580–1585. https://doi.org/10.1016/J.CUB.2007.08.048

Schyns, P. G., Thut, G., & Gross, J. (2011). Cracking the code of oscillatory

activity. *PLoS Biology*, *9*(5), e1001064. https://doi.org/10.1371/journal.pbio.1001064

Shafto, M. A., Tyler, L. K., Dixon, M., Taylor, J. R., Rowe, J. B., Cusack, R., … Cam-CAN, F. E. (2014). The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC Neurology*, *14*, 204. https://doi.org/10.1186/s12883-014-0204-1

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., Gosselin, F., & Schyns, P. G. (2004). Receptive fields for flexible face categorizations. *Psychological Science*, *15*(11), 753–761. https://doi.org/10.1111/j.0956-7976.2004.00752.x

Smith, M. L., Gosselin, F., & Schyns, P. G. (2007). From a face to its category via a few information processing states in the brain. *NeuroImage*, *37*(3), 974–984. https://doi.org/10.1016/j.neuroimage.2007.05.030

Smith, T., Gildeh, N., & Holmes, C. (2007). The Montreal Cognitive Assessment: Validity and Utility in a Memory Clinic Setting. *The Canadian Journal of Psychiatry*, *52*(5), 329–332. https://doi.org/10.1177/070674370705200508

Smith, Gosselin, F., & Schyns, P. G. (2012). Measuring internal representations from behavioral and brain data. *Current Biology*, *22*(3), 191–196. https://doi.org/10.1016/j.cub.2011.11.061

Sullivan, S., Campbell, A., Hutton, S. B., & Ruffman, T. (2017). What's good for the goose is not good for the gander: Age and gender differences in scanning emotion faces. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *72*(3), 441–447. https://doi.org/10.1093/geronb/gbv033

Sullivan, S., & Ruffman, T. (2004). Emotion recognition deficits in the elderly. *International Journal of Neuroscience*, *114*(3), 403–432.

https://doi.org/10.1080/00207450490270901

Sur, S., & Sinha, V. K. (2009). Event-related potential: An overview. *Industrial Psychiatry Journal*, *18*(1), 70–3. https://doi.org/10.4103/0972-6748.57865

Tanaka, H. (2016). Facial cosmetics exert a greater influence on processing of the mouth relative to the eyes: Evidence from the N170 event-related potential component. *Frontiers in Psychology*, *7*, 1359. https://doi.org/10.3389/fpsyg.2016.01359

Taylor, M. J., Itier, R. J., Allison, T., & Edmonds, G. E. (2001). Direction of gaze effects on early face processing: eyes-only versus full faces. *Cognitive Brain Research*, *10*(3), 333–340. https://doi.org/10.1016/S0926-6410(00)00051-3

Thierry, G., Martin, C. D., Downing, P., & Pegna, A. J. (2007). Controlling for interstimulus perceptual variance abolishes N170 face selectivity. *Nature Neuroscience*, *10*(4), 505–511. https://doi.org/10.1038/nn1864

Thomas, C., & Joy, T. (2006). *Elements of Information Theory* (2nd ed). Hoboken, NJ: Wiley-Interscience.

Van Rijsbergen, N. J., & Schyns, P. G. (2009). Dynamics of trimming the content of face representations for categorization in the brain. *PLoS Computational Biology*, *5*(11). https://doi.org/10.1371/journal.pcbi.1000561

Ventura, P. (2014). Let's face it: reading acquisition, face and word processing. *Frontiers in Psychology*, *5*, 787. https://doi.org/10.3389/fpsyg.2014.00787

Wang, S. (2018). Face size biases emotion judgment through eye movement. *Scientific Reports*, *8*(1), 317. https://doi.org/10.1038/s41598-017-18741-9

Wiese, H., Schweinberger, S. R., & Hansen, K. (2008). The age of the beholder: ERP evidence of an own-age bias in face memory. *Neuropsychologia*, *46*(12), 2973–2985. https://doi.org/10.1016/J.NEUROPSYCHOLOGIA.2008.06.007

Wilcox, R. (2013). *Introduction to Robust Estimation and Hypothesis Testing* (3rd ed.). Academic Press.

Yardley, L., McDermott, L., Pisarski, S., Duchaine, B., & Nakayama, K. (2008). Psychosocial consequences of developmental prosopagnosia: A problem of recognition. *Journal of Psychosomatic Research*, *65*(5), 445–451. https://doi.org/10.1016/j.jpsychores.2008.03.013

Yi, F. (2018). *Robust Eye Coding Mechanisms in Humans during Face Detection*. The University of Glasgow.

Yip, A. W., & Sinha, P. (2002). Contribution of color to face recognition. *Perception*, *31*(8), 995–1003. https://doi.org/10.1068/p3376

Zanto, T. P., Hennigan, K., Östberg, M., Clapp, W. C., & Gazzaley, A. (2010). Predictive knowledge of stimulus relevance does not influence top-down suppression of irrelevant information in older adults. *Cortex*, *46*(4), 564–574. https://doi.org/10.1016/j.cortex.2009.08.003

Zhang, J. (2018). The Human Anger Face Likely Carries a Dual-Signaling Function. *Frontiers in Behavioral Neuroscience*, *12*, 26. https://doi.org/10.3389/fnbeh.2018.00026