Information Processing in the Social Brain

Isobel Louise Ward

A thesis submitted to Cardiff University

for the degree of Doctor of Philosophy

March 2021

Summary of Thesis

This thesis explores how facial expressions are perceived in the context of an expressive body posture. Previous work has demonstrated that facial expressions can be biased by the emotion portrayed by an affective body posture. It remains unclear how these cues are combined in the human brain to form a whole-person percept of emotion. Additionally, the role of individual differences, and the developmental trajectories of facial expression and body posture integration remain unknown.

In **Chapter 3**, I developed a novel psychophysical paradigm to quantify the influence of body posture on facial expression perception. Body context significantly biased the perception of facial expressions, but the magnitude of this bias was highly variable between individuals. This variability was highly correlated with observers' ability to perceive isolated facial expressions; better facial expression recognition resulted in less influence of body context.

In **Chapter 4**, I assessed how body posture influenced perception of facial expressions in children. I also quantified microstructure of white matter connecting cortical nodes responsible for face and body processing and related these metrics to children's perceptual ability. With increasing age, children's ability to recognise facial expressions improved, and the biasing influence of the body context decreased. Microstructural properties of functionally-defined white matter were found to predict children's perceptual abilities.

In **Chapter 5**, I applied a mathematical model to conceptualise the integration of facial expression and body posture cues under a cue integration framework. My results provide novel insights into the integration of face and body cues, such that the integration was found to be a weighted average of the reliability of observers'

facial expression and body posture representations.

In summary, the experimental work presented in this thesis has important implications for understanding real-world social perception, where multiple social signals must be integrated to create a coherent experience of the social world.

Impact of Thesis

The experimental work in this thesis has been presented at national and international conferences.

Chapter 3

Precision of isolated facial-expression and body-posture representations determines integrated whole-person perception of emotion. Poster presentation at *42nd Edition of the European Conference on Vision Perception (ECVP), Leuven, Belgium, August 2019* and *21st Conference of the European Society for Cognitive Psychology (ESCoP), Tenerife, September 2019.* Prize awarded at ECVP for poster.

Chapter 4

Face and body emotion perception across development and associated white matter microstructure. Poster presentation at *Organisation of Human Brain Mapping (OHBM), OHBM, Virtual event, July 2020.*

Functionally-defined changes in white matter microstructure: Face and Body emotion perception across development. Platform presentation at *International Society for Magnetic Resonance in Medicine (ISMRM): The British and Irish Chapter Postgraduate special, September 2020.*

Chapter 5

Perceptual processing of facial expression and body posture: A social cue-integration framework. Platform presentation at *Experimental Psychology Society (EPS) January 2021 Meeting, January 2021.*

Contents

List of Figures

List of Tables

List of Abbreviations

AIC akaike information criterion

ASD autism spectrum disorder

ATFP anterior temporal face patch

ATL anterior temporal lobe

BIC bayesian information criterion

BOLD blood oxygenation level dependant

CSF cerebrospinal fluid

CST corticospinal tract

dMRI diffusion magnetic resonance imaging

DTI diffusion tensor imaging

DWI diffusion-weighted imaging

EBA extrastriate body area

EEG electroencephalography

EPI echo-planar imaging

FA fractional anisotropy

FBA fusiform body area

FDWM functionally-defined white matter

FFA fusiform face area

fMRI functional magnetic resonance imaging

fODF fibre orientation distribution function

FTD frontotemporal dementia

GLM general linear model

GM grey matter

IFOF inferior frontal occipital fasiculus

ILF inferior longitudinal fasiculus

LEAS Levels of Emotional Awareness Scale

LGN lateral geniculate nucleus

MD mean diffusivity

MLE maximum likelihood estimation

MNI montreal neurological institute

MR magnetic resonance

MRI magnetic resonance imaging

MVPA multivoxel pattern analysis

OFA occipital face area

OT oxytocin

pDev proportion of deviance values

PDS pubertal development scale

PDSS PDS-Shirtcliff

PF psychometric function

PSE point of subjective equality

pSTS posterior superior temporal sulcus

RDM representational dissimilarity matrix

ROI region of interest

RSA representational similarity analysis

RT reaction time

SD semantic dementia

SNR signal-to-noise ratio

T tesla

TD typically developing

TMS transcranial magnetic stimulation

UF uncinate fasciculus

UHF ultra-high-field

VIF variance inflation factor

WM white matter

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr Elisabeth von dem Hagen, who has been an excellent advisor throughout the course of my PhD research. I am privileged to have had the opportunity to learn from Elisabeth and evolve as a researcher. I would also like to thank Dr Christoph Teufel for his supervision, particularly for his input and wisdom with the psychophysics and modelling aspects of my PhD. In addition to my supervisory team, I would like to give a special mention to Dr Erika Raven; her mentorship has been invaluable, and I am constantly inspired by her resilience and passion for science. I am grateful to all the volunteers who took part in my research, especially the fantastic families who dedicated their time to take part in multiple visits to CUBRIC. Of course, there are a number of other people who have academically supported me; thank you to each one of you for your insight, feedback and guidance.

I am proud to have had the opportunity to undertake a scientific PhD when many women around the world are not encouraged to pursue scientific subjects or are denied access to education. I hope through the opportunities I have been given I may have inspired just one more young girl to believe she can be a scientist too.

Statement of collaborative work

All the work in thesis is the result of my own work. The data presented in Chapter 4, however, was part of a larger neurodevelopmental study at CUBRIC. I acquired all 7T MRI scans (fMRI), all behavioural data, and the majority of 3T MRI scans (DWI) in the children; the remainder of the 3T MRI scans were acquired by collaborators on this larger study. Furthermore, the DWI data was pre-processed by members of staff on the CUBRIC Kids project team to ensure optimal processing in the cohort, and consistency across outputs and publications. The conceptual design and all other processing and analyses of the children's data presented in Chapter 4 were conducted by me for the purposes of this thesis.

Chapter 1

Introduction

Interpreting social cues is an integral component of human communication and behaviour. A social cue is defined as a signal in either the verbal or non-verbal domain that guides an individual's social behaviour (Leekam et al., 2010). In a rich and dynamic sensory world, the ability to integrate multiple social cues is imperative, and is a sophisticated skill (Zaki, 2013). Facial expressions are one of the most researched social cues; such signals produced by contraction of facial muscles, signify an emotional state, and upon recognition will guide an observer's behaviour and subsequent actions (Calder, 2011). Importantly, facial expressions are typically encountered with a body posture. A growing consensus illustrates the relevance of the body posture for conveying emotional expression, in addition to its role in action-orientated behaviours (de Gelder et al., 2015). The work in this thesis explores visual processing of high-level social cues, specifically how facial expression and body posture cues are perceived and integrated.

1.1 The organisation of the visual system

Vision is a complex and important sensory modality in humans (Hutmacher, 2019). The ability to operate in a complex sensory visual world relies on an extensive network of neural regions responsible for processing incoming visual information. Understanding the organisational principles of the visual system is fundamental for the applied research conducted as part of my PhD, exploring how facial expression and body posture cues are combined. Given my research focuses on 'higher-level' visual cues, emphasis in the following section is on how the low-level visual systems integrate with the brain networks

known to be critical in face and body processing.

Visual information from the world reaches the eye as light which is converted into electrical signals by photoreceptors (Goldstein et al., 2005). This signal is transmitted to the primary visual cortices and association cortices of the brain (Lu & Dosher, 2014). The translation of features to form an integrated coherent precept subsequently guides behaviours and actions.

One dominant model of visual processing describes the transmission of information as hierarchical (Hubel & Wiesel, 1962). Information processing starts at the retina, proceeds to the lateral geniculate nucleus (LGN) in the thalamus, before arriving at the primary visual cortices (V1-V3) (Schyns & Oliva, 1999). The information is fed-forward for further processing in association regions (V4-V5/MT) and the inferior temporal area. The types of information that are extracted at different levels of the visual processing stream are dependent on the neural architecture of the receptive field size in each region [Figure 1.1] (Herzog & Clarke, 2014). Neurons in lower visual areas (V1 - V3) are sensitive to basic visual features such as lines and edges and have small receptive fields. Early processing regions project to later stages of the processing hierarchy, which encode increasingly complex features (Goldstein et al., 2005). At the inferior temporal stage of the ventral stream, higher-level processing takes place such as whole object recognition.

An influential model for higher-level visual processing is the dual-pathway architecture (Goodale & Milner, 1992). This posits that there are two interacting streams of ventral and dorsal information processing beyond V4 linked to object recognition and spatial localisation respectively (Ungerleider, 1994). These extended areas of the visual processing system are highly specialised and hierarchically organised.

1.2 Face perception

Faces provide key information about identity, expression, gaze direction and visual cues about speech, amongst other things (Calder, 2011). Faces are a special category of visual stimuli, conveying a vast amount of information to the viewer and facilitating social communication (Haxby et al., 2000). In addition to faces being functionally special, the human brain also has specialised visual processing streams for processing faces, most likely a con-

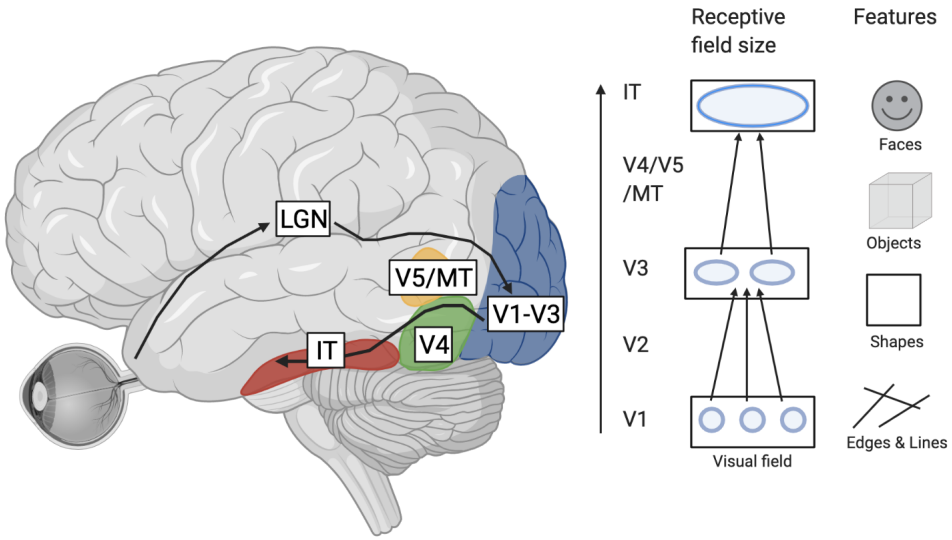


Figure 1.1: **Hierarchical model of visual processing** The schematic illustrates the hierarchical feed forward model of visual processing. Information processing proceeds from the retina to the LGN, and then V1 and later visual areas. Early visual areas have smaller receptive fields and code features such as edges and lines. Higher-level areas have gradually increasing receptive field sizes and integrate information over lager portions of the visual field. At higher-levels, such as IT, these regions receive lower level inputs and encode more complex features, such as faces. Figure recreated from Herzog and Clarke, 2014.

sequence of their social importance (Calder, 2011). Behavioural markers of face-specificity in neurologically healthy participants has been demonstrated through inverting faces (Kanwisher & Yovel, 2006). One of the most robust findings in the face processing literature is that inverted faces are not recognised as rapidly as upright faces (Kanwisher et al., 1998; Rossion & Gauthier, 2002). In contrast to other objects, the diminished recognition of faces when inverted is significantly greater compared to non-face stimuli. This is, in part, due to faces being processed configurally, unlike other object categories. Configural processing refers to processing that involves perception of the relationship between features in a stimulus (Minnebusch & Daum, 2009). Further evidence to illustrate the importance of faces is the presence of face-specific neural representations (Kanwisher & Yovel, 2006). Face perception is often referred to as 'domain-specific', as the processing modules in the brain are thought to be specialised for encoding faces (Kanwisher, 2000). Research from individuals with prosopagnosia, a disorder characterised by the inability to recognise the identity of faces, provides further evidence to support the notion for specialised visual processing regions for faces. Prosopagnosics display domain-specific face recognition deficits, whilst other categories of object recognition are preserved, providing evidence for clusters

of face-specific neurons in the brain (Duchaine et al., 2006).

1.2.1 Facial expression processing

A key feature of effective face processing is the ability to extract emotional expression. Facial expressions provide information to help guide an observer's behaviour during social interactions (Rajhans et al., 2016). Ekman and colleagues pioneered the modern era of research on facial actions in emotion (Ekman, 1965). They proposed the concept of 'Basic Emotion Theory', which stated that humans have a limited number of biologically and psychologically basic emotions which they can recognise. Within this framework, it is assumed that expressions of emotion are brief and coherent patterns of facial muscle contraction. These basic emotions fall broadly into six categories: anger, happiness, sadness, fear, disgust and surprise, and are believed to be universally recognised (Calder, 2011; Ekman, 1992). Reliable individual differences have been reported in individual's ability to recognise facial expressions across the general population (Palermo et al., 2018). The mechanisms that have been attributed to many factors such as gaze pattern, personal traits, and anxiety scores (Alharbi et al., 2020; Green & Guo, 2018).

1.2.2 Models for face recognition

The idea that facial identity and facial expression are processed by separate visual routes has dominated face research for over 20 years and is at the heart of an early prominent model of face perception. Bruce and Young (1986) were the first to document an account of face recognition by combining and extending several early models. The term 'recognition' was used very broadly to refer to any kind of stored information from faces. The framework proposed by Bruce and Young (1986) was purely a functional cognitive account and did not extend to the localisation of processing in the brain. The cognitive model posited that seven different types of visual information were extracted from faces: pictorial, structural, visually derived semantic, identity-specific semantic, name, expression and facial speech codes (Bruce & Young, 1986). Representations of the face were compared to stored face-recognition units, where the face was recognised as familiar if there was a matching of the encoded representation and the stored structural code. Since the emergence of neuroimaging has become commonplace in cognitive psychology, the early

cognitive account has been developed and built upon with subsequent research.

A face selective area of cortex is defined as a cortical region that responds more to faces than any other stimulus category. This 'response' amplitude is often measured in humans using the haemodynamic blood oxygenation level dependant (BOLD) response to spatially localise cortical regions from functional magnetic resonance imaging (fMRI). Seminal work from Kanwisher (1997) and colleagues revealed an area in the fusiform gyrus which was significantly more active when subjects viewed faces compared to objects (Kanwisher et al., 1997). Further research demonstrated that category specific regions for faces showed increased activation only when presented upright, demonstrating their 'gestalt-like' processing. Regions of cortex which are selective to faces in the ventral fusiform regions are thought to receive inputs from earlier retinotopic regions (V1-V4) through axonal projections (Grill-Spector et al., 2018).

Building upon the cognitive account from Bruce and Young (1986), Haxby and colleagues (2006) outlined a neuroanatomical framework of face perception which distinguished aspects of face processing and their neural representations (Haxby et al., 2000). This model divides neural face selective regions into a 'core' and 'extended' system [Figure 1.2]. Regions in the core network include the fusiform face area (FFA) located in the lateral fusiform gyrus, the occipital face area (OFA) and the posterior superior temporal sulcus (pSTS). The OFA and FFA were proposed to be involved with processing the invariant aspects of faces (e.g., facial identity), and the pSTS for changeable aspects (e.g., facial expression) (Haxby et al., 2000). The extended system incorporates additional neural regions, such as the anterior temporal regions, amygdala and the limbic system, to process different kinds of social information from faces. In contrast to the core system, the network of areas in the extended system has been said to be responsible for encoding person-specific semantic knowledge and assessing the value of facial expressions.

Recent work investigating the face-selective regions of the brain challenges aspects of the Haxby model (Duchaine & Yovel, 2015). In 2015, a revised model for face processing [Figure 1.3] was published which modified the framework proposed by Haxby and colleagues. Research challenged the idea that the FFA was not involved in processing emotion, with the revised model updating this view to state that the FFA contributes to the perception of changeable aspects of faces (Wegrzyn et al., 2015). Another key mod-

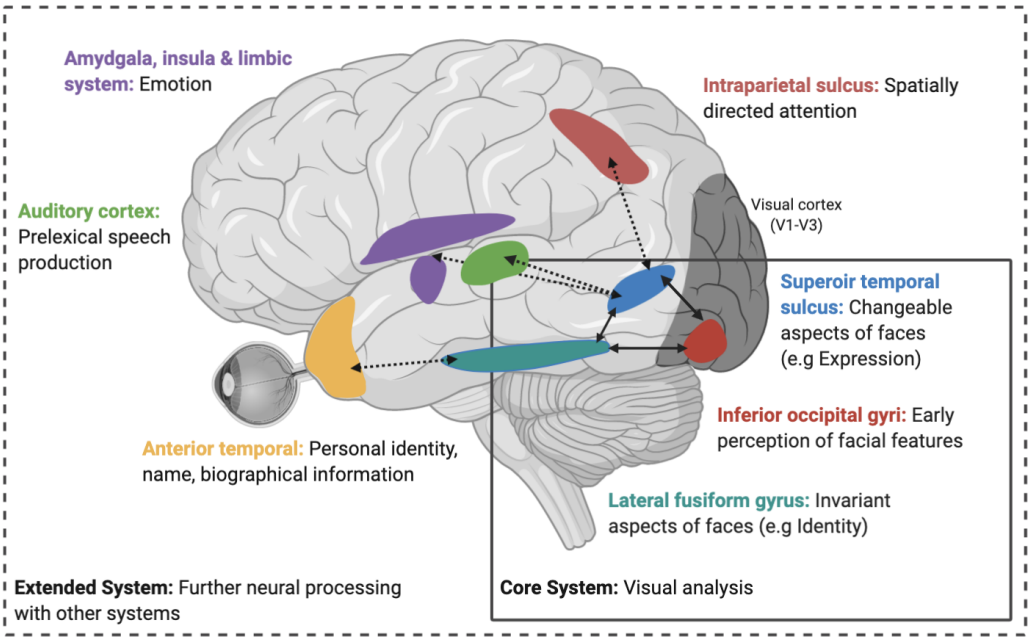


Figure 1.2: **Haxby model of face perception**
The figure highlights the core and extended systems in the Haxby model of face perception. The regions early in the visual hierarchy (V1-V3) are illustrated in darker grey. Solid arrows depict relationships between nodes in the core face processing network, and dashed lines illustrate the connections between the extended network.

ification from the Haxby model was that information from early visual areas is able to enter the face-processing system via multiple parallel routes, as opposed to just via the OFA. Notwithstanding the tentative evidence from Haxby and colleagues that the OFA was the gateway to the face-processing network, patient work revealed that despite lacking a right OFA, face-selectivity in right FFA and right pSTS was retained and comparable to controls (Sorger et al., 2007). Using diffusion tensor imaging (DTI) and reconstructing fibres connecting cortical regions has also revealed direct anatomical connections between early visual areas and the FFA, further supporting the idea that information can bypass the OFA to enter the face-processing network (Gschwind et al., 2012; Pyles et al., 2013). Despite the early evidence for dissociable parallel pathways for identity and emotion, it is now evident that there is interdependence between the neural processing of these features. The revised neural framework for face processing led to the understanding that information extracted by faces is processed by distributed and interacting modules (Dima et al., 2018).

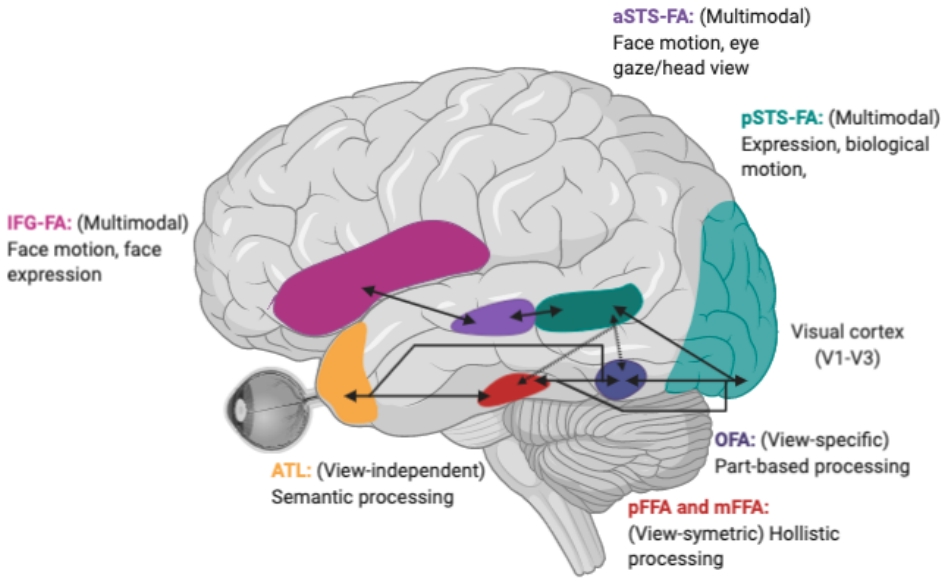


Figure 1.3: **Revised model of face perception**
A revised model of face perception based on the model proposed by Duchaine & Yovel (2015). The regions early in the visual hierarchy (V1-V3) are illustrated in turquoise.

In addition to the cortical regions previously noted to be involved in face processing, the revised model identifies the role of anterior temporal regions in humans. Evidence from macaque research highlights the importance of the anterior temporal face patch (ATFP) which displays strong neural responses to presentation of faces in contrast to other stimuli (Fisher & Freiwald, 2015). In contrast to the facial processing nodes earlier in the processing hierarchy, the neurons in the macaque ATFP respond invariantly to different face views, suggesting that the encoding is more of a higher-level representation in this region (Meyers et al., 2015). In humans, functionally defined ATFP bilaterally evoke more activity in response to faces than any other category (Harry et al., 2016). In the right ATFP, Harry and colleagues (2016) also demonstrated an equally strong neural response to headless bodies. Probing this further, multivoxel pattern analysis (MVPA) revealed that the voxels encoding face and whole-body responses were significantly overlapping in the right ATFP. However, this relationship was not shown with faces and headless bodies. Taken together, these findings support the idea that as information travels through the ventral visual stream towards the temporal pole, there is increasing convergence of information streams.

1.3 Body processing

Many insights into emotion processing in humans have focused on facial expressions. However, body postures can convey emotional information, and are typically understudied (de Gelder, 2006). Despite this, research has demonstrated that adults identify emotions conveyed in body postures with comparable accuracy to performance on facial expression recognition (Zieber et al., 2014). Importantly, bodies and faces are most frequently encountered together in space and time and are both biologically relevant, conveying information about conspecifics (Taylor et al., 2007).

1.3.1 Body posture processing

In 1965 Ekman conducted early investigations into bodies as affective communicators (Ekman, 1965). Early research exploring body perception sought to uncover if they were processed in a configurable manner, in a similar way to the face, and if there was a dedicated functional neural circuitry responsible for this processing (de Gelder, 2006). A robust methodology utilised to assess configural processing is the inversion effect; recognition of inverted bodies was found to be significantly impaired in contrast to upright bodies (Reed et al., 2003; Robbins & Coltheart, 2012). Recent evidence from magnetic resonance imaging (MRI) research shows that the selective areas of cortex which respond to bodies are preferentially tuned to whole-bodies over body-parts (Brandman & Yovel, 2014). However, bodies convey a rich amount of social information about an individual from the posture, movement and emotional expression and much of the research exploring the network responsible for this processing highlights that the neural representation is dependent on the attributes being processed (de Gelder, 2006). A framework with interrelated brain networks has been proposed by de Gelder (2006) for the processing of emotional body language, referred to as the 'two-systems' model of emotion-behaviour connectivity. Information is thought to be processed in parallel through a subcortial reflex-like network, and a cortical visuomotor perceptual route. These two systems have connections with the more frontal regions involved in the body awareness system.

1.3.2 Neural processing of bodies

Using fMRI in humans, category-selective brain regions have been identified for processing bodies; namely, the fusiform body area (FBA), extrastriate body area (EBA), and pSTS (Downing, 2001; de Gelder et al., 2015). The FBA in the mid-fusiform gyrus partially overlaps with the FFA and has been shown to respond to headless or faceless bodies. Despite the overlap, on a subject-by-subject basis the peaks of activation to faces and bodies occupy two close, but distinct, locations (Peelen & Downing, 2005). In the occipitotemporal cortex the EBA has been shown to respond more to bodies and body parts, than other object categories and faces (de Gelder et al., 2015). The response in EBA has been shown to linearly increase in relation to the proportion of a body posture displayed, consistent with part-based processing in EBA, whilst FBA shows a as a step-wise increase in response amplitude to whole bodies compared to body parts (Taylor et al., 2007). The pSTS has shown responsiveness to biological motion, but also static images of bodies implying motion, highlighting its importance for social perception. Application of transcranial magnetic stimulation (TMS) to the pSTS results in diminished ability of an observer to distinguish between affective body postures, demonstrating the importance of pSTS in detecting emotionally salient body postures (Candidi et al., 2011).

1.4 Facial expression processing in context

In our everyday interactions we typically encounter faces in conjunction with a body. The research on simultaneous perception of face and body cues is still sparse, which limits our understanding of how facial expressions are encountered in the 'real world'. The term 'context' is used broadly in the literature to refer to any social cue that is external to the face. For example, it could refer to an environmental scene or an expressive body posture. In the work that I will discuss and have conducted in this thesis, the term 'context' will be used solely when referring to an affective body posture.

Given the early notion of 'basic emotions' it was thought that a facial expression was readily judged in to a discrete category based on unique muscle signatures. As a consequence, much research focusing on facial expression perception has relied on decontextualised images of faces, as the facial expression itself was believed to be diagnostic.

Typically, the expressions shown are extreme, and not reflective of the typical displays of emotion encountered in everyday social interactions. In 2005, Meeren and colleagues published research demonstrating that observers' judgements of facial expressions became biased towards the emotion expressed by a non-congruent body posture (Meeren et al., 2005). Participants were asked to categorise facial expressions, presented with either a congruent (e.g., angry facial expression with angry body posture) or incongruent body posture (e.g., angry facial expression with fearful body posture), as rapidly and accurately as possible. Concurrent electroencephalography (EEG) was recorded from participants throughout the task. The behavioural results revealed congruency affected participants' accuracy and reaction times. When the facial and bodily expressions were incongruent, participants' judgement of the facial expressions became biased to the emotion conveyed from the body posture, and their reaction times were slower. Electrophysiologically, the P1 amplitude was found to be significantly larger for the incongruent face-body compound stimuli. The authors claimed that this enhanced P1 component points to rapid and early extraction of information about the congruency of simultaneously presented facial and bodily emotional expressions. The P1 component is believed to originate from 'early' extrastriate visual areas, which are regions known to be involved with processing changes in low-level properties (Schindler et al., 2021). Importantly, the congruent and incongruent stimuli did not differ in their low-level properties, just their emotional content. Taken together, these behavioural and electrophysiological results indicate that an incongruent body posture can result in biasing of the facial expression towards the emotion conveyed by the body posture. Furthermore, their results suggest that detection of congruency occurs early in the visual processing hierarchy.

More recent work by Aviezer *et al.* (2008) replicated Meereen's (2005) finding, demonstrating facial expression recognition was modulated by an affective body posture (Aviezer et al., 2008). The use of stimuli expressing anger, disgust, sadness and fear revealed that the greater the similarity between the emotions presented in the face and body, the larger the contextual influence of the body; this has been referred to as the *confusability effect*. When observers were asked to judge a disgusted facial expression, the participant was more likely to miscategorise the facial expression as the context, when the body posture was angry, as opposed to sad or fearful, due to the similarity of the anger and disgust

expressions. Other work by Aviezer *et al.* (2012a) presented ecologically valid facial expressions of defeat and victory in combination with tennis players bodies. When judging the valence of isolated faces observers failed to rate the affective valence of winners as more positive than the affective valence of losers. With the addition of the tennis player's body, participants' valence ratings of the facial expressions were biased towards the valence of the body posture (Aviezer et al., 2012a). This research demonstrates that body context has a clear influence on how facial expressions are perceived.

A limited amount of published work investigates how facial expressions are influenced by body posture during development. Mondloch (2012) investigated how emotion from the face and body are processed in children. 8-year-old children and adults were presented with fearful and sad facial expressions, on either a congruent or incongruent body posture, and they were required to make categorisation judgements (Mondloch, 2012). Comparing congruent trials (sad facial expression and sad body posture), with incongruent trials (sad facial expression and fearful body), revealed that both children and adults were more accurate and had faster reaction times when categorising the facial expression in the congruent task. When incongruent stimuli were presented, the influence of body posture on categorisation of the facial expression was larger in 8-year-olds than adults (Mondloch, 2012). This provides evidence to suggest that the influence of body posture on facial expression perception varies with age, and provides a rationale to further explore the mechanisms underlying facial expression and body posture integration in a developmental cohort.

1.4.1 Mechanisms for facial expression and body posture processing

It remains unclear exactly how facial expression and body posture cues are integrated. Aveizer *et al.* (2011) asked a group of adult observers to categorise facial expressions (anger, happiness, sadness, disgust, fear and surprise) presented in the context of an affective body posture, whilst they were explicitly instructed to ignore the body. Despite these instructions, observers' categorisation of the facial expression was still biased by the body posture. In a follow-up study, participants completed the same emotion categorisation task, but whilst completing a concurrent memory task, which added a 'cognitive load'.

Despite the additional cognitive load, the body posture was still found to influence emotion recognition from the face (Aviezer et al., 2011), implying that face-context integration occurs unintentionally. In combination with Meeren and colleagues (2005) electrophysiological findings, these results suggest that the integration of emotion information from the face and body may be 'automatic' and occurring early in the visual processing stream (Aviezer et al., 2012a).

In a study published by Aviezer and colleagues (2012b), they attempted to discern if the face and body were processed separately and then integrated, or if the whole-person was perceived as one gestalt. Facial expressions were paired with incongruent body postures to form 'whole-people'; however, the face and body were aligned in some trials and were manipulated to be misaligned in other cases. Their results show that misalignment resulted in a decreased influence of body context on facial expression perception, which the authors interpreted as a disruption in the holistic person gestalt (Aviezer et al., 2012b). This demonstrates that misalignment of the face and body weakens the 'perceptual unit', which lessened the influence of the body context on the face. In addition, they manipulated the physical distance between the face and body in the misalignment trials and no graded relationship was found between the influence of body context on facial expression judgement and the distance. This strongly supports the notion that the face and body are processed as one gestalt unit perceptually.

Recent work challenges the 'automatic' and 'early processing' view for integration of face and body stimuli (Teufel et al., 2019). Using a psychophysical adaptation procedure to emotional faces, Teufel and colleagues (2019) demonstrated that facial expression adaptation is unaffected by body posture. This suggests that integration of the facial expression and body posture happens later in the visual processing stream than has previously been proposed, and downstream of those regions involved in adaptation (Aviezer et al., 2011; Meeren et al., 2005).

1.4.2 Integration of social cues in the brain

Much of the research exploring facial expression and body processing in the brain highlights cortical regions which display specificity in prescribed domains. Only recently has research begun exploring the interplay between the neural regions in face and body processing.

Fisher and Freiwald (2015) explored the face and body responsive areas using fMRI in the macaque brain. The aim of this work was to uncover if face-selectivity is purely driven by the visual attributes of the stimulus, or if the activation of neural regions is a consequence of faces indicating the presence of a larger social agent. This study localised face patches in four monkeys using fMRI which revealed that some face patches responded to both faces and bodies, whilst the body selective regions were much more category specific repsonding only to images of bodies (Fisher & Freiwald, 2015). This provides evidence that the neural activity in certain face areas can be synergistically increased with the addition of bodies. Cross-species work highlights that similar regions exist in human brains, and with quantitative topographic modelling, homologous regions can be identified (Rajimehr et al., 2009). In humans, the ATFP responds strongly to the presentation of faces and whole bodies. However, 'parts' of faces and bodies do not elicit a strong response (Harry et al., 2016). These findings allude to a hierarchical stream of processing for faces and bodies along the ventro-occipital temporal cortex.

Distilling literature from sub-disciplines of facial expression and body posture processing, Hu and colleages (2020) recently proposed a theoretical framework for the combined processing of both cues. At its crux the framework states that the integration of faces and bodies is mediated by the goal of the processing. A hierarchical functional neural architecture has been proposed for integration of these cues that retains a degree of separation between the dorsal and ventral processing streams (Hu et al., 2020). Within this architecture two centres of integration have been proposed: a ventral semantic integration hub that is the culmination of posterior to anterior face-body integration, and a social agent integration hub in the dorsal stream pSTS. This model fits with the ideas proposed in this thesis regarding the visual integration of facial expression and body posture cues, and indicates that the anterior temporal lobe and pSTS are important candidates for the site for integration. Furthermore, this theoretical model supports evidence from Teufel *et al.* (2019) for a later site for integration of facial expression and body posture along the extended ventral visual processing stream (Teufel et al., 2019).

1.5 Cue integration

A perceptual challenge for the human brain is to combine multiple sensory estimates about the world to guide an observer's perception and interactions (Trommershauser et al., 2011). These estimates are always associated with uncertainty due to the noisiness of neural information-processing (Whiteley & Sahani, 2008). Therefore, to minimise the uncertainty in sensory cue measurement, observers often combine multiple cues to improve the reliability of their estimates (Trommershauser et al., 2011). The human brain is thought to arrive at integrated estimates from sensory cues in an efficient and optimal manner (Helbig & Ernst, 2007).

Optimal cue integration, based on the maximum-likelihood principle (MLE), predicts that the relative reliability of the sensory cues determines how much they will contribute to the integrated representation, and that the integrated estimate should be more reliable than the individual sensory estimates (Ernst and Banks, 2002). Put simply, more reliable information is given a stronger role in the final integrated estimate, which is more reliable than either of the individual cues. These predictions are supported by empirical evidence from low-level vision research (Bejjanki et al., 2011; Dekker et al., 2015; Ernst & Banks, 2002; Martin, 2016).

In order to understand how humans effortlessly exist in a complex social world, one must consider one of the most basic perceptual challenges: how does our brain integrate more complex social signals, like facial expression and body posture? In the real world, the goal of perception is not to distinguish isolated cues, but to process multiple cues and form a coherent percept. Based on the principles of cue integration, that have been applied to understand low-level cue combination, proposals have been put forward that similar mathematical models may be applied to understand how social cues are integrated (Zaki, 2013). However, these proposals have not yet been explicitly tested, with part of the challenge being the difficulty in achieving the necessary control over complex social stimulus properties.

1.6 Aims and overview of thesis

The research in this thesis aims to provide a mechanistic account for the integration of facial expression and body posture cues [Figure 1.4]. As highlighted, exactly how these two social cues are integrated to form whole-person representations of emotion still remains unknown. In Chapter 3, I developed a novel psychophysical paradigm to quantify how much body posture influenced perception of facial expressions between observers. In addition, observers' ability to categorise isolated facial expressions and body postures was also quantified. A key question I attempted to resolve in this chapter was how observer's individual facial expression and body posture representations related to their whole-person perception of emotion. I found individual differences in the magnitude of the body biasing effect, such that some observers were more influenced by body context than others. Those observers' who were less precise in their ability to discriminate facial expression relied more on body posture for their facial expression judgments in the whole-person condition, and vice versa.

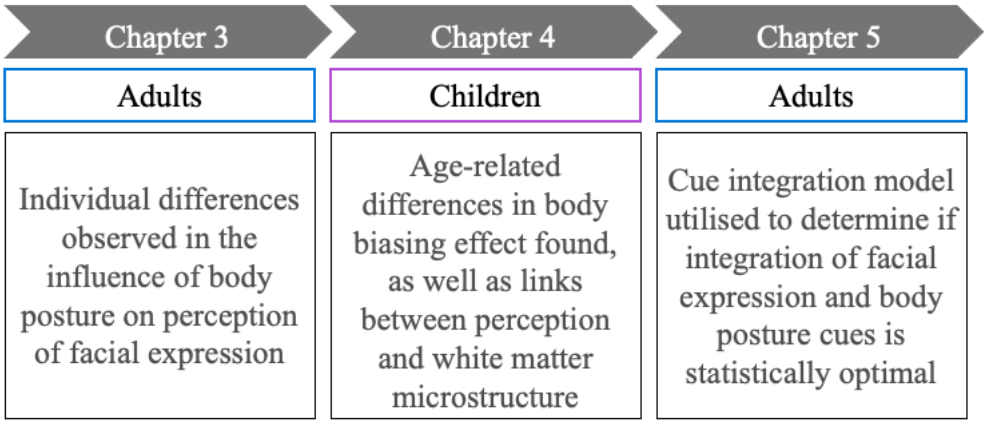


Figure 1.4: **Thesis summary and main findings**

In Chapter 4, I aimed to uncover how changes in facial expression and body posture recognition across development affect the integration of these signals into a whole-person percept. Furthermore, in the developing brain, children undergo a prolonged period of white matter maturation. I assessed if microstructure of white matter tracts connecting functional nodes of the brain involved in face and body processing was linked to perception of facial expression, body posture and whole-person cues. I found children's ability to recognise facial expressions improved with increasing age, and that the influence of body

context on the perception of facial expression decreased as children got older. I also found that children's facial expression recognition reliability could be predicted by the fractional anisotropy of white matter tracts connecting OFA and FFA. The research in this chapter combines psychophysics with functional and diffusion MRI to provide insights into the emergence of facial expression and body posture perception across development.

In Chapter 5, I applied a cue integration model to determine if facial expression and body posture are integrated in a statistically optimal manner. The results of this chapter provide novel insights into the mechanisms underlying integration of facial expression and body posture cues, suggesting that the integrated representation is a weighted average of the isolated cues. These results have important implications for our real world understanding of social perception and the principles that drive the integration of multiple social signals.

Chapter 2

Methods

2.1 Psychophysical approach to perception

The term psychophysics was first coined in 1860 by Gustav Theodor Fechner and is a sub-discipline of psychology that quantifies the relationship between physical stimuli and their subjective percepts (Kingdom & Prins, 2010). The use of psychophysical methodologies facilitates quantification of the relationship between the physical world and human perception (Lu & Dosher, 2014). As a methodology, psychophysics can be applied to many domains, but the 'classic' work focuses on early sensory systems (Read, 2015). The examples I discuss, and the application of these methods in my thesis, are specific to the visual system.

One focus in psychophysics is measuring the sensitivity of the perceptual system. Sensitivity can be defined as the minimal physical stimulus change that is detectable by the observer. By manipulating the physical stimulus, and measuring the resulting changes in perception, psychologists can characterise the link between input and percept. This relationship is typically summarised in the form of a psychometric function (PF), which relates quantitative stimulus properties to the probability of a particular percept (Read, 2015). Typically, parameters of interest are extracted from the PF that summarise an observer's sensory processing. A PF can be fully described by four parameters: the threshold (α), slope (β), guess rate and lapse rate (Prins, 2013). Two parameters of interest for characterising the underlying sensory mechanisms are α and β. α specifies the location of the function along the stimulus dimension and is usually defined by a

point where a specific level of performance is achieved (Prins, 2013). One example of a location measure adopted in appearance-based, forced-choice matching tasks, is the point of subjective equality (PSE). The point of subjective equality (PSE) is defined as the point at which the stimulus appears perceptually equal to another stimulus or to an internal standard. At this point, the observer will be as likely to indicate either response option for a given stimulus level (Kingdom & Prins, 2010). The parameter β determines the rate of change in the observer's responses. In the adaptive procedures detailed in the experimental work in this thesis the guess rate and lapse rate are referred to as nuisance parameters and are therefore set as fixed values (Kontsevich & Tyler, 1999).

In order to acquire the data to which PFs were fitted for each observer, I used the psi method (Kontsevich & Tyler, 1999). This adaptive procedure is a Bayesian approach, optimised for acquisition of both the threshold and slope of the PF (Kontsevich & Tyler, 1999). Essentially, on each trial the psi algorithm selects the stimulus intensity that maximises the expected information to be gained by completion of that trial. The efficiency gained by this approach facilitates estimation of both the slope and threshold with equivalent precision to conventional methods with significantly fewer trials (Kontsevich & Tyler, 1999).

In conclusion, utilising psychophysical methods is advantageous as they allow us to get closer to the underlying neuronal mechanisms responsible for perception (Read, 2015). They allow us to determine how sensory information is encoded, and how this information is converted into perceptual judgements. Changes in perceived stimulus intensity are believed to be directly related to changes in how sensory neurons encode physical stimuli. As a methodology, psychophysics is a key tool for quantifying human perception of the visual world and can provide a window to brain function.

2.2 Face and body processing in the brain: Magnetic resonance imaging

MRI is a non-invasive scanning technique which utilises a strong static magnetic field (B_0) to acquire images of the human body (Huettel et al., 2014). While important for many clinical applications, MRI has evolved into an important tool for modern cognitive

neuroscience as it allows us to explore relationships between brain structure, function and behaviour.

2.2.1 Basic concepts of MRI

Hydrogen (H_1) atoms are composed of a single proton and are highly abundant in the human body. Each nucleus has an intrinsic 'spin', which means when a person is placed in the magnetic field of an MRI scanner the protons begin to *precess*. The frequency at which a proton precesses is referred to as the Larmor frequency (Tubridy & McKinstry, 2000). Most protons will precess in a low-energy state, aligned with the applied magnetic field. However, a few will precess in the opposite direction, in a high-energy state. Radiofrequency pulses are applied at the resonance frequency of hydrogen to excite the aligned H_1 atoms. This causes protons in the low-energy state to absorb energy and move to a higher-energy state. Following an excitation pulse, the equilibrium of the state must be restored. The process of the higher-energy spins returning to the lower-energy state causes a release of radio-waves, which is what constitutes the magnetic resonance (MR) signal. Different tissue types in the brain relax at different rates which results in differences in image intensities for different tissues. Localisation of the signal for image construction relies on the application of magnetic field gradients, in addition to the B_0. These are secondary magnetic fields which vary as a function of position, causing the resonance frequency of protons to change according to location. This spatial 'tagging' of spins is used to selectively excite a targeted slab of tissue, as well as create spatially-dependent differences in frequency and phase within the excited slice of tissue. This in turn allows for the reconstruction of a three-dimensional image that is sensitive to the local magnetic environment, which is reflective of the different tissue properties.

2.2.2 Functional MRI (fMRI)

The development of fMRI in the 1990s was an exciting advance, facilitating imaging of brain activity to identify neural responses to specific stimuli or tasks (Ogawa et al., 1990). It became apparent that oxygenation levels influenced the magnetic properties of the blood, and these changes could be quantified using MRI, birthing a new era of neuroimaging (Huettel et al., 2014).

Blood oxygenation level dependent (BOLD) fMRI

One of the most common approaches for fMRI, first described in the early-1990s, is BOLD imaging, which relies on regional differences in cerebral blood flow (Ogawa & Lee, 1990). By capitalising on the differences in magnetic susceptibility of oxygenated and deoxygenated blood, the applications of BOLD fMRI are now vastly utilised as an effective method to non-invasively capture spatial changes in neural activity in the brain.

Oxyhaemoglobin is formed during respiration when oxygen binds to the haem component of the protein haemoglobin in red blood cells. It has no unpaired electrons, however, when oxygen is released, and deoxyhaemoglobin is formed, the unpaired electrons result in the molecule becoming strongly paramagnetic. This causes a difference in the magnetic susceptibility of the blood and the surrounding tissues, which result in differences in signal decay. Upon neural activation within a local region of the brain, changes in both the local cerebral blood flow and oxygenation concentration occur, resulting in differences in the MR signal decay (Glover, 2011). It is important to note that the generation of the BOLD contrast is much more complex than just relying on blood oxygenation, as it is also dependent on physiological factors such as blood flow, volume and vasculature (Arthurs & Boniface, 2002). Early work has indicated that changes in BOLD responses linearly relate to the underlying neural activity (Logothetis, 2003). In addition, work has been undertaken to simultaneously record electrophysiological data and fMRI, which revealed that the BOLD changes reflect the neural changes induced by a stimulus (Hillman, 2014).

High-field fMRI

With MRI being a dominant methodology utilised for investigating the functioning human brain, advances have seen the advent of higher-field strength scanners. Ultra-high-field (UHF) MRI refers to static magnetic fields of 7 Tesla (T) and above (Chang et al., 2016). With UHF scanning, significant advances in spatial resolution have been possible. Typical voxel dimensions on a 3T system are around 3x3x3mm^3 for BOLD fMRI, whilst at UHF the resolution can reach below 1x1x1mm^3. In addition to the improved signal-to-noise ratio (SNR) at UHF, there are differences in the magnetic susceptibility of the blood resulting in a concomitant increase in BOLD signal at higher static magnetic field strengths (Ladd et al., 2018). At 7T, the BOLD signal is more tightly related to the concentration of

deoxyhaemoglobin in capillaries rather than larger veins, which is thought to more closely reflect of the neural activity (Chang et al., 2016).

2.2.3 Diffusion MRI

Diffusion magnetic resonance imaging (dMRI) is an established methodology used to non-invasively characterise tissue properties at the microscopic level (Brun et al., 2019). Diffusion-weighted imaging (DWI) is based on the principles of random Brownian motion which describes how molecules are driven by thermal energy and are constantly undergoing small random fluctuations (Einstein, 1905). The rate of molecular diffusion is different in different tissues of the brain; in the cerebrospinal fluid (CSF), diffusion of water molecules is free or isotropic, in contrast to the more restricted, or anisotropic, diffusion of the grey and white matter (Jones, 2010). By using MRI sequences that are sensitive to differences in the diffusion of water, images can be produced that have distinct contrasts for different tissue types, providing insight into the underlying structure of the brain.

The acquisition of DWI requires the application of gradient pulses, which essentially magnetically labels 'spins' carried by diffusing molecules (Jones, 2010). As diffusion driven displacement of water molecules occurs, the MRI signal becomes attenuated, providing a quantitative measure for diffusion along a particular gradient direction, referred to as the diffusion coefficient (D) (Le Bihan & Johansen-Berg, 2012). Diffusion of water is highly anisotropic in the myelinated white matter, and therefore the changes in D along specific directions can allow inferences to be made about long-range white matter pathways of the brain (Le Bihan & Johansen-Berg, 2012).

Diffusion tensor imaging (DTI)

DTI is one of the most commonly utilised applications in neuroscience research to characterise the diffusion of water in the brain (Basser & Pierpaoli, 1996; Jones, 2010). Fitting the tensor model allows the three-dimensional shape of diffusion to be quantified across three principle directions within each voxel, thus providing insights into the microstructural organisation of the brain (Jones, 2010). The three-dimensional shape, and magnitude of diffusion, is dependent on the underlying brain architecture and physiology (Huisman, 2010). Within the CSF where diffusion is free and equal in all directions, the diffusion

can be described as 'isotropic' and the diffusion tensor represented as a sphere. In white matter tracts, where the diffusion of water is mainly in the direction parallel to the long axis of the tract, diffusion is 'anisotropic' and can be graphically represented as an ellipsoid (Huisman, 2010). Diffusion tensor imaging allows anisotropy in the tissue to be accounted for by using at least six different gradient directions to acquire images (Jones, 2010). This allows the diffusivities along the principle directions to be calculated. Metrics such as mean diffusivity (MD), fractional anisotropy (FA), radial diffusivity (RD) and axial diffusivity (AD) can be recovered by fitting the tensor model. One of the most widespread measures of anisotropy is FA, as it is highly sensitive to microstructural changes and has shown good correspondence with myelination (Alexander et al., 2007; Beaulieu, 2002; Choe et al., 2012). Despite FA being referred to as a measure of 'microstructural integrity', it is very non-specific, making it sensitive to multiple microstructural attributes, including axon orientational dispersion, axon density, membrane permeability and myelination (Basser & Pierpaoli, 1996; Winston, 2012).

Tractography

Tractography is an extension of DTI, where directional information from each voxel about the diffusivity of water is used to generate three-dimensional white-matter maps (Yogarajah et al., 2009). Water molecules are less 'hindered' along fibres, compared to perpendicular orientations, which means it is possible to infer the long-range white-matter pathways in the brain based on neighbouring local fibre orientations (Jeurissen et al., 2017). The advent of tractography allowed the non-invasive study of white-matter tracts in-vivo for the first time, and is now the methodology of choice to investigate quantitative MRI parameters in specific fibre bundles (Jeurissen et al., 2017; Yogarajah et al., 2009).

Two predominant tractography algorithms exist: deterministic and probabilistic (Descoteaux et al., 2009). Deterministic approaches assume that each voxel has a unique orientation estimate which can be tracked between voxels resulting in a single pathway emanating from a starting point (Jeurissen et al., 2017). However, the reality is that each of these voxel-wise estimates on orientation is subject to a degree of error which is not considered in deterministic tracking. On the other hand, probabilistic tractography allows this measurement uncertainty to be characterised resulting in a distribution of possible

trajectories from a seed region (Jeurissen et al., 2017). This provides an estimate of confidence regarding the tracking of the route with the least hindered diffusion, which can improve the reliability of tractography in regions with multiple crossing fibres and high uncertainty (Jbabdi & Johansen-Berg, 2011).

As highlighted above, DTI is one of the most commonly used models for characterising the diffusivity within a voxel. Despite its widespread application, it is inherently limited as it is only capable of distinguishing a single fibre population per voxel (Jeurissen et al., 2017). Hence, voxels with crossing fibre configurations have a low anisotropy index due to the non-gaussian diffusion of crossing fibres (Descoteaux et al., 2009). Estimates indicate that up to a third of all white matter voxels in the brain contain multiple fibre crossings as a consequence of kissing, branching or fanning axons, just to highlight a few examples (Behrens et al., 2007). To overcome these limitations, 'higher-order' fibre modelling methods have been developed with the ability to estimate multiple fibre orientations per voxel. Instead of each voxel being represented by a unidirectional ellipsoid, the fibre orientation distribution function (fODF) represents fibre orientations as a continuous function of a sphere (Jeurissen et al., 2017). Derivation of fODFs is dependent on advanced imaging acquisitions which employ a large number of gradient directions at high b-values, where b-values are proportional to the gradient strength, duration and diffusion time (Berman et al., 2013). These recent advancements in modelling, in combination with new hardware and acquisitions, overcome the limitations of the crossing fibre problem which makes tractography more robust.

High-gradient dMRI

The diffusion data reported in Chapter 4, was acquired on a 3TM Connectom Siemens system (Siemens Healthcare, Erlangen, Germany) with gradients of 300 mT/m. The stronger diffusion weightings per unit time result in an improvement in the SNR (Jones et al., 2018). In comparison to a standard 3T system, where the gradients are typically 80 mT/m the Connectom's increased gradient strength improves SNR by approximately 50%. An advantage of the Connectom system is that stronger diffusion gradients, up to approximately b=6000s/mm^2 can be achieved. Standard research MR systems typically utilise b-values in the range of 500-1200. By using higher b-values it is possible to get closer to the un-

derlying biophysical processes, as the signal is predominately coming from within axons or glial processes at such values (McNab et al., 2013).

Chapter 3

Individual differences in adult's perception of facial expressions in context

3.1 Introduction

A key feature of effective face detection is the ability to extract emotional information from the expression (Calder, 2011). Facial expressions provide information to help guide an observer's behaviour during a social interaction and can aid prediction of conspecifics actions (Rajhans et al., 2016). However, facial expressions are rarely encountered in isolation, and are typically perceived in conjunction with a body posture. Previous research on the perception of facial expressions has highlighted the importance of context (Aviezer et al., 2012a; Aviezer et al., 2008; Hassin et al., 2013; Meeren et al., 2005; Teufel et al., 2019).

Observers' judgements of facial expressions have been shown to be biased towards the emotion expressed by a non-congruent body posture (Meeren et al., 2005). Electrophysiological differences have also been reported in early visual regions when viewing congruent versus incongruent facial expression and body postures. Meeren *et al.* (2005) interpreted this difference as the presence of a rapid and automatic evaluation to assess if the biologically important emotional information in the face and body was congruent. More recent work by Aviezer and colleagues (2008) found that an affective body posture influenced perception of facial expression, such that observers were more likely to perceive a disgusted face as 'angry' when presented in the context of an angry body. When participants were instructed to ignore the body posture when making facial expression categorisations,

or were given a concurrent working memory task to add cognitive load, body posture was still found to influence the perception of the facial expression (Aviezer et al., 2011). This work supports the idea that integration of facial expression and body postures is automatic. However, more recent work challenges the 'early integration' view of face and body stimuli (Teufel et al., 2019). Using an adaptation paradigm, Teufel and colleagues (2019) provide evidence for later integration of facial expression and body posture, downstream of early core face perception regions. The argument for integration occurring early in the visual processing stream conflicts with established models of face recognition which describe expression processing taking place further downstream.

Despite current attempts to understand at which stage in the processing stream integration of facial expression and body posture cues takes place, there is a gap in the literature exploring exactly how these signals are combined. To establish exactly how humans make judgements about facial expressions in the context of a whole-person, accurate measurements are required of how the individual face and body cues are represented. To my knowledge, no published research has quantified how well observers are able to categorise individual facial expressions and body postures and related this to whole-person categorisation of facial expressions. Furthermore, much of the current literature relies on the use of maximum intensity prototype facial expressions; although these faces are validated from databases, they are frequently not akin to the typical expressions of emotion we encounter in the real world. In the present study, in order to create a graded continuum of body postures, novel 3D body avatars were created that could be morphed between different postures. To address limitations in the literature, the current study adopts psychophysical methods to accurately quantify how observers judge isolated facial expressions, body postures and facial expressions in the context of an affective body posture.

In order to get a better understanding about what factors lead to integration of facial expression and body posture, research must take into consideration both the individual representations of the two cues, and other higher-order processing differences. Differences in an individual's acquired emotional knowledge and experience could play an important role in integration of emotional social cues. To gain insight into individual differences in propositional knowledge for emotion I administered the Levels of Emotional Awareness

Scale (LEAS) questionnaire. Individuals low in emotional complexity, as reflected by a low total LEAS score, encode and represent emotion in action-oriented terms, whereas individuals high in emotional complexity are able to process multiple emotional concepts with more normative recognition of emotion cues in others (Lane et al., 1998; Tavares et al., 2011). Based on previous research using the LEAS questionnaire, I predicted that how an individual performs on this questionnaire may be reflective of perceptual abilities when categorising facial expressions with an expressive body posture (Tavares et al., 2011). In observers who have lower emotional complexity and are therefore thought to be more action-oriented, one would expect them to be more influenced by the body posture when judging the facial expression in the context of a whole-person. In contrast, one would hypothesise an observer with a higher LEAS score, reflective of more complex emotion concepts, is less likely to have their judgements of the facial expression affected by an expressive body posture.

The primary aim of this study was to assess the feasibility of a novel psychophysical paradigm to quantify how much body posture influenced perception of facial expressions. In addition, I aimed to determine observers' ability to categorise isolated facial expressions and body postures, and relate these representations to whole-person perception of emotion. Furthermore, I also aimed to capture differences in higher-order emotional awareness between observers using the LEAS questionnaire and determine if these differences were related to an observer's perceptual ability for categorisation of social cues.

3.2 Methods

3.2.1 Observers

A total of 43 naïve observers (9 male) were recruited from both the general population and Cardiff University School of Psychology undergraduate students. Only neurologically and physically healthy participants were recruited. All participants had normal or corrected to normal vision. All participants were over 18 years of age (mean age = 20.68 ± 2.84 years, range = 18 – 29). Observers provided written informed consent prior to participating. Experimental protocols were approved by Cardiff University School of Psychology Ethics Committee, and were in line with the Declaration of Helsinki. All participants were fully debriefed at the end of the testing session and payment was provided in the form of cash or course credits.

3.2.2 Stimuli

Throughout my thesis, angry and disgusted expressions were used for the following reasons: firstly, previous research exploring the influence of body posture on facial expressions has shown that due to the high 'confusability' of these expressions, the contextual influence is greatest (Aviezer et al., 2008). In the current feasibility study, I wanted to relate the reliability of the individual facial expression and body posture representations to the body biasing effect, hence, I aimed to maximally induce this effect with the expressions selected. Secondly, given that the use of body posture morphs in my study is novel, I selected expressions that are easily identifiable from a body posture, and would lend themselves to morphing between postures. Previous research has shown that angry and disgusted body postures are two of the most readily recognised distinct static postures from the range of basic expressions (Lopez et al., 2017). Presentation of all stimuli was in grey-scale with a grey background.

Facial Expressions

Facial expression stimuli were generated using photographs of male actors from the Radboud and NimStim validated sets of facial expressions (Langner et al., 2010; Tottenham et al., 2009). Four Caucasian male identities were selected. The angry and disgusted

facial expressions were morphed together for each identity using FantaMorph software [FantaMorph Pro, Version 5]. This was accomplished using the 'Face Locator' feature to map out the main features of each face. This procedure generated morph continua for each identity. The morphs changed in increments of 5% between the angry and disgusted facial expressions for each identity. In total, there were 21 morph levels generated for each identity between anger and disgust. Each morph was exported from Fantamorph as an individual image file. For categorisation of the facial expressions in isolation, a mask blending into the mean grey background was used to remove any external facial features, such as the hair and ears [Figure 3.1a].

Body Postures

Body posture morphs were weighted averages of two motion captured 3D photorealistic body avatars expressing anger and disgust (unpublished stimuli from collaborators at Max Plank Institute Tubingen, 2018). These were created by a male adult actor in a motion capture suit with motion trackers distributed over the whole body. The poses for the angry and disgusted postures were based on the stimulus set used in the Aviezer *et al.* (2008) study. Visualisation of these postures was achieved using Unity 3D game engine [Unity, 2018]. Four unique angry and disgusted body postures were generated by the actor adopting slightly different poses. For each of the identities the individuals body composition and clothing was different. This resulted in 4 identities, each having one disgusted and one angry posture. The avatars were morphed between the expressive postures producing a range of angry and disgusted body posture morphs changing in increments of 5%, resulting in a total of 21 morph levels per identity. Each body posture was combined with a facial identity to make a photorealistic 'whole-person', and a mean grey oval was placed centrally over the face to conceal the distinguishing features of the facial expression [Figure 3.1b].

Whole-person stimuli

To create the whole-person composite stimuli the individual morphed emotional faces were manually pasted onto the body postures using GIMP [GNU Image Manipulation Program, Version 2.10]. In order to maximise the influence of body posture on categorisation of

facial expressions, the morphed facial expressions for each identity were merged with both fully-angry and fully-disgusted body postures [Figure 3.1c].

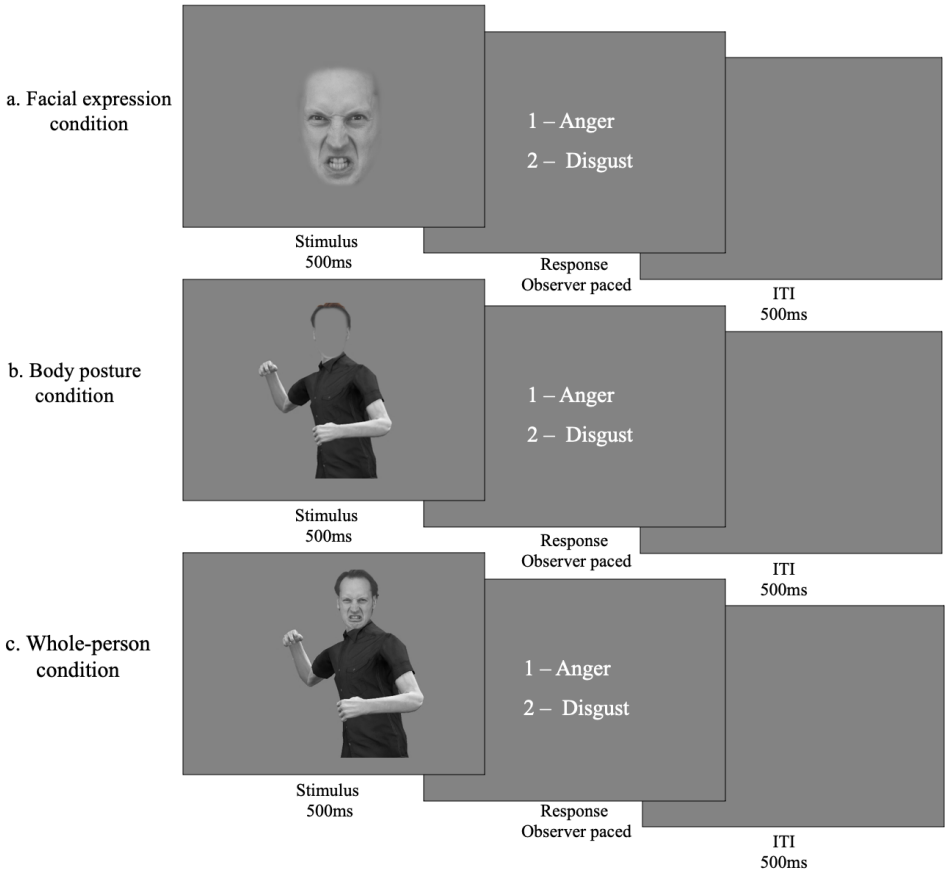


Figure 3.1: **Experimental conditions**
The figure depicts the three experimental conditions: a. Facial expression, b. Body posture, c. Whole-person conditions. On each trial the stimulus was presented centrally for 500ms. Observers were instructed to categorise the facial expression in the [a.] facial expression and [c.] whole-person conditions, and the body posture in the [b.] body posture condition. Following presentation of the stimuli, response options appeared on screen and remained until an observer made a response. The response was made with a key press; observers pressed 1 for anger and 2 for disgust. The next trial commenced following a 500ms inter-trial interval.

3.2.3 Procedure

Presentation of the task was controlled by custom-written MATLAB [Version 2016b, The MathWorks, Natwick, MI, USA] code using the Psychophysics Toolbox (Version 3.0.14) (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Observers took part in three conditions: 1. Facial expression categorisation, 2. Body posture categorisation, 3. Whole-person categorisation [Figure 3.1]. For all conditions the stimulus presentation procedure was identical. The test stimulus was presented centrally on the screen for 500ms. The images subtended approximately 15° visual angle (vertically) by 10° visual angle (horizontally). Following stimulus presentation, the response options appeared and remained on screen until the participant made a response. The participants were instructed to make a categorisation of the stimuli as disgusted or angry, indicating their decision with a key press. Participants were explicitly instructed to categorise the facial expression, and ignore the body posture in the whole-person condition, in order to determine how much the body context biased perception of the facial expression. In the absence of eye-tracking data, to ensure observers were not just indicating the expression shown in the body posture, participants were specifically instructed to categorise the emotion presented in the facial expression. This instruction was the same as other published research in this field (Aviezer et al., 2008, 2011; Meeren et al., 2005; Teufel et al., 2019). After the key press the next trial commenced following a 500ms inter-trial interval.

The stimulus level of the morphed facial expressions, body postures and whole-person stimuli was determined using the adaptive 'psi' algorithm (Kingdom & Prins, 2010). The psi method is considered to be one of the most efficient adaptive methods used for estimating full psychometric functions, thus minimising the overall number of trials required (Kingdom & Prins, 2010). The total number of trials for the facial expression and body posture conditions was 100. For the whole-person condition there were 496 trials. Following in-house piloting of the whole-person task, it appeared to be more perceptually challenging than categorisation of the isolated face and body stimuli, hence the number of trials was increased to ensure an adequate fitting of psychometric functions. Each of the four identities were displayed an equal number of times in each condition. The total task took approximately 30 minutes to complete.

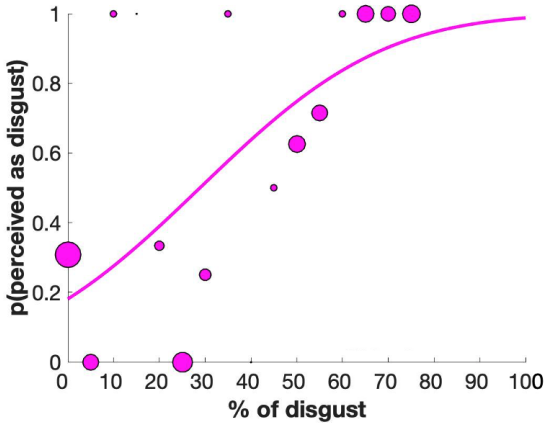
The study design was a repeated-measures block design. The order of the three experi-

mental conditions was counterbalanced across participants. For each condition the participant underwent a short training phase at the beginning of the block. For the whole-person task, observers were instructed to categorise the facial expression and explicitly told to ignore the body posture. A subset of observers (n=23) completed the LEAS questionnaire (Lane et al., 1990). The LEAS questionnaire is a written performance measure that requires the subject to articulate in written form how a range of scenarios would make themselves and another person feel. The questionnaire is composed of 20 vignettes and a structural scoring criterion was used to identify the emotional terminology utilised by the participant (Lane et al., 1998). Specific training was undertaken by the experimenter prior to response scoring on the questionnaire. Observers completed the questionnaire electronically using Qualtrics [Qualtrics, Provo, UT]. The questionnaire took approximately 20 minutes to complete.

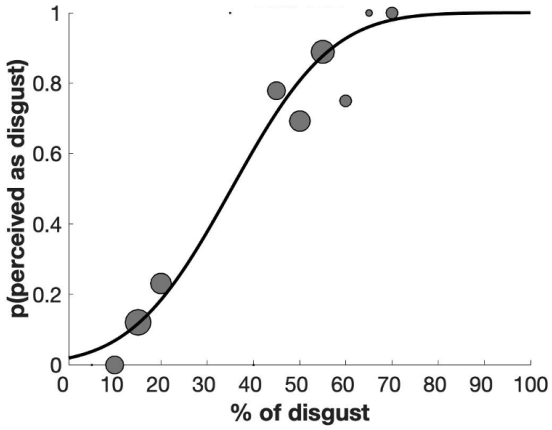
3.2.4 Data Analysis

Using custom-written MATLAB code with the Palamedes toolbox (Prins & Kingdom, 2018), PFs were fitted based on a cumulative Gaussian to estimate the PSE and slope value for each observer, for each condition [Figure 3.2]. Lapse rate was fixed at 0.03; guess rate was determined by the experimental procedure and was fixed at 0. The steeper the slope of the psychometric function, the more reliably the observer distinguished between the morphed stimuli (Kingdom & Prins, 2010). The PSE is the point at which the observer was equally likely to respond disgust or anger to a particular stimulus. The PSE change in the whole-person condition between facial expression morphs shown with fully-angry and fully-disgusted body postures reflected the modulation of the facial expression judgement due to the contextual influence of the body posture [Figure 3.2c]. Goodness-of-fit of the PFs was assessed using the method described in Wichmann and Hill (2001) and implemented in the Palamedes toolbox in MATLAB (Prins & Kingdom, 2018). The procedure allowed estimation of the proportion of deviance values (pDev) obtained from bootstrapping, that were greater than the deviance value of the original data (Wichmann & Hill, 2001). Furthermore, visual inspection of PF fitting was assessed independently by two researchers. Nine participants were excluded from the facial expression condition, one from the body posture condition and nine were excluded from the whole-person condition

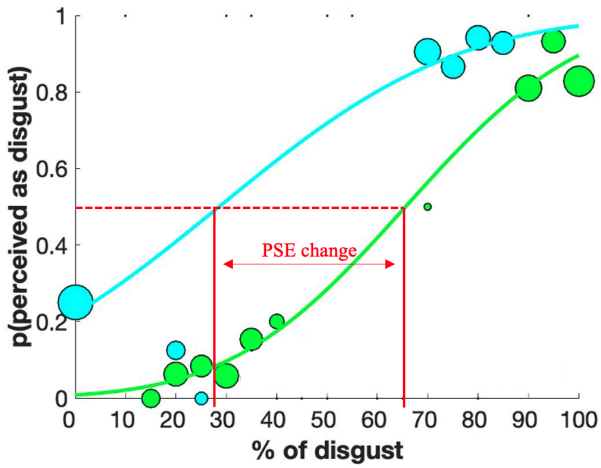
based on poor fits of the PF. For subsequent analyses Spearman's rank correlations were performed as the data followed a non-normal distribution. To ensure robustness of the results, Cook's distance was implemented to ensure data points with large residuals were removed from the correlations reported.



(a) Facial expression condition



(b) Body posture condition



(c) Whole-person condition

Figure 3.2: **Example of psychometric function fitting**
For each condition, for every observer, a PF was fitted to their data. The figure shows an example of one observer's data for the [a.] facial expression, [b.] body posture and [c.] whole-person categorisation. In the whole-person condition the green PF shows the facial expression shown with a 100% angry body posture and the blue PF shows the facial expression shown with 100% disgusted body posture. The PSE change is shown with the red arrow. The x-axis reflects the percentage of disgust present in the stimulus presented. Therefore, at less than 50% disgust, there is a higher percentage of anger present in the stimulus than disgust. The y-axis depicts the probability of the stimulus being categorised as disgusted by the observer. The size of the plotted data points reflects the number of trials at the specific stimulus level, with a larger size indicating more trials at a particular morph level.

3.3 Results

3.3.1 Isolated facial expression and body posture categorisation

The slope values of the estimated PF provided a measure of how accurately an observer distinguished between the morphed stimuli; the steeper the slope, the more precise the individual's performance. I found a significant positive correlation between the estimated slopes of the facial expression and body posture conditions ($r_{(s)}$=0.54, p<0.01) [Figure 3.3]. This suggests that observers who were better at categorising facial expressions were better at categorising body postures, and vice versa.

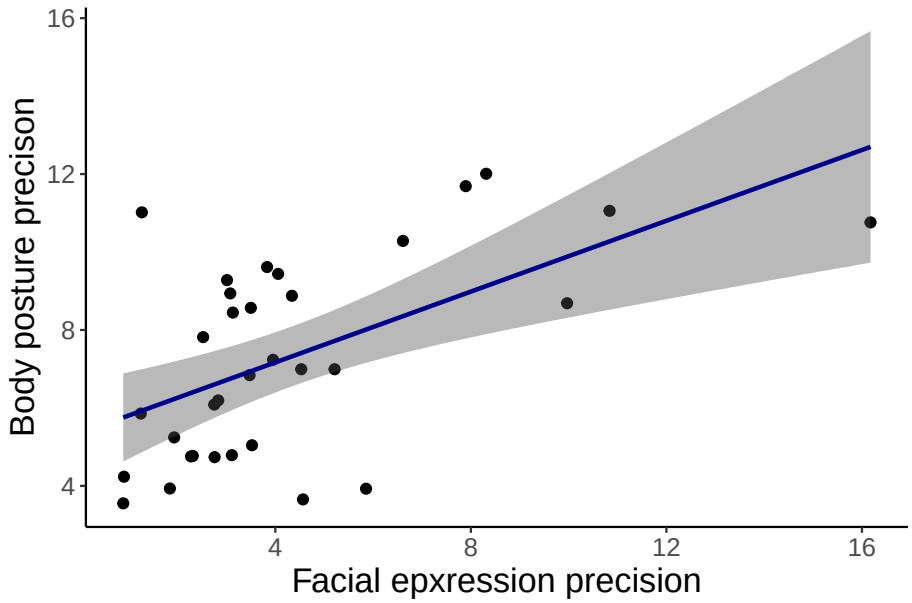


Figure 3.3: **Relationship between facial expression and body posture precision** A significant positive correlation was seen between observers facial expression and body posture precision ($r_{(s)}$=0.54, p<0.01). The precision for each condition was indexed by the slope estimate of the individuals PF for each condition. Each point represents one observer. The 95 % confidence interval is shown with grey shading.

Despite the strong correlation between observers' performance on facial expression and body posture conditions, the estimated slopes for the body posture only condition, were significantly steeper than the facial expression slopes, as determined with a Wilcox Signed Rank test (z=1.06, p<0.001) [Figure 3.4]. This indicates that observers performance when categorising body postures was more reliable than facial expression categorisation.

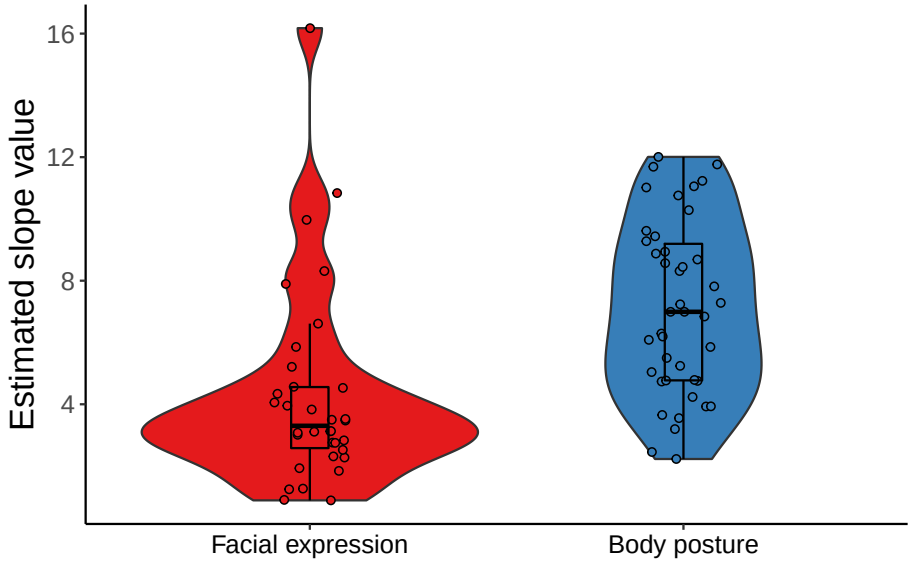


Figure 3.4: **Estimated slope values from PFs of facial expression and body posture categorisation**
The violin plot displays the estimated slope values for each observer in the facial expression and body posture conditions. There was a significant difference between the estimated slope values of the facial expression and body posture conditions (z=1.06, p<0.001). Each point represents one observer. The distribution of the values is illustrated by the shaded area, with the overlaid boxplot indicating the median and the interquartile range.

3.3.2 Whole-person categorisation

Body posture was found to significantly influence the perception of facial expressions. The greater the difference in PSE between psychometric functions from the whole-person condition, the larger the contextual influence the body posture had on perception of the facial expressions. When a facial expression was judged with a 100% angry body posture, a greater amount of disgust was required in the facial expression to perceive the face as disgusted due to the influence of the affective body [Figure 3.2]. A paired t-test revealed a significant difference in the PSE for facial expressions presented on 100% angry and 100%

disgusted body postures (t(33)= 7.21, p<0.001), indicating overall there was a significant influence of body posture on facial expression perception. Furthermore, I observed large individual differences in the magnitude of body posture influence on perception of the facial expression across participants [Figure 3.5]. The PSE change for some observers was small, indicating that body posture did not influence the perception of facial expressions much, however for other observers the PSE change was much larger.

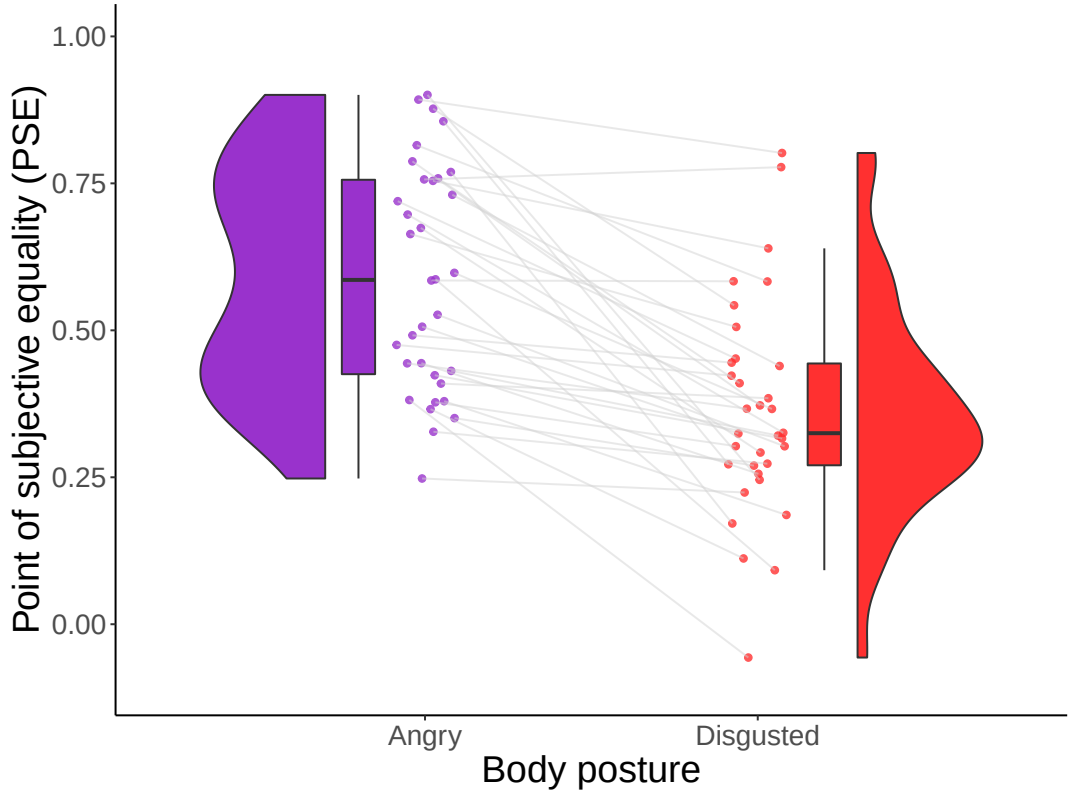


Figure 3.5: **PSE change between categorisation of facial expressions on a 100% angry and a 100% disgusted body posture**
The raincloud plot displays the PSE values for the whole-person condition when facial expressions were presented with a fully-angry or fully-disgusted body posture. A significant difference was observed between the PSE values when the facial expression was categorised with a 100% angry and 100% disgusted body posture (t(33)= 7.21, p<0.001). Each line represents one observer and depicts the change in PSE. The distribution of the values is illustrated by the shaded area, with the boxplot indicating the median and the interquartile range.

A significant negative relationship was found between isolated facial expression precision and the influence of body posture, as indexed with the PSE change ($r_{(s)}$= -0.397, p=0.041)[Figure 3.6]. Observers who had more precise facial expression representations showed less influence of body posture on perception of facial expressions in the whole-

person condition. No relationship between isolated body posture precision and the magnitude of the PSE change was observed ($r_{(s)}$= -0.293, p=0.1).

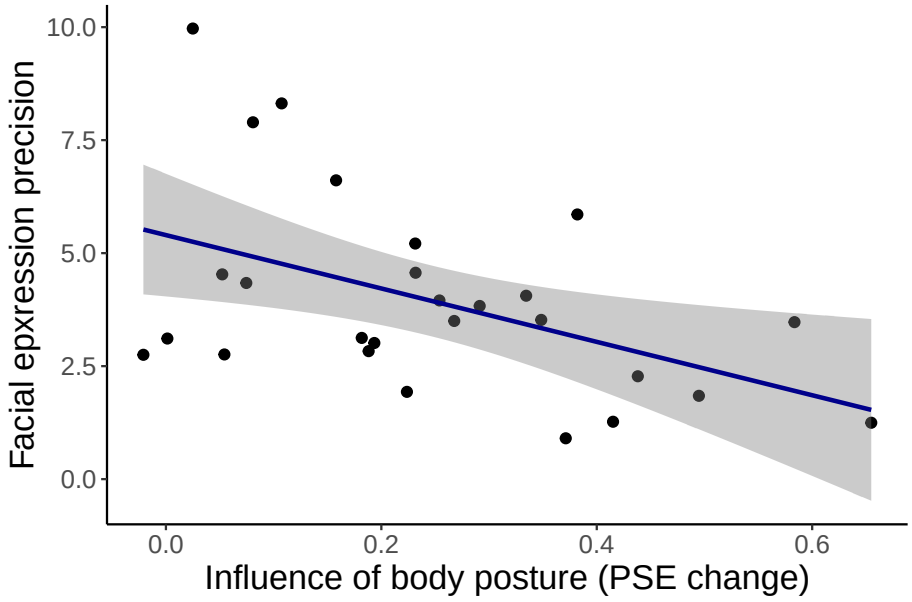


Figure 3.6: **Relationship between isolated facial expression precision and PSE change**
A significant negative correlation was seen between observers facial expression precision and PSE change ($r_{(s)}$= -0.397, p<0.05). Each point represents one observer. The 95 % confidence interval is shown with grey shading.

In order to further understand the relationship between the isolated facial expression and body posture precision with the magnitude of the body biasing effect, a linear regression model was fitted. Model 1 [Equation 3.1] explained the PSE change simply in terms of the facial expression precision. Model 2 [Equation 3.2] included the body posture precision component in the model, as well as facial expression precision.

$$\text{Model 1:} \quad \text{PSE change} \sim \text{Facial expression precision} \tag{3.1}$$

$$\text{Model 2:} \quad \text{PSE change} \sim \text{Facial expression precision} + \text{Body posture precision} \tag{3.2}$$

Comparing the Akaike information criterion (AIC) for the two models (Model 1 = -18.84, Model 2 = -15.33) showed that Model 1 fits the data better, as it has a lower AIC value. As Model 2 has a different number of parameters, I also used Bayesian information criterion (BIC) to confirm the findings. BIC penalises model complexity more heavily than AIC. Model 1 has a lower BIC than model 2 (Model 1 = -14.95, Model 2 = -10.30)

therefore, Model 1 is considered more likely to be the better fit to, or more representative of, the data. This indicates that adding body posture precision to this simple linear model did not improve the prediction of the PSE change.

3.3.3 LEAS

The total LEAS scores ranged from 59 – 85, which is in line with other published work that administers this questionnaire (Lane et al., 1998). I found no relationship between PSE change and total LEAS score ($r_{(s)}$= 0.261, p=0.228). No relationship was observed between LEAS score and isolated facial expression ($r_{(s)}$= 0, p=1) or body posture precision ($r_{(s)}$= 0.127, p=0.518).

3.4 Discussion

The results of this study demonstrate the feasibility of applying psychophysical methods to quantify the influence of body context on facial expression perception. I provide data quantifying body posture precision, as well as individual facial expression precision, and the influence of body posture on facial expression perception in whole-person emotion perception. My results demonstrate that an affective body posture influences the perception of a facial expression; this finding replicates previous published research (Aviezer et al., 2011; Aviezer et al., 2008; Meeren et al., 2005; Teufel et al., 2019). However, in addition, I found individual differences in the magnitude of the influence of body context on the perception of facial expressions. The development of the psychophysical methodology to assess whole-person perception facilitated the quantification of the extent of body posture influence on perception of the facial expression through the PSE change. Some observers' perception of the facial expression was minimally affected by the body context, whilst others' perception was heavily biased by the body posture.

The results of the present study show that the precision of the isolated face cue is related to how much body posture influences the judgement of the facial expression in whole-person perception of emotion. Observers who had more precise isolated facial expression representations, displayed less of an influence of body posture on perception of facial expression in the context of a whole-person. Similarly, observers' who had less precise facial expression representations were more influenced by the affective body context. One potential explanation for this finding is that observers who display poor facial expression recognition may more readily rely on other social cues to support their categorisation of faces in the real world. Therefore, in the whole-person condition these observers' were more biased by body posture.

Other research has shown that the influence of body posture on facial expression perception is reduced after administration of oxytocin (OT) (Perry et al., 2013). Conversely, recognition of facial expressions is known to be improved following administration of OT (Shahrestani et al., 2013). Perry and colleagues (2013) found that the influence of body posture on perception of facial expression was less in a group of adults that had nasal administration of OT compared to a placebo group. Although it is unclear what the underlying mechanism for this is, one possibility is that it could be a consequence of an

overall enhancement in facial expression recognition. This explanation is supported by my current findings, where enhancement of facial expression recognition causes observers to rely less on body posture when perceiving the facial expression in whole-person emotion perception.

Previous research with older adults has demonstrated that they give the body more 'weight' when judging incongruent whole-person displays of emotion (Abo Foul et al., 2018). This finding has been attributed to an optimal social-expertise strategy (Abo Foul et al., 2018). It has been suggested that real world facial expressions are typically ambiguous, and the contextual aid of a body posture provides additional information to support an observer's judgement. Abo Foul (2018) and colleagues suggest, that in older adults, adopting this more holistic approach to emotion perception is an advantageous strategy. My findings in the present study challenge this hypothesis. With normal ageing comes a gradual decline in cognitive and perceptual abilities, and facial expression recognition is no exception (Lott et al., 2005; Sullivan & Ruffman, 2004). In the present study, I found that observers who were most influenced by body posture had poorer facial expression recognition. Therefore, I would expect that the increased reliance on body context in older adults is not evidence of social-expertise but could be a consequence of a decline in perceptual ability of isolated facial expression recognition, resulting in greater reliance on body posture. To further explore this idea between perceptual ability of the individual cues and the contextual influence of the body posture, Chapter 4 explores whole-person recognition of emotion across development in children and adolescents. Children are still developing and refining their ability to perceive facial expressions (Dalrymple et al., 2017). This provides an excellent opportunity to probe individual differences, due to developmental stage, in the relationship between facial expression perception and the contextual influence of body posture in whole-person emotion recognition.

Individual differences provide a unique opportunity to understand the cognitive mechanisms involved in the integration of social information from facial expressions and body postures (Boogert et al., 2018). Typically, differences between observers are treated as measurement error, and this 'noise' is deemed to be a nuisance parameter (Mollon et al., 2017). However, the individual differences reported between observers often arise from real differences in optical, neural and perceptual differences that mediate perception. These

differences can provide a plethora of information about perception and warrant thorough investigation. Reporting individual differences here provides a further opportunity to explore how face and body emotion cues are integrated under a more formal mathematical framework [Chapter 5]. Findings from the simple linear regression model reported in the present chapter indicate the PSE change is best predicted by the model that contained the facial expression precision only. This approach is limited as it does not truly account for the individual differences in perception reported. Adopting a more formal mathematical model to assess how facial expression and body posture cues are combined to form whole-person representations, utilising the naturally existing variability between individuals, motivated subsequent work in Chapter 5 of this thesis.

In addition to perceptual factors, I also captured differences in cognitive factors that could be affecting perception of expressive stimuli in individuals. In contrast to my hypothesis, no relationship was found between PSE change and observers' total LEAS score. Previous work has demonstrated that the higher an individual's emotional awareness, as indexed by a higher total LEAS score, the greater one's capacity to appreciate the complexity of a scenario both in themselves and another (Bajgar et al., 2005; Lane et al., 1998). One possible explanation for the absence of a relationship in my results could be a consequence of the small sample size, therefore any interpretation must be made with caution. Ultimately, however, in the current limited sample the factors captured from the LEAS questionnaire regarding individuals' emotional complexity were not related to observers' perceptual ability in distinguishing social cues.

One limitation of the present study is the high number of exclusions as a consequence of poorly fitting PFs. One reason could be due to the specific facial expressions selected. To probe the mechanism for facial expression and body posture integration under a formal mathematical framework in subsequent research [Chapter 5], obtaining reliable estimates of facial expression perception in isolation is critical. Therefore, to ensure observers reach a certain threshold in performance for future research into cue integration, the research in Chapter 5 utilises different facial expressions identities. Another advantage of incorporating novel stimuli is that it allows the subsequent findings to translate to different facial identities and therefore be generalised further. An alternative explanation for the high level of exclusions could be a due to the sample of participants. A subset of the participants

were recruited from the Undergraduate School of Psychology population and were paid in course credits for their participation. Given the adaptive psychometric methodology adopted in this study, and the large number of trials, it required participants to be highly motivated. The performance of some participants', in comparison to staff that were used for piloting, and even children using the same task and stimuli [Chapter 4], indicated that some of these observers' may not have performed to the best of their ability. To mitigate these potential confounds from the participant pool in future, for my subsequent research in adults [Chapter 5], I targeted recruitment to a more motivated sample who received cash payment.

3.5 Chapter Summary

To summarise, this chapter demonstrates the feasibility of using novel psychophysical methodology to explore facial expression and body posture integration. I found a significant influence of body posture on facial expression judgements; however, I also discovered large individual differences in the magnitude of this effect. Observers with poorer facial expression recognition displayed a greater influence of body posture on facial expression perception when categorising faces in context. In contrast, observers' who showed a more precise representation of facial expressions were influenced less by the affective body posture. These findings motivated the subsequent research in Chapters 4 and 5 of my thesis. These results have important implications for understanding how humans combine facial expression and body posture, and provide the foundation for the subsequent research in my thesis.

Chapter 4

Functionally-defined white matter microstructure: Facial expression and body posture perception in children

4.1 Introduction

The ability to recognise facial expressions develops across childhood. With increasing age, children become more accurate at detecting subtle changes in facial expressions (Herba et al., 2006; Thomas et al., 2007). Detecting emotions from facial expressions is thought to be an essential component of children developing emotional knowledge (Widen & Russell, 2010a). From birth, newborns show a preference for faces over other stimuli (Simion & Giorgio, 2015), however the process of acquiring an 'adult-like' ability to recognise emotion from facial expressions is a protracted developmental process (Gao & Maurer, 2009). Children are believed to start with two broad categories of positive and negative emotions, which during development will differentiate into more complex, mature adult categories of emotion (Widen & Russell, 2010b). Accuracy decoding of the basic expressions has revealed different developmental trajectories for different expressions (Herba et al., 2006; Rodger et al., 2015; Widen & Russell, 2010a, 2010b, 2013). Recognition of happiness and fearfulness is adult-like from early childhood, whilst sadness and surprise show gradual improvements with increasing age (Rodger et al., 2015). Disgust, anger and neutral expressions show the steepest improvement in recognition from early childhood (Rodger et al., 2015). Disgusted facial expressions are the last of the basic expressions to be reliably

recognised by children, with children younger than 8 years of age mislabelling disgusted faces as angry (Widen & Russell, 2010b, 2013).

There is sparse research exploring recognition of emotion from body postures across development (Heck et al., 2018; Ross et al., 2012; Vieillard & Guidetti, 2009). Ross and colleagues (2012) presented dynamic expressive body posture videos to children and found steep improvements in performance up to 8.5 years of age, with slower improvements through late childhood and adolescence. This work highlights that emotion recognition from body posture is not fully mature by adolescence, similar to facial expression recognition maturation (Ross et al., 2012; Vieillard & Guidetti, 2009). Mondloch *et al.* (2013) found that 4-year-olds could categorise fearful or sad body postures above chance, whilst in 6-year-olds, approximately 70% of the participants performed above chance identifying the body postures presented. Importantly, in the same task adults were able to identify the sad and fearful body postures without error (Mondloch et al., 2013).

A limited amount of research has examined the role of body context during facial expression recognition in children (Leitzke & Pollak, 2016; Mondloch, 2012; Mondloch et al., 2013; Nelson & Mondloch, 2017; Rajhans et al., 2016). Mondloch (2012) found when children were asked to categorise facial expressions presented with an incongurent body posture, their categorisation of the face was more affected by the body posture compared to adults (Mondloch, 2012). This suggests that children are more influenced by body context than adults when making facial expression judgements.

In Chapter 3, I found that individual differences in adults' recognition of isolated facial expressions determined the extent to which body posture influenced facial expression perception, such that those who were better at distinguishing facial expressions, were less influenced by body posture, and vice versa. Given the developmental patterns of facial expression recognition outlined above, I wanted to assess if the improvement in recognition could be linked to the influence of body context on perception of facial expressions in children. Given younger children are less reliable at categorising emotions from facial expressions, one would expect the categorisation of facial expression in the context of a whole-person to be heavily biased by the body posture. In contrast, older children and adolescents are thought to be more reliably able to categorise isolated facial expressions, therefore, I would expect the influence of body posture on perception of a facial expression

in older children to be reduced. In addition, I wanted to assess how body posture recognition develops across childhood and adolescence given the limited developmental research exploring how visual perception of body postures develops.

At the cortical level, the development of face-specific and body-specific processing regions in the ventral temporal cortex is a protracted process (Golarai et al., 2009). The proportion of the fusiform cortex with face-selectivity has been shown to be substantially smaller in adolescents than adults (Golarai et al., 2010; Haist & Anzures, 2017). Interestingly the increase in the spatial extent and intensity of neural responses in the FFA has been positively correlated with face recognition memory in late childhood (Golarai et al., 2010), suggesting that development of face recognition ability emerges in parallel with development of face-selective cortical activity. Work in children uncovering the evolving cortical changes in body posture processing is limited; BOLD changes in FBA, EBA and pSTS have been found to be higher in adults compared to children when viewing dynamic body stimuli (Ross et al., 2019). Therefore, the evolving cortical response to bodies across development appears to draw some parallels with the neural development of face processing in the brain.

Accepted models of face and body processing, specifically for facial expression and body postures, illustrate the importance of a network of cortical areas for processing these cues (Duchaine & Yovel, 2015; de Gelder, 2006; Haxby et al., 2000). Studying the white matter connectivity between cortical regions is particularly important in children, as we know there is a prolonged period of development. Myelination continues into adolescence and alters the architecture of white matter tracts (Lebel & Beaulieu, 2011). Changes have been observed in both volume and microstructural properties of white-matter with age (Mills et al., 2016; Paus, 2010). Across development the fatty myelin sheath forms around axons, increasing conduction velocity and refining the synchrony of electrical impulses (Buzsaki, 2004; Fields, 2008). Whilst this myelination is initially accelerated in early childhood, coinciding with the advent of cognitive skills and abilities, it continues until 20 to 30 years of age, and undergoes continuous refinements throughout the lifespan (Bartzokis et al., 2012; Yakovlev & Lecours, 1967).

Ventral face-selective regions are highly interconnected to one another through direct white matter connections between the cortical regions (Grill-Spector et al., 2017; Gschwind

et al., 2012). These connections act as the building blocks for transmitting signals to different cortical nodes in the visual hierarchy which are responsible for extracting distinct features of faces. Not only do these connections reflect a key information transfer pathway, the microstructural properties of white matter underpinning face processing have been linked to face recognition ability. Tavor et al. (2014) reported that face recognition ability was highly associated with the fractional anisotropy (FA) of the anterior part of the inferior longitudinal fasiculus (ILF) in the right hemisphere of adults (Tavor et al., 2014). The ILF is one of the main association bundles projecting through the occipito-temporal cortex, which connects multiple face perception nodes (Wang, 2018). Specific to emotion, the FA of the inferior frontal occipital fasiculus (IFOF) has been shown to significantly predict facial expression recognition ability in adults (Unger et al., 2016). The microstructure of the uncinate fasciculus (UF) has also been linked to adults ability to decode emotions from facial expressions (Coad et al., 2017). This clearly highlights that inter-individual differences in perceptual ability can be reflected in microstructural properties of associative white matter tracts.

Neurodevelopmental research exploring white matter microstructure and face processing is lacking. However, work has shown that the FA and mean diffusivity (MD) of the ILF in a cohort of 6 to 23-year-olds was tightly coupled with age-related increases in the size of functional face selective regions (Scherf et al., 2014). This work highlights the importance between the structural-functional coupling of the brain regions underpinning face perception, and illustrates the need for further research linking white matter to perception. However, linking the microstructure of large associative anatomical tracts to behaviour is inherently limited due to the lack of specificity that can be deduced between cognition and large white matter bundles. For example, while the FA of ILF has been found to be correlated with face recognition abilities (Tavor et al., 2014), ILF microstructure has also been linked to broader social cognition skills and reading ability (Ortibus et al., 2012; Yeatman et al., 2014). This is unsurprising given the pathway traverses many functional domains from occipital cortex to the temporal pole. In order to make specific inferences about the role white-matter microstructure plays in specific aspects of perception, exploration of the structural connectivity between functionally connected regions of cortex is required. Gomez *et al.* (2015) combined measurements of white-matter

connectivity, behaviour and functional selectivity, within subjects, to reveal the interplay between the structure-function relationships of the brain and visual behaviour. By investigating functionally-defined white matter (FDWM) tracts local to the fusiform face region, Gomez *et al.* (2015) found that adults accuracy on a face recognition task was positively correlated with the local FA of the right hemisphere FDWM tract (Gomez et al., 2015), demonstrating the usefulness of a targeted approach for identifying functionally-relevant tracts.

In this chapter, I specifically focus on the white-matter connections between cortical nodes involved in face and body processing. Based on models of visual processing for facial expression and body posture, and the expected site of integration of the face and body cues, this work explores connectivity between functionally-defined category specific ventral visual regions (OFA, EBA, FFA, FBA), posterior STS (pSTS) and the anterior temporal lobe (ATL). Research exploring the influence of body posture on facial expression perception has alluded to the site for integration of face and body cues being late in the visual processing hierarchy, and anterior to the core face processing regions in the temporal lobe (Fisher & Freiwald, 2015; Harry et al., 2016; Teufel et al., 2019). Given the limited research exploring body posture processing in the brain, inferences from macaques suggest that the regions processing faces and bodies in the early inferotemporal cortex are largely independent parallel networks (Premereur et al., 2016). To this end, I independently localised the OFA, FFA and pSTS-face in the face network and EBA, FBA and pSTS-body in the body processing network to explore white matter connectivity between these functional nodes. Measures of microstructure to quantify differences in underlying tissue properties were selected to be reflective of different cellular features: fractional anistropy (FA), mean diffusivity (MD) and spherical mean at high b-value ($\mathring{S}_{\mu}(b = 6000)$). FA is a highly sensitive but non-specific marker of anisotropy within the microstructural environment. In complement to FA, MD provides a measure of the average diffusivity within a voxel (Alexander et al., 2007). In addition, I utilised ultra-strong diffusion gradients in this study to assess diffusion-weighted signal at high b-values (i.e., b=6000) (Mirzaalian et al., 2016). Extracellular signal is suppressed at high b-value, resulting in a more feature-specific marker of axons and/or glial processes in the microstructure (Hagmann et al., 2010; McNab et al., 2013; Raven et al., 2020; Veraart et al., 2019).

To summarise, in the current study my aims were twofold; firstly, to examine individual differences in facial expression processing in context as a function of development, and secondly, to link changes in face and body perception of emotion to microstructural differences in the white matter of the ventral visual system. To address the behavioural question, I indexed the precision with which observers were able to categorise individual facial expressions and body posture cues, and related those to the body biasing effect observed, similar to Chapter 3. In order to specifically target changes in the brain networks processing face and body cues, I adopted a functionally-defined approach to extract white matter pathways connecting cortical nodes known to process faces and bodies. In turn, I then related the microstructural metrics of these tracts to individual differences in children's perceptual performance.

4.2 Methods

The study described herein was a subset of a larger neurodevelopmental project at Cardiff University Brain Research Imaging Centre (CUBRIC), assessing the feasibility of scanning children with high-gradient and ultra-high-field MRI scanners. In addition to the methods described below, additional diffusion and functional scans were acquired at 3T and 7T. Children visited CUBRIC on two separate days; on the first visit children were given the opportunity to acclimatise to the scanner environment in a mock MRI scanner (de Bie et al., 2010). Following the practice, they were scanned on a 3T Siemens Connectom system with 300 mT/m gradients. On the second visit children were scanned on a 7T Siemens Magnetom and took part in cognitive testing outside the scanner. The average time between the two visits was 20 days (SD $\pm$ 35 days).

4.2.1 Observers

A total of 45 typically-developing children (22 female) between 8 to 18 years of age (mean age = 12.96 $\pm$ 3.1) were recruited to take part in the present study. Children were recruited from the local Cardiff community through schools and public outreach activities. One child did not complete the 7T MRI component of the study. All children had normal or corrected to normal vision and were screened to exclude major neurological disorders. All children underwent an IQ assessment using the 2-subtest Weschler Abbreviated Scale Intelligence and had IQ values above 86 (Mean score $\pm$ SD = 107 $\pm$ 14.26, range = 86-145) (Wechsler, 1999). Pubertal stage was determined using parental report on the pubertal development scale (PDS) (Petersen et al., 1988).

Primary caregivers of children provided written informed consent prior to participating. Children aged 16 and older also provided their own written consent. Experimental protocols were approved by Cardiff University School of Psychology Ethics Committee and were in line with the Declaration of Helsinki. All participants were fully debriefed at the end of the testing and children were thanked for their participation with Amazon vouchers and a certificate of participation which included a picture of their own brain.

4.2.2 Cognitive testing

Stimuli

Children completed the same psychophysical behavioural task as detailed in Section 3.2.2 of Chapter 3. The stimuli used were identical to those previously described, and were morphed facial expressions, body postures and facial expressions with 100% angry and 100% disgusted body postures.

Procedure

Presentation of the psychophysical task was in in MATLAB [Version 2018b, The MathWorks, Natwick, MI, USA] with the Psychophysics (Version 3.0.15) Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). The adaptive psychophysical tasks procedure is as detailed in Section 3.2.3. Children were asked to categorise facial expressions morphed between anger and disgust, body postures morphed between angry and disgusted body postures, and morphed facial expressions presented in the context of a whole-person with a fully-angry or fully-disgusted body posture. In a subset of 8 to 10-year-old participants the experimenter pressed the key to respond after the child indicated their response verbally.

4.2.3 MRI acquisition

fMRI acquisition

Whole-brain echo-planar imaging (EPI) gradient echo data was acquired on a 7T Siemens Magnetom (Siemens Healthcare, Erlangen, Germany) (TR/TE=2000/30ms; resolution= 1.5x1.5x1.5mm; 87 slices; multi-band factor=3; TA=9min 18sec). Acquisition was angled along the anterior commissure and posterior commissure to minimise signal drop out from the temporal regions. In addition, a B0 field map (TR/TE 1/TE 2=560/5.1/6.12ms; resolution=3x3x3mm; 44 slices; TA=1min 7sec) was acquired to unwarp the high-field EPI data. A high-resolution MP2RAGE structural scan (TR/TE=6000/2.7ms; resolution =0.65x0.65x0.65mm; TA=10min 46secs) was also acquired (Marques et al., 2010). A SA2RAGE B1 map (TR/TE=2400/0.72ms; resolution=3.25x3.25x6mm; TA=1min 26sec) was acquired to bias-field correct the MP2RAGE (Eggenschwiler et al., 2012). Retrospective correction of head movement was used during image acquisition of the MP2RAGE to

mitigate blurring of the high-resolution data at 7T (Gallichan et al., 2016).

Functional localiser task

For the functional localiser task, participants were presented with grey-scale images of faces, houses, bodies, and chairs projected onto a mirror mounted on the MRI head coil. Both male and female faces and bodies were presented. All expressions were neutral. Images were shown in blocks by stimulus category, and there were two orders of the task which were counterbalanced across participants. Within each block 15 images were displayed; each being presented for 800ms with a 200ms inter-stimulus interval. A block of fixation followed the stimulus category block, and the participant was asked to focus on the fixation cross displayed centrally on a mean grey screen (15s). For the last second of the fixation block, the fixation cross turned red to indicate that the next block of images would commence. In total there were four blocks of each stimulus category. Therefore, for each stimulus category, 64 trials per condition were presented overall. The participants were instructed to respond using a key press if the same image was presented twice in succession (1-back task). The number of repeated trials per block varied between 0 and 3. The average accuracy on the 1-back task was 90% (SD $\pm$ 0.18). Button responses were recorded using a right-hand MR compatible button box. All participants underwent a short practice of the task before entering the 7T MRI.

dMRI acquisition

Diffusion MRI (dMRI) data was acquired on a 3T Connectom scanner (Siemens Healthcare, Erlangen, Germany) with 300mT/m gradients and a 32-channel radiofrequency coil (Nova Medical, Wilmington, MA, USA). dMRI data were acquired using a multi-shell diffusion-weighted EPI sequence with an anterior-to-posterior phase encoding direction (TR/TE=2600/48ms; resolution=2x2x2mm; 66 slices; b-values=0 (14 vols), 500; 1200 (30dirs), and 2400; 4000; 6000 (60dirs)s/mm^2; TA=16min 14sec). One additional volume was acquired with a posterior-to-anterior phase encoding direction for the purpose of EPI distortion correction. T1-weighted anatomical images were acquired using a 3D Magnetization Prepared Rapid Gradient Echo (MP-RAGE) sequence (TR/TE=2300/2ms; resolution=1x1x1mm; 192 slices; TA=5min 32sec).

4.2.4 Analysis

fMRI pre-processing

fMRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) in FMRIB (Functional Magnetic Resonance Imaging of the Brain) Software Library (FSL) (Jenkinson et al., 2012). High-field functional data was unwarped using the B0 field map generated from a brain extracted magnitude and phase image in FSL (Woolrich et al., 2001). Registration to high-resolution MP2RAGE and Montreal Neurological Institute (MNI) space was carried out using FMRIB's Linear Image Registration Tool (FLIRT) (Jenkinson et al., 2002; Jenkinson & Smith, 2001). Functional data was also registered to pre-processed (skull stripped, bias corrected, neck cropped) $0.65mm^3$ MP2RAGE structural images, that were registered to each subject's DWI from the 3.0T Connectom scanner using Advanced Normalization Tools (ANTs) (Avants et al., 2011). Two children were excluded due to failed registration. The registered MP2RAGE images were re-sampled to $1mm^3$. This registration step translated the functional data into subject specific dMRI space for subsequent tractography. Functional data was registered to MNI $1mm^3$ brain template to validate the neuroanatomical regions of interest activated in the fMRI task. Motion correction was performed with FSL's Motion Correction using FMRIB's Linear Image Registration Tool (MCFLIRT); one participant was removed from the analysis due to excessive motion during the fMRI task. This was identified as movement of more than 2 voxels (3mm) over the scan. The average absolute motion for all participants was 0.74mm $\pm$ 0.7. To mitigate the effects of motion in the functional imaging analysis, the estimated motion traces from MCFLIRT were added to the general linear model (GLM) as nuisance regressors (Jenkinson et al., 2002). Skull-stripping and removal of non-brain tissue was completed using Brain Extraction Tool (BET) (Smith, 2002). All data was high-pass temporal filtered with a gaussian-weighted least-squares straight line fitting (sigma=50.0s). On an individual subject basis spatial smoothing with a Gaussian smoothing kernel with a Full Width at Half Maximum (FWHM) 4mm was applied to preserve high spatial resolution (Woolrich et al., 2004). For second-level group analysis of the functional data, spatial smoothing with a Gaussian smoothing kernel with a FWHM 10mm was applied (Woolrich et al., 2004). Time-series statistical analysis was carried out using FMRIB's Improved

Linear Model (FILM) with local autocorrelation correction (Woolrich et al., 2001).

fMRI data analysis

A univariate GLM was implemented to examine the BOLD response associated with face and body stimuli. Faces > Houses and Bodies > Chairs contrasts were used to localise cortical regions involved in face and body processing respectively. Region of interests (ROIs) involved in face processing (FFA, OFA and pSTS-face) and body processing (FBA, EBA, pSTS-body) were identified on a subject-by-subject basis in subject space. Z-statistic images were uncorrected and thresholded at p=0.1. The number of subjects where the ROIs could be reliably identified is listed in the appendix [Appendix: Table 1]. The ROIs were identified within-subjects to facilitate generation of tracts connecting functionally defined nodes. As a validation step, the individual ROIs were translated into MNI space and an average coordinate was determined for each ROI [Table 4.1]. The average coordinates were comparable to other studies where these regions have been extracted in MNI space [Appendix: Table 2] (Bona et al., 2015; Harry et al., 2016; Schobert et al., 2018; Spiridon et al., 2006; Taylor et al., 2007; Vocks et al., 2010). The Euclidean distance was calculated between each ROI and the average coordinates in MNI space for every child to provide a measure of how variable the locations were across participants. ROIs were inflated to 10mm in diameter into surrounding WM to be used for functionally-defined white matter tractography. In addition to these functionally-defined ROIs an anatomical ROI in the right ATL was manually drawn based on anatomical landmarks because activation in this region was not reliably identified in subject-specific or group-level analyses. A coronal plane was drawn in the right temporal lobe extending from the lateral fissure to the ventral surface of the brain (Hodgetts et al., 2015). The anterior-posterior location along the temporal lobe was determined by positioning the plane just anterior of the position where the central sulcus meets the lateral fissure.

ROI	x	y	z	σ distance
FFA	38	-43	-20	7.31
OFA	38	-76	-7	11.35
pSTS-face	46	-57	12	6.07
FBA	38	-43	-18	8.08
EBA	41	-78	0	7.65
pSTS-body	44	-59	11	8.98

Table 4.1: **Average coordinates for ROIs**
The coordinates reported are the average coordinates for the individual ROIs that were registered to MNI space, allowing the average position to be compared to coordinates reported in the literature. The Euclidean distance was calculated between each ROI and the average coordinate for each observer. The standard deviation (σ) is reported to provide a measure of how variable the ROIs were across observers.

Second-level analysis was carried out using a fixed-effects model in FMRIB's Local Analysis of Mixed Effects (FLAME) (Woolrich et al., 2004). Z-statistic images, as defined from the contrasts of interest, were thresholded using a cluster defining threshold of $Z>2.3$, $p<0.01$, and a corrected cluster significance threshold of $p<0.05$ (Worsley, 2001). The FFA and pSTS-face were identified at the corrected cluster threshold. For clusters not detectable at the corrected cluster significance of $p<0.05$, uncorrected voxel coordinates are reported [Table 4.2]. BOLD activation in the left hemisphere did not robustly activate the cortical ROIs so the subsequent analysis focused on the right-hemisphere.

ROI	x	y	z
FFA	35	-49	-18
OFA*	38	-63	-9
pSTS-face	48	-44	8
FBA*	36	-43	-21
EBA*	44	-80	8
pSTS-body	43	-58	5

Table 4.2: **Group-level peak cluster coordinates for functionally-defined ROIs**
The coordinates are reported in MNI space for the ROIs. The clusters were defined with a cluster defining threshold of $Z>2.3$, $p<0.01$, and the alpha threshold of $p<0.05$. The asterisk denotes the clusters detected at the uncorrected alpha threshold level.

DWI pre-processing

DWI data quality assurance was performed on the raw diffusion volumes using slicewise outlier detection (SOLID) to control for motion (Sairanen et al., 2018). DWI data were pre-processed to reduce thermal noise and image artefacts which included image denoising (Veraart et al., 2016), correction for signal drift (Vos et al., 2017), motion, eddy current, and susceptibility-induced distortion correction (Andersson & Sotiropoulos, 2016), gradient non-linearities, and Gibbs ringing (Kellner et al., 2016). The pre-processing pipeline was implemented in MATLAB, but depended on open-source software packages from MRtrix (Tournier et al., 2019) and FSL (Jenkinson et al., 2012).

Tractography

Multi-shell multi-tissue constrained spherical deconvolution (Jeurissen et al., 2014) was applied to the pre-processed images to obtain voxel-wise estimates of fODFs (Descoteaux et al., 2009; Seunarine & Alexander, 2014; Tournier et al., 2007; Tournier et al., 2004) with maximal spherical harmonics order lmax = 8. FDWM tracts were generated between ROIs within face (OFA to FFA, FFA to ATL, pSTS-face to ATL) [Figure 4.1] and body networks (EBA to FBA, FBA to ATL and pSTS-body to ATL) [Figure 4.2] in each subject. Streamlines were generated using a probabilistic algorithm in MRtrix using one ROI as seeding mask and the second as an inclusion region, following the organisation of the visual processing hierarchy. All generated FDWM tracts were visually inspected, and spurious fibres manually removed. Any tracts with less than 20 streamlines were removed from the subsequent analysis (FFA-ATL: n=1, pSTS-face-ATL: n=1, FBA-ATL: n=1, EBA-FBA: n=1). Fractional anisotropy (FA), mean diffusivity (MD) and spherical mean ($\mathring{S}_{\mu}(b = 6000)$) metrics were extracted from each tract by averaging over streamlines, resulting in one metric per tract in each child. FA and MD were derived from diffusion kurtosis imaging using three shells (b=500, 1200, 2400) to improve the accuracy of the diffusion tensor metrics (Jensen et al., 2005; Veraart et al., 2011). A unique feature of the present study is the sensitivity to axonal morphology at b=6000 s/mm^2, achieved by exploiting the ultra-strong gradients, of the Connectom scanner. The spherical mean of the dMRI signal was computed with high b-value data (Mirzaalian et al., 2016; Veraart et al., 2019).

In addition to the functionally defined white matter tracts generated, several anatomically defined tracts were extracted in the right hemisphere: ILF, IFOF, UF and corticospinal tract (CST). TractSeg segmentation software (Wasserthal et al., 2019) was used to automatically extract the fibre tracts using the b=6000 s/mm^2 shell for each participant. FA, MD and $\mathring{S}_\mu(b = 6000)$ was determined for each of the anatomically defined tracts.

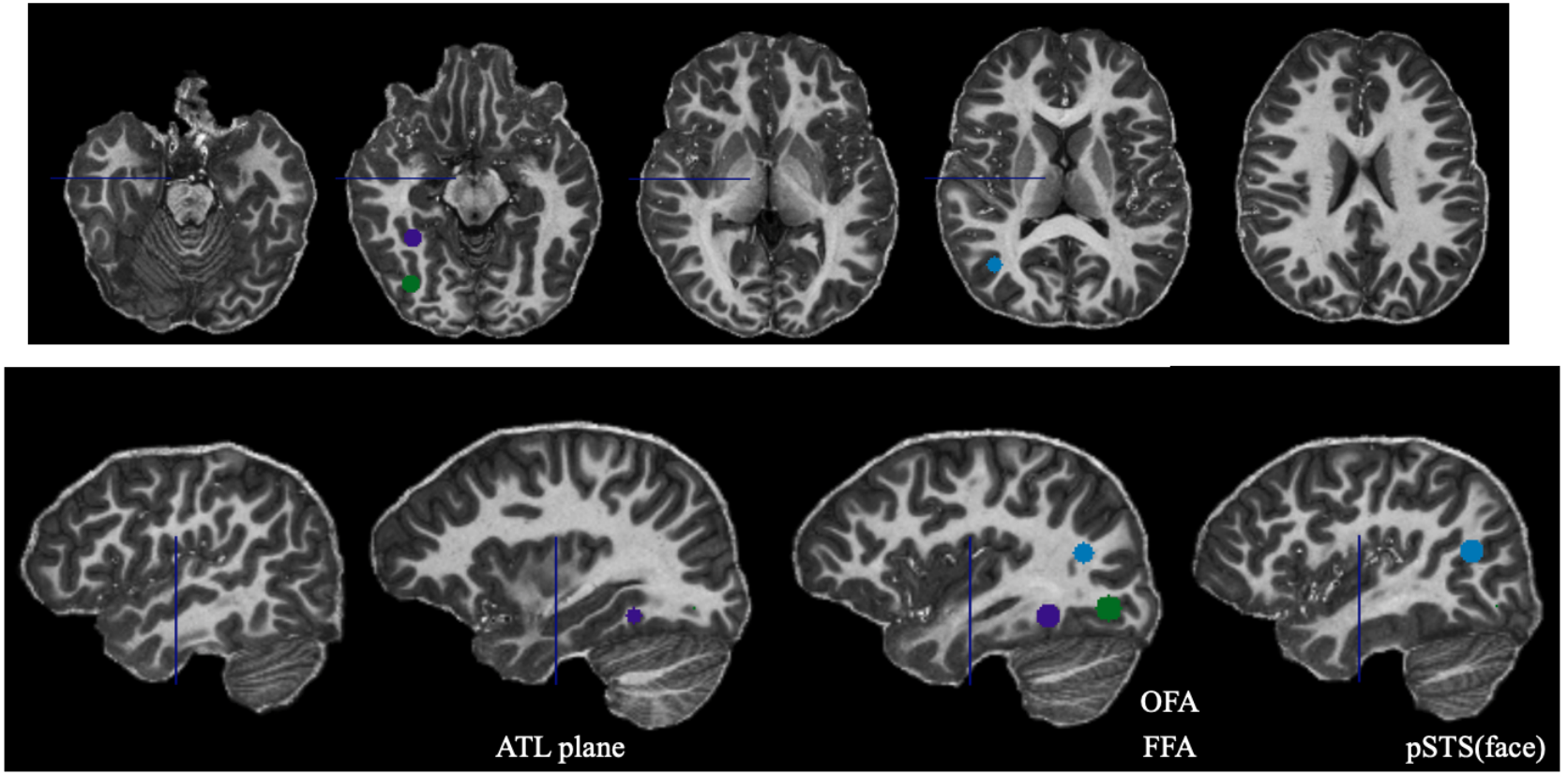


Figure 4.1: **ROIs in face processing network: OFA, FFA, pSTS(face) and ATL**
This figure shows the regions in the face network that were used as seeds and/or inclusion regions for tractography in one child's native space. The OFA is shown in green, FFA in purple, pSTS(face) in light-blue and ATL plane in dark blue. The ROIs shown are 5mm in diameter for illustrative purposes and located in the right hemisphere. The top row is an axial cross section moving from inferior to superior slices, from left to right of the figure. The bottom row illustrates a sagittal view moving from lateral to medial slices from left to right of the figure.

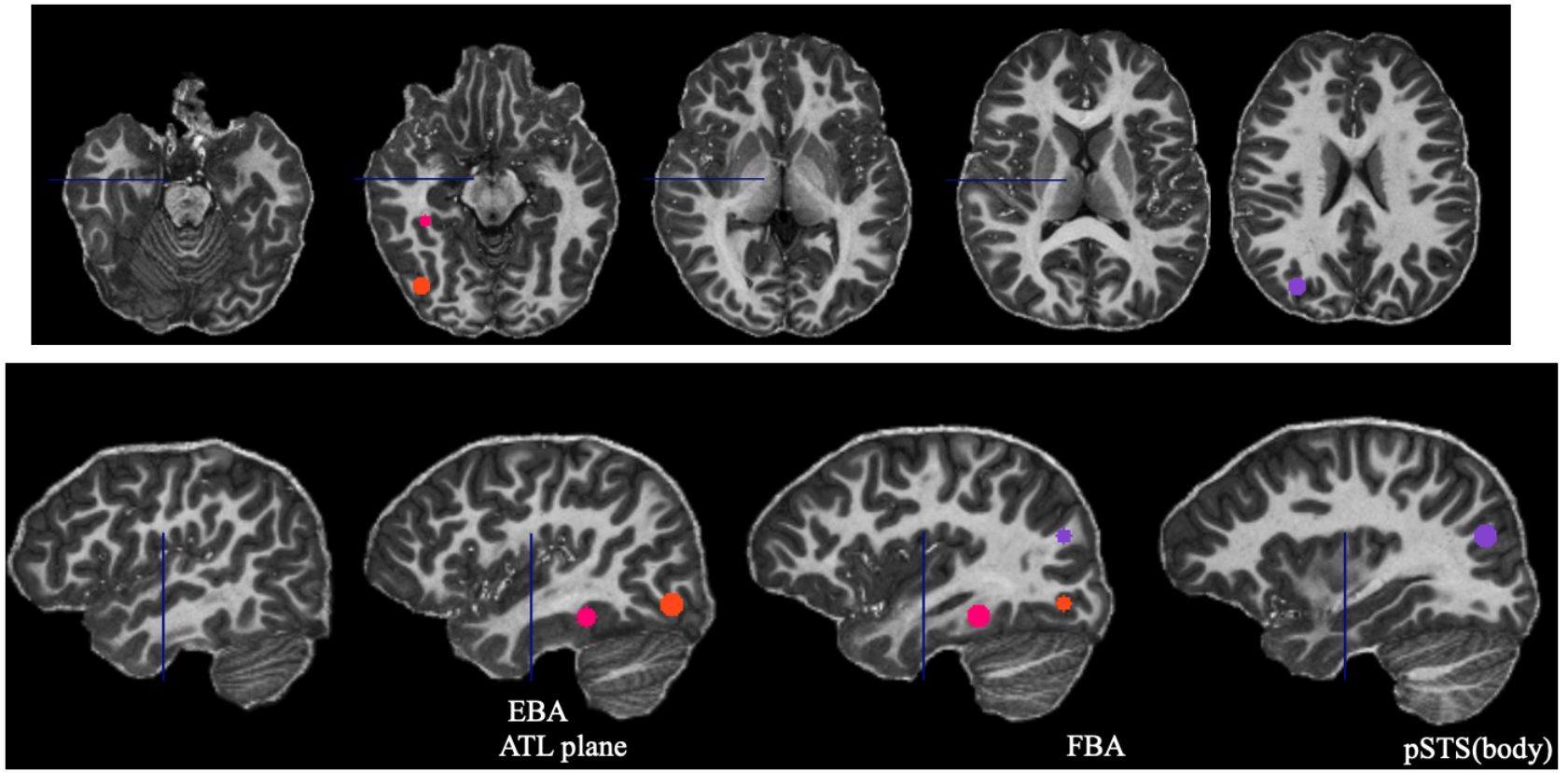


Figure 4.2: **ROIs in body processing network: EBA, FBA, pSTS(body) and ATL**
This figure shows the regions in the face network that were used as seeds and/or inclusion regions for tractography in one child's native space. The EBA is shown in orange, FBA in pink, pSTS(body) in purple and ATL plane in dark blue. The ROIs shown are 5mm in diameter for illustrative purposes and located in the right hemisphere. The top row is an axial cross section moving from inferior to superior slices, from left to right of the figure. The bottom row illustrates a sagittal view moving from lateral to medial slices from left to right of the figure.

Statistical analysis

Psychophysical analysis

Using custom-written MATLAB code with the Palamedes toolbox (Prins & Kingdom, 2018), PFs were fitted based on a cumulative Gaussian to estimate a PSE and slope value for each observer, for each condition. Lapse rate was fixed at 0.03; guess rate was determined by the experimental procedure and was fixed at 0. The steeper the slope of the psychometric function, the more reliably the observer distinguished between the morphed stimuli (Kingdom & Prins, 2010). The PSE is the point at which the observer was equally likely to respond disgust or anger to a particular stimulus. The PSE change reflected the modulation of the facial expression judgement due to the contextual influence of body posture. Goodness-of-fit of the PFs was assessed as detailed in Section 3.2.4. A total of 12 children were excluded from the facial expression condition, 6 children from the body posture condition, and 17 children from the whole-person condition due to poor fitting of the PFs.

Pubertal Development

Puberty stage was determined by calculating a combined PDS-Shirtcliff (PDSS) score (Shirtcliff et al., 2009). A tight positive correlation between age and PDSS score was found ($r(46)=0.85$, $p<0.0001$) [Appendix: Figure 1]. By reporting such a high correlation between age and pubertal development in the current cohort, adjusting subsequent statistical analysis to account for age captured a significant proportion of developmental variability.

FDWM with behaviour

To assess the relationship between behavioural metrics and FDWM tracts, a series of Spearman correlations were performed as data was non-normally distributed. Bonferroni correction was used to control for multiple comparisons of the six FDWM tracts with a corrected value of $p<0.008$ (0.05/6).

Subsequently, the relationship with age between tracts that showed a significant correlation with cognitive performance was disentangled using multiple linear regression. For the multiple regressions the age variable was mean centred. Given this stage was a secondary analysis the p-value accepted as significant was $p<0.05$.

4.3 Results

4.3.1 Behavioural Results

Isolated facial expression and body posture categorisation

The slope values of the estimated PFs provided a measure of how accurately an observer distinguished between the morphed stimuli; the steeper the slope, the more precise the individual's performance. Performing a Wilcox signed rank test revealed the estimated face and body slopes were significantly different ($z=0.861$, $p<0.001$), indicating that children were much better at categorising the body postures than the isolated facial expressions [Figure 4.3]. A positive relationship was found between the estimated slopes of the facial expression and the body posture conditions ($r_{(s)}=0.280$, $p=0.14$), where children who were better at categorising facial expressions showed better body posture categorisation, however, this relationship was not significant.

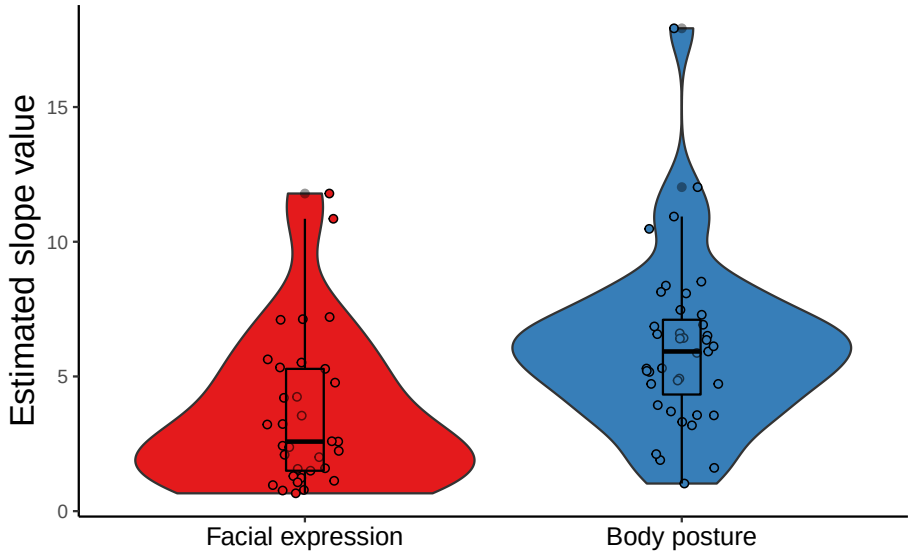


Figure 4.3: **Estimated slope values from PFs of facial expression and body posture categorisation**
The violin plot displays the estimated slope values for each observer in the facial expression and body posture conditions. There was a significant difference between the estimated slope values of the facial expression and body posture conditions ($z=0.861$, $p<0.001$). Each point represents one observer. The distribution of the values is illustrated by the shaded area, with the overlaid boxplot indicating the median and the interquartile range.

Whole-person categorisation

Body posture was found to have a significant influence on perception of facial expressions in children. The greater the difference in PSE between psychometric functions in the whole-person condition, the more influence the body posture had on the perception of the facial expression in that individual. A paired t-test revealed a significant difference in the PSE between the facial expressions presented on 100% angry and 100% disgusted body postures ($t(27)= 5.737$, $p<0.0001$), suggesting that overall there was a significant influence of body posture on facial expression perception in children [Figure 4.4]. Furthermore, I also observed large individual differences in how much body posture influenced the perception of the facial expression across observers, similar to my findings in Chapter 3. The PSE change for some observers was small, indicating that body posture did not influence the perception of facial expressions much, however for other observers the PSE change was much larger.

Similar to findings in adults [Chapter 3], I found a negative relationship between PSE change and facial expression precision, however this did not reach significance ($r_{(s)}= -0.322$, $p=0.109$). Controlling for age in this analysis, still did not result in a significant relationship ($r_{(s)}=-0.322$, $p=0.124$). No relationship was observed between PSE change and body posture precision ($r_{(s)}=-0.126$, $p=0.546$), nor when age was controlled for ($r_{(s)}=-0.126$, $p=0.566$).

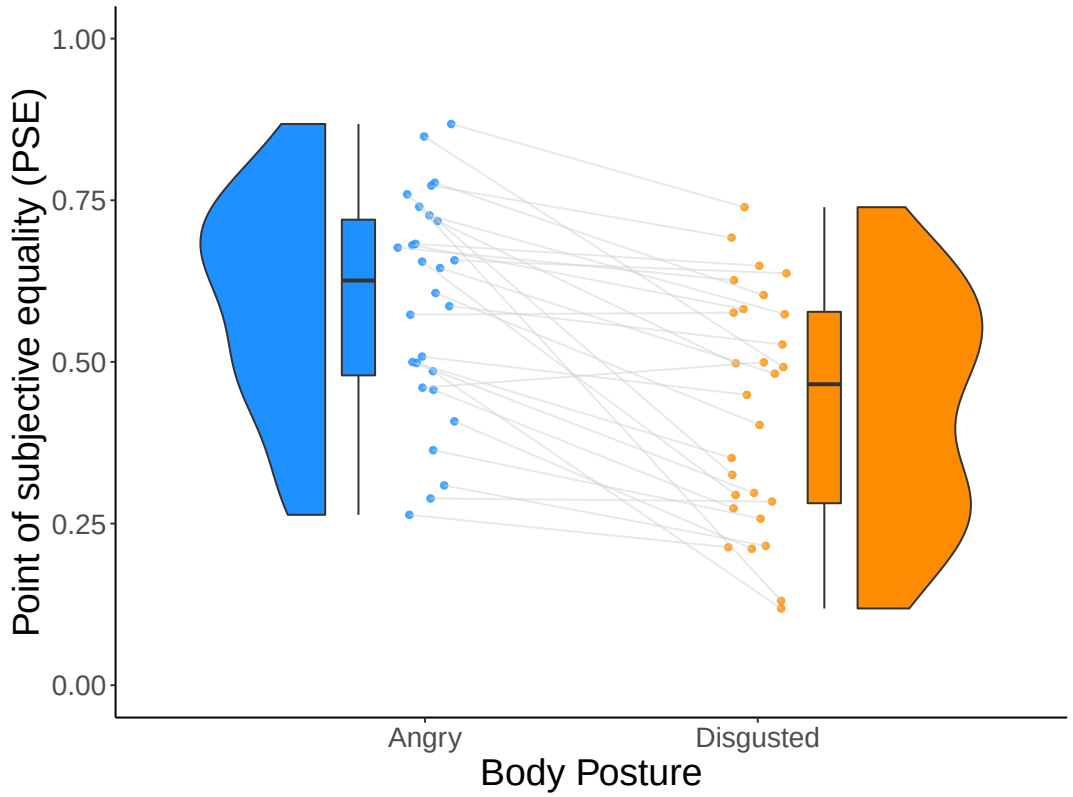
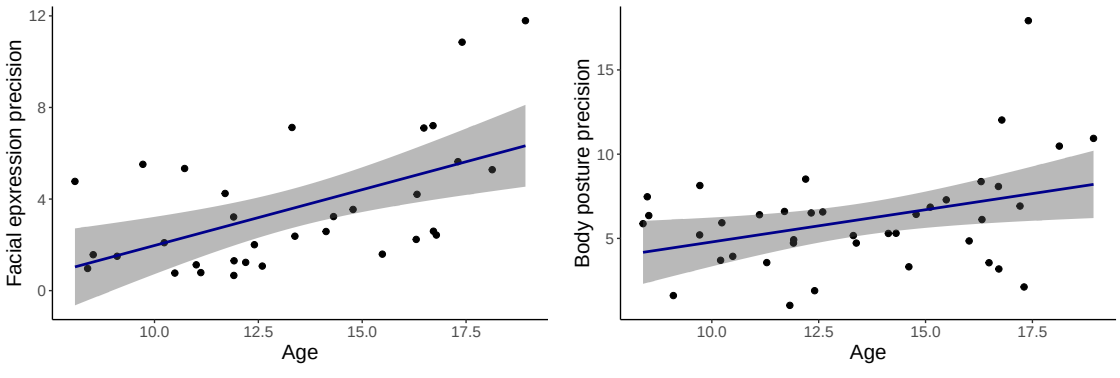


Figure 4.4: **PSE change between categorisation of facial expressions on a 100% angry and a 100% disgusted body posture**
The raincloud plot displays the PSE values for the whole-person condition when facial expressions were presented with a fully-angry or fully-disgusted body posture. A significant difference was observed between the PSE values when the facial expression was categorised with a 100% angry and 100% disgusted body posture ($t(27)= 5.737$, $p<0.0001$). Each line represents one observer and depicts the change in PSE. The distribution of the values is illustrated by the shaded area, with the boxplot indicating the median and the interquartile range.

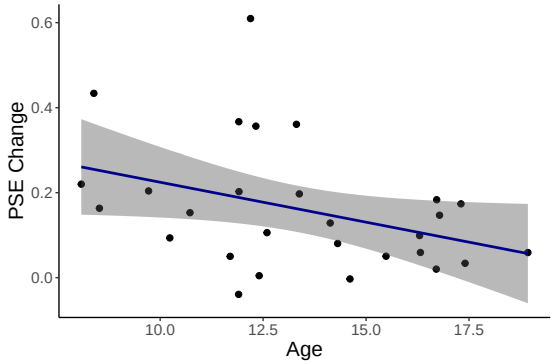
Age related changes in performance

A significant positive relationship was found between age and facial expression precision ($r_{(s)}= 0.513$, $p<0.01$) [Figure 4.5a], suggesting that isolated facial expression discrimination improved with age. Similarly, a borderline significant positive correlation was found between body posture and age ($r_{(s)}= 0.286$, $p=0.072$), again suggesting that isolated body posture discrimination improved with age [Figure 4.5b]. In contrast, I found a significant negative relationship between the magnitude of the PSE change and age ($r_{(s)}= -0.3923$ $p=0.0395$) [Figure 4.5c]. This suggests that younger children were more influenced by body posture in their categorisation of facial expressions.



(a) Facial expression (*)

(b) Body posture



(c) PSE change (*)

Figure 4.5: **Relationship between age and facial expression precision, body posture precision and PSE change**
The correlation between age and facial expression [4.5a], body posture [4.5b] and PSE change [4.5c] are shown in this figure. Significant relationships ($p<0.05$) are indicated with an asterisk. Age is plotted in years. The precision for each condition was indexed by the slope estimate of the individual's PF for each condition. The PSE change was calculated from the different PSE's in the whole-person condition when a face morph was presented with either a fully-angry or fully-disgusted facial expression. Each point represents one observer. The 95 % confidence interval is shown with grey shading.

4.3.2 Functionally-defined white matter

Age-related changes in FDWM

There were no significant relationships between FA and any of the FDWM tracts with age [Figure 4.6]. MD had a significant negative correlation with age in tracts between rEBA-rFBA (r(33)=-0.44, p=0.0079), pSTSbody-ATL r(36)=-0.47, p=0.0026) and pSTSface-ATL (r(37)=-0.460, p=0.0032). Moving beyond the tensor measures of microstructure, the $\mathring{S}_{\mu}(b = 6000)$ showed significant positive relationships with age in five of the six functionally-defined tracts (FBA-ATL: r(36)=0.423, p=0.0078; pSTSbody-ATL: r(36)=0.635, p<0.001; FFA-ATL: r(39)=0.457, p=0.003; OFA-FFA: r(33)=0.473, p=0.004; pSTSface-ATL: r(37)=0.555, p<0.001). Similar to work already published on this cohort, where spherical mean was found to be a sensitive marker of age-related changes in microstructure of large anatomically-defined tracts (Raven et al., 2020), here I demonstrate this is also the case in targeted functionally-defined white matter tracts.

Perceptual abilities related to FDWM

To look at the relationship between microstructural changes and perceptual abilities, behavioural metrics (facial expression precision, body posture precision and PSE change) were correlated with FA, MD and $\mathring{S}_{\mu}(b = 6000)$ of the extracted FDWM tracts from the face and body network [Figures 4.7, 4.8, 4.9]. A significant positive correlation was found between facial expression precision and FA of the OFA-FFA tract ($r_{(s)}$=0.584, p=0.005) [Figure 4.7]. Furthermore, body posture precision and $\mathring{S}_{\mu}(b = 6000)$ of OFA-FFA were found to be significantly positively correlated ($r_{(s)}$=0.623, p<0.001) [Figure 4.8]. Finally, a significant negative correlation was found between $\mathring{S}_{\mu}(b = 6000)$ pSTSbody-ATL and PSE change ($r_{(s)}$=-0.593, p=0.003) [Figure 4.9].

As a validation measure, I related the perceptual performance for facial expression precision, body posture precision and PSE change to several large anatomical tracts that traverse the ventral visual stream (ILF [Appendix: Figure 2]; IFOF [Appendix: Figure 3]; UF [Appendix: Figure 4]) and the CST [Appendix: Figure 5] as an additional control tract outside the visual processing stream. No significant relationships were reported between the microstructural metrics of these large tracts and any of my behavioural measures (all

p>0.09) [Appendix: Figures 6, 7 and 8], demonstrating the specificity of FDWM tract metrics to identify individual differences related to perception.

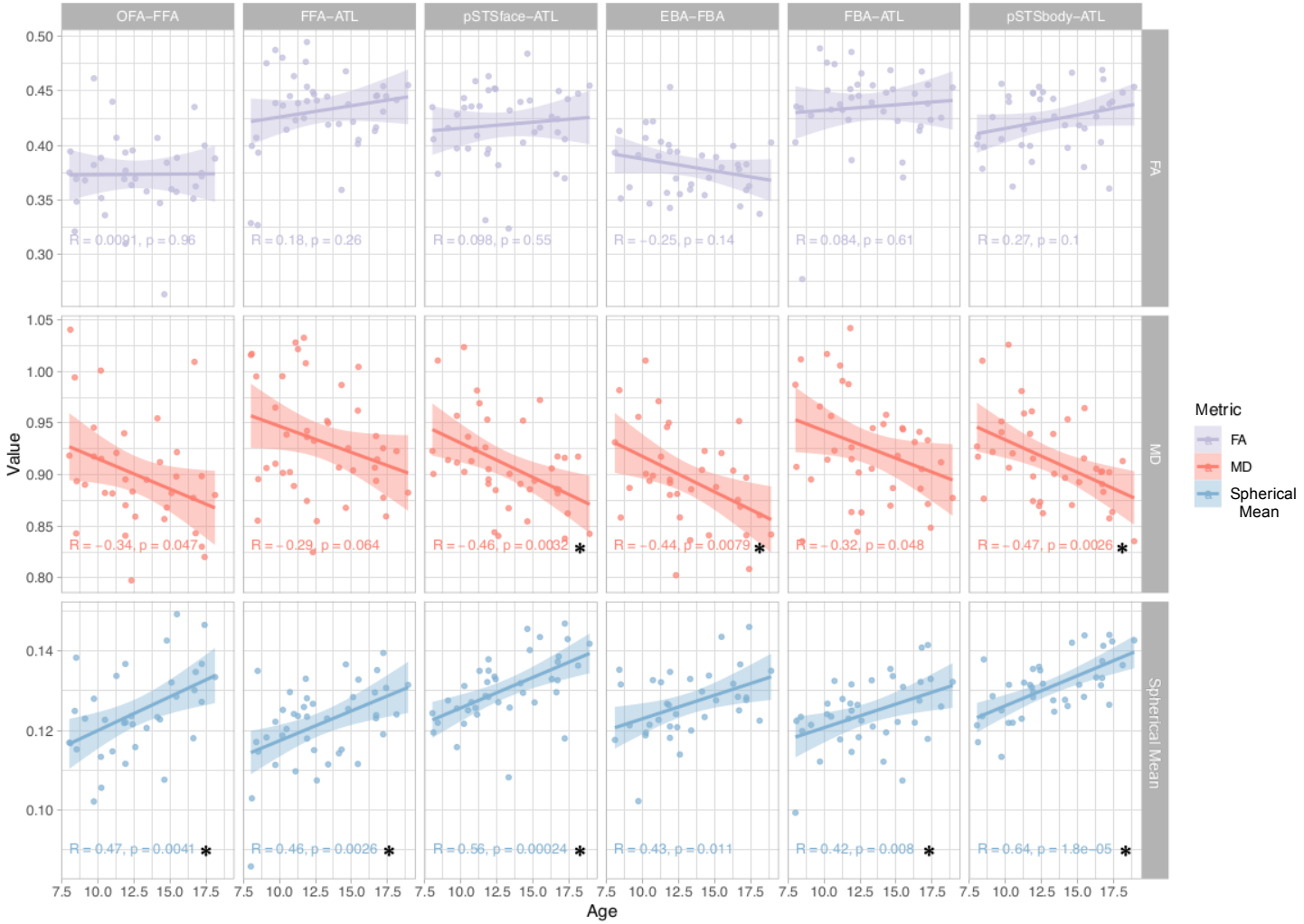


Figure 4.6: **Relationship between microstructure and age, grouped by FDWM tract and coloured by metric**
For each of the correlation plots the age is plotted in years along the x-axis, and the y-axis shows the value of the metric. FA is shown in the top panel in purple, MD in the middle panel in orange and $\mathring{S}_\mu(b = 6000)$ in the bottom panel in blue. The tracts from the face network are shown on the left panels, and body on the right. The Pearson's rank correlation coefficient and p-value are indicated for each relationship. Significant relationships are indicated with an asterisk ($p<0.008$, following Bonferroni correction). The 95 % confidence interval is shown with shading.

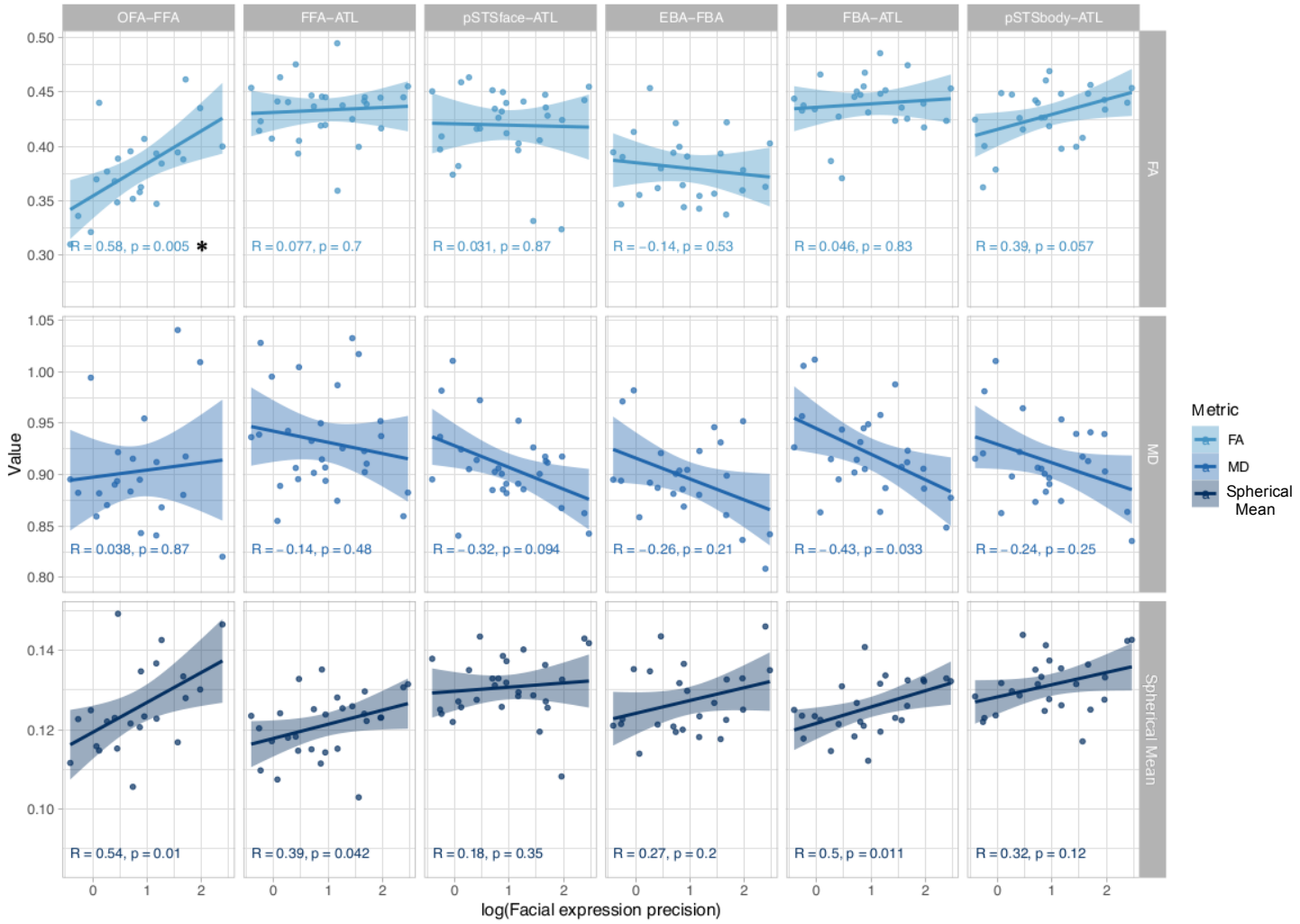


Figure 4.7: **Relationship between microstructure and facial expression precision, grouped by FDWM tract and coloured by metric** For each of the correlation plots the x-axis displays the logged facial expression precision, and the y-axis shows the value of the metric. FA is shown in the top panel in light-blue, MD in the middle panel in medium-blue and $\mathring{S}_{\mu}(b = 6000)$ in the bottom panel in dark-blue. The tracts from the face network are shown on the left panels, and body on the right. The Spearman's rank correlation coefficient and p-value are indicated for each relationship. Significant relationships are indicated with an asterisk ($p<0.008$, following Bonferroni correction). The 95 % confidence interval is shown with shading.

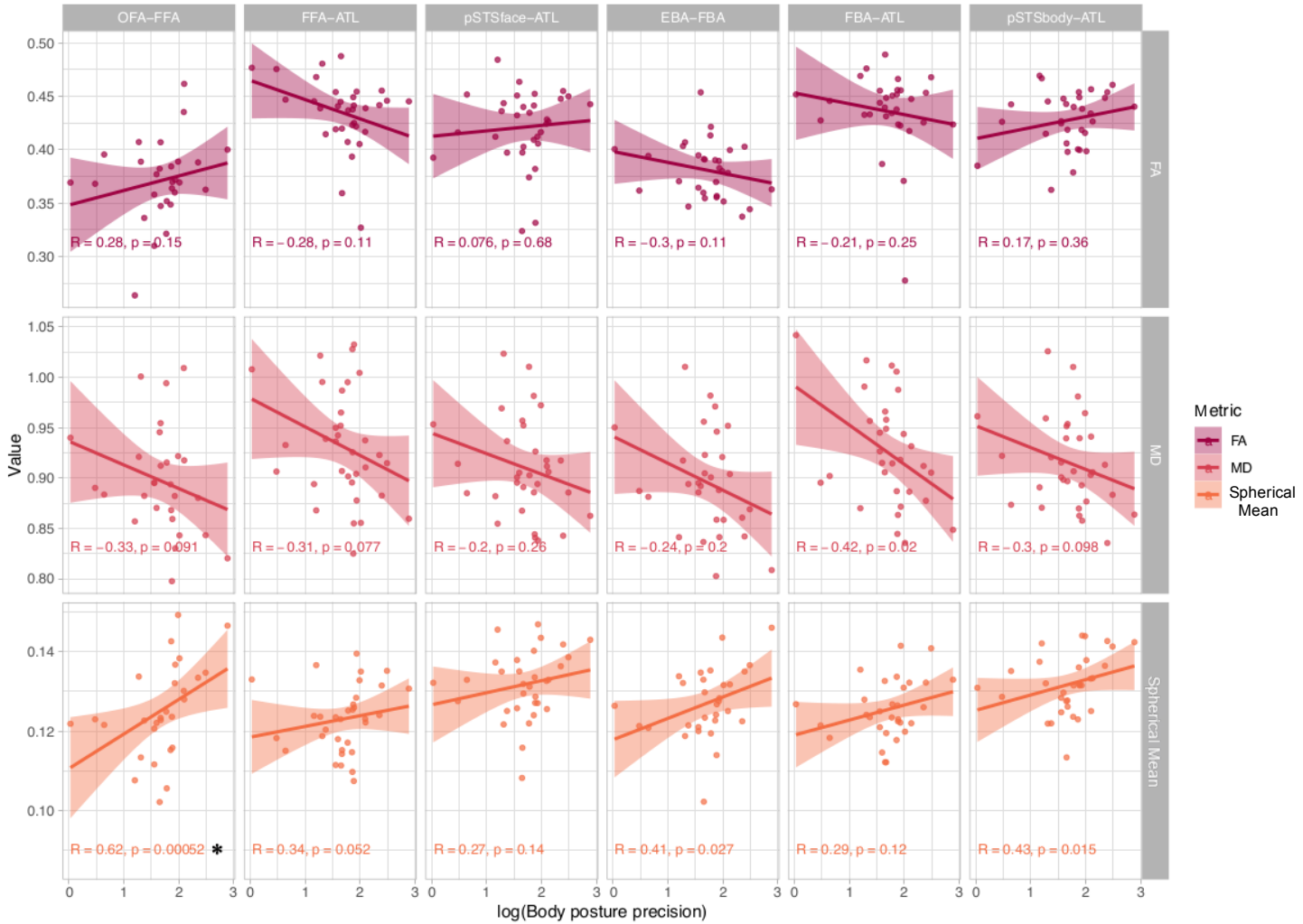


Figure 4.8: **Relationship between microstructure and body posture precision, grouped by FDWM tract and coloured by metric** For each of the correlation plots the x-axis displays the logged body posture precision, and the y-axis shows the value of the metric. FA is shown in the top panel in dark-red, MD in the middle panel in medium-red and $\mathring{S}_\mu(b = 6000)$ in the bottom panel in orange. The tracts from the face network are shown on the left panels, and body on the right. The Spearman's rank correlation coefficient and p-value are indicated for each relationship. Significant relationships are indicated with an asterisk (p<0.008, following Bonferroni correction). The 95 % confidence interval is shown with shading.

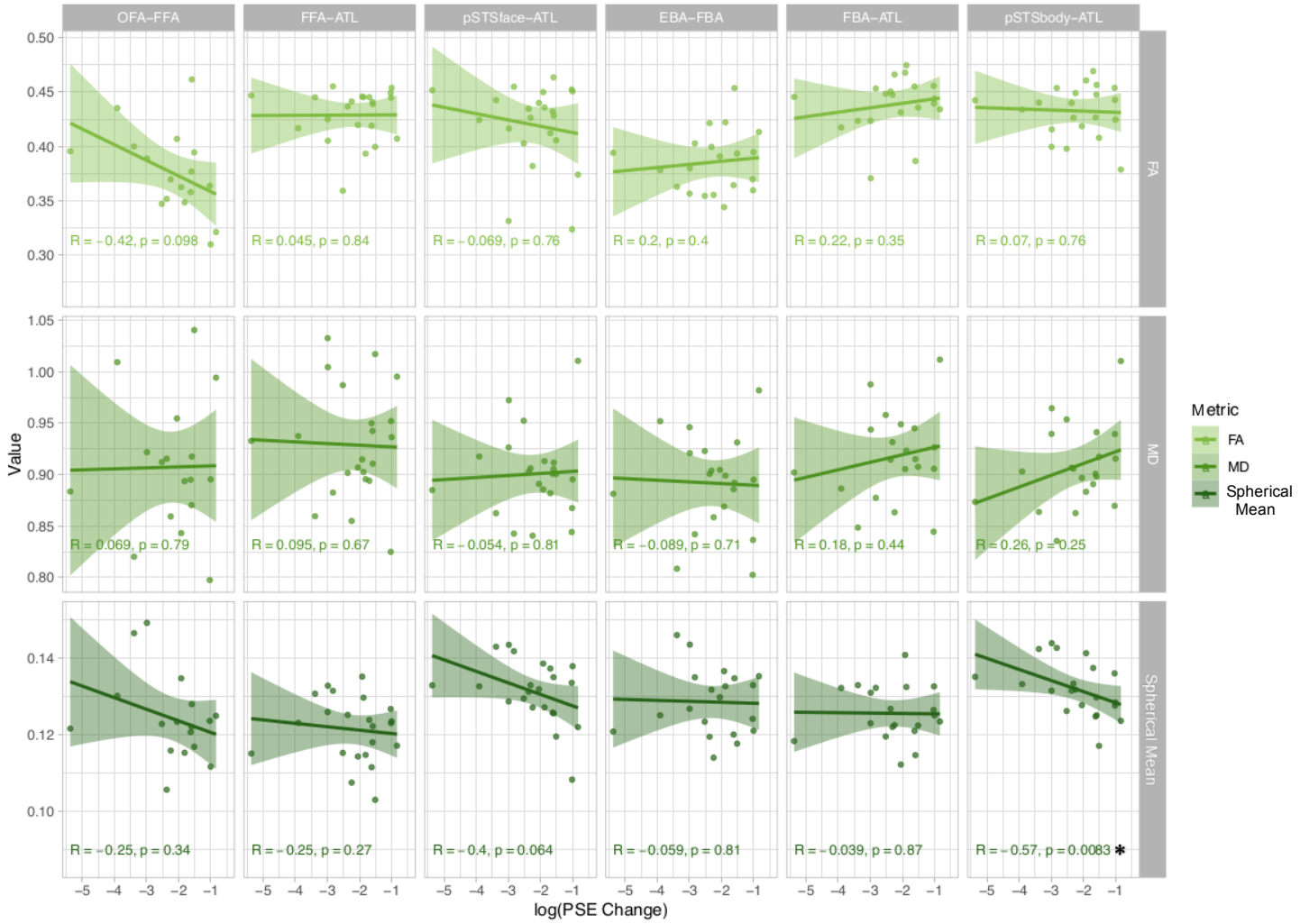


Figure 4.9: **Relationship between microstructure and PSE change, grouped by FDWM tract and coloured by metric**
For each of the correlation plots the x-axis displays the logged PSE change, and the y-axis shows the value of the metric. FA is shown in the top panel in light-green, MD in the middle panel in medium-green and $\mathring{S}_{\mu}(b = 6000)$ in the bottom panel in dark-green. The tracts from the face network are shown on the left panels, and body on the right. The Spearman's rank correlation coefficient and p-value are indicated for each relationship. Significant relationships are indicated with an asterisk ($p<0.008$, following Bonferroni correction). The 95 % confidence interval is shown with shading.

To disentangle age-related effects in the FDWM tracts whose metrics showed significant correlations with behavioural performance, I ran separate multiple linear regressions. The three relationships assessed were: facial expression precision predicted by OFA-FFA(FA), body posture precision predicted by OFA-FFA($\mathring{S}_{\mu}(b = 6000)$) and PSE change predicted by pSTSbody-ATL($\mathring{S}_{\mu}(b = 6000)$). The results of the regression analysis allowed for exploration of both age-related changes in brain microstructure, and specific microstructural changes reflecting perceptual abilities.

Facial expression precision

The FA of OFA-FFA [Figure 4.10] was found to be a significant predictor of facial expression precision ($F_{1,20}$=10.95, p<0.01), with the model explaining 32.2% of the variance as indicated with the adjusted R^2. When age was added as a predictor to the model, the model improved to explain 43.7% of the variance ($F_{2,19}$=9.14, p<0.01). This suggests that having OFA-FFA (FA) and age in the model explains more of the variance in facial expression precision, than was predicted by the microstructural metric alone. Both age (B_{age}=0.09, p=0.036) and OFA-FFA(FA) (B_{tract}=10.18, p<0.01) were found to significantly contribute to the model [Table 4.3]. This demonstrates that the FA of the OFA-FFA tract significantly predicts an individual's facial expression precision, even when adjusted for age.

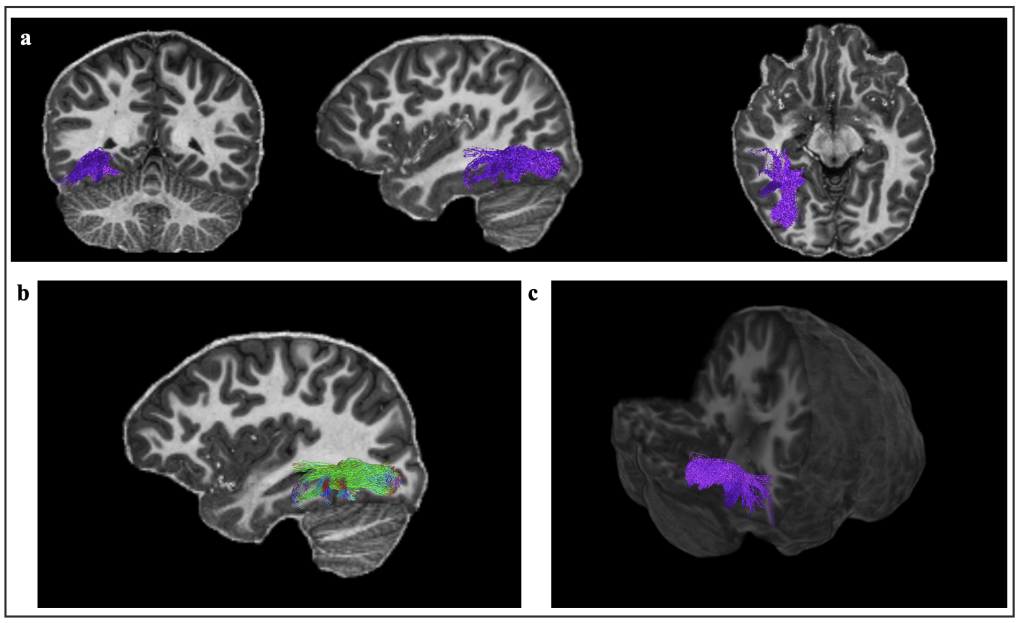


Figure 4.10: **Example of FDWM tracts: OFA-FFA**
Examples of functionally-defined tracts connecting OFA and FFA are shown in one child. In the top panel the tracts can be visualised in the coronal, sagittal and axial planes [4.10a]. The tracts can be seen with directional colour encoding in the sagittal plane, in the bottom left panel [4.10b]. A 3D view of the tracts can be seen in the bottom right panel [4.10c].

Table 4.3: **Regression output: Facial expression precision, fractional anisotropy of OFA-FFA, and age**
The table shows the regression outputs for the model with facial expression precision and tract microstructure (OFA-FFA(FA)) in the first column, and the facial expression precision predicted by the tract microstructure (OFA-FFA(FA)) and age in the second column. The coefficient estimates, confidence intervals and significance levels for each of the predictors in the models are shown. The significant p-values ($p<0.05$) are indicated in bold text. The R^2 and adjusted R^2 were used to assess the model fit.

	Facial expression precision			**Facial expression precision**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-3.65	-6.48 – -0.82	**0.014**	-2.98	-5.64 – -0.33	**0.030**
OFA-FFA(FA)	11.79	4.36 – 19.23	**0.003**	10.18	3.23 – 17.14	**0.006**
Age				0.09	0.01 – 0.18	**0.036**
Observations	22			22		
R^2 / R^2 adjusted	0.354 / 0.322			0.490 / 0.437		

Body posture precision

The spherical mean of the OFA-FFA tract significantly predicted the body posture precision ($F_{1,26}$=6.00, p=0.021), however the model only explained 15.6% of the variance. Adding age to the model improved the overall fit of the model to explain 19.3% of the variance and was significant ($F_{2,25}$=4.22, p=0.026) [Table 4.4]. However, in this model $\mathring{S}_\mu(b = 6000)$ of OFA-FFA (B_{tract}=14.11, p=0.170) and age (B_{age}=0.06, p=0.153) were not found to be significant predictors. One explanation for the fact that the spherical mean of OFA-FFA and age were not significant predictors when both in the model, could be an artefact of their high correlation with one another ($r_{(s)}$= 0.470, p=0.004). Given both of these predictors showed a significant correlation, multicollinearity was assessed using the variance inflation factor (VIF); the value for each of the predictors was 1.35, suggesting collinearity was present in the model, therefore these results should be interpreted with caution.

Table 4.4: **Regression output: Body posture precision, spherical mean of OFA-FFA and age**
The table shows the regression outputs for the model with body posture precision and tract microstructure (OFA-FFA($\mathring{S}_\mu(b = 6000)$)) in the first column, and the body posture precision predicted by the tract microstructure (OFA-FFA($\mathring{S}_\mu(b = 6000)$)) and age in the second column. The coefficient estimates, confidence intervals and significance levels for each of the predictors in the models are shown. The significant p-values (p<0.05) are indicated in bold text. The R^2 and adjusted R^2 were used to assess the model fit.

	Body posture precision			**Body posture precision**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	-1.03	-3.31 – 1.24	0.360	-0.07	-2.67 – 2.53	0.956
OFA-FFA(Spherical mean)	21.58	3.47 – 39.68	**0.021**	14.11	-6.47 – 34.69	0.170
Age				0.06	-0.02 – 0.14	0.153
Observations	28			28		
R^2 / R^2 adjusted	0.188 / 0.156			0.253 / 0.193		

PSE change

PSE change was found to be significantly predicted from the $\mathring{S}_{\mu}(b = 6000)$ of the pSTS(body)-ATL tract [Figure 4.11] ($F_{1,19}$=4.85, p=0.04), with the $\mathring{S}_{\mu}(b = 6000)$ of pSTS(body)-ATL tract accounting for 16.1% of the explained variability in PSE change. For this relationship adding age to the model did not improve the overall model fit and it was no longer found to be significant ($F_{2,18}$=2.34, p=0.13), with the adjusted R^2 decreasing by 4.3 % [Table 4.5]. Again, there was a high correlation between both predictors ($r_{(s)}$= -0.475, p=0.015). The VIF for both predictors was 2.44, which is high enough to cause concern for multicollinearity. Therefore, these results must be interpreted with caution due to the high correlation between spherical mean of pSTS(body)-ATL and age. To interpret the contribution of both microstructural predictors and age in this instance may be more appropriate independently. I found a significant negative correlation between PSE change and age ($r_{(s)}$= 0.3923, p=0.0395), a significant negative correlation between PSE change and pSTS(body)-ATL ($\mathring{S}_{\mu}(b = 6000)$ ($r_{(s)}$=-0.57, p=0.008) and a significant positive correlation between pSTS(body)-ATL ($\mathring{S}_{\mu}(b = 6000)$ and age ($r_{(s)}$=0.62, p<0.0001).

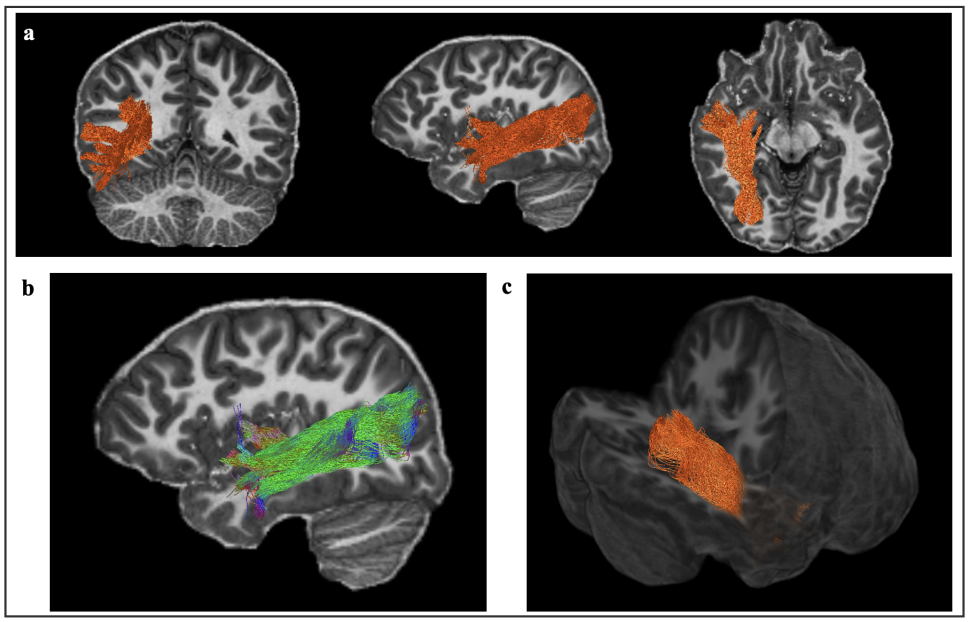


Figure 4.11: **Example of FDWM tracts: pSTS(body)-ATL**
Examples of functionally-defined tracts connecting pSTS(body) and ATL are shown in one child. In the top panel the tracts can be visualised in the coronal, sagittal and axial planes [4.11a]. The tracts can be seen with directional colour encoding in the sagittal plane, in the bottom left panel [4.11b]. A 3D view of the tracts can be seen in the bottom right panel [4.11c].

Table 4.5: **Regression output: PSE change, spherical mean of pSTS(body)-ATL and age**
The table shows the regression outputs for the model with PSE change and tract (pSTS(body)-ATL($\mathring{S}_\mu(b = 6000)$)) in the first column, and the PSE change predicted by the tract microstructure (pSTS(body)-ATL($\mathring{S}_\mu(b = 6000)$)) and age in the second column. The coefficient estimates, confidence intervals and significance levels for each of the predictors in the models are shown. The significant p-values ($p<0.05$) are indicated in bold text. The R^2 and adjusted R^2 were used to assess the model fit.

	PSE Change			**PSE Change**		
Predictors	*Estimates*	*CI*	*p*	*Estimates*	*CI*	*p*
(Intercept)	7.04	-1.78 – 15.86	0.111	5.78	-8.34 – 19.91	0.401
pSTSbody-ATL(Spherical mean)	-70.26	-137.01 – -3.50	**0.040**	-60.68	-168.11 – 46.76	0.251
Age				-0.03	-0.28 – 0.22	0.810
Observations	21			21		
R^2 / R^2 adjusted	0.203 / 0.161			0.206 / 0.118		

4.4 Discussion

In the present chapter, children's perception of facial expressions was found to be significantly influenced by the presence of a body posture; younger children showed a greater biasing effect than older children. Furthermore, isolated facial expression recognition ability improved with increasing age. Individual differences in the microstructural properties of functionally-defined white matter tracts underpinning face and body processing networks were found to be predictive of children's perceptual abilities.

4.4.1 Behavioural results

Children's facial expression recognition was found to significantly improve with increasing age. These results support previous research showing that with increasing age, children become better at detecting subtle differences between facial expressions (Dalrymple et al., 2017). For the first time, I show that body posture recognition also increases with increasing age. However, it is important to note that this relationship was only borderline significant and subsequent research is required to confirm this finding.

Body posture was found to significantly influence the perception of facial expressions in children. These results demonstrate that children's facial expression judgements were biased by an affective body posture; the magnitude of the body biasing effect diminished with increasing age. To my knowledge, this is the first work to look at the influence of body posture on facial expressions across development. Previous work has focused on children of a particular age and compared their performance to adults (Mondloch, 2012; Mondloch et al., 2013). In contrast, my current results illustrate how the body biasing effect changes across the course of development. My findings support previous developmental work exploring facial expression processing in context; Mondloch (2012) found that children's judgements of facial expressions were more affected than adults by the presence of an incongruent body posture. A limitation of the Mondloch (2012) work is that isolated facial expression and body posture recognition ability in the children was not reported. In the current chapter, I demonstrate that younger children were more influenced by context than older children, and that there are improvements in how well the individual cues are recognised with increasing age. This pattern of results echoes the

findings reported in Chapter 3. In adults, I found that observers with less precise facial expression recognition were affected more by the body posture when categorising facial expressions in the context of a whole-person. The current results in children show the same pattern across development, highlighting that the representation of the individual cues is important in determining how they are integrated.

The current approach has several limitations, one being the use of only anger and disgust expressions. Therefore, the conclusions drawn regarding development of isolated facial expression and body posture processing must be interpreted with caution. In order to generalise these findings further, subsequent research is required with additional expressions. Despite this, the selection of angry and disgusted expressions was based upon previous research (Aviezer et al., 2008), in order to understand how body context influenced perception of facial expressions in children. Furthermore, previous research has shown that angry and disgusted body postures are two of the most readily recognised distinct static postures from the range of basic expressions (Lopez et al., 2017), hence their selection for this novel paradigm in children.

4.4.2 Functionally-defined white-matter

In the current work, individual differences in microstructure of functionally-defined white matter tracts were found to be related to individual differences in children's perceptual abilities. Facial expression precision was found to be predicted by the FA of OFA-FFA tract and age. Early models of face processing in the brain segregated processing of faces into dorsal and ventral streams, with changeable and invariant aspects being encoded in each respectively (Haxby et al., 2000). Facial expressions are classified as changeable features of faces, therefore the predominant view has been that the neural processing of facial expressions is restricted to dorsal regions, such as pSTS. However, more recent fMRI research has highlighted the importance of regions in the ventral stream, namely the FFA in processing of facial expressions (Duchaine & Yovel, 2015; Ganel et al., 2005). The revised model of face processing proposed by Duchaine and Yovel (2015) suggests the FFA also plays a role in the processing of changeable aspects of faces. More recently the distinction between dorsal and ventral streams has been suggested to be better represented by the processing of 'motion' and 'form' information respectively (Bernstein & Yovel, 2015).

The ventral stream, which includes the FFA, is suggested to be important in extracting 'form' from static face images, which includes expression. Furthermore, recent multivariate functional MRI has provided evidence for encoding of facial expression information in face responsive regions of the fusiform cortex (Dima et al., 2018; Muukkonen et al., 2020). My current finding, where microstructural properties of the fibre tracts connecting OFA and FFA were predictive of facial expression recognition ability in children, provides support for these revised models of face perception in the brain, with the ventral stream playing a key role in the processing of expression.

Clinical research has also demonstrated the importance of white-matter pathways in the ventral temporal region for facial expression processing. In a large sample of patients with brain lesions, damage to the IFOF was found to significantly predict the overall facial expression recognition impairments in patients (Philippi et al., 2009). This work further illustrates the functional importance of the ventral temporal stream for visual processing of facial expressions, highlighting the role of white matter tracts in facial expression perception in these regions. Although I found no evidence for a relationship between microstructural metrics of ILF or IFOF and children's perceptual abilities, my results show that a more targeted, functionally-defined approach to extracting white matter tract microstructural metrics provides evidence for a strong relationship with facial expression perception.

Finding that the microstructure (FA) of tracts connecting OFA and FFA, is reflective of individual differences in children's facial expression recognition, adds to the growing consensus of research appreciating the importance of the white matter pathways underpinning face processing more generally in the ventral visual system. This work builds upon Gomez *et al.* (2015) where a correlation between the FA of FDWM local to face-selective cortex in the right temporal lobe and facial identity recognition was reported in adults. Taken together, this research highlights the importance of a targeted approach to explore local white-matter connectivity, with functional relevance to cognition. Furthermore, my current work highlights the feasibility of a targeted approach to explore developmentally relevant changes in white matter microstructure that are predictive of perception.

The spherical mean of OFA-FFA tracts was found to predict individual differences in children's body posture recognition. In contrast to the emerging field researching white

matter underpinning face perception, to my knowledge no published literature has sought to uncover the relationship between white matter and body posture recognition. Given the proximity of the cortical regions involved in face and body processing, it is to be expected that the white matter connections involved in face and body perception may be partially overlapping (Peelen & Downing, 2005). This does not conflict with the idea that face and body processing remain segregated until later in the ventral visual stream, rather it just highlights the limitations in the resolution and specificity of dMRI. Tractography is ultimately limited as it is a modelling approach used to estimate the underlying white matter architecture based on the movement of water in the brain. It may be methodologically impossible, at present, to disentangle the unique face and body networks in-vivo due to their proximal cortical locations. I found no evidence for a relationship between children's body posture precision and any of the tracts connecting body processing regions. Given I found a relationship between body posture precision and tracts connecting OFA-FFA, it is likely a consequence of overlapping white matter architecture, or could support the notion that networks underpinning face and body perception are not as distinct as hypothesised (Arcaro et al., 2020). Importantly, facial expression precision was shown to correlate with FA, and body posture precision with $\mathring{S}_{\mu}(b = 6000)$ of OFA-FFA; demonstrating that the relationships observed were reflective of different features of the underlying architecture. Previous work has highlighted that spherical mean captures age related changes in microstructure (Raven et al., 2020), therefore caution must be made when interpreting the relationship with body posture and spherical mean of OFA to FFA tracts.

An alternative hypothesis emerging from non-human primate research posits that potentially face and body processing regions are not as segregated as previously thought. Recent monkey work has highlighted that the neurons involved in category specific processing, are in fact less tuned to individual categories but are also sensitive to object relationships (Arcaro et al., 2020). Using direct neuronal recordings Arcaro *et al.* (2020) demonstrated that the cells in face processing regions of the middle lateral, and posterior lateral, inferior temporal cortex respond to images of complex scenes, that lacked faces but indicated where the face ought to be. This resulted in the discovery that face-selective neurons could respond to stimuli where contextual cues, such as a body, indicated a face should be. Although currently in macaques, parallel non-invasive research in humans is

warranted to explore the neural influence of context in category specific regions responsible for face and body processing.

The results presented here suggest the integration of facial expression and body posture cues occurs downstream of pSTS. I found that the spherical mean of tracts extending from the pSTS to anterior temporal lobe were predictive of how much body context biased children's perception of facial expressions. The role of the pSTS has been attributed to both face and body processing, with a key role in dynamic face perception and motion detection (Basil et al., 2017; Downing, 2001; Duchaine & Yovel, 2015; Haxby et al., 2000). Furthermore, the pSTS has been highlighted as critical in visuomotor perception of emotional body language (de Gelder et al., 2015). O'Toole *et al.* (2002) proposed that the pSTS plays an important role in identification of social context when perceiving dynamic facial expressions. This work, in combination with the current results identifying a relationship between the microstructure of the pSTS to ATL and influence of body context, allude to the importance of the connectivity in this region for processing the contextual relevance of faces and bodies. This finding aligns with behavioural theories for integration of facial expression and body posture; Teufel *et al.* (2019) proposed that integration of facial expression and body posture cues occurs downstream of core face perception regions like pSTS and FFA.

Compilation of human and non-human primate research has culminated with a recent proposal for a third visual pathway, in addition to the established notion of dorsal and ventral streams (Pitcher & Ungerleider, 2021). The structural connectivity of the third pathway is believed to extend laterally from early visual cortex, to V5, pSTS and anterior STS. The 'third pathway' has been proposed to be specialised for for the dynamic aspects of social perception (Pitcher & Ungerleider, 2021). Causal evidence for a direct anatomical connection between pSTS and anterior portions of STS has been shown using concurrent TMS and fMRI (Pitcher et al., 2017). In adults, following application of TMS to the pSTS, the neural response to moving faces was diminished in both the pSTS and anterior STS. The results in the current study, finding that the influence of body context on facial expression perception could be predicted by the $\mathring{S}_{\mu}(b = 6000)$ of pSTS-ATL(body), supports the proposal from Pitcher *et al.* (2019) that this pathway plays an important role in social perception, and adds to the consensus that integration of facial expression

and body posture cues takes place later in the visual processing hierarchy.

A unique feature of the current study is the sensitivity to axonal morphology and glial processes gained using specialised hardware (3T Connectom scanner), which I assessed in combination with tensor measures of microstructure. Myelination is known to contribute to the functional characteristics and information relay speed between particular regions, which is typically thought to be reflected in FA changes (Etxeberria et al., 2016). However, myelination, axonal packing, axon coherence and fibre diameter can all contribute to the FA, highlighting limitations in the microstructural specificity of changes associated with FA (Beaulieu, 2002; Jones et al., 2013). With this in mind, it is unsurprising that age-related changes in the FA of functionally-defined white matter were not observed, particularly due to the cross-sectional design. Many studies that report changes in FA, with increasing age, are longitudinal large-scale studies with more power to detect these changes (Lebel & Beaulieu, 2011). In contrast to the FA metrics, I utilised high b-value diffusion data to provide an insight into axonal properties of the tissue. At high b-values, such as b=6000s/mm^2, the spherical mean relates strongly to specific information about intracellular microstructure, including the axon diameter (Veraart et al., 2020). The spherical mean metric for the majority of the FDWM tracts was found to be positively correlated with age. Other developmental research comparing classic DTI measures with models of intra-neurite space (Neurite density index (NDI)), revealed a strong correlation between NDI and age (Chang et al., 2015); NDI has been found to explain more variance in white-matter microstructure with age than fractional anisotropy (Genc et al., 2017; Mah et al., 2017).

Given intracellular measures appear to be reflective of age-related changes, subsequent work should account for additional tissue changes, such as myelin, to enable researchers to further disentangle general developmental changes from individual differences. Recent empirical work has provided evidence of microstructural proliferation driving changes in high-level visual areas involved in face processing across development (Gomez et al., 2017). Gomez *et al.* (2017) found improvements in face recognition that were correlated with tissue development in the fusiform face-selective regions of cortex of children. In contrast to the dominant model of synaptic pruning thought to be driving developmental changes in brain tissue in infants (Huttenlocher & de Courten, 1987), Gomez *et al.*'s (2017) research

shows that in later childhood differential patterns of maturation across the cortex emerge, which are likely driven by microstructural proliferation in multiple cortical compartments (Gomez et al., 2017). Potential candidates for this proliferation are cell bodies, dendritic structures and myelin sheaths. It is a well established finding that cortical thickness decreases over childhood whilst there is a concurrent increase in the total volume of white matter (Mills et al., 2021). In deep cortical layers and at the cortical boundary, the quantity of myelin has been reported to increase with age across adolescence (Paquola et al., 2019; Whitaker et al., 2016). These dynamic deep layer cortical regions are the targets of axonal projections which are known to play a critical role in connecting regions of the cortex via white matter pathways. In future, by combining measures about myelin, with intracellular measures and more non-specific markers as indexed by FA, it may be possible to gain additional insight into what neuronal changes are driving changes in performance. Several factors are likely to be at play, as suggested recently by Grill-Spector *et al.* (2017), where face perception in the brain is likely dependent on a combination of functional activity, local white-matter properties and the microstructural properties of the grey-matter cortical regions (Grill-Spector et al., 2017).

4.5 Chapter Summary

In the present chapter I provide data quantifying the influence of body context on facial expression perception across development. I found that with increasing age, children were significantly less biased by an affective body posture when judging a facial expression. An accompanying increase in isolated facial expression recognition was also found with increasing age. This finding supports the results discussed in Chapter 3, where adults who had less precise facial expression representation, were more influenced by body posture in the whole-person emotion perception. The results in the current chapter echo this finding in a developmental cohort, suggesting that changes in the influence of body context on facial expression perception across development are largely driven by age-related changes in isolated facial expression perception. These findings have important implications for our understanding of visual processing of facial expression and body postures, in addition to having wider reaching impacts for social cue processing more broadly across childhood

and adolescence.

In addition to the novel behavioural findings, I also adopted a targeted functionally defined white matter analysis strategy to specifically explore the connectivity between cortical nodes involved in face and body processing in the brain. Microstructural changes in functionally-defined face and body networks were found to be predictive of children's perceptual ability; specifically, facial expression precision was related to FA of tracts linking OFA and FFA. This result demonstrates the importance of fusiform regions for facial expression processing and highlights feasibility of exploring individual difference in perception that are related to white matter microstructure. Body posture precision and $\mathring{S}_{\mu}(b = 6000)$ of tracts linking OFA and FFA, and PSE change with $\mathring{S}_{\mu}(b = 6000)$ of tracts connecting pSTS(body) and ATL, were also found to be related. Furthermore, I also reported specific developmental sensitivity to intracellular morphology in FDWM. Taken together, this research uncovers developmental differences in facial expression and body posture processing and their integration, and also highlights the functional specificity of white matter connectivity underpinning face and body processing in the developing brain. This research paves a way for subsequent work to explore the interplay between behaviour and microstructure in a hypothesis driven manner.

Chapter 5

Perceptual processing of facial expression and body posture cues: A social cue integration framework

5.1 Introduction

The human brain must combine multiple sensory estimates about the world to guide perception (Trommershauser et al., 2011). Our estimates about the world are always associated with uncertainty because neural information processing is noisy (Whiteley & Sahani, 2008). In order to minimise the uncertainty in sensory cue measurement, observers often combine multiple cues to improve the reliability of their estimates (Trommershauser et al., 2011). This combination is typically conceptualised within a cue integration framework that allows us to formally assess how our brains utilise multiple sensory estimates to arrive at a stable and robust percept of the world (Bejjanki et al., 2011; Dekker et al., 2015; Ernst & Banks, 2002; Martin, 2016). Optimal cue integration, based on the maximum-likelihood principle, predicts that the relative reliability of the sensory cues determines how much they will contribute to the integrated representation, and that the integrated estimate should be more reliable than the individual sensory estimates (Ernst & Banks, 2002). These predictions are supported by empirical evidence from low-level vision research (Ernst & Bülthoff, 2004; Landy & Kojima, 2001). Cue integration has recently been proposed as a key feature of complex social cognition, as this relies on the interaction between multiple environmental signals and information processing streams (Zaki, 2013).

Despite the theoretical proposal, to my knowledge, no published research demonstrates the application of cue integration modelling to higher-order social cues. The principles of understanding how we are able to make social inferences from complex social cues, are akin to the senses. The interplay between incoming sensory signals and information processing streams is an essential component of social cognition (Zaki, 2013). In low-level vision the use of formal models has been successful in providing a deep and mechanistic understanding of information-processing (Backus et al., 1999; Jacobs, 1999). Applying the models and insights from low-level vision to processing of complex stimuli such as faces, holds the promise of providing similar insights into the processes underpinning high-level social perception.

In Chapters 3 and 4 of this thesis, I demonstrated the important role of body context in perception of a facial expression. Typical encounters with facial expressions in the real-world are in conjunction with a body posture. Therefore, it is important we understand how our brain effortlessly integrates these social cues to form whole-person representations. In the present chapter, facial expression and body posture integration is conceptualised under a cue integration framework to determine if human behaviour is consistent with maximum likelihood estimation (MLE) predictions. I make two key quantitative predictions: firstly, observers' who have more precise facial expression representations than body posture representations will be less affected by the body context in whole person perception. This is because more reliable cues contribute more to the integrated representation. Secondly, the observer's integrated whole person representation should be more reliable than either of the individual sensory cues. Within this framework, optimality has to be defined perceptually (Cormack, 2005), where optimality is defined as the precision of the response the observer is making based on the incoming sensory evidence.

The cue integration framework has been developed as a model to explain perceptual information-processing, rather than post-perceptual decision-making stages. Therefore, prior to fitting the model, I aimed to provide evidence to support the idea that the integration of facial expression and body posture is truly perceptual. In addition to a task in which observers were asked to categorise facial expressions using verbal labels, as previously described [Chapter 3 & 4], I also tested performance in a task, in which response options were non-verbal. Instead of selecting the words 'disgust' or 'anger', observers had

to choose between angry or disgusted facial expressions or body postures. Previous research has suggested that the use of a non-verbal response option is a better way to assess if a process is perceptual, as observers do not need to rely on their verbalisation ability (Palermo et al., 2013).

To summarise, the goal in the present work was to determine whether integration of facial expression and body posture cues was consistent with predictions from a formal cue integration framework, based on the maximum likelihood principle. In addition, I provide new evidence to support the hypothesis that the influence of body posture on perception of facial expressions is perceptual.

5.2 Methods

The behavioural paradigm used in this work was a modified version of the task detailed in Chapter 3. Modifications to the methodology described in Section 3.2 are detailed below.

5.2.1 Observers

A total of 38 naïve observers (14 male) were recruited from the general population. Only neurologically and physically healthy participants were recruited. All participants had normal or corrected to normal vision. All participants were over 18 years of age (mean age = 29.3 ± 8.15, range = 18-54). Observers provided written informed consent prior to participating. Experimental protocols were approved by Cardiff University School of Psychology Ethics Committee, and were in line with the Declaration of Helsinki. All participants were fully debriefed at the end of the testing session and payment was provided in the form of cash.

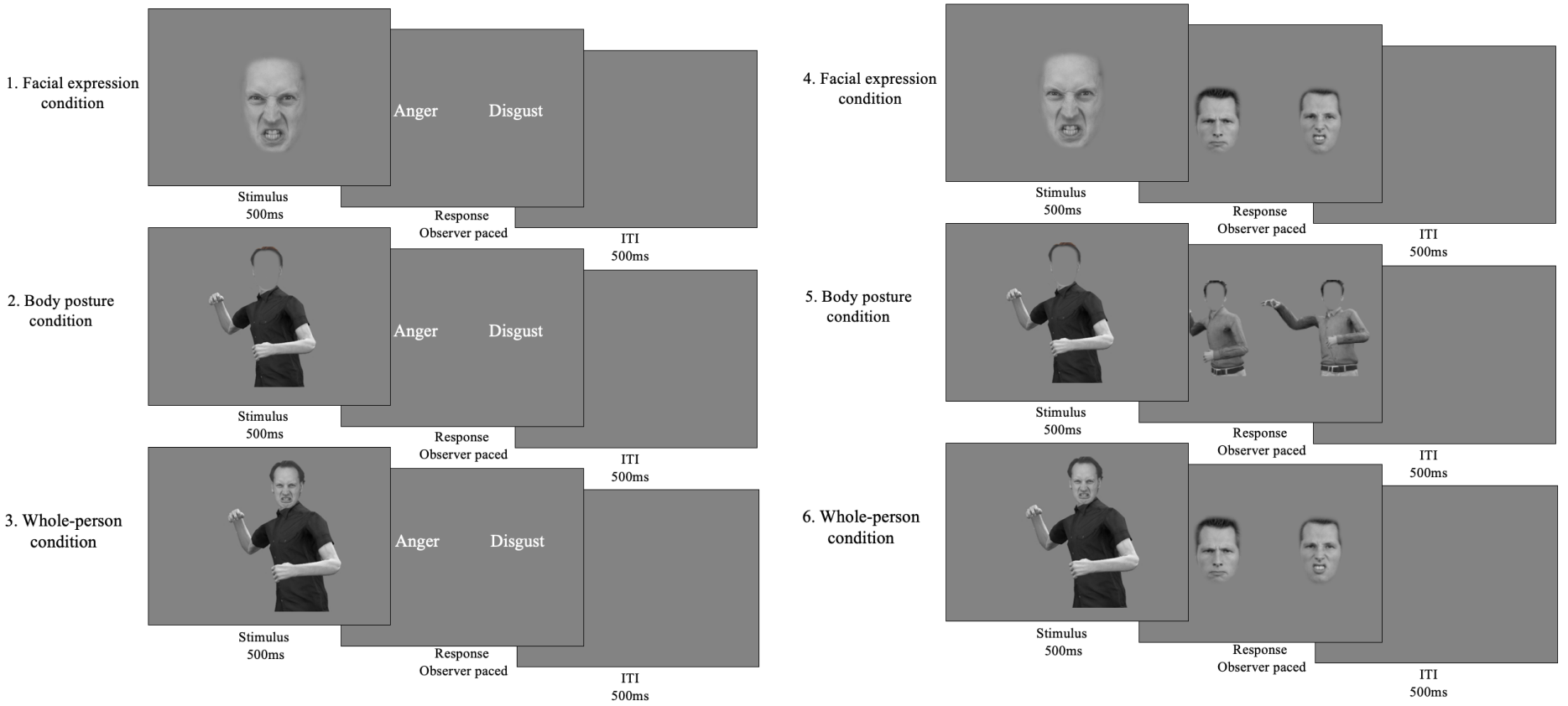
5.2.2 Stimuli

Facial expression stimuli were generated using photographs of male actors from the Radboud and Karolinska Directed Emotional Faces validated sets of facial expressions (Langner et al., 2010; Lundqvist et al., 1998). Angry and disgusted expressions were selected for four Caucasian male identities. These identities differed from those used in Chapters 3 and 4. For each identity, the fully-angry and fully-disgusted facial expressions were morphed together using FantaMorph software [FantaMorph Pro, Version 5]. This procedure generated morph continua for each identity. The morphs changed in increments of 5% between the angry and disgusted facial expressions resulting in 21 morph levels for each identity. For categorisation of the facial expressions in isolation, a mask blending into the mean grey background was used to remove any external facial features, such as the hair and ears. The body postures used were identical to those detailed in Section 3.2.2 of Chapter 3. Whole-person stimuli were created by manually pasting the individual morphed emotional faces onto fully-angry and fully-disgusted body postures using GIMP [GNU Image Manipulation Program, Version 2.10].

5.2.3 Procedure

Presentation of the task was controlled by custom-written MATLAB (Version 2016b) code using the Psychophysics Toolbox (Version 3.0.14) (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Observers took part in three conditions twice, once with verbal response options, once with non-verbal response options: 1.Facial expression categorisation, 2.Body posture categorisation, 3.Whole-person categorisation [Figure 5.1]. Observers were asked to categorise a test stimulus as either angry or disgusted in facial expression, body posture and whole-person conditions. For the whole-person condition, observers were explicitly instructed to judge the facial expression and ignore the body posture. The stimulus presentation was as detailed in Section 3.2.3, however the response options were different for the present task. For the verbal response options, following test stimulus presentation, the words 'Anger' and 'Disgust' were presented on either the left or right hand side of the screen [Figure 5.1a]. The observer was instructed to select the word that categorised the test stimulus presented on that trial. To respond, observers pressed a key corresponding to the location of the word. The location of the words appearing on the left or right hand side of the screen was counterbalanced within each task. For the non-verbal response options the procedure was the same, except the response options displayed were images of fully-angry and fully-disgusted stimuli [Figure 5.1b]. The response choice stimuli options displayed for the facial expression and whole-person conditions were isolated facial expressions, and body postures for the body posture condition. The identity displayed for the response choice stimuli was different from the identity of the test stimulus.

The study design was a repeated-measures block design. The order of the six experimental conditions was counterbalanced across participants. For each task the participant underwent a short training phase at the beginning of the block.



(a) Verbal response option conditions

(b) Non-verbal response option conditions

Figure 5.1: **Experimental conditions**
The figure depicts the six experimental conditions: 1. Facial expression categorisation (verbal), 2. Body posture categorisation (verbal), 3. Whole-person categorisation (verbal), 4. Facial expression categorisation (non-verbal), 5. Body posture categorisation (non-verbal), 6. Whole-person categorisation (non-verbal). On each trial the stimulus was presented centrally for 500ms. Observers were instructed to categorise the facial expressions in the facial expression and whole-person conditions, and the body posture in the body posture conditions. Following presentation of the stimuli, response options appeared on screen and remained until an observer made a response. The verbal response options were the words 'Anger' and 'Disgust' [5.1a], and the non-verbal response options were angry and disgusted expressions of the image category the observer was required to categorise [5.1b]. The next trial commenced following a 500ms inter-trial interval.

5.2.4 Analysis

Using custom-written MATLAB code with the Palamedes toolbox (Prins & Kingdom, 2018), PFs were fitted based on a cumulative Gaussian to estimate the PSE and slope value for each observer, for each condition. Lapse rate was fixed at 0.03; guess rate was determined by the experimental procedure and was fixed at 0. The steeper the slope of the psychometric function, the more reliably the observer distinguished between the morphed stimuli (Kingdom & Prins, 2010). The PSE is the point at which the observer was equally likely to respond disgust or anger to a particular stimulus. The PSE change reflected the modulation of the facial expression judgement due to the contextual influence of body posture. Goodness-of-fit of the PFs was assessed visually and by using the method described in Wichmann and Hill (2001) and implemented in the Palamedes toolbox in MATLAB (Prins & Kingdom, 2018). This resulted in three observers being excluded from the facial expression verbal and non-verbal conditions, one from the body posture verbal and non-verbal conditions, six from the verbal whole-person, and five from the non-verbal whole-person condition. For the MLE model fitting, observers were required to have complete data across all conditions, therefore a total of 26 observers were included in the modelling analysis.

Cue integration model

From the psychophysical data, I calculated how each of the individual face and body cues contributed to the whole-person percept, and how reliable this integrated whole-person representation was.

The estimated slopes of the psychometric functions for the face- and body-only conditions provided a measure of the reliability of the isolated representations, on a subject-by-subject basis. Given the high correlation between observers' performance on verbal and non-verbal versions of the facial expression, body posture and whole-person conditions, the data was pooled across the two experiments. The average estimated slope and PSE was determined for each PF for each condition on an individual subject basis. The amount of sensory evidence associated with a particular stimulus can be represented as a Gaussian curve, with a tall, narrow distribution indicating a more reliable cue (Jeffery et al., 2016).

The PSE is the point at which an observer is as likely to indicate that the stimuli

shown was angry or disgusted. In my observer model, this point is considered to index an observer's fixed internal criterion. If the amount of sensory evidence exceeded the internal criterion, the observer indicated that the stimulus was disgusted. However, if it fell below this criterion the stimuli was categorised as angry. The discrimination threshold is the difference between the PSE, and the observer's performance when the probability of categorising the response as disgust equates to 0.76. This difference corresponds to the standard deviation of the underlying representation [Figure 5.2].

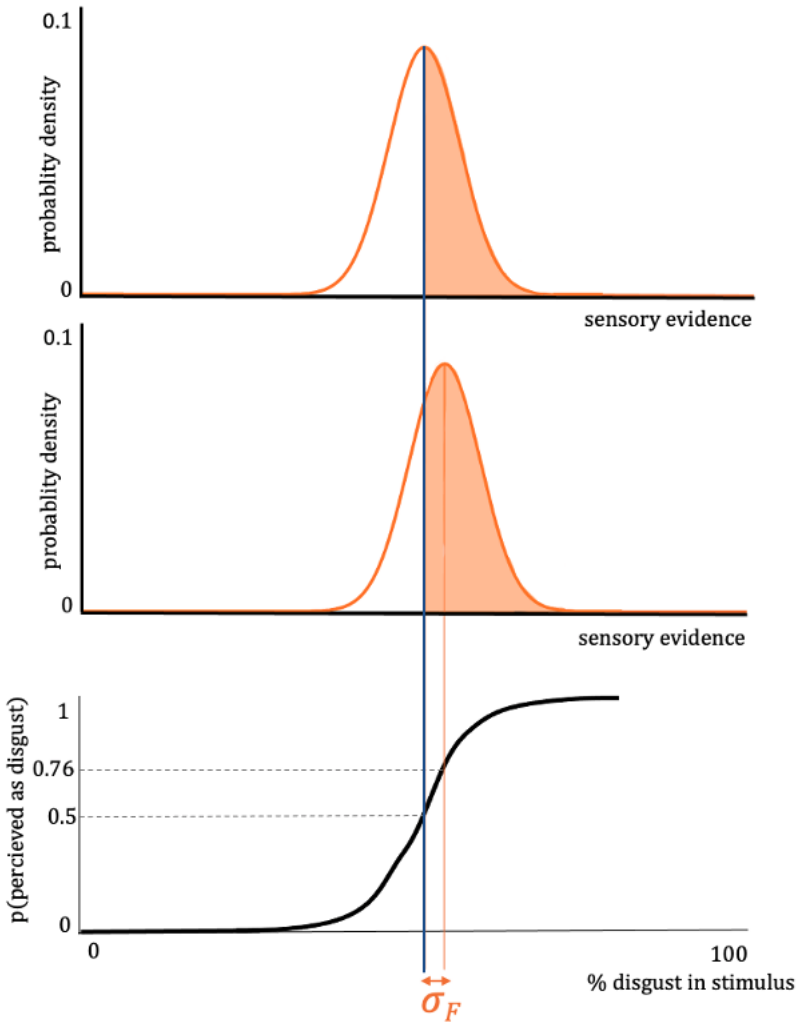


Figure 5.2: **Facial expression condition: Observer model and link to performance**
The figure depicts the perceptual process and how it is linked to performance for one participant. When viewing a facial expression morph the visual system generates a representation that contains 'sensory evidence' for the amount of disgust versus anger in the stimulus. The more disgust a stimulus contains the greater the amount of sensory evidence for disgust, as indicated by the shaded portion of the distribution. The top panel is an example of a morph with approximately equal amounts of anger and disgust in, and the middle panel shows the sensory evidence associated with a stimulus containing a higher percentage of disgust. The observer is thought to compare the amount of sensory evidence against an internal criterion, illustrated by the solid blue line that extends across all three panels. If the sensory evidence exceeded the criterion, the observer would categorise the stimulus as disgusted. If the evidence did not exceed the criterion, the stimulus would be categorised as angry. The observer's resulting psychometric function, shown in the bottom panel, allows the estimation of the underlying facial expression representations: the difference threshold at 0.76 performance is equivalent to the standard deviation (σ) of the underlying facial expression representation. σ_F is the standard deviation of the facial expression representation. This is illustrated in the figure with the orange arrow.

For each observer, a theoretical face weight ($W_{F(T)}$) value was calculated to predict how much the facial expression cue contributed to the whole-person integration. Based on the maximum likelihood principle, the influence of facial expression on the whole-person representation was theoretically derived from the standard deviation of the empirically measured isolated face and body cues:

$$W_{F(T)} = \frac{\frac{1}{\sigma_B^2}}{(\frac{1}{\sigma_B^2} + \frac{1}{\sigma_F^2})} \tag{5.1}$$

The term σ_F denotes the standard deviation of the facial expression representation, and σ_B of the body posture representation. The more reliable the representation, the steeper the PF, and the smaller the standard deviation. The theoretical body weight ($W_{B(T)}$) was calculated using an equivalent approach, but as the combined weights of the face and body equate to 1 ($W_{F(T)} + W_{B(T)} = 1$), for the remainder of the chapter, I refer to the $W_{F(T)}$.

Based on performance in the whole-person condition, the theoretical predictions for the face weights were compared to the empirical data [Figure 5.3]. To calculate the empirical weights, which were the weights that the observers actually used in the whole-person condition, the visual system was assumed to be an unbiased estimator, in line with the majority of work in this area (Scarfe, 2020). This means that the average estimate for a given cue was equivalent to the stimulus level that was presented. The empirical weighting of the face cue in the whole-person representation was calculated by:

$$W_{F(E)} = \frac{PSE_F - E_B}{d} \qquad \text{with} \qquad d = E_F - E_B \tag{5.2}$$

Where PSE_F is the point-of-subjective-equality of the isolated face condition, E_F is the stimulus level of the face morph at the PSE in the whole-person condition, and E_B is the stimulus level of the body posture, which was fixed at either 100% angry or 100% disgusted in the whole-person condition. The empirical weighting of the face cue was first calculated for the averaged (verbal and non-verbal) face presented with an angry body posture, and then for the face presented with the disgusted body. An average was then taken of these two values to determine the theoretical face weight from the empirical whole-person condition, shown above.

From the isolated face and body cues, the theoretically predicted reliability of the whole-person percept was estimated by determining the standard deviation of the integrated whole-person representation (σ_{WP}).

$$\sigma_{WP(T)} = \sqrt{\frac{\sigma_F^2 \times \sigma_B^2}{(\sigma_F^2 + \sigma_B^2)}} \tag{5.3}$$

Therefore, the theoretically predicted reliability is:

$$R_{WP(T)} = \frac{\sigma_F^2 + \sigma_B^2}{(\sigma_F^2 \times \sigma_B^2)} \tag{5.4}$$

Finally, the empirically-measured reliability of the whole-person representation was derived directly from the estimated slope of the whole-person PF. The empirical whole-person reliability was defined as:

$$R_{WP(E)} = \frac{1}{\sigma_{WP}^2} \tag{5.5}$$

This facilitated comparison between the predicted reliability of the whole-person integration, calculated from the face and body cues, with the empirical values.

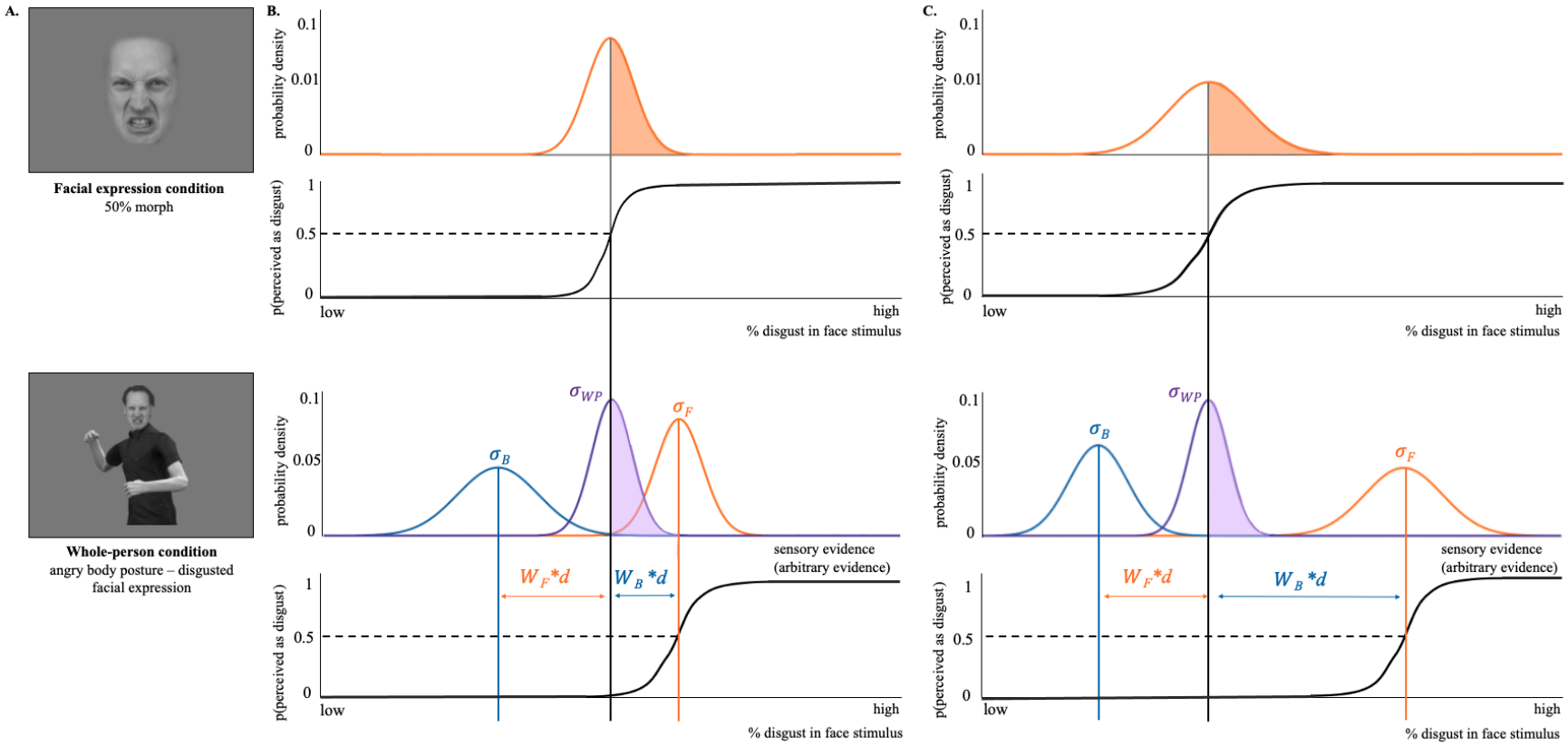
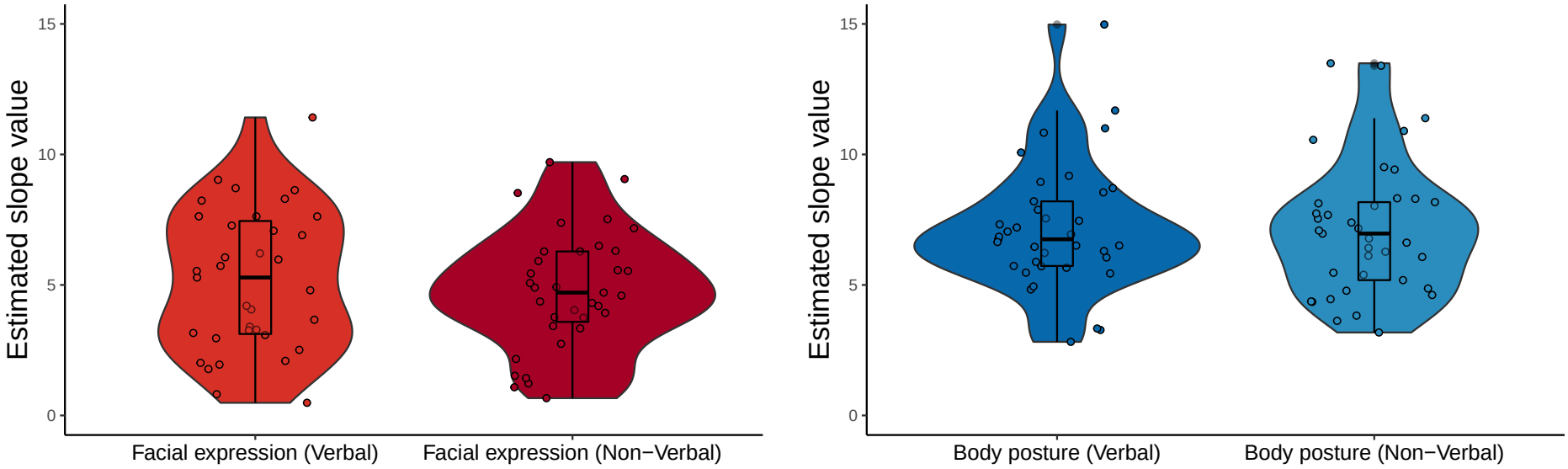


Figure 5.3: **Whole-person condition: Observer model and link to performance**
Panel A displays example stimuli for the facial expression condition, and the whole-person condition. Panel B and C illustrate the perceptual process, and the link with performance in two separate observers. The observer in panel B has a more reliable facial expression (orange) than body posture representation (blue), the observer in panel C shows the opposite relationship. The sensory evidence associated with the whole-person representation (purple) is achieved by integrating face and body representations according to the MLE principle. The face weight is represented by W_F, and the body weight by W_B. The weights were both multiplied by a constant (d) [Equation 5.2].

5.3 Results

5.3.1 Isolated facial expression and body posture categorisation

Observers could reliably distinguish between angry and disgusted facial expressions and body postures, as indexed by the estimated PF slope values, for both the verbal and non-verbal conditions. A paired t-test revealed no significant difference in the facial expression precision between the verbal and non-verbal response tasks ($t(34)=1.119$, $p=0.27$), or between the body posture verbal and non-verbal response tasks ($t(36)=0.087$, $p=0.931$), suggesting observers' reliability was not different for the verbal and non-verbal response tasks [Figure 5.4]. Observers' precision, when categorising facial expressions in isolation, was highly correlated between the verbal and non-verbal response tasks ($r_s=0.634$, $p<0.001$). For the body posture condition, there was also a strong positive correlation between observers' performance for the verbal and non-verbal response tasks ($r_s=0.540$, $p<0.001$). These strong positive correlations indicate that observers' performance in categorising the individual facial expression and body posture cues was consistent, regardless of response option type.



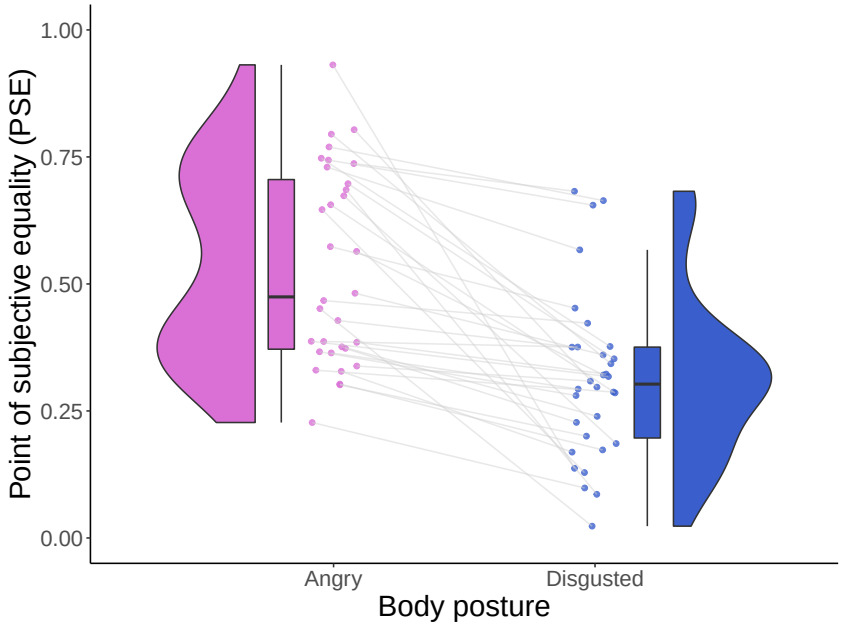
(a) Facial expression precision

(b) Body posture precision

Figure 5.4: **Estimated slope values from PFs for facial expression and body posture categorisation for verbal and non-verbal response option tasks** The violin plots display the estimated slope values for the facial expression [5.4a] and body posture conditions [5.4b] for the verbal and non-verbal response option tasks. No difference was observed between the facial expression precision for the verbal and non-verbal response tasks, or for the body posture precision. Each point represents one observer. The distribution of the values is illustrated by the shaded area, with the overlaid boxplot indicating the median and the interquartile range.

5.3.2 Whole-person categorisation

Body posture was found to have a significant influence on facial expression perception in both the verbal and non-verbal whole-person conditions. A Wilcoxon signed rank test revealed a significant difference in the PSE values between the facial expressions presented on 100% angry and 100% disgusted body postures for the verbal response option condition ($z=1.39$, $p<0.0001$), suggesting that overall, there was a significant influence of body posture on facial expression perception [Figure 5.5a]. For the non-verbal condition, a paired t-test also revealed a significant difference between the PSE of facial expressions presented with a 100% angry and 100% disgusted body postures ($t(32)= 4.923$, $p<0.0001$) [Figure 5.5b]. This clearly illustrates that regardless of response option type, body posture significantly biased observers' perception of the facial expression. These results suggest that the biasing effect is truly perceptual, as even when facial expressions were presented as response options following the test stimulus, observers selected the expression that captured the perceived facial expression, biased by the body posture. The PSE change was highly correlated between the verbal and non-verbal response option for the whole-person condition ($r_s=0.499$, $p<0.01$) [Figure 5.6].



(a) Verbal change in PSE

(b) Non-verbal change in PSE

Figure 5.5: **Change in PSE between facial expression judgements on a 100% angry and a 100% disgusted body posture** The raincloud plots display the PSE for the facial expression morphs presented with an angry or disgusted body posture from the whole-person condition for the verbal [5.5a] and [5.5b] non-verbal response option tasks. The distribution of the values is illustrated by the shaded area, with the boxplot indicating the median and the interquartile range. Each line represents one observer and depicts the change in PSE.

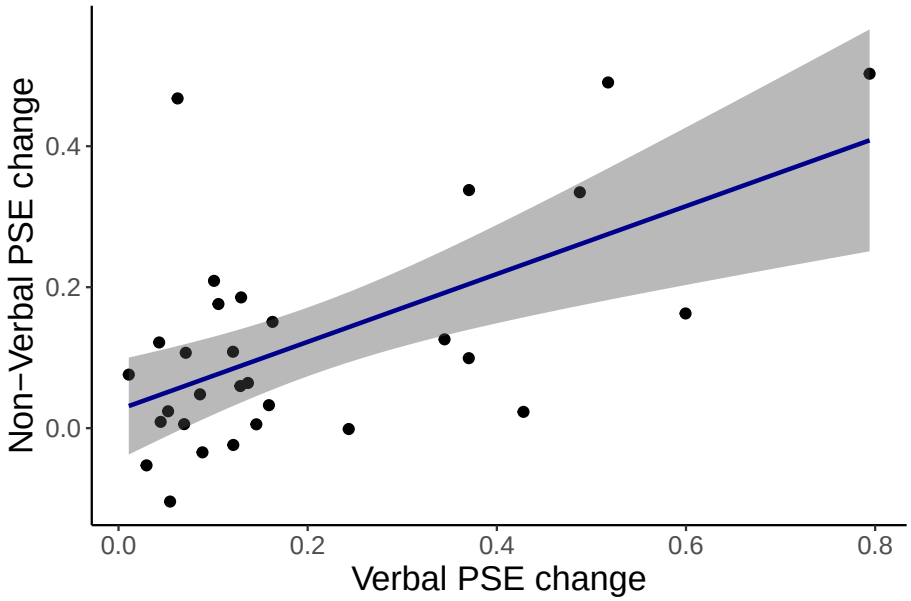


Figure 5.6: **Relationship between PSE change for verbal and non-verbal response options**
The correlation plot displays the relationship between the PSE change for the verbal and non-verbal response options. The PSE change was highly correlated between the verbal and non-verbal response option tasks (r_s=0.499, p<0.01). Each point denotes one observer, and a smoothed linear regression line is shown in blue with a 95% confidence interval shaded.

A significant negative relationship was found between isolated facial expression precision and the influence of body posture on facial expression perception for the verbal response option tasks (r_s=-0.494, p<0.01). This result echoes my findings from Chapter 3, demonstrating that observers' who had very precise facial expression representations were less influenced by the body posture when judging the facial expression in the whole-person condition, and vice versa. For the non-verbal conditions, a negative trend was found between facial expression precision and PSE change, however it did not reach significance (r_s=-0.297, p=0.09).

No relationship was found between body posture precision and PSE change for the verbal conditions (r_s=-0.159, p=0.38), but a significant negative relationship was found for the non-verbal conditions (r_s=-0.426, p=0.014).

5.3.3 Reaction times

For each observer, the average reaction time (RT) was calculated per condition by averaging the RTs over each trial. The RTs were found to be slower for the non-verbal conditions, compared to the equivalent verbal conditions (Face: $t(34)=-8.802$, $p<0.0001$; Body: $t(36)=-4.374$, $p<0.0001$; Whole-person: $t(29)=-7.336$, $p<0.0001$) [Figure 5.7]. This suggests that task demands for the non-verbal response options were more challenging, thus requiring more time for the observers to respond. However, even with the slower RTs, and a greater 'decision time', a significant influence of body posture on perception of facial expressions is still seen. This result adds additional support to the argument that the body context effect is truly perceptual. Comparing the RTs between the facial expression condition and whole-person condition for the verbal response options, there was no significant difference in RTs ($t(30)=-1.367$, $p=0.182$). The same was also true for the non-verbal facial expression and whole-person condition ($t(32)=0.607$, $p=0.548$). This finding indicates, that on average, observers' judgement of a facial expression was not significantly slower in the presence of a body posture within the verbal or non-verbal response option tasks. If a difference had been observed, it could have indicated that the overall effect of body posture influencing perception of the facial expression was due to differences in RTs. The fact that a difference is not observed further supports the argument that body posture biasing facial expression perception is truly perceptual.

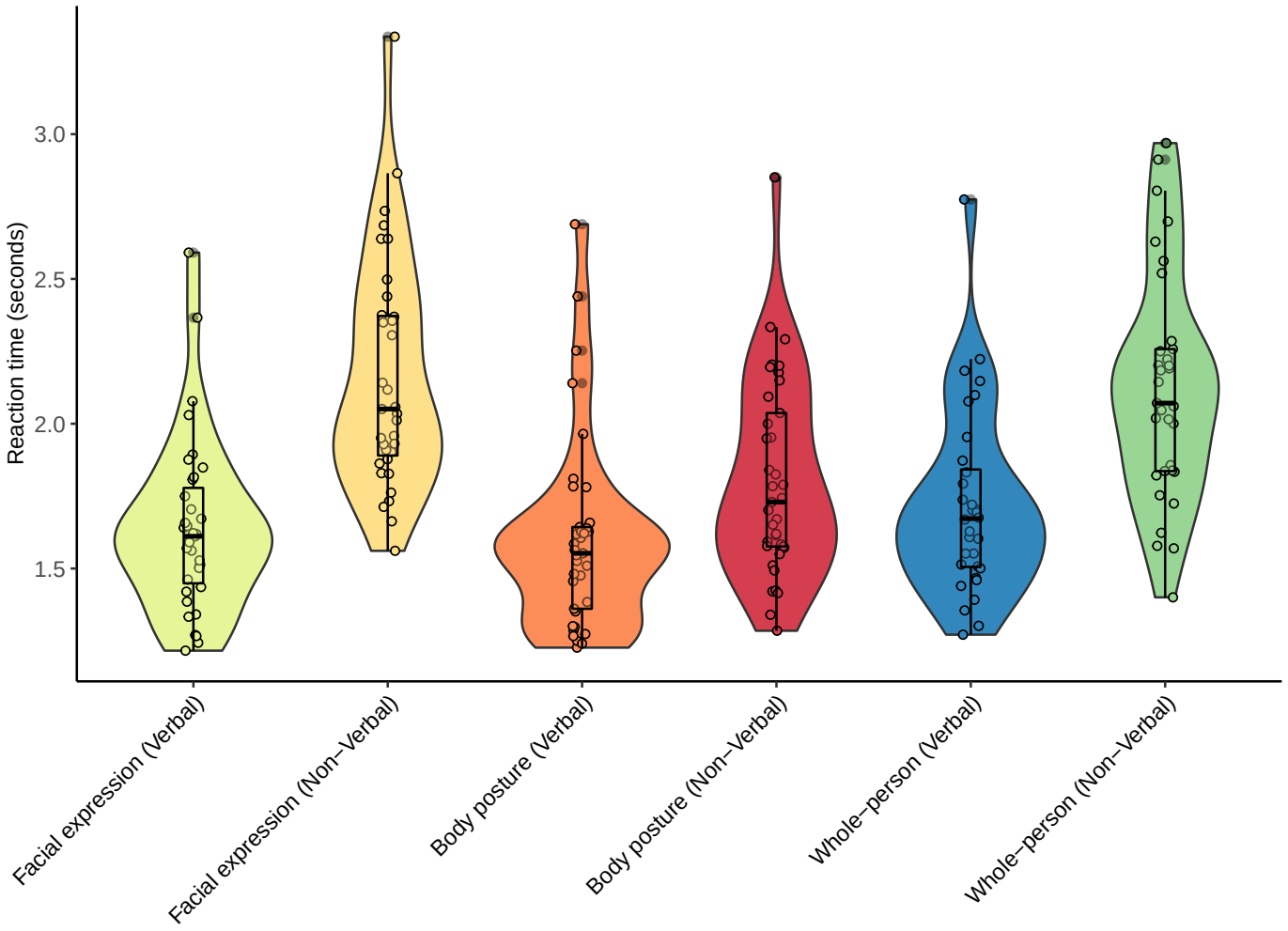
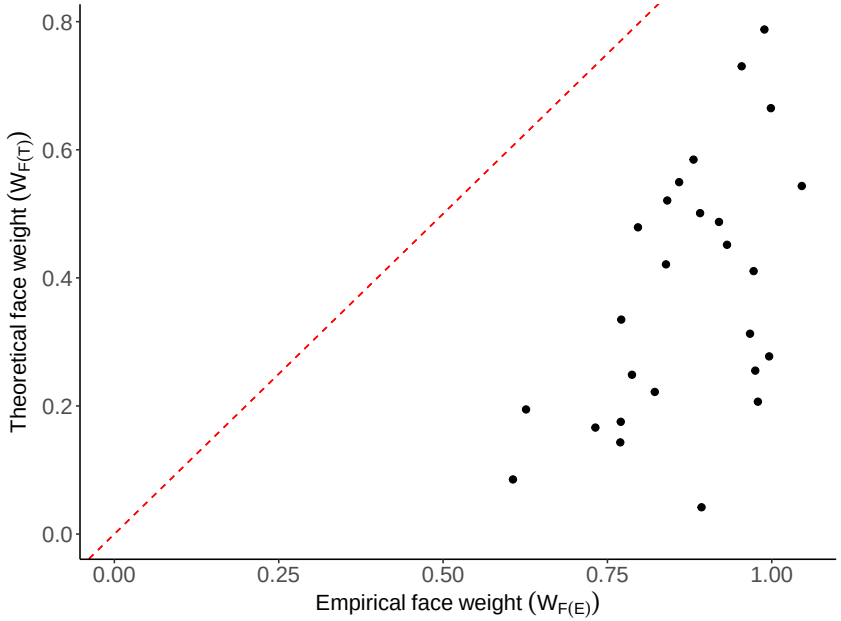


Figure 5.7: **Average reaction times**
The violin plot displays the average reaction time for each observer, calculated by averaging over all trials. Each condition is shown in a separate violin with the verbal and non-verbal response option tasks displayed side-by-side. Each point represents one observer. The distribution of the values is illustrated by the shaded area, with the overlaid boxplot indicating the median and the interquartile range.

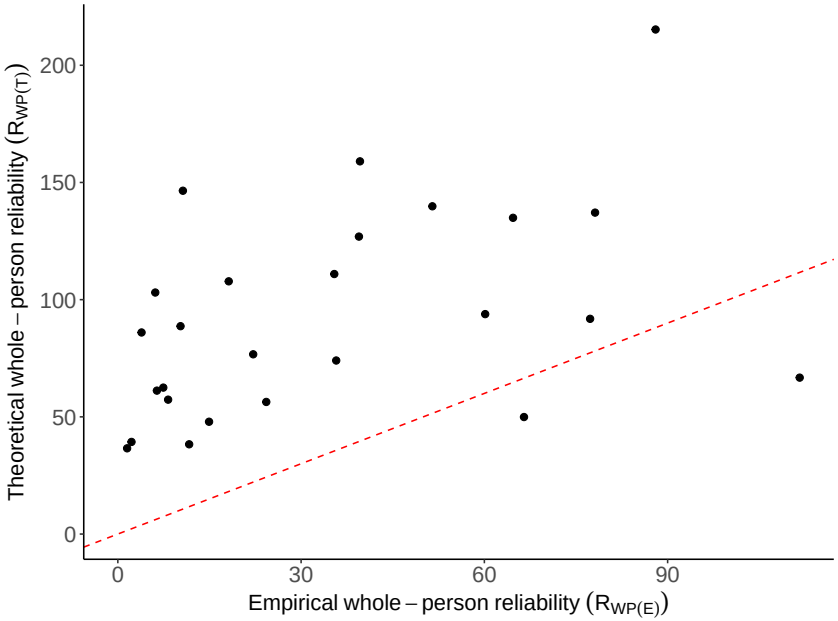
5.3.4 Cue integration modelling

I found a significant positive correlation between the theoretically calculated face weights, and empirically determined face weights, on a subject-by-subject basis ($r(24)=0.5309$, $p<0.01$) [Figure 5.8a]. This suggests that observers performed in a manner predicted by a MLE cue integration model. However, there was a significant difference between the theoretical and empirical values ($t(25)=14.71$, $p<0.0001$), with the empirical face weights being higher than the theoretical weights. This suggests that there was a deviation from optimality, such that observers were weighting the face cue as more reliable (or the body cue as less reliable) in real-life than would be predicted by MLE for an optimal observer.

For the whole-person precision, the empirical and theoretical calculations were found to be significantly positively correlated ($r(24)=0.4151$, $p=0.0350$) [Figure 5.8b]. Furthermore, the values were found to be significantly different from one another ($t(25)=-7.17$, $p<0.001$). Again, this indicates that observers deviated from optimality, as the reliability of the empirically measured integrated representation was lower than predicted. Under a MLE cue integration model, the integrated representation is predicted to be more reliable, as it makes use of the uncertain sensory cues to form a more precise estimate about the integrated percept. However, in the current application of the model, this is not the case. The whole-person precision was not always found to be greater than either of the individual cues [Figure 5.9].



(a) Relationship between theoretical and empirical face weights



(b) Relationship between theoretical and empirical whole-person precision

Figure 5.8: **Comparison of MLE cue integration theoretical predictions with empirical values**
The correlation plots displayed show the relationship between the predicted and empirical face weights [5.8a] and whole-person reliability [5.8b] The dashed red line in both plots illustrates the expected relationship for an optimal observer. Each point represents one observer.

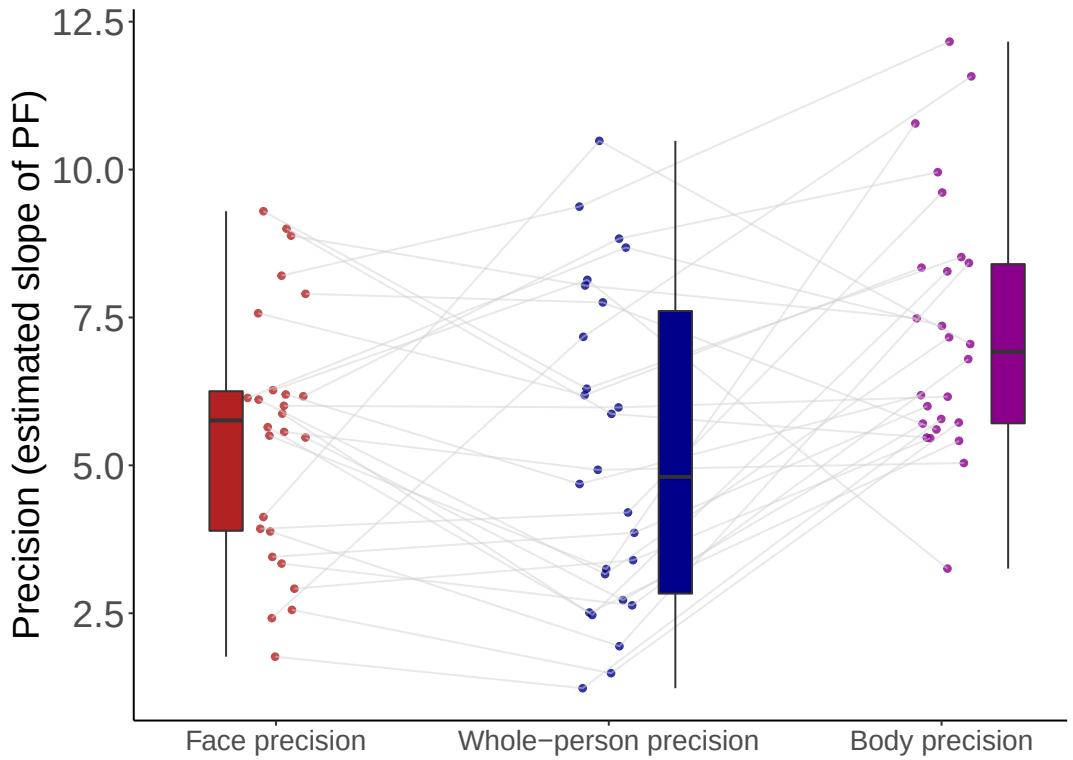


Figure 5.9: **Reliability of face, body and whole-person cues**
The estimated slope values are shown for each observer for the face, body and whole-person. The steeper the slope of the PF, the more reliable the estimation, therefore the smaller the standard deviation (σ). Each line connects one observer's reliability across for each of the representations. Based on MLE predictions, the whole-person precision (integrated representation) should be more reliable than either of the individual cues. In the current results, the whole-person representation is not always more reliable than the individual face and body cues. A box and whisker plot illustrates the median and interquartile range of the data for each representation.

In summary, the predictions made about the weighting of the face cue, and the whole-person reliability were highly correlated with the empirical calculations. However, these results also demonstrate that observers' performance was not optimal as defined by MLE.

5.4 Discussion

In the present chapter I demonstrate the successful application of a formal mathematical cue integration model for facial expression and body posture integration. Finding a relationship between the theoretical predictions and the empirical values for both the face weights and whole person reliability suggests that the model captures some aspects of this integration. However, a systematic deviation between the model and empirical data was observed, suggesting the model is missing crucial components of how humans integrate these social signals. These results reveal that observers' typically weight the facial expression cue as more reliable than predicted, or conversely treat the body posture cue as less reliable. Taken together, this means that observers' deviate from optimality when integrating facial expression and body posture cues. In addition, I provide evidence to support the claim that the influence of body posture on facial expression perception is perceptual. Adopting a within-observer design, I show that regardless of response option (verbal or non-verbal), the body biasing effect emerges, and there is a tight relationship between the verbal and non-verbal response option tasks.

Matching tasks are believed to be more perceptual than labelling tasks, as they simply require participants to discriminate between expressions on the basis of the visual properties alone (Palermo et al., 2013). Using verbal labels as response options requires additional verbalization of responses, which is thought to be more cognitively demanding due to the requirement of additional vocabulary skills. Previous empirical work comparing response option types in expression categorisation tasks has been mixed; some work finds no correlation between verbal and non-verbal facial expression labelling abilities (Croker & McDonald, 2005), whilst other work finds a tight correlation between facial expression labelling with verbal and non-verbal options (Addington & Addington, 1998). In the present study, I found a tight relationship between the verbal and non-verbal response options for isolated facial expression and body posture conditions, suggesting that observers' categorisation ability was unaffected by the format of the response option. Interestingly, the reaction times for the non-verbal tasks were significantly slower than the verbal equivalent, however no difference was observed in the categorisation of the expressions. In addition, I also found a strong positive correlation between the verbal and non-verbal response options for the influence of body context on facial expression perception. This

within-observer evidence suggests the claim made by previous authors that the body biasing effect is truly perceptual. However, the magnitude of PSE change appeared to be smaller in the non-verbal iteration of the task. One explanation for this finding could be a consequence of the task demands. Observers were required to select a facial expression that best captured the emotion displayed from the facial expression presented with an expressive body posture. Whilst a different identity was shown in the response choice stimuli to the test stimulus in order to minimise observers using a matching strategy, this cannot completely mitigate the possibility participants were still trying to use some element of matching between the choice and response options. One indication of this could be seen with the slower RTs observed in the non-verbal condition. To further mitigate this possibility, subsequent research could consider using non-verbal stimuli that are more different to the test stimuli, for example, expressions which are non-Caucasian or non-male.

It has been proposed that there is a distinction between two types of visual processing; a perceptual process where information about the visual input is extracted, followed by a higher-level decision making process evaluating the relevance of the visual information related to goals and expectations, and then generation of an appropriate behavioural response (VanRullen & Thorpe, 2001). Empirically disentangling the stages underlying conscious perception is one of the challenges modern psychology faces. In the current work the influence of body posture on facial expression perception is believed to be perceptual due to the evidence presented in this chapter and previous work (Aviezer et al., 2011; Meeren et al., 2005; Teufel et al., 2019), however it should be highlighted that other attention or decision making processes could be at play and should be assessed in subsequent research.

Despite the observed deviation from optimality in the current work, I found a significant positive correlation between the theoretical model predictions and the empirical values for the face weights and whole-person precision. Therefore, conceptualising whole person integration under a MLE framework does provide some insight into how face and body cues are combined, suggesting that the integrated whole-person representation is a weighting of the reliability of the isolated face and body cues. One advantage of the current model fitting was the use of naturally occurring variability between participants for the MLE calculations. Typically, cue integration modelling relies on experimentally manipu-

lating the stimuli within observers to determine how differences in reliability contribute to perception. In Chapters 3 and 4, large individual differences in the magnitude of the body biasing effect were found. Individual differences reported between observers often arise from real differences in optical, neural and perceptual differences that mediate perception (Mollon et al., 2017). Therefore, using naturally occurring variability between observers', the model captures real world individual differences between individuals' perceptual experiences, thus aligning closer to the underlying ecological mechanism of whole-person emotion perception.

Multisensory information is believed to be integrated in the brain according to the incoming sensory cues' relative reliability (Helbig & Ernst, 2007; Meijer et al., 2019). Single unit recordings in macaques have shed light onto the neuronal correlates of cue integration (Fetsch et al., 2013). Fetsch and colleagues (2013) report that neurons in the dorsal medial superior temporal area of the macaque were modulated by changes in cue reliability across time. Although in Festsch's paradigm cue integration was across visual and vestibular senses, this work provides direct evidence for a neuronal mechanism mediating the mathematical combination of cues. Therefore, the fact that the theoretical predictions for the relative contribution of the face cue to the whole-person representation were highly correlated with the empirical measurements, provides some insight into the statistical inference of the brain when solving this sensory challenge.

One explanation for the empirical face weighting being higher than theoretically predicted by MLE integration, could be a consequence of the task. Despite the model conceptualising integration as a 'whole-person' cue, observers' were specifically instructed to categorise the facial expression and ignore the body posture. Therefore, the fact that the empirical face weights were higher than the theoretical predictions, could be a reflection of the task demands. This could indicate that observers' have cognitive access to re-weighting cues, despite their sensory reliability, as a function of the task demands. One approach to further explore this proposal would involve a modification of the current study, where the whole-person condition would consist of body posture morphs presented with 100% angry or 100% disgusted facial expressions. Observers would be explicitly instructed to categorise the body posture in this whole-person task. If observers were able to re-weight the incoming information depending on the task demands, as I have suggested from the

current results above, I would expect the empirical body posture cue to be given a higher weighting in this iteration of the study.

An alternative explanation for the observed deviation from optimality could be that my definition of optimality does not capture the real complexity of the task. In contrast to low level visual cues, where MLE modelling is commonplace, the information which must be integrated from higher-order social cues is perceptually rich. In the instance of facial expression and body posture cues, each individual has distinct previous experience with such cues. Given the complexity of assessing these additional cognitive factors, it may be more suitable to assess the integration of these cues under another framework that allows explicit modelling of such prior assumptions.

Furthermore, there are some disparities between classical descriptions of cue integration in the literature and its implementation in the present chapter to model higher order-social cues. In adults, typically under a MLE cue integration framework, the integrated estimate is a third entity and a consequence of 'fusion'. This fusion results in loss of access to the individual cues (Hillis, 2002; Nardini et al., 2010). A challenge with the application of a cue integration framework to whole-person integration of facial expression and body posture, is that the integrated percept is not a separate third entity; participants were instructed to judge the face in the context of a whole-person. I also found that the integrated precision of the whole-person estimate was not always higher than the reliability of the individual cues. One of the predictions of MLE cue integration is that the combined representation should be more reliable than either cue alone (Ernst & Banks, 2002). In order to overcome some of the limitations of the present work, future research should assess the suitability of other frameworks to model the integration of facial expression and body posture. One alternative framework to consider is a causal inference model. Causal inference is the process of inferring whether one event was caused by another and is a problem the human perceptual system must solve (Shams & Beierholm, 2010). Cue integration models assume that different signals are all caused by the same source, therefore model how the nervous system combines these signals under assumed unity. To this end, MLE models estimate the integration based on a weighted average of the estimate for each cue and does not consider any prior information. However, in the real world multiple sensory signals can have multiple sources and do not 'fuse' as cue

integration depicts. To overcome these limitations a Bayesian framework, such as causal inference, should be explored for future research.

5.5 Chapter Summary

In the present chapter, I applied a MLE model of cue integration to determine if the integration of facial expression and body posture cues was statistically optimal. I illustrate that the integrated emotion percept of a whole-person can be predicted by combining the weighted reliability of the individual face and body cues of emotion. In Chapter 3, I reported a relationship between facial expression precision and the magnitude of the body biasing effect. In the current work, I developed this further to illustrate the reliability of the face cue not only correlates with whole-person perception, but contributes to the integrated representation as a weighted average of the cues underlying sensory reliability on a subject-by-subject basis. Taken together, the results in this chapter have impacts for understanding the neural processing of facial expression and body posture cues and appreciating the complexity of higher order social cue processing in the brain. Future work is needed to optimise the framework for combination of higher-order social cues, considering the complexity of the signals. This work is the first of its kind, to my knowledge, to apply a mathematical framework to the integration of facial expression and body posture cues and furthers the current state of the literature to provide a more mechanistic account for the integration of facial expression and body posture cues.

Chapter 6

General Discussion

In the following discussion I will highlight the key findings from each of my experimental chapters, discuss the relevance of the findings, and highlight their implications in the wider context of emotion perception more generally.

6.1 Overview of thesis results

In Chapter 3, I provided empirical evidence demonstrating that facial expression perception can be biased by an affective body posture, and the magnitude of this bias is variable between observers. The reliability with which an observer could categorise facial expression morphs between anger and disgust was found to be negatively correlated with the magnitude of the body biasing effect, such that observers' who were more reliably able to categorise isolated facial expressions were less influenced by body posture, and vice versa. In turn, observers' who had less precise isolated facial expression representations were more biased by the body posture in the whole-person condition. This research quantified the influence of body posture on facial expression perception, as well as measuring the reliability of isolated facial expression and body posture representations. The results highlight evidence for a relationship between individual cue reliability and whole-person representation.

In Chapter 4, I explored whole-person perception across development. Previous research has highlighted that children's ability to categorise facial expressions improves with increasing age (Herba et al., 2006; Thomas et al., 2007). Specifically, the ability to detect subtle differences in morphed facial expressions improves with age (Dalrymple et al.,

2017). In Chapter 4, I replicated this finding, showing that in 8 to 18-year-olds the ability to recognise emotion from morphed facial expressions of anger and disgust significantly improved with increasing age. Interestingly, I also found that with increasing age, children's judgements of facial expressions were less biased by an affective body posture. These findings echo my results from Chapter 3; as isolated facial expression recognition improved, the influence of body context on facial expression perception decreased. In Chapter 4, I also provide evidence for the first time, to my knowledge, that children's recognition of body postures improves with increasing age.

The work in Chapter 4 also uncovers microstructural changes in functionally-relevant white matter that were predictive of children's individual differences in facial expression, body posture and whole-person perception. It has previously been established that face and body processing is dependent on an extended network of brain areas. Therefore, to explore changes in white-matter specifically related to the perception of these cues, I delineated the tracts connecting key cortical nodes involved in face and body processing. This resulted in a network of functionally-defined white matter tracts for face and body processing encompassing regions in the visual processing hierarchy that have been associated with having category-specific neural responses to faces or bodies. The fractional anisotropy of tracts connecting occipital face area and fusiform face area was predictive of individual differences in facial expression recognition in children. This work provides support for recent functional MRI accounts indicating that the ventral visual stream plays an important role in facial expression processing, and is not exclusively dedicated to identity processing (Dima et al., 2018; Duchaine & Yovel, 2015). In addition to changes in fractional anisotropy in relation to facial expression perception, a metric more sensitive to intracellular signal, the spherical mean, was found to reflect developmental changes of tracts in both the face and body networks. The spherical mean provides some insight into the underlying biophysical changes that could be driving development in these functional circuits across adolescence.

In Chapter 5, I applied a formal mathematical model for cue integration to determine if observers integrated facial expression and body posture cues in a statistically optimal manner. Based on the principle of MLE, it is thought that more reliable cues contribute more to the integrated representation. To determine if this was the case in whole-person

perception, I generated theoretical predictions about how much the face and body cues should contribute to the integrated representation under an ideal observer model, and compared these to my empirical data. I found a tight correlation between the theoretical and empirical calculations for both how much the face cue contributed to the integrated percept, and for the precision of the whole-person representation. These results indicate that a cue-integration framework goes some way to conceptualise how the human brain makes sense of the noisy social world, and integration of face and body cue appears to be dependent on a weighted average of the reliability of the incoming sensory cues. However, despite this observed relationship, a deviation from optimality was found. Namely, observers typically weighted the face cue as more reliable empirically, than would be predicted from cue integration modelling. Alternatively, observers may have been weighting the body posture as less reliable in real life than theoretically predicted. In addition, the empirical whole-person precision was found to be lower than theoretically predicted. Therefore, observers were not found to be performing in an 'optimal' manner, with this work raising interesting questions regarding the mechanisms responsible for integration of face and body cues. The work in Chapter 5 capitalises on individual differences observed to provide a mechanistic understanding of how the human brain integrates facial expression and body posture cues, and to my knowledge is the first formal mathematical framework which has attempted to model the perception of higher-order visual social cues in this way.

6.2 Clinical relevance

Difficulties in facial expression perception are characteristic of several psychiatric and neurological conditions (Calder, 2011). Brewer and colleagues (2017) explored how contextual cues from body postures influenced facial expression perception in a group of individuals diagnosed with Autism spectrum disorder (ASD). People with ASD are believed to employ more 'local processing' strategies to cope with incoming information, to the detriment of global processing. Brewer *et al.* (2017) hypothesised that as a consequence of local processing strategies, observers with ASD should be less affected by body posture in their perception of the face. Surprisingly, despite explicit instructions to disregard the

body context, the ASD group were biased by the body posture to a similar extent as typically developing (TD) controls (Brewer et al., 2017). The work in my thesis may provide insights into why no differences were observed between ASD and TD participants in the Brewer *et al.* (2017) study. Throughout my thesis, I have consistently found an association between the reliability of the isolated facial expression representations and the influence of body context. My work would suggest a reinterpretation of Brewer *et al.*'s (2017) result to be a consequence of the equivalent reliability of isolated facial expression cues in ASD effectively 'cancelling' any reduction in the influence of body posture due to their enhanced local processing. The literature regarding isolated facial expression processing in autistic individuals is limited, with some research reporting deficits, whilst other work does not (Keating & Cook, 2021). In order to understand whole-person processing of emotion in ASD, it will be important to characterise individual differences in responses to the individual cues, as well as to the whole-person cues. Such an approach is more likely to provide insights into the mechanisms underlying integration of these social signals in ASD.

Another very different clinical condition with known difficulties in emotion perception is frontotemporal dementia (FTD). FTD is a relatively uncommon type of dementia that causes a broad range of symptoms including changes in behaviour, language, executive control and motor symptoms (Olney et al., 2017). Kumfor *et al.* (2018) explored contextual processing of faces in patients diagnosed with two variants of FTD: behavioural variant FTD (bvFTD) and semantic variant, also referred to as semantic dementia (SD). Facial expression recognition in both bvFTD and SD was found to be impaired relative to a control group (Kumfor et al., 2018). When presented with an incongruent facial expression and body posture, the bvFTD group were found to make more miscategorisations of the facial expression than either SD patients or controls. This shows an interesting behavioural distinction in contextual processing between subgroups of FTD patients, with bvFTD patients seemingly over-reliant on the contextual cues of the body posture despite having similarly impaired facial expression perception as the SD group. In contrast to my results where isolated facial expression recognition was found to be predictive of the contextual influence of body context, these clinical findings suggest a different mechanism could be at play between different types of FTD. In bvFTD, atrophy of the right temporal

lobe is associated with changes in behaviour, personality and emotion processing (Olney et al., 2017). In SD patients, temporal lobe atrophy is left hemisphere dominant with observed deficits being language-based with a progressive loss of semantic knowledge. The fact that observers with bvFTD are more affected by the context in face perception as reported by Kumfor *et al.* (2018), and are known to have atrophied temporal lobes, predominantly in the right hemisphere, provides support for a key role of the right anterior temporal lobe in integration of face and body cues. To summarise, work with FTD patients indicates that the mechanism for integration of facial expression and body posture may be different in this clinical presentation and subsequent research exploring contextual processing of facial expressions in a range of neurological conditions is warranted.

6.3 Integration of facial expression and body posture cues

There remains conflict in the literature regarding exactly how facial expression and body posture cues are integrated. Some researchers claim that the face and body are processed as 'gestalt' unit, and this processing is automatic and occurs early in the visual system (Aviezer et al., 2011; Meeren et al., 2005), whilst more recently Teufel (2019) and colleagues have challenged this early integration account. Using an adaptation paradigm, Teufel *et al.* (2019) found evidence for late integration of facial expression and body posture cues, downstream of early core face perception regions, such as FFA and pSTS.

Work exploring both spatial and temporal alignment disruption of facial expression and body posture cues highlights that the body biasing effects can be affected (Aviezer et al., 2012b; Lecker et al., 2017). When a facial expression was misaligned from the body posture, the affective posture was no longer found to bias the perception of the face (Aviezer et al., 2012b). Lecker *et al.* (2017) introduced a short temporal delay between facial expression and body posture presentation, and even with immediate sequential presentation of the cues, the body context effect was diminished (Lecker et al., 2017). Taken together, these studies indicate that face and body emotion cues are processed as one whole-perceptual unit, and breaking this form disrupts processing. However, despite work exploring if the face and body are processed as one-unit, limited work aims to resolve if the

body context effect is truly perceptual or is a consequence of post-perceptual processes. In Chapter 5, I provide empirical within-observer evidence to suggest that the body biasing effect is truly perceptual. Modifying the response options in one version of the task to include non-verbal options, thought to be a better marker of perception (Palermo et al., 2013), I found that body context biased observers' perception of facial expression perception regardless of response option type. Even with slower response times in the non-verbal whole-person task, observers were still found to be affected by the body posture in their perception of the facial expression. These findings combine to support the notion that the effect of body context on facial expression perception is truly perceptual.

A recent review paper from Hu *et al.* (2020) proposed a theoretical model for facial expression and body posture processing in the brain [Figure 6.1]. They proposed that there are two sites for integration in the brain, a visual semantic integration hub and a social agent hub, which depend on the task to be accomplished and the stage of neural processing (Hu et al., 2020). The dorsal social agent hub encompasses the pSTS and anterior STS and is thought to play a role in biological motion and multi-modal person integration. The ventral visual semantic hub extends from early regions such as EBA and OFA to face and body selective regions in the fusiform cortex, culminating in high-level person perception in the ATL. These ideas align with the Hub-and-Spoke model for semantic memory and conceptual knowledge, which suggests the ATL plays a critical role in representing information in an amodal manner (Patterson & Lambon Ralph, 2016).

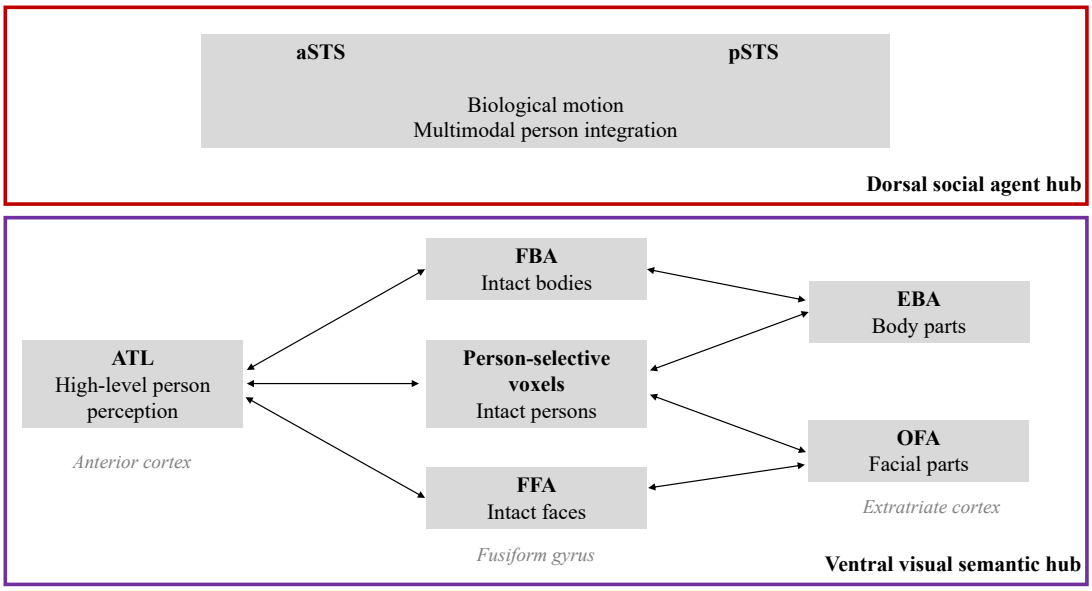


Figure 6.1: **Proposed model of the neural hierarchical structure that underlies face and body integration**
A theoretical model proposed by Hu and colleagues (2020) for the neural processing of faces and bodies. The two proposed sites for integration are illustrated with the dorsal social agent hub shown on the top panel and the visual semantic hub on the bottom. Figure recreated from Hu *et al.*, 2020.

Despite these ideas regarding the site for integration of facial expression and body posture being alluded to in the literature, the mechanism for integration remains unknown. In Chapter 3, I found that the worse an observer's isolated facial expression representation, the greater the biasing influence of the body context on their perception the face. These results suggest that the reliability of the individual cue encoding may be important in determining the integrated whole-person representation. One proposed mechanism for individual differences in categorisation of isolated facial expressions has been attributed to an observers' gaze patterns and personal traits (Green & Guo, 2018). One striking finding is that higher anxiety traits have been associated with improved recognition of emotion across a range of expressions. In addition, personal traits have also been shown to underlie differences in facial expression recognition, for example it has been shown that individuals who have more aggressive tendencies are more likely to categorise an ambiguous facial expression as angry (Penton-Voak et al., 2013). Surprisingly, in Chapter 3, I did not find a relationship between individuals' ability to recognise isolated facial expression cues, body posture cues or the magnitude of the influence of body posture on

facial expression. Further work should aim to resolve if there are links between personal traits and individual differences in whole-person perception.

In Chapter 5, I formally assessed how the reliability of the individual facial expression and body posture cues are combined in whole-person perception. Adopting a MLE cue integration framework, I sought to assess if observers' integration of facial expression and body posture was as predicted from this model. Whilst this framework reflected some features of the integration, a deviation from optimality was observed, suggesting the model was missing crucial components of how humans integrate these social signals. Importantly, the MLE model only accounts for the incoming sensory information in the final perceptual estimate. Given the complexity of social cues, additional frameworks should be considered that account for the additional demands of processing facial expression and body posture cues. A Bayesian approach to person perception has recently been proposed to combine expectations about the environment with sensory evidence (Clifford et al., 2015). Such a framework is useful in considering additional factors, beyond sensory input, that affect perception.

6.4 Processing of facial expression and body posture cues in the brain

Functional analogies have been drawn between key cortical processing nodes in the face and body processing system. For example, the EBA involved in body processing, and face sensitive OFA, are both relatively early processing regions in the face and body networks. Interestingly, there are regions of overlap between neural regions that process bodies, and those that have been associated with the perception of faces (Downing, 2001). Areas of cortex which are functionally specialised for face processing have also been shown to be modulated by body stimuli (Cox, 2004). Cox and colleagues (2004) presented blurred faces with body postures and found that the neural response in the FFA was comparable to the response evoked when viewing high-resolution non-blurred faces. When the blurred face stimuli were presented in isolation the neural response was attenuated, clearly highlighting the importance of the contextual information, for the detection of a 'face'. More recent research has demonstrated that representation of face and body stimuli along the ventral

visual stream to the anterior temporal lobe (ATL) is hierarchical (Collins & Olson, 2014; Harry et al., 2016; Song et al., 2013). This research supports a convergence of face and body information along a visual hierarchy extending from the occipital cortex to the anterior temporal pole.

In addition to fMRI evidence, direct recording from implanted electrodes in the ventro-occipital temporal cortex in humans has confirmed face selectivity along the hierarchy, even in ATL (Jonas et al., 2016). Non-human primate work has also provided further insight into the networks of face and body processing in the brain, revealing hierarchical networks responsible for the processing of faces and bodies, supporting the recent findings in humans (Arcaro et al., 2020; Fisher & Freiwald, 2015; Premereur et al., 2016). This work is important due to the methodological challenges associated with scanning anterior temporal regions of the brain in humans and supports the recent proposal from Hu *et al.* (2020) depicting the importance of the anterior temporal regions for whole-person representation.

6.5 Development of facial expression and body posture recognition

In Chapter 4, I found that between the ages of 8 to 18 years, children's ability to recognise facial expressions significantly improved. When presented with facial expressions morphed between anger and disgust, older children were able to detect a difference in the expression from smaller changes in the stimulus. These findings are in line with other research exploring the development of facial expression perception (Dalrymple et al., 2017; Herba et al., 2006; Thomas et al., 2007). In addition, I also found an increase in the precision with which children were able to categorise body postures; with increasing age, children became better at categorising body postures. Very limited previous work has explored the development of body posture recognition (Mondloch, 2012; Ross et al., 2019). My results suggest that emotion recognition from face and body cues appear to show similar developmental patterns, however additional work is required to determine at which age children reach 'adult-like' abilities.

Furthermore, I demonstrate how the influence of body posture on facial expression

perception changes across development. In Chapter 4, I report that the influence of body context on facial expression categorisation decreases with increasing age. In other words, younger children's facial expression judgements were found to be more biased by an affective body posture than older children. Taken together with the increase in isolated facial expression recognition observed with increasing age, this work supports the hypothesis proposed in Chapter 3, suggesting that the more precise the isolated facial expression representation, the smaller the body biasing effect. Albeit, in children, this relationship appears to be driven by age. My results in Chapter 4 challenge the proposal put forward by Abo Foul *et al.* (2018). They posited that older people adopt an optimal social expertise strategy when making whole person judgements of emotion, and are therefore more biased by the body posture, as they have learnt that facial expressions are typically ambiguous, hence, strategically give the body posture more weight. If this social expertise strategy was an adequate explanation of the body context effect, in my current work in children, one would predict younger children would be less biased by body posture, as they have not yet learnt to rely on the body posture cue. In fact, I show the opposite relationship, suggesting the social expertise hypothesis cannot explain the influence of body context across human development and ageing. Instead, as detailed in my thesis, the reliability of the isolated facial expression representation appears to be the important factor in determining the magnitude of the body biasing effect, not an observer's experience with whole-person cues per se.

Fractional anisotropy of white matter tracts connecting OFA and FFA was found to significantly predict facial expression recognition abilities in children [Chapter 4]. Even when accounting for age, the microstructure of tracts between occipital and fusiform face areas was found to be predictive of individual differences in facial expression ability. This work has important implications for understanding complex relationships between white matter microstructure and perception, using a targeted methodology. In contrast to assessing perceptual performance in relation to large white matter tracts, the approach taken in Chapter 4 targeted white matter connectivity between cortical nodes known to be involved in face and body processing; this is a novel approach to explore face perception in the developing brain. In addition to the methodological novelty, these results provide important insights into facial expression processing networks in children. In contrast to

the early models of face processing, where features were thought to be segregated into changeable and invariant features, via the dorsal and ventral streams respectively (Haxby et al., 2000), my results highlight the importance of the ventral visual stream for facial expression processing. My results align with contemporary models of facial expression processing, highlighting the importance of the fusiform cortex for facial expression perception (Duchaine & Yovel, 2015; Ganel et al., 2005).

In addition to DTI metrics traditionally used to quantify microstructure of white matter, in Chapter 4, I also reported the spherical mean, a measure thought to be more sensitive to intracellular features. In previous work, the spherical mean has been shown to reflect age-related changes in the microstructure of large white matter tracts (Raven et al., 2020), and here I extend this finding to demonstrate its sensitivity to age in functionally-defined white matter tracts. Due to the tight relationship between the spherical mean of FDWM and age, interpretation of the importance of these tracts in relation to age-related changes in behaviour must be made with caution. For instance, the spherical mean of tracts connecting pSTS to ATL was found to significantly predict individual differences in the influence of body posture on facial expression perception. This result aligns with other published work discussing the site for integration of facial expression and body posture (Hu et al., 2020; Teufel et al., 2019), and suggests that connectivity from pSTS to anterior STS may be important for contextual integration. However, due to relationships with age observed in both the spherical mean of these tracts from pSTS to ATL, as well as between age and contextual influence, one cannot be conclusive about the behavioural relevance of this tract. In order to further disentangle the relationships between age-sensitive markers of white matter with developmental changes in perception, additional features, such as myelin, should be accounted for in subsequent research. Furthermore, exploring FDWM in adults would mitigate the additional developmental factors, and changes in microstructure would be more likely to be reflective of individuals' perceptual processes.

6.6 Limitations and future directions

One of the limitations of the work in my thesis is the use of only angry and disgusted displays of emotion. Anger and disgust were explicitly selected due to their high 'confus-

ability' as determined from previous research (Aviezer et al., 2008), therefore generating reliable and measurable influences of body posture on perception of facial expressions. Given that I aimed to quantify the influence of body context on facial expression perception and relate this to the reliability of individual differences in face and body emotion encoding, it was particularly important that expressions were chosen to maximise these effects. Furthermore, the ability to reliably recognise emotions from static body posture displays is challenging, however, anger and disgust have been found to be reliably recognised (Lopez et al., 2017). Therefore, the selection of anger and disgust was also driven by the ability to create reliably recognisable body posture displays which could be morphed with one another and not create unnatural postures within the morph range. In Chapter 4, I report improvements in children's ability to recognise facial expressions and body postures with increasing age. In order to make broader claims about overall emotion perception development across childhood, subsequent research should utilise additional facial expressions and body postures, such as fear and surprise, for example. However, for the research herein, to uncover how context influences facial expression perception in children this is not an issue.

Another limitation relating to the stimuli used was the use of all male expressions of anger and disgust in this thesis. Male facial expressions were selected for several reasons: previous work in this field opted for the use of male facial expressions (Aviezer et al., 2008; Aviezer et al., 2012b; Mondloch, 2012; Teufel et al., 2019), which allowed comparison between my research and other published literature. Secondly, different expressions can be embodied in different ways between genders. A male body posture displaying anger is thought to be more 'stereotypically' angry than a female body posture (Kret et al., 2013), with similar effects being found for facial expressions of anger (Harris et al., 2016). Future research should explore the influence of body context on facial expression with both male and female displays of emotion. In the same vein, in the current work in children [Chapter 4] the stimuli used were adults' facial expressions and body postures. Previous work has revealed an 'own-age bias' such that children recognise children's faces more accurately than adult faces (Hills & Lewis, 2011). Subsequent work disentangling the body biasing effect should consider the use of children's facial expressions to determine if children's performance is different when perceiving emotions of other children in the

context of whole person perception.

The work in this thesis focused on the influence of body context on facial expression perception. In Chapter 5, I refer to this as the 'whole-person' representation, however it is important to note that observers were explicitly instructed to identify the facial expression displayed, ignoring the body. As demonstrated from cue integration modelling, observers typically weighed the face cue as more reliable in their integrated whole-person representation than theoretically predicted. One explanation for these results, could be a consequence of the task requiring observers to categorise facial expression in the whole-person task. To further understand how facial expression and body posture cues integrate, research should assess how facial expression influences perception of body postures. Scarce research explores the effect of facial expression on the perception of a body posture (Reed et al., 2018). Recently it has been shown that when observers are asked to judge body postures, the facial expression can influence the perception of the body postures in a bidirectional manner, with the confusability of the expressions in the face and body determining how likely the contextual effect is to occur (Lecker et al., 2020). A limitation of the Lecker *et al.* (2020) study is that 100% expressions were used, instead of graded morphs. In order to fully appreciate the complexity of real-world social interactions, subsequent work should aim to uncover the interplay between the face and body cues further by exploring if and how, facial expressions can influence body posture perception using morphed expressions.

In the comparison between facial expression precision and body posture precision, as reported in Chapter 3 and 4, observers' performance has been found to be highly correlated. This does indicate that observers who had high facial expression precision also had high body posture precision; However, caution must be taken when interpreting these results due to the differences in image properties of the face and body stimuli. Despite the morph levels in the two conditions being physically equated, the perceptual expression in these cues may not be equivalent. For instance, while a physically 70% disgusted face morph may be perceived as disgusted the majority of the time, a 70% disgusted body morph may be perceived as disgusted 100% of the time. In order to further explore similarities and differences in how individuals perceive facial expressions and body postures, additional research should be undertaken to 'scale' the relative perceived inten-

sities of facial expression and body posture cues in order to equate the stimuli for further comparison. One approach that could be taken to explore this is magnitude estimation. Magnitude estimation is a psychophysical approach in which subjects assign a numerical value to physical stimuli (Stevens, 1956). The magnitude estimation procedure requires subjects to judge and assign numerical estimates to the perceived strength of a stimulus. Highly reliable judgments can be achieved for a whole range of sensory modalities, such as brightness, loudness, or tactile stimulation. For the application herein to facial expression and body posture, observers would be required to assign values proportional to the intensity of the expressions perceived across the stimuli enabling the perceived intensity of the expressions to be scaled to one another.

Importantly, in the real-world facial expressions and body postures are dynamic. Research has shown that the influence of body posture on facial expression is enhanced with the use of dynamic incongruent stimuli in both children and adults (Nelson & Mondloch, 2017). Therefore, to mirror the complexity of whole-person processing in research settings, future work should consider the use of dynamic facial expressions and body postures.

An open question which follows on naturally from the work in this thesis is the transformation and integration of facial expression and body posture cues at the cortical level, as well as the time course of integration of these stimuli. Although previous research has suggested ATL as a probable location encoding whole-person representations, to my knowledge none of these studies have looked at the integration of emotional information from body and face. Future research could employ powerful multivariate fMRI analysis techniques to uncover the site for integration of facial expression and body posture cues in the brain. Utilising the psychophysical methodology I developed, in combination with multivariate fMRI, may determine how facial expression and body posture signals are transformed along the visual processing pathway. In contrast to typical univariate fMRI analysis, multivariate analysis enables detection of patterns of activity across all voxels in different regions of the brain (Haxby, 2001). Representational similarity analysis (RSA) is a multivariate technique where brain-activity measurements (across multiple modalities) can be compared to models of behaviour by comparing activity-pattern dissimilarity matrices (Kriegeskorte, 2008). For each brain region a dissimilarity matrix is constructed comparing the neural activity patterns associated with pairs of conditions. In addition, the

behavioural similarity between observers' responses can also be constructed to produce a representational dissimilarity matrix (RDM). Comparison of RDMs between behavioural performance and different neural ROIs could subsequently determine the neural basis for perception along the visual ventral stream. Based on the proposed models for face and body integration (Harry et al., 2016; Hu et al., 2020), one would hypothesise neural regions early in the visual processing stream, such as the OFA, will be sensitive primarily to the physical features of facial expressions irrespective of context; different facial expressions will evoke distinct neural patterns irrespective of body posture. Regions found later in the visual hierarchy, such as anterior temporal lobes, will instead display patterns of activity capturing the perceived emotion biased by the body posture, indicating that the information has been integrated into a whole-person global percept. An advantage of RSA is the projection of data into the same geometric space, so subsequent research could also consider the temporal evidence associated with integration of face and body cues, in conjunction with the spatial data, by acquiring EEG or magnetoencephalography data. Furthermore, this design could be employed in a developmental cohort to provide insight into the development of facial expression and body posture integration in children.

6.7 Conclusion

In summary, the research in this thesis provides novel insights into the processing and integration of facial expression and body posture cues. Spanning from children to adults, my results uncover individual differences in the perception of facial expressions in the presence of a body posture. Furthermore, variation between observers was found to be attributed to an individual's reliability in encoding isolated facial expressions. Representations of facial expression, body posture and whole-person reliability were found to be related to microstructure of functionally relevant white matter tracts in face and body networks in children. Taken together, this thesis supports the hypothesis suggesting the influence of body context on facial expression perception is a truly perceptual effect and integration of these cues is dependent on the reliability of the isolated face and body cues, providing a more mechanistic understanding of the integration of facial expression and body posture. In turn, the work in this thesis paves the way for subsequent research to

further explore the hierarchical nature of whole-person processing along the ventral visual stream. In conclusion, this thesis has important implications for our understanding of real-world individual differences in social perception.

Appendices

ROI	Number
OFA	37
FFA	43
pSTS(face)	41
EBA	39
FBA	40
pSTS(body)	39

Table 1: **Number ROIs for OFA, FFA, pSTS(face), EBA, FBA and pSTS(body) identified**
For each ROI the number of participants this region could be identified in is listed.

ROI	x	y	z	Source
OFA	46	-75	-3	Bona et al., 2015
	17.8	-88.9	-11.3	Spiridon et al., 2006
	44	-80	-14	Schobert et al., 2018
	44	-76	-12	Harry et al., 2016
FFA	31.6	-57.2	-10.4	Spiridon et al., 2006
	40	-55	-12	Schobert et al., 2018
	42	-52	-24	Neurosynth: *Fusiform face*
	43.3	-49.6	-20.5	Harry et al., 2016
EBA	32.6	-77.3	2.4	Spiridon et al., 2006
	32	-67.6	6	Spiridon et al., 2006
	-42	-64	6	Vocks et al., 2010
	46	-70	0	Neurosynth: *Body*
FBA	40	-48	-22	Vocks et al., 2010
	38	-41.2	-17	Taylor et al., 2007
pSTS	54	-42	6	Neurosynth: *Posterior superior*
	50	-47	13	Schobert et al., 2018

Table 2: **Coordinates from literature for ROIs**
The table lists coordinates for OFA, FFA, EBA, FBA and pSTS extracted from publications and Neurosynth [https://neurosynth.org/analyses/]. The search terms used in the Term-based meta-analyses are listed in the source column. All coordinates listed are in MNI space.

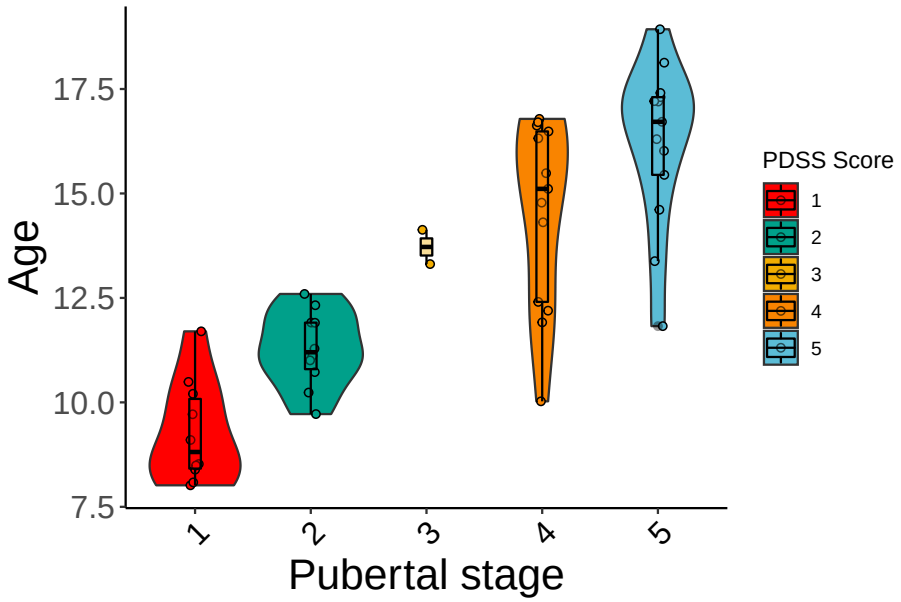


Figure 1: **Relationship between age and PDSS score**
Children's age shown with Pubertal Development Scale Shirtcliff score. Age is shown on the y-axis in years. A tight positive correlation between age and PDSS score was found ($r(46)=0.85$, $p<0.0001$). Each point represents one observer. The distribution of the values is illustrated by the shaded area, with the overlaid boxplot indicating the median and the interquartile range.

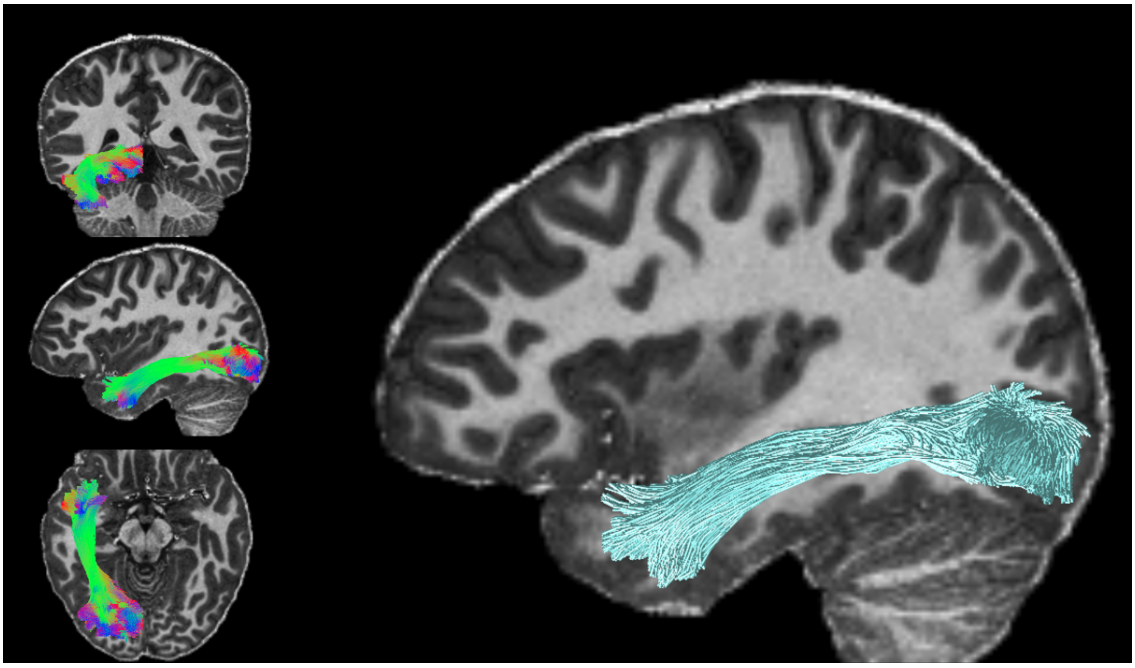


Figure 2: **Inferior longitudinal fasiculus**
Example of the anatomically defined ILF in one child. In the left panel the tract can be seen in coronal, sagittal and axial planes with directional colour encoding.

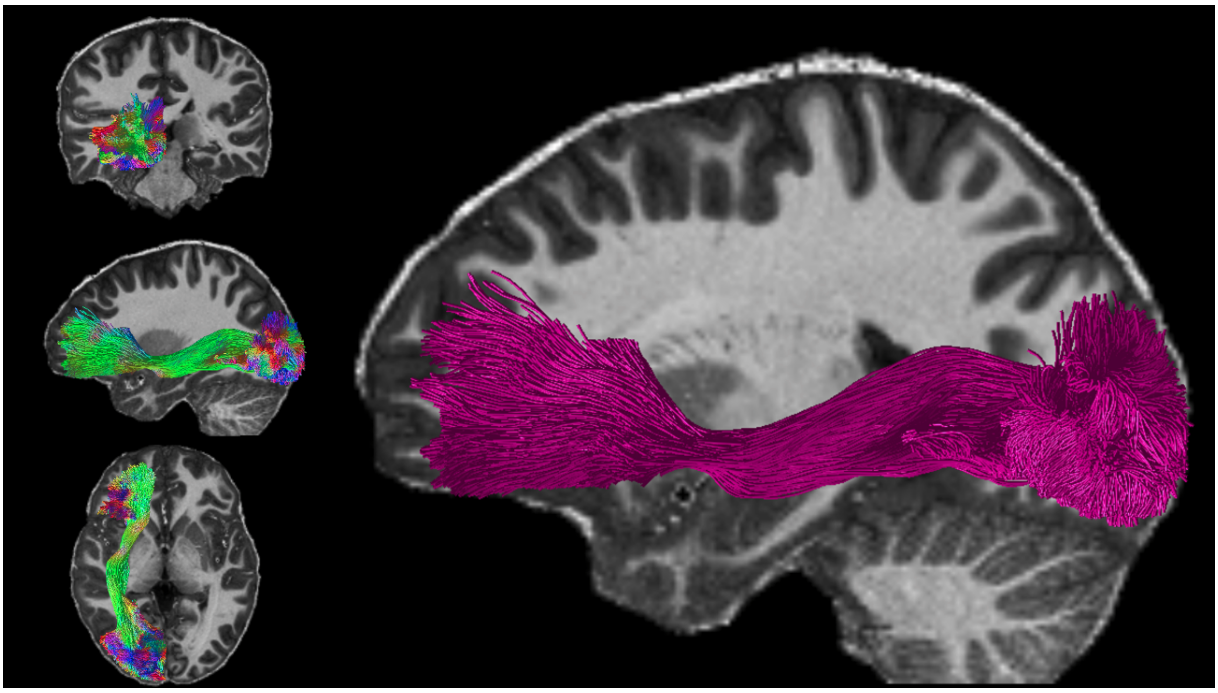


Figure 3: **Inferior frontal occipital fasiculus**
Example of the anatomically defined IFOF in one child. In the left panel the tract can be seen in coronal, sagittal and axial planes with directional colour encoding.

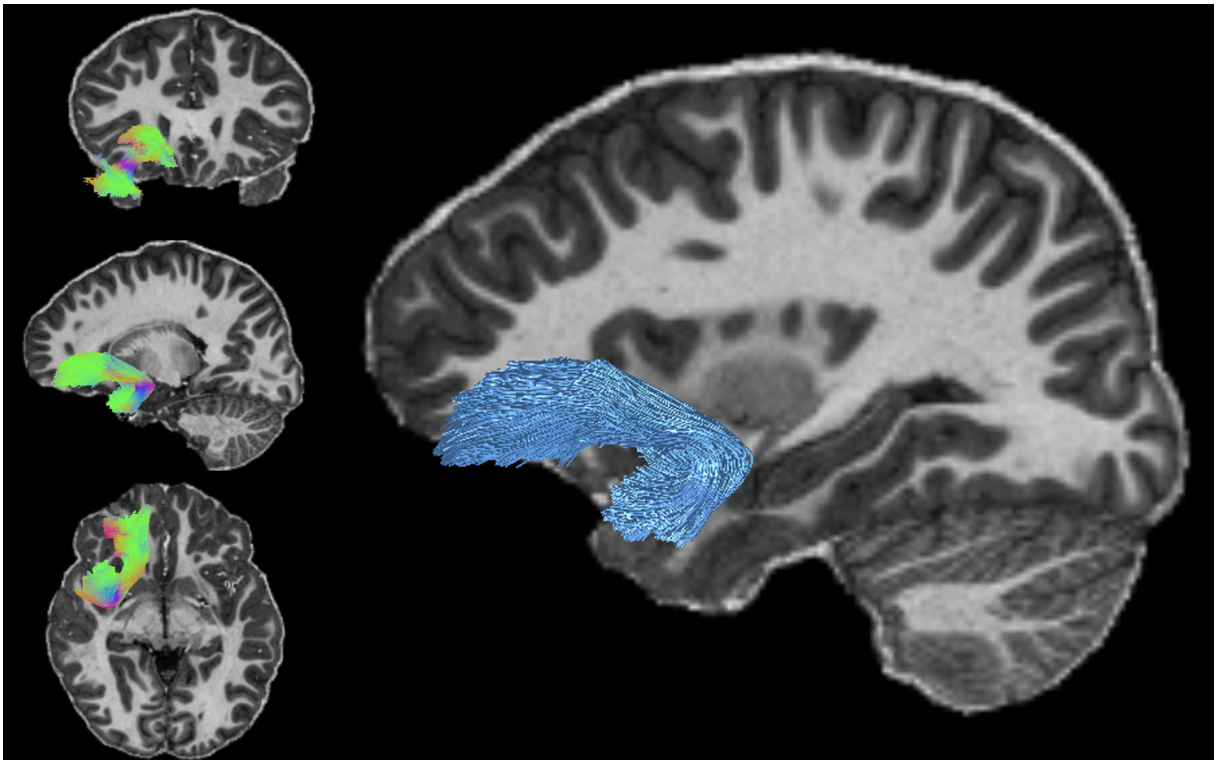


Figure 4: **Uncinate fasiculus**
Example of the anatomically defined UF in one child. In the left panel the tract can be seen in coronal, sagittal and axial planes with directional colour encoding.

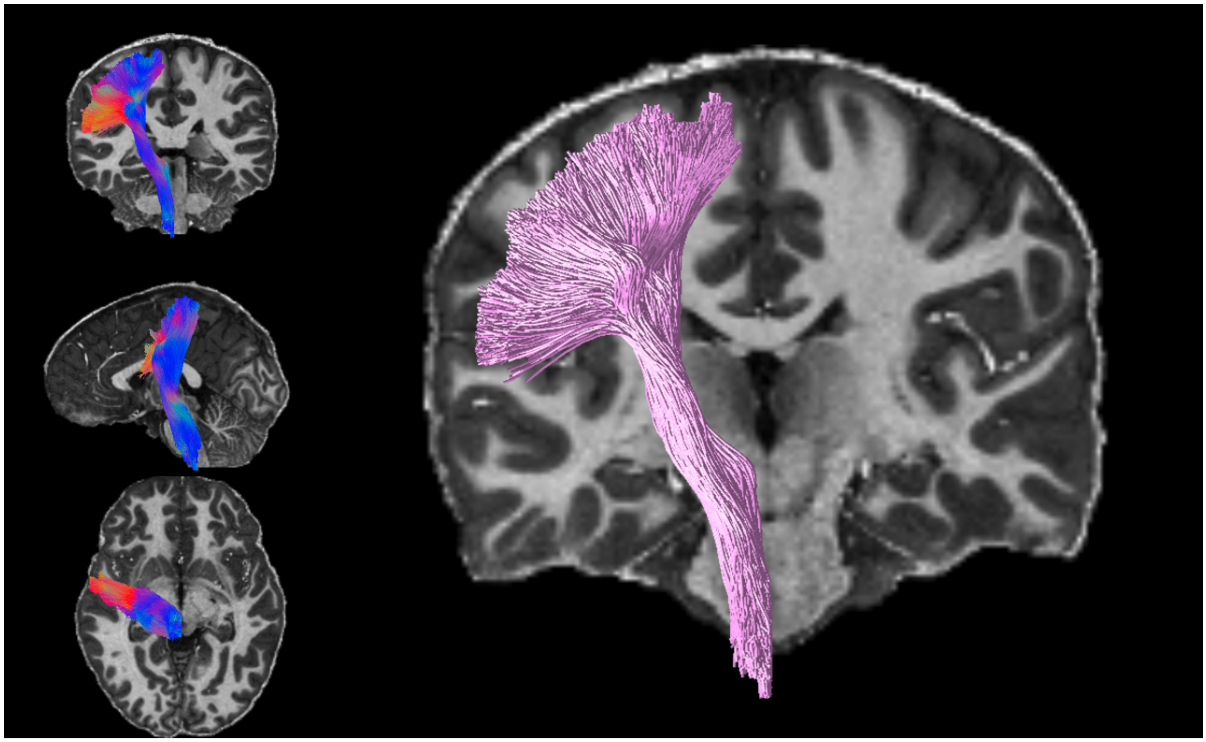


Figure 5: **Corticospinal tract**
Example of the anatomically defined CST in one child. In the left panel the tract can be seen in coronal, sagittal and axial planes with directional colour encoding.

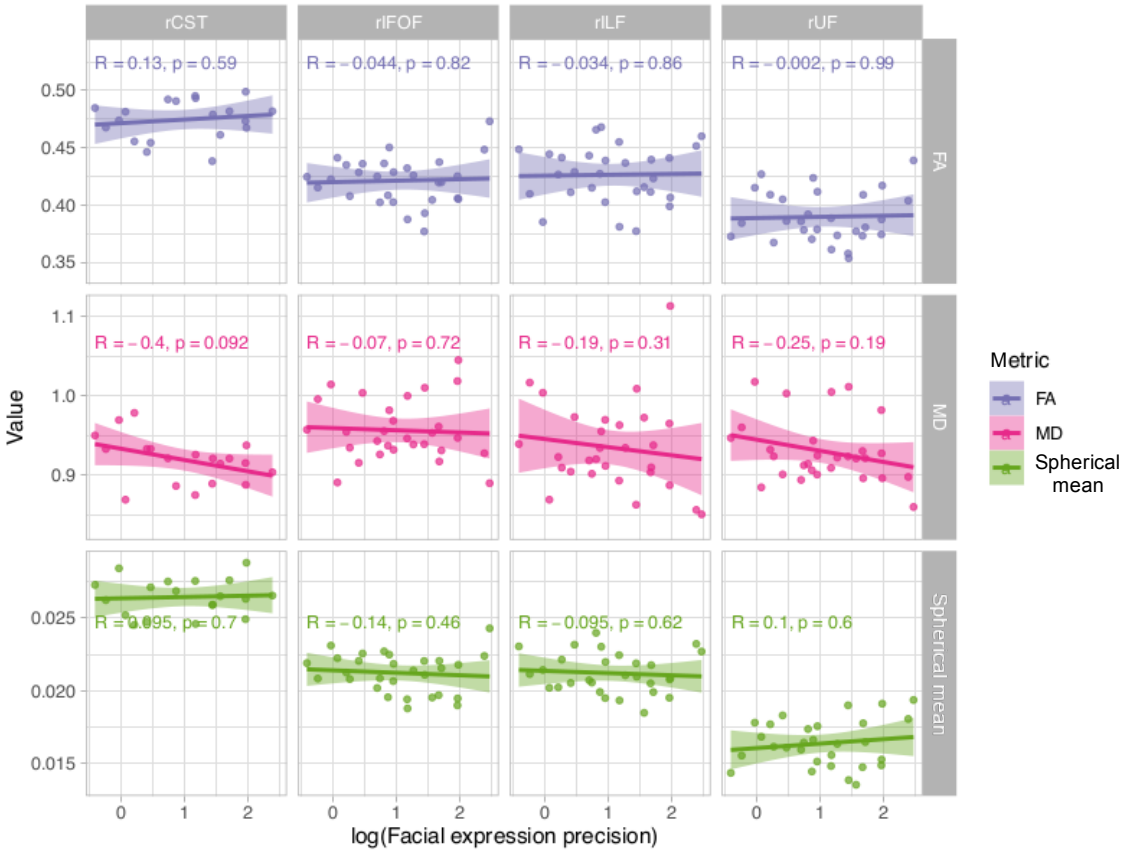


Figure 6: **Relationship between microstructure and facial expression precision, grouped by white matter tract and coloured by metric**
For each of the correlation plots the x-axis displays the logged facial expression precision, and the y-axis shows the value of the metric. FA is shown in the top panel in purple, MD in the middle panel in pink and $\mathring{S}_{\mu}(b = 6000)$ in the bottom panel in green. The Spearman's rank correlation coefficient and p-value are indicated for each relationship. The 95 % confidence interval is shown with shading.

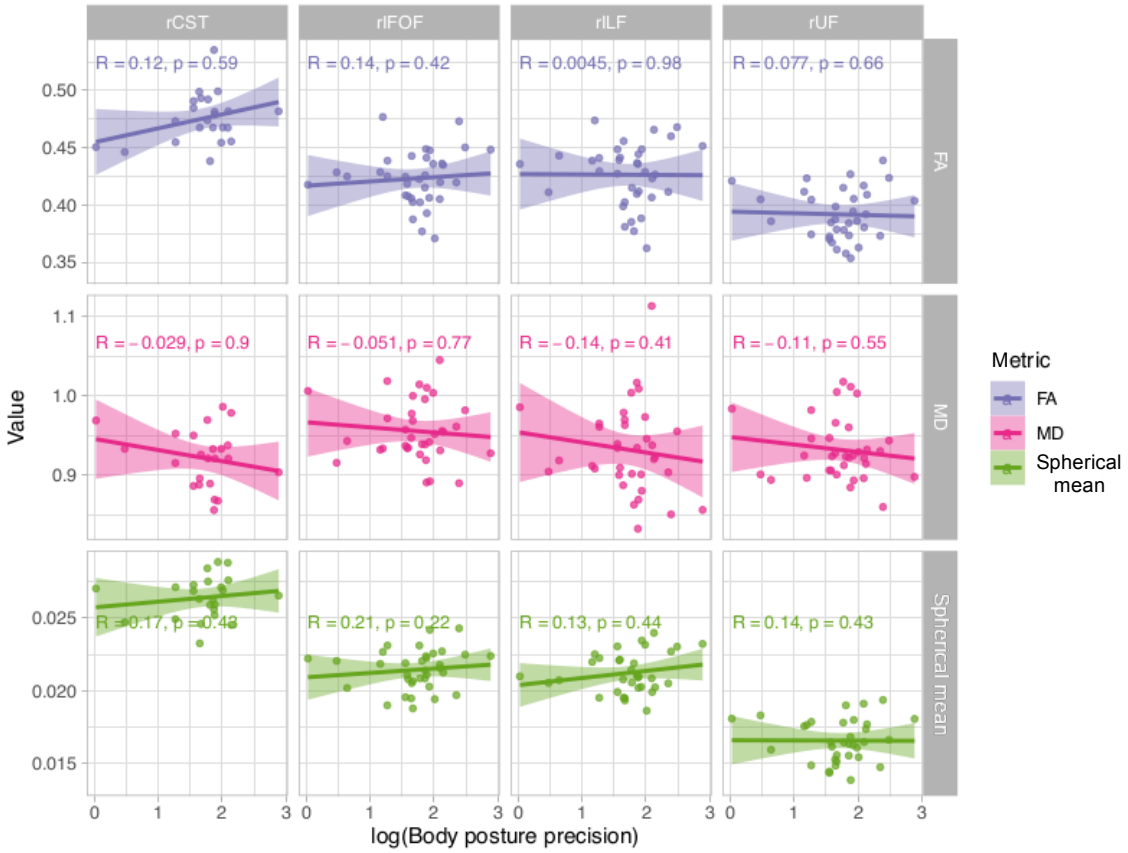


Figure 7: **Relationship between microstructure and body posture precision, grouped by white matter tract and coloured by metric**
For each of the correlation plots the x-axis displays the logged body posture precision, and the y-axis shows the value of the metric. FA is shown in the top panel in purple, MD in the middle panel in pink and $\mathring{S}_\mu(b = 6000)$ in the bottom panel in green. The Spearman's rank correlation coefficient and p-value are indicated for each relationship. The 95 % confidence interval is shown with shading.

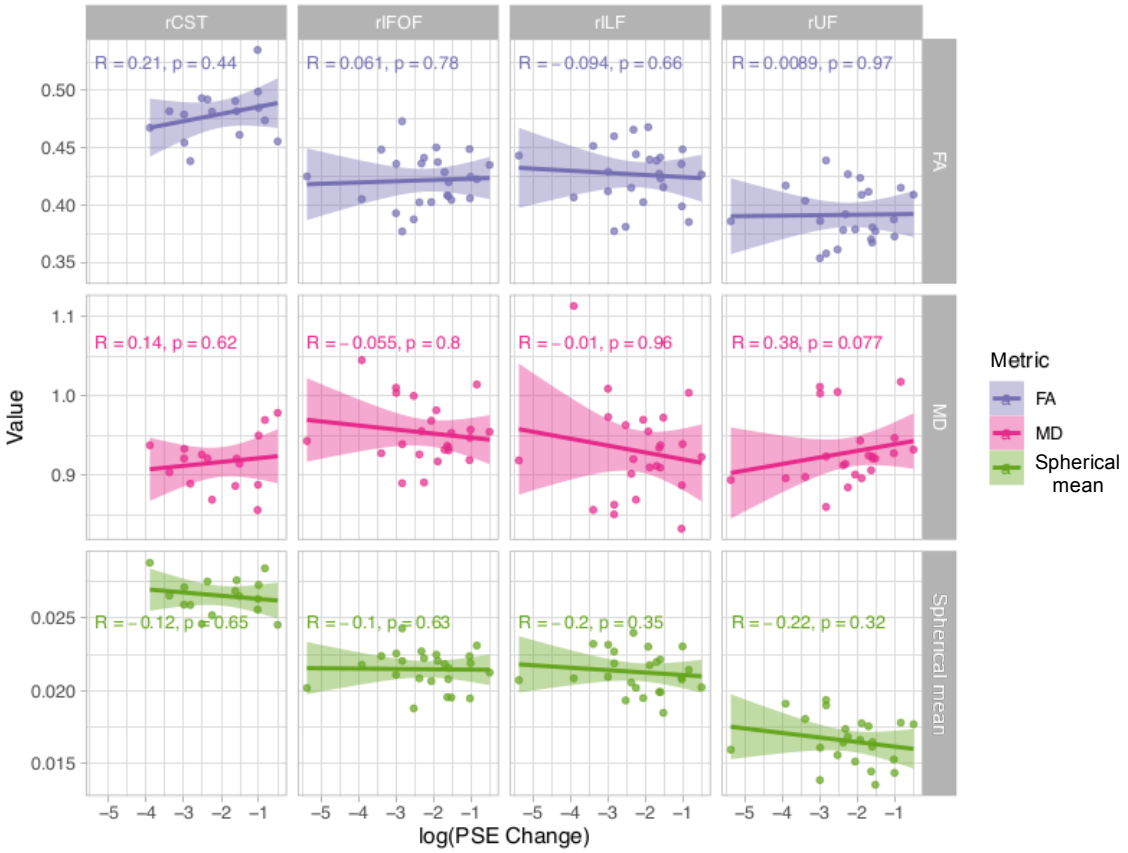


Figure 8: **Relationship between microstructure and PSE change, grouped by white matter tract and coloured by metric**
For each of the correlation plots the x-axis displays the logged PSE change, and the y-axis shows the value of the metric. FA is shown in the top panel in purple, MD in the middle panel in pink and $\mathring{S}_\mu(b = 6000)$ in the bottom panel in green. The Spearman's rank correlation coefficient and p-value are indicated for each relationship. The 95 % confidence interval is shown with shading.

Bibliography

Abo Foul, Y., Eitan, R., & Aviezer, H. (2018). Perceiving emotionally incongruent cues from faces and bodies: Older adults get the whole picture. *Psychology and Aging*, *33*(4), 660–666.

Addington, J., & Addington, D. (1998). Facial affect recognition and information processing in schizophrenia and bipolar disorder. *Schizophrenia Research*, *32*(3), 171–181.

Alexander, A. L., Lee, J. E., Lazar, M., & Field, A. S. (2007). Diffusion tensor imaging of the brain. *Neurotherapeutics: The Journal of the American Society for Experimental NeuroTherapeutics*, *4*(3), 316–329.

Alharbi, S. A. H., Button, K., Zhang, L., O'Shea, K. J., Fasolt, V., Lee, A. J., DeBruine, L. M., & Jones, B. C. (2020). Are affective factors related to individual differences in facial expression recognition? *Royal Society Open Science*, *7*(9), 190699.

Andersson, J. L., & Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage*, *125*, 1063–1078.

Arcaro, M. J., Ponce, C., & Livingstone, M. (2020). The neurons that mistook a hat for a face. *eLife*, *9*, e53798.

Arthurs, O. J., & Boniface, S. (2002). How well do we understand the neural origins of the fMRI BOLD signal? *Trends in Neurosciences*, *25*(1), 27–31.

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, *54*(3), 2033–2044.

Aviezer, H., Trope, Y., & Todorov, A. (2012a). Body Cues, Not Facial Expressions, Discriminate Between Intense Positive and Negative Emotions. *Science*, *338*(6111), 1225–1229.

Aviezer, H., Bentin, S., Dudarev, V., & Hassin, R. R. (2011). The automaticity of emotional face-context integration. *Emotion*, *11*(6), 1406–1414.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, Disgusted, or Afraid? Studies on the Malleability of Emotion Perception. *Psychological Science*, *19*(7), 724–732.

Aviezer, H., Trope, Y., & Todorov, A. (2012b). Holistic person processing: Faces with bodies tell the whole story. *Journal of Personality and Social Psychology*, *103*(1), 20–37.

Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, *39*(6), 1143–1170.

Bajgar, J., Ciarrochi, J., Lane, R. D., & Deane, F. (2005). Development of the Levels of Emotional Awareness Scale for Children (LEAS-C). *British Journal of Developmental Psychology*, *23*, 569–586.

Bartzokis, G., Lu, P. H., Heydari, P., Couvrette, A., Lee, G. J., Kalashyan, G., Freeman, F., Grinstead, J. W., Villablanca, P., Finn, J. P., Mintz, J., Alger, J. R., & Altshuler, L. L. (2012). Multimodal Magnetic Resonance Imaging Assessment of White Matter Aging Trajectories Over the Lifespan of Healthy Individuals. *Biological Psychiatry*, *72*(12), 1026–1034.

Basil, R. A., Westwater, M. L., Wiener, M., & Thompson, J. C. (2017). A Causal Role of the Right Superior Temporal Sulcus in Emotion Recognition From Biological Motion. *Open Mind*, *2*(1), 26–36.

Basser, P. J., & Pierpaoli, C. (1996). Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance. Series B*, *111*(3), 209–219.

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR in Biomedicine*, *15*(7-8), 435–455.

Behrens, T., Berg, H. J., Jbabdi, S., Rushworth, M., & Woolrich, M. (2007). Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? *NeuroImage*, *34*(1), 144–155.

Bejjanki, V. R., Clayards, M., Knill, D. C., & Aslin, R. N. (2011). Cue Integration in Categorical Tasks: Insights from Audio-Visual Speech Perception (D. G. Pelli, Ed.). *PLoS ONE*, *6*(5), e19812.

Berman, J., Lanza, M., Blaskey, L., Edgar, J., & Roberts, T. (2013). High Angular Resolution Diffusion Imaging Probabilistic Tractography of the Auditory Radiation. *American Journal of Neuroradiology*, *34*(8), 1573–1578.

Bernstein, M., & Yovel, G. (2015). Two neural pathways of face processing: A critical evaluation of current models. *Neuroscience & Biobehavioral Reviews*, *55*, 536–546.

de Bie, H. M. A., Boersma, M., Wattjes, M. P., Adriaanse, S., Vermeulen, R. J., Oostrom, K. J., Huisman, J., Veltman, D. J., & Delemarre-Van de Waal, H. A. (2010). Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans. *European Journal of Pediatrics*, *169*(9), 1079–1085.

Bona, S., Cattaneo, Z., & Silvanto, J. (2015). The Causal Role of the Occipital Face Area (OFA) and Lateral Occipital (LO) Cortex in Symmetry Perception. *The Journal of Neuroscience*, *35*(2), 731–738.

Boogert, N. J., Madden, J. R., Morand-Ferron, J., & Thornton, A. (2018). Measuring and understanding individual differences in cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1756), 20170280.

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

Brandman, T., & Yovel, G. (2014). Bodies are Represented as Wholes Rather Than Their Sum of Parts in the Occipital-Temporal Cortex. *Cerebral Cortex*, bhu205.

Brewer, R., Biotti, F., Bird, G., & Cook, R. (2017). Typical integration of emotion cues from bodies and faces in Autism Spectrum Disorder. *Cognition*, *165*, 82–87.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*(3), 305–327.

Brun, L., Pron, A., Sein, J., Deruelle, C., & Coulon, O. (2019). Diffusion MRI: Assessment of the Impact of Acquisition and Preprocessing Methods Using the BrainVISA-Diffuse Toolbox. *Frontiers in Neuroscience*, *13*.

Buzsaki, G. (2004). Neuronal Oscillations in Cortical Networks. *Science*, *304*(5679), 1926–1929.

Calder, A. J. (Ed.). (2011). *The Oxford handbook of face perception*. Oxford University Press.

Candidi, M., Stienen, B. M. C., Aglioti, S. M., & de Gelder, B. (2011). Event-Related Repetitive Transcranial Magnetic Stimulation of Posterior Superior Temporal Sulcus Improves the Detection of Threatening Postural Changes in Human Bodies. *Journal of Neuroscience*, *31*(48), 17547–17554.

Chang, C., Raven, E. P., & Duyn, J. H. (2016). Brain–heart interactions: Challenges and opportunities with functional magnetic resonance imaging at ultra-high field. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2067), 20150188.

Chang, Y. S., Owen, J. P., Pojman, N. J., Thieu, T., Bukshpun, P., Wakahiro, M. L. J., Berman, J. I., Roberts, T. P. L., Nagarajan, S. S., Sherr, E. H., & Mukherjee, P. (2015). White Matter Changes of Neurite Density and Fiber Orientation Dispersion during Human Brain Maturation (G. Gong, Ed.). *PLOS ONE*, *10*(6), e0123656.

Choe, A. S., Stepniewska, I., Colvin, D. C., Ding, Z., & Anderson, A. W. (2012). Validation of diffusion tensor MRI in the central nervous system using light microscopy: Quantitative comparison of fiber properties: VALIDATION OF DIFFUSION TENSOR MRI. *NMR in Biomedicine*, *25*(7), 900–908.

Clifford, C., Mareschal, I., Otsuka, Y., & Watson, T. (2015). A Bayesian approach to person perception. *Consciousness and Cognition*, *36*, 406–413.

Coad, B. M., Postans, M., Hodgetts, C. J., Muhlert, N., Graham, K. S., & Lawrence, A. D. (2017). Structural connections support emotional connections: Uncinate Fasciculus microstructure is related to the ability to decode facial emotion expressions. *Neuropsychologia*.

Collins, J. A., & Olson, I. R. (2014). Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. *Neuropsychologia*, *61*, 65–79.

Cormack, L. K. (2005). Computational Models of Early Human Vision. *Handbook of Image and Video Processing* (pp. 325–IX). Elsevier.

Cox, D. (2004). Contextually Evoked Object-Specific Responses in Human Visual Cortex. *Science*, *304*(5667), 115–117.

Croker, V., & McDonald, S. (2005). Recognition of emotion from facial expression following traumatic brain injury. *Brain Injury*, *19*(10), 787–799.

Dalrymple, K. A., Visconti di Oleggio Castello, M., Elison, J. T., & Gobbini, M. I. (2017). Concurrent development of facial identity and expression discrimination (P. J. Hills, Ed.). *PLOS ONE*, *12*(6), e0179458.

Dekker, T. M., Ban, H., van der Velde, B., Sereno, M. I., Welchman, A. E., & Nardini, M. (2015). Late Development of Cue Integration Is Linked to Sensory Fusion in Cortex. *Current Biology*, *25*(21), 2856–2861.

Descoteaux, M., Deriche, R., Knosche, T., & Anwander, A. (2009). Deterministic and Probabilistic Tractography Based on Complex Fibre Orientation Distributions. *IEEE Transactions on Medical Imaging*, *28*(2), 269–286.

Dima, D. C., Perry, G., Messaritaki, E., Zhang, J., & Singh, K. D. (2018). Spatiotemporal dynamics in human visual cortex rapidly encode the emotional content of faces. *Human Brain Mapping*.

Downing, P. E. (2001). A Cortical Area Selective for Visual Processing of the Human Body. *Science*, *293*(5539), 2470–2473.

Duchaine, B., & Yovel, G. (2015). A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, *1*(1), 393–416.

Duchaine, B. C., Yovel, G., Butterworth, E. J., & Nakayama, K. (2006). Prosopagnosia as an impairment to face-specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology*, *23*(5), 714–747.

Eggenschwiler, F., Kober, T., Magill, A. W., Gruetter, R., & Marques, J. P. (2012). SA2RAGE: A new sequence for fast B1+-mapping. *Magnetic Resonance in Medicine*, *67*(6), 1609–1619.

Einstein, A. (1905). ber die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, *322*(8), 549–560.

Ekman, P. (1965). Differential communication of affect by head and body cues. *Journal of Personality and Social Psychology*, *2*(5), 726–735.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition & Emotion*, *6*(3/4), 169–200.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

Ernst, M. O., & Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends in Cognitive Sciences*, *8*(4), 162–169.

Etxeberria, A., Hokanson, K. C., Dao, D. Q., Mayoral, S. R., Mei, F., Redmond, S. A., Ullian, E. M., & Chan, J. R. (2016). Dynamic Modulation of Myelination in Response to Visual Stimuli Alters Optic Nerve Conduction Velocity. *Journal of Neuroscience*, *36*(26), 6937–6948.

Fetsch, C. R., Pouget, A., DeAngelis, G. C., & Angelaki, D. E. (2013). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature Neuroscience*, *15*(1), 146–154.

Fields, R. D. (2008). White matter in learning, cognition and psychiatric disorders. *Trends in Neurosciences*, *31*(7), 361–370.

Fisher, C., & Freiwald, W. A. (2015). Whole-agent selectivity within the macaque face-processing system. *Proceedings of the National Academy of Sciences*, *112*(47), 14717–14722.

Gallichan, D., Marques, J. P., & Gruetter, R. (2016). Retrospective correction of involuntary microscopic head movement using highly accelerated fat image navigators (3D FatNavs) at 7T. *Magnetic Resonance in Medicine*, *75*(3), 1030–1039.

Ganel, T., Valyear, K. F., Goshen-Gottstein, Y., & Goodale, M. A. (2005). The involvement of the "fusiform face area" in processing facial expression. *Neuropsychologia*, *43*(11), 1645–1654.

Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology*, *102*(4), 503–521.

de Gelder, B., de Borst, A., & Watson, R. (2015). The perception of emotion in body expressions: Emotional body perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(2), 149–158.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience*, *7*(3), 242–249.

Genc, S., Malpas, C. B., Holland, S. K., Beare, R., & Silk, T. J. (2017). Neurite density index is sensitive to age related differences in the developing brain. *NeuroImage*, *148*, 373–380.

Glover, G. H. (2011). Overview of Functional Magnetic Resonance Imaging. *Neurosurgery Clinics of North America*, *22*(2), 133–139.

Golarai, G., Ghahremani, D. G., Grill-Spector, K., & Gabrieli, J. D. E. (2010). Evidence for maturation of the fusiform face area (FFA) in 7 to 16 year old children. *Journal of Vision*, *5*(8), 634–634.

Golarai, G., Liberman, A., Yoon, J. M. D., & Grill-Spector, K. (2009). Differential development of the ventral visual cortex extends through adolescence. *Frontiers in Human Neuroscience*.

Goldstein, E. B., Humphreys, G. W., Shiffrar, M., & Yost, W. A. (Eds.). (2005). *Blackwell handbook of sensation and perception*. Blackwell Pub.

Gomez, J., Barnett, M. A., Natu, V., Mezer, A., Palomero-Gallagher, N., Weiner, K. S., Amunts, K., Zilles, K., & Grill-Spector, K. (2017). Microstructural proliferation in human cortex is coupled with the development of face processing. *Science*, *355*(6320), 68–71.

Gomez, J., Pestilli, F., Witthoft, N., Golarai, G., Liberman, A., Poltoratski, S., Yoon, J., & Grill-Spector, K. (2015). Functionally Defined White Matter Reveals Segregated Pathways in Human Ventral Temporal Cortex Associated with Category-Specific Processing. *Neuron*, *85*(1), 216–227.

Goodale, M. A., & Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20–25.

Green, C., & Guo, K. (2018). Factors contributing to individual differences in facial expression categorisation. *Cognition and Emotion*, *32*(1), 37–48.

Grill-Spector, K., Weiner, K. S., Gomez, J., Stigliani, A., & Natu, V. S. (2018). The functional neuroanatomy of face perception: From brain measurements to deep neural networks. *Interface Focus*, *8*(4), 20180013.

Grill-Spector, K., Weiner, K. S., Kay, K., & Gomez, J. (2017). The Functional Neuroanatomy of Human Face Perception. *Annual Review of Vision Science*, *3*(1), 167–196.

Gschwind, M., Pourtois, G., Schwartz, S., Van De Ville, D., & Vuilleumier, P. (2012). White-Matter Connectivity between Face-Responsive Regions in the Human Brain. *Cerebral Cortex*, *22*(7), 1564–1576.

Hagmann, P., Sporns, O., Madan, N., Cammoun, L., Pienaar, R., Wedeen, V. J., Meuli, R., Thiran, J. P., & Grant, P. E. (2010). White matter maturation reshapes structural connectivity in the late developing human brain. *Proceedings of the National Academy of Sciences*, *107*(44), 19067–19072.

Haist, F., & Anzures, G. (2017). Functional development of the brain's face-processing system. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1-2), e1423.

Harris, D. A., Hayes-Skelton, S. A., & Ciaramitaro, V. M. (2016). What's in a Face? How Face Gender and Current Affect Influence Perceived Emotion. *Frontiers in Psychology*, *7*.

Harry, B. B., Umla-Runge, K., Lawrence, A. D., Graham, K. S., & Downing, P. E. (2016). Evidence for Integrated Visual Face and Body Representations in the Anterior Temporal Lobes. *Journal of Cognitive Neuroscience*, *28*(8), 1178–1193.

Hassin, R. R., Aviezer, H., & Bentin, S. (2013). Inherently Ambiguous: Facial Expressions of Emotions, in Context. *Emotion Review*, *5*(1), 60–65.

Haxby, J. V. (2001). Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science*, *293*(5539), 2425–2430.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233.

Heck, A., Chroust, A., White, H., Jubran, R., & Bhatt, R. S. (2018). Development of body emotion perception in infancy: From discrimination to recognition. *Infant Behavior and Development*, *50*, 42–51.

Helbig, H. B., & Ernst, M. O. (2007). Optimal integration of shape information from vision and touch. *Experimental Brain Research*, *179*(4), 595–606.

Herba, C. M., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: Effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry*, *47*(11), 1098–1106.

Herzog, M. H., & Clarke, A. M. (2014). Why vision is not both hierarchical and feedforward. *Frontiers in Computational Neuroscience*, *8*.

Hillis, J. M. (2002). Combining Sensory Information: Mandatory Fusion Within, but Not Between, Senses. *Science*, *298*(5598), 1627–1630.

Hillman, E. M. (2014). Coupling Mechanism and Significance of the BOLD Signal: A Status Report. *Annual Review of Neuroscience*, *37*(1), 161–181.

Hills, P. J., & Lewis, M. B. (2011). Rapid communication: The own-age face recognition bias in children and adults. *Quarterly Journal of Experimental Psychology*, *64*(1), 17–23.

Hodgetts, C. J., Postans, M., Shine, J. P., Jones, D. K., Lawrence, A. D., & Graham, K. S. (2015). Dissociable roles of the inferior longitudinal fasciculus and fornix in face and place perception. *eLife*, *4*.

Hu, Y., Baragchizadeh, A., & O'Toole, A. J. (2020). Integrating faces and bodies: Psychological and neural perspectives on whole person perception. *Neuroscience & Biobehavioral Reviews*, *112*, 472–486.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154.

Huettel, S. A., Song, A. W., & McCarthy, G. (2014). *Functional magnetic resonance imaging* (Third edition). Sinauer Associates, Inc., Publishers.

Huisman, T. (2010). Diffusion-weighted and diffusion tensor imaging of the brain, made easy. *Cancer Imaging*, *10*(1A), S163–S171.

Hutmacher, F. (2019). Why Is There So Much More Research on Vision Than on Any Other Sensory Modality? *Frontiers in Psychology*, *10*.

Huttenlocher, P. R., & de Courten, C. (1987). The development of synapses in striate cortex of man. *Human Neurobiology*, *6*(1), 1–9.

Jacobs, R. A. (1999). Optimal integration of texture and motion cues to depth. *Vision Research*, *39*(21), 3621–3629.

Jbabdi, S., & Johansen-Berg, H. (2011). Tractography: Where Do We Go from Here? *Brain Connectivity*, *1*(3), 169–183.

Jeffery, K. J., Page, H. J. I., & Stringer, S. M. (2016). Optimal cue combination and landmark-stability learning in the head direction system: Cue combination in the head direction system. *The Journal of Physiology*, *594*(22), 6527–6534.

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage*, *17*(2), 825–841.

Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W., & Smith, S. M. (2012). FSL. *NeuroImage*, *62*(2), 782–790.

Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, *5*(2), 143–156.

Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H., & Kaczynski, K. (2005). Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic Resonance in Medicine*, *53*(6), 1432–1440.

Jeurissen, B., Descoteaux, M., Mori, S., & Leemans, A. (2017). Diffusion MRI fiber tractography of the brain. *NMR in Biomedicine*, *32*(4), e3785.

Jeurissen, B., Tournier, J.-D., Dhollander, T., Connelly, A., & Sijbers, J. (2014). Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage*, *103*, 411–426.

Jonas, J., Jacques, C., Liu-Shuang, J., Brissart, H., Colnat-Coulbois, S., Maillard, L., & Rossion, B. (2016). A face-selective ventral occipito-temporal map of the human brain with intracerebral potentials. *Proceedings of the National Academy of Sciences*, *113*(28), E4088–E4097.

Jones, D. K. (Ed.). (2010). *Diffusion MRI: Theory, methods, and application*. Oxford University Press.

Jones, D. K., Knösche, T. R., & Turner, R. (2013). White matter integrity, fiber count, and other fallacies: The do's and don'ts of diffusion MRI. *NeuroImage*, *73*, 239–254.

Jones, D., Alexander, D., Bowtell, R., Cercignani, M., Dell'Acqua, F., McHugh, D., Miller, K., Palombo, M., Parker, G., Rudrapatna, U., & Tax, C. (2018). Microstructural imaging of the human brain with a 'super-scanner': 10 key

advantages of ultra-strong gradients for diffusion MRI. *NeuroImage*, *182*, 8–38.

Kanwisher, N. (2000). Domain specificity in face perception. *Nature Neuroscience*, *3*(8), 759–763.

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.

Kanwisher, N., Tong, F., & Nakayama, K. (1998). The effect of face inversion on the human fusiform face area. *Cognition*, *68*(1), B1–B11.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109–2128.

Keating, C. T., & Cook, J. L. (2021). Facial Expression Production and Recognition in Autism Spectrum Disorders. *Psychiatric Clinics of North America*, *44*(1), 125–139.

Kellner, E., Dhital, B., Kiselev, V. G., & Reisert, M. (2016). Gibbs-ringing artifact removal based on local subvoxel-shifts: Gibbs-Ringing Artifact Removal. *Magnetic Resonance in Medicine*, *76*(5), 1574–1581.

Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: A practical introduction* (1. ed). Elsevier.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception 36 ECVP Abstract Supplement*.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*(16), 2729–2737.

Kret, M. E., Roelofs, K., Stekelenburg, J. J., & de Gelder, B. (2013). Emotional signals from faces, bodies and scenes influence observers' face expressions, fixations and pupil-size. *Frontiers in Human Neuroscience*, *7*.

Kriegeskorte, N. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*.

Kumfor, F., Ibañez, A., Hutchings, R., Hazelton, J. L., Hodges, J. R., & Piguet, O. (2018). Beyond the face: How context modulates emotion processing in frontotemporal dementia subtypes. *Brain*, *141*(4), 1172–1185.

Ladd, M. E., Bachert, P., Meyerspeer, M., Moser, E., Nagel, A. M., Norris, D. G., Schmitter, S., Speck, O., Straub, S., & Zaiss, M. (2018). Pros and cons of ultra-high-field MRI/MRS for human application. *Progress in Nuclear Magnetic Resonance Spectroscopy*, *109*, 1–50.

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America A*, *18*(9), 2307.

Lane, R. D., Reiman, E. M., Axelrod, B., Yun, L. S., Holmes, A., & Schwartz, G. E. (1998). Neural correlates of levels of emotional awareness. Evidence of an interaction between emotion and attention in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *10*(4), 525–535.

Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A., & Zeitlin, S. B. (1990). The Levels of Emotional Awareness Scale: A Cognitive-Developmental Measure of Emotion. *55*(1&2), 124–134.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition & Emotion*, *24*(8), 1377–1388.

Le Bihan, D., & Johansen-Berg, H. (2012). Diffusion MRI at 25: Exploring brain tissue structure and function. *NeuroImage*, *61*(2), 324–341.

Lebel, C., & Beaulieu, C. (2011). Longitudinal Development of Human Brain Wiring Continues from Childhood into Adulthood. *Journal of Neuroscience*, *31*(30), 10937–10947.

Lecker, M., Dotsch, R., Bijlstra, G., & Aviezer, H. (2020). Bidirectional contextual influence between faces and bodies in emotion perception. *Emotion*, *20*(7), 1154–1164.

Lecker, M., Shoval, R., Aviezer, H., & Eitam, B. (2017). Temporal integration of bodies and faces: United we stand, divided we fall? *Visual Cognition*, *25*(4-6), 477–491.

Leekam, S. R., Solomon, T. L., & Teoh, Y.-S. (2010). Adults' social cues facilitate young children's use of signs and symbols: Social cues and symbols. *Developmental Science*, *13*(1), 108–119.

Leitzke, B. T., & Pollak, S. D. (2016). Developmental changes in the primacy of facial cues for emotion recognition. *Developmental Psychology*, *52*(4), 572–581.

Logothetis, N. K. (2003). The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal. *The Journal of Neuroscience*, *23*(10), 3963–3971.

Lopez, L. D., Reschke, P. J., Knothe, J. M., & Walle, E. A. (2017). Postural Communication of Emotion: Perception of Distinct Poses of Five Discrete Emotions. *Frontiers in Psychology*, *8*, 710.

Lott, L. A., Haegerstrom-Portnoy, G., Schneck, M. E., & Brabyn, J. A. (2005). Face Recognition in the Elderly: *Optometry and Vision Science*, *82*(10), 874–881.

Lu, Z.-L., & Dosher, B. (2014). *Visual psychophysics: From laboratory to theory*. The MIT Press.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces*. Department of Clinical Neuroscience, Psychology section, Karolinska Institutet.

Mah, A., Geeraert, B., & Lebel, C. (2017). Detailing neuroanatomical development in late childhood and early adolescence using NODDI. *PLOS ONE*, *12*(8), e0182340.

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.-F., & Gruetter, R. (2010). MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *NeuroImage*, *49*(2), 1271–1281.

Martin, A. E. (2016). Language Processing as Cue Integration: Grounding the Psychology of Language in Perception and Neurophysiology. *Frontiers in Psychology*, *7*.

McNab, J. A., Edlow, B. L., Witzel, T., Huang, S. Y., Bhat, H., Heberlein, K., Feiweier, T., Liu, K., Keil, B., Cohen-Adad, J., Tisdall, M. D., Folkerth, R. D., Kinney, H. C., & Wald, L. L. (2013). The Human Connectome Project and beyond: Initial applications of 300mT/m gradients. *NeuroImage*, *80*, 234–245.

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, *102*(45), 16518–16523.

Meijer, D., Veselič, S., Calafiore, C., & Noppeney, U. (2019). Integration of audiovisual spatial signals is not consistent with maximum likelihood estimation. *Cortex*, *119*, 74–88.

Meyers, E. M., Borzello, M., Freiwald, W. A., & Tsao, D. (2015). Intelligent Information Loss: The Coding of Facial Identity, Head Pose, and Non-Face Information in the Macaque Face Patch System. *Journal of Neuroscience*, *35*(18), 7069–7081.

Mills, K. L., Goddings, A.-L., Herting, M. M., Meuwese, R., Blakemore, S.-J., Crone, E. A., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., & Tamnes, C. K. (2016). Structural brain development between childhood and adulthood: Convergence across four longitudinal samples. *NeuroImage*, *141*, 273–281.

Mills, K. L., Siegmund, K. D., Tamnes, C. K., Ferschmann, L., Wierenga, M., Bos, M. G. N., Luna, B., & Herting, M. M. (2021). Individual variability in structural brain development from late childhood to young adulthood, 31.

Minnebusch, D. A., & Daum, I. (2009). Neuropsychological mechanisms of visual face and body perception. *Neuroscience & Biobehavioral Reviews*, *33*(7), 1133–1144.

Mirzaalian, H., Ning, L., Savadjiev, P., Pasternak, O., Bouix, S., Michailovich, O., Grant, G., Marx, C., Morey, R., Flashman, L., George, M., McAllister, T., Andaluz, N., Shutter, L., Coimbra, R., Zafonte, R., Coleman, M., Kubicki, M., Westin, C., ... Rathi, Y. (2016). Inter-site and inter-scanner diffusion MRI data harmonization. *NeuroImage*, *135*, 311–323.

Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, *141*, 4–15.

Mondloch, C. J. (2012). Sad or fearful? The influence of body posture on adults' and children's perception of facial displays of emotion. *Journal of Experimental Child Psychology*, *111*(2), 180–196.

Mondloch, C. J., Horner, M., & Mian, J. (2013). Wide eyes and drooping arms: Adult-like congruency effects emerge early in the development of sensitivity to emotional faces and body postures. *Journal of Experimental Child Psychology*, *114*(2), 203–216.

Muukkonen, I., Ölander, K., Numminen, J., & Salmela, V. (2020). Spatio-temporal dynamics of face perception. *NeuroImage*, *209*, 116531.

Nardini, M., Bedford, R., & Mareschal, D. (2010). Fusion of visual cues is not mandatory in children. *Proceedings of the National Academy of Sciences*, *107*(39), 17041–17046.

Nelson, N. L., & Mondloch, C. J. (2017). Adults' and children's perception of facial expressions is influenced by body postures even for dynamic stimuli. *Visual Cognition*, *25*(4-6), 563–574.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(24), 9868–9872.

Ogawa, S., & Lee, T.-M. (1990). Magnetic resonance imaging of blood vessels at high fields:In vivo andin vitro measurements and image simulation. *Magnetic Resonance in Medicine*, *16*(1), 9–18.

Olney, N. T., Spina, S., & Miller, B. L. (2017). Frontotemporal Dementia. *Neurologic Clinics*, *35*(2), 339–374.

Ortibus, E., Verhoeven, J., Sunaert, S., Casteels, I., De Cock, P., & Lagae, L. (2012). Integrity of the inferior longitudinal fasciculus and impaired object recognition in children: A diffusion tensor imaging study: ILF and Object Recognition in Children. *Developmental Medicine & Child Neurology*, *54*(1), 38–43.

Palermo, R., Jeffery, L., Lewandowsky, J., Fiorentini, C., Irons, J. L., Dawel, A., Burton, N., McKone, E., & Rhodes, G. (2018). Adaptive face coding contributes to individual differences in facial expression recognition independently of affective factors. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(4), 503–517.

Palermo, R., O'Connor, K. B., Irons, J., & McKone, E. (2013). New Tests to Measure Individual Differences in Matching and Labelling Facial Expressions of Emotion, and Their Association with Ability to Recognise Vocal Emotions and Facial Identity. *PLoS ONE*, *8*(6), e68126.

Paquola, C., Bethlehem, R. A., Seidlitz, J., Wagstyl, K., Romero-Garcia, R., Whitaker, K. J., Vos de Wael, R., Williams, G. B., NSPN Consortium, Vértes, P. E., Margulies, D. S., Bernhardt, B., & Bullmore, E. T. (2019). Shifts in myeloarchitecture characterise adolescent development of cortical gradients. *eLife*, *8*, e50482.

Patterson, K., & Lambon Ralph, M. A. (2016). The Hub-and-Spoke Hypothesis of Semantic Memory. *Neurobiology of Language* (pp. 765–775). Elsevier.

Paus, T. (2010). Growth of white matter in the adolescent brain: Myelin or axon? *Brain and Cognition*, *72*(1), 26–35.

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the Human Body in the Fusiform Gyrus. *Journal of Neurophysiology*, *93*(1), 603–608.

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

Perry, A., Aviezer, H., Goldstein, P., Palgi, S., Klein, E., & Shamay-Tsoory, S. G. (2013). Face or body? Oxytocin improves perception of emotions from facial expressions in incongruent emotional body context. *Psychoneuroendocrinology*, *38*(11), 2820–2825.

Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133.

Philippi, C. L., Mehta, S., Grabowski, T., Adolphs, R., & Rudrauf, D. (2009). Damage to Association Fiber Tracts Impairs Recognition of the Facial Expression of Emotion. *Journal of Neuroscience*, *29*(48), 15089–15099.

Pitcher, D., Japee, S., Rauth, L., & Ungerleider, L. G. (2017). The Superior Temporal Sulcus Is Causally Connected to the Amygdala: A Combined TBS-fMRI Study. *The Journal of Neuroscience*, *37*(5), 1156–1161.

Pitcher, D., & Ungerleider, L. G. (2021). Evidence for a Third Visual Pathway Specialized for Social Perception. *Trends in Cognitive Sciences*, *25*(2), 100–110.

Premereur, E., Taubert, J., Janssen, P., Vogels, R., & Vanduffel, W. (2016). Effective Connectivity Reveals Largely Independent Parallel Networks of Face and Body Patches. *Current Biology*, *26*(24), 3269–3279.

Prins, N. (2013). The psi-marginal adaptive method: How to give nuisance parameters the attention they deserve (no more, no less). *Journal of Vision*, *13*(7), 3–3.

Prins, N., & Kingdom, F. A. A. (2018). Applying the Model-Comparison Approach to Test Specific Research Hypotheses in Psychophysical Research Using the Palamedes Toolbox. *Frontiers in Psychology*, *9*, 1250.

Pyles, J. A., Verstynen, T. D., Schneider, W., & Tarr, M. J. (2013). Explicating the Face Perception Network with White Matter Connectivity (M. Ptito, Ed.). *PLoS ONE*, *8*(4), e61611.

Rajhans, P., Jessen, S., Missana, M., & Grossmann, T. (2016). Putting the face in context: Body expressions impact facial emotion processing in human infants. *Developmental Cognitive Neuroscience*, *19*, 115–121.

Rajimehr, R., Young, J. C., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences*, *106*(6), 1995–2000.

Raven, E., Veraart, J., Kievit, R., Genc, S., Ward, I., Cunningham, A., Doherty, J., Bree, M. v. d., & Jones, D. (2020). *In vivo evidence of microstructural hypoconnectivity of brain white matter in 22q11.2 deletion syndrome* (preprint). In Review.

Read, J. (2015). The place of human psychophysics in modern neuroscience. *Neuroscience*, *296*, 116–129.

Reed, C. L., Bukach, C. M., Garber, M., & McIntosh, D. N. (2018). It's Not All About the Face: Variability Reveals Asymmetric Obligatory Processing of Faces and Bodies in Whole-Body Contexts. *Perception*, *47*(6), 626–646.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The Body-Inversion Effect. *Psychological Science*, *14*(4), 302–308.

Robbins, R. A., & Coltheart, M. (2012). The effects of inversion and familiarity on face versus body cues to person recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1098–1104.

Rodger, H., Vizioli, L., Ouyang, X., & Caldara, R. (2015). Mapping the development of facial expression recognition. *Developmental Science*, *18*(6), 926–939.

Ross, P., de Gelder, B., Crabbe, F., & Grosbras, M.-H. (2019). Emotion modulation of the body-selective areas in the developing brain. *Developmental Cognitive Neuroscience*, 10.

Ross, P. D., Polson, L., & Grosbras, M.-H. (2012). Developmental Changes in Emotion Recognition from Full-Light and Point-Light Displays of Body Movement (M. Lappe, Ed.). *PLoS ONE*, *7*(9), e44815.

Rossion, B., & Gauthier, I. (2002). How Does the Brain Process Upright and Inverted Faces? *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 63–75.

Sairanen, V., Leemans, A., & Tax, C. (2018). Fast and accurate Slicewise OutLIer Detection (SOLID) with informed model estimation for diffusion MRI data. *NeuroImage*, *181*, 331–346.

Scarfe, P. (2020). *Experimentally disambiguating models of sensory cue integration* (preprint). Neuroscience.

Scherf, K. S., Thomas, C., Doyle, J., & Behrmann, M. (2014). Emerging Structure-Function Relations in the Developing Face Processing System. *Cerebral Cortex*, *24*(11), 2964–2980.

Schindler, S., Bruchmann, M., Gathmann, B., Moeck, R., & Straube, T. (2021). Effects of low-level visual information and perceptual load on P1 and N170 responses to emotional expressions. *Cortex*, *136*, 14–27.

Schobert, A.-K., Corradi-Dell'Acqua, C., Frühholz, S., van der Zwaag, W., & Vuilleumier, P. (2018). Functional organization of face processing in the human superior temporal sulcus: A 7T high-resolution fMRI study. *Social Cognitive and Affective Neuroscience*, *13*(1), 102–113.

Schyns, P., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.

Seunarine, K. K., & Alexander, D. C. (2014). Multiple Fibers. *Diffusion MRI* (pp. 105–123). Elsevier.

Shahrestani, S., Kemp, A. H., & Guastella, A. J. (2013). The Impact of a Single Administration of Intranasal Oxytocin on the Recognition of Basic Emotions in Humans: A Meta-Analysis. *Neuropsychopharmacology*, *38*(10), 1929–1936.

Shams, L., & Beierholm, U. R. (2010). Causal inference in perception. *Trends in Cognitive Sciences*, *14*(9), 425–432.

Simion, F., & Giorgio, E. D. (2015). Face perception and processing in early infancy: Inborn predispositions and developmental changes. *Frontiers in Psychology*, *6*.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, *17*(3), 143–155.

Song, Y., Luo, Y. L. L., Li, X., Xu, M., & Liu, J. (2013). Representation of Contextually Related Multiple Objects in the Human Ventral Visual Pathway. *Journal of Cognitive Neuroscience*, *25*(8), 1261–1269.

Sorger, B., Goebel, R., Schiltz, C., & Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *NeuroImage*, *35*(2), 836–852.

Spiridon, M., Fischl, B., & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping*, *27*(1), 77–89.

Stevens, S. S. (1956). The Direct Estimation of Sensory Magnitudes: Loudness. *The American Journal of Psychology*, *69*(1), 1–25.

Sullivan, S., & Ruffman, T. (2004). Emotion Recognition Deficits in the Elderly. *International Journal of Neuroscience*, *114*(3), 403–432.

Tavares, P., Barnard, P. J., & Lawrence, A. D. (2011). Emotional Complexity and the Neural Representation of Emotion in Motion. *Social Cognitive and Affective Neuroscience*, *6*(1), 98–108.

Tavor, I., Yablonski, M., Mezer, A., Rom, S., Assaf, Y., & Yovel, G. (2014). Separate parts of occipito-temporal white matter fibers are associated with recognition of faces and places. *NeuroImage*, *86*, 123–130.

Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI Analysis of Body and Body Part Representations in the Extrastriate and Fusiform Body Areas. *Journal of Neurophysiology*, *98*(3), 1626–1633.

Teufel, C., Westlake, M. F., Fletcher, P. C., & von dem Hagen, E. (2019). A hierarchical model of social perception: Psychophysical evidence suggests late rather than early integration of visual information from facial expression and body posture. *Cognition*, *185*, 131–143.

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, *10*(5), 547–558.

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249.

Tournier, J.-D., Calamante, F., & Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: Non-negativity constrained super-resolved spherical deconvolution. *NeuroImage*, *35*(4), 1459–1472.

Tournier, J.-D., Calamante, F., Gadian, D. G., & Connelly, A. (2004). Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *NeuroImage*, *23*(3), 1176–1185.

Tournier, J.-D., Smith, R., Raffelt, D., Tabbara, R., Dhollander, T., Pietsch, M., Christiaens, D., Jeurissen, B., Yeh, C.-H., & Connelly, A. (2019). MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation. *NeuroImage*, *202*, 116137.

Trommershauser, J., Kording, K., & Landy, M. S. (Eds.). (2011). *Sensory cue integration*. Oxford University Press.

Tubridy, N., & McKinstry, C. S. (2000). Neuroradiological history: Sir Joseph Larmor and the basis of MRI physics. *Neuroradiology*, *42*(11), 852–855.

Unger, A., Alm, K. H., Collins, J. A., O'Leary, J. M., & Olson, I. R. (2016). Variation in White Matter Connectivity Predicts the Ability to Remember Faces and Discriminate Their Emotions. *Journal of the International Neuropsychological Society*, *22*(02), 180–190.

Ungerleider, L. (1994). What and where in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157–165.

VanRullen, R., & Thorpe, S. J. (2001). The Time Course of Visual Processing: From Early Perception to Decision-Making. *Journal of Cognitive Neuroscience*, *13*(4), 454–461.

Veraart, J., Fieremans, E., & Novikov, D. S. (2016). Diffusion MRI noise mapping using random matrix theory: Diffusion MRI Noise Mapping. *Magnetic Resonance in Medicine*, *76*(5), 1582–1593.

Veraart, J., Fieremans, E., & Novikov, D. S. (2019). On the scaling behavior of water diffusion in human brain white matter. *NeuroImage*, *185*, 379–387.

Veraart, J., Nunes, D., Rudrapatna, U., Fieremans, E., Jones, D. K., Novikov, D. S., & Shemesh, N. (2020). Noninvasive quantification of axon radii using diffusion MRI. *eLife*, *9*, e49855.

Veraart, J., Poot, D. H. J., Van Hecke, W., Blockx, I., Van der Linden, A., Verhoye, M., & Sijbers, J. (2011). More accurate estimation of diffusion tensor parameters using diffusion kurtosis imaging: Improved Estimation of DTI Parameters with DKI. *Magnetic Resonance in Medicine*, *65*(1), 138–145.

Vieillard, S., & Guidetti, M. (2009). Children's perception and understanding of (dis)similarities among dynamic bodily/facial expressions of happiness, pleasure, anger, and irritation. *Journal of Experimental Child Psychology*, *102*(1), 78–95.

Vocks, S., Busch, M., Grönemeyer, D., Schulte, D., Herpertz, S., & Suchan, B. (2010). Differential neuronal responses to the self and others in the extrastriate body area and the fusiform body area. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(3), 422–429.

Vos, S. B., Tax, C. M. W., Luijten, P. R., Ourselin, S., Leemans, A., & Froeling, M. (2017). The importance of correcting for signal drift in diffusion MRI. *Magnetic Resonance in Medicine*, *77*(1), 285–299.

Wang, S. (2018). Face size biases emotion judgment through eye movement. *Scientific Reports*, *8*(1).

Wasserthal, J., Neher, P. F., Hirjak, D., & Maier-Hein, K. H. (2019). Combined tract segmentation and orientation mapping for bundle-specific tractography. *Medical Image Analysis*, *58*, 101559.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. The Psychological Corporation: Harcourt Brace & Company.

Wegrzyn, M., Riehle, M., Labudda, K., Woermann, F., Baumgartner, F., Pollmann, S., Bien, C. G., & Kissler, J. (2015). Investigating the brain basis of facial expression perception using multi-voxel pattern analysis. *Cortex*, *69*, 131–140.

Whitaker, K. J., Vértes, P. E., Romero-Garcia, R., Váša, F., Moutoussis, M., Prabhu, G., Weiskopf, N., Callaghan, M. F., Wagstyl, K., Rittman, T., Tait, R., Ooi, C., Suckling, J., Inkster, B., Fonagy, P., Dolan, R. J., Jones, P. B., Goodyer, I. M., the NSPN Consortium, & Bullmore, E. T. (2016). Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proceedings of the National Academy of Sciences*, *113*(32), 9105–9110.

Whiteley, L., & Sahani, M. (2008). Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *Journal of Vision*, *8*(3), 2.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313.

Widen, S. C., & Russell, J. A. (2010a). The "disgust face" conveys anger to children. *Emotion*, *10*(4), 455–466.

Widen, S. C., & Russell, J. A. (2010b). Differentiation in preschooler's categories of emotion. *Emotion*, *10*(5), 651–661.

Widen, S. C., & Russell, J. A. (2013). Children's recognition of disgust in others. *Psychological Bulletin*, *139*(2), 271–299.

Winston, G. P. (2012). The physical and biological basis of quantitative parameters derived from diffusion MRI. *Quantitative Imaging in Medicine and Surgery*, *2*(4), 254–265.

Woolrich, M. W., Behrens, T. E., Beckmann, C. F., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, *21*(4), 1732–1747.

Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *NeuroImage*, *14*(6), 1370–1386.

Worsley, K. J. (2001). Statistical analysis of activation images. In P. Jezzard, P. M. Matthews, & S. M. Smith (Eds.), *Functional Magnetic Resonance Imaging* (pp. 251–270). Oxford University Press.

Yakovlev, P., & Lecours, A. (1967). *The myelogenetic cycles of regional maturation of the brain* (Minkowski A, eds). Oxford: Blackwell.

Yeatman, J. D., Weiner, K. S., Pestilli, F., Rokem, A., Mezer, A., & Wandell, B. A. (2014). The vertical occipital fasciculus: A century of controversy resolved by in vivo measurements. *Proceedings of the National Academy of Sciences*, *111*(48), E5214–E5223.

Yogarajah, M., Duncan, J., & Koepp, M. (2009). IMAGING — Structure and Function: Imaging Connectivity Using Tractography. *Encyclopedia of Basic Epilepsy Research* (pp. 1586–1592). Elsevier.

Zaki, J. (2013). Cue Integration: A Common Framework for Social Cognition and Physical Perception. *Perspectives on Psychological Science*, *8*(3), 296–312.

Zieber, N., Kangas, A., Hock, A., & Bhatt, R. S. (2014). Infants' Perception of Emotion From Body Movements. *Child Development*, *85*(2), 675–684.