EMOTION RECOGNITION IN THE HUMAN FACE AND VOICE

A thesis submitted for the degree of Doctor of Philosophy

By

Lisa Katharina Kuhn

Department of Psychology, Brunel University London

December 2014

I, Lisa Katharina Kuhn, declare that this thesis is written by me and that the data presented in it stems from my original research work. Wherever contributions of others are involved, this is indicated clearly such as by providing references to other literature.

Abstract

At a perceptual level, faces and voices consist of very different sensory inputs and therefore, information processing from one modality can be independent of information processing from another modality (Adolphs & Tranel, 1999). However, there may also be a shared neural emotion network that processes stimuli independent of modality (Peelen, Atkinson, & Vuilleumier, 2010) or emotions may be processed on a more abstract cognitive level, based on meaning rather than on perceptual signals. This thesis therefore aimed to examine emotion recognition across two separate modalities in a within-subject design, including a cognitive Chapter 1 with 45 British adults, a developmental Chapter 2 with 54 British children as well as a cross-cultural Chapter 3 with 98 German and British children, and 78 German and British adults. Intensity ratings as well as choice reaction times and correlations of confusion analyses of emotions across modalities were analysed throughout. Further, an ERP Chapter investigated the time-course of emotion recognition across two modalities. Highly correlated rating profiles of emotions in faces and voices were found which suggests a similarity in emotion recognition across modalities. Emotion recognition in primary-school children improved with age for both modalities although young children relied mainly on faces. British as well as German participants showed comparable patterns for rating basic emotions, but subtle differences were also noted and Germans perceived emotions as less intense than British.

Overall, behavioural results reported in the present thesis are consistent with the idea of a general, more abstract level of emotion processing which may act independently of modality. This could be based, for example, on a shared emotion brain network or some more general, higher-level cognitive processes which are activated across a range of modalities. Although emotion recognition abilities are already evident during childhood, this thesis argued for a contribution of ‘nurture’ to emotion mechanisms as recognition was influenced by external factors such as development and culture.

Acknowledgments

My thanks and gratitude goes to my excellent supervisors Prof. Taeko Wydell and Dr. Lucia Garrido for their extraordinary support, for encouraging me and for believing in me. Thank you for all your feedback, ideas, keeping me on the right track and proof-reading this thesis. It was a pleasure being your student. I would also like to especially thank Prof. Michael Wright for his patience in teaching me the wonders of the EEG world.

Thanks to the Economics and Social Research Council who made this thesis possible by funding this research and for giving me the opportunity to attend various conferences. Thank you also to my brilliant previous supervisor Dr. Maria Uther who guided me through my early years at university and who encouraged me to go for the 1+3 ESRC studentship in the first place.

A big thank you to all colleagues from the PG office who turned into close friends. I will miss the time we had together! Thanks also to the rest of the Psychology department for feedback and to the Psychology technicians who were always happy to help.

Most of all, thank you for all the lovely people that have walked the long way with me. It’s not where you go; it’s who you travel with. The biggest ‘Danke’ to my family back in Germany – always there in my thoughts - as well as my adopted English family Jackie and Klaus; thanks for the support! Without you, none of this would have ever been possible. A big ‘thank you’ also goes to Krysia, Peter and my darling Max who never fails to show me the truly important things in life. Thank you to my QMSC windsurfing crew for providing me with a great work-life balance and to always push me to break out of my comfort zone. A special thanks also goes to my friends in England and back home for keeping me sane.

Thank you to each and every colleague I have met at conferences along the way, I hope we can continue to exchange ideas in the future.

Lastly, thanks to all my participants who helped to make this project possible and especially to Mrs. J.M. and the local primary school, my parents and Jackie for organising access to participants for me.

Thank you for believing in me and being there for me. It was a great journey.

Contents

List of Tables

List of Figures

1

Introduction

There's no art to find the mind's construction in the face. He was a gentleman on whom I built an absolute trust.

Shakespeare, Macbeth (1.4.14-5, 2010/1699)

In Shakespeare's famous play Macbeth (2010/1699), King Duncan got betrayed by Thane of Cawdor and subsequently concluded that one cannot read another person's mind or trustfulness from the face alone. This quote may still apply today to some every-day situations such as communicative misunderstandings. On the other hand, it is also evident that we are actually very well adapted to read other's emotions in the face – and even in the voice. To what extend voices and faces indeed signal specific emotions that can be reliably recognised and whether both modes of communication are equally useful to read another person's emotional state will be the focus of this particular work. In everyday life, information from voice and face are usually combined simultaneously in order to deliver a more detailed picture of social situations. However, the present thesis aims to investigate vocal as well as facial information separately in order to examine how we recognise emotions when we only have one modality available such as in darkness or very noisy situations. The aim of this thesis is to demonstrate to the reader that –contrary to Shakespeare's well-known words – humans are to some degree able to see the mind's construction in the face and even in the voice. Further, the goal of this thesis is to investigate whether this emotion recognition ability is already evident in primary-school aged children and what role cultural influences play during emotion recognition across modalities.

1.1 The Six Basic Emotions

According to Hess and Thibault (2009), '*emotions are considered to be relatively short-duration intentional states that entrain changes in motor behavior, physiological changes, and cognitions'* (p. 120). The ability to produce as well as recognise emotions has often been contributed to the concept of social and emotional intelligence (Goleman, 1995; Salovey, & Mayer, 1990) and is thought to be a separate construct to general IQ or personality traits (Mayer, 2000). Emotional competence (Scherer, 2007) can be used as an umbrella term for emotion production, regulation and perception.

Whilst all three mechanisms are related and needed for high levels of emotional intelligence, the present thesis focuses on the ability to perceive and categorise emotions. The six basic emotions covered in this piece of research are the emotions of happiness, sadness, fear, anger, surprise and disgust as described by Ekman and Friesen (1975).

Although the focus of the present thesis is not emotion expression but rather emotion recognition, it is important to understand how the basic emotions (Ekman & Friesen, 1975) can be produced in order to be recognised by others. In past literature on emotion production, competing hypotheses have either claimed emotion expressions that are based on categorical emotion-specific responses (*e.g*. Ekman & Friesen, 1976) or emotion expressions that are based on the subjective appraisal of events which are not necessarily specific to one particular emotion (*e.g*. Frijda, Kuipers, & ter Schure, 1989; Scherer, 1984). In more detail, accounts of emotion expressions based on discrete categories have been supported by innate and universal recognition of prototypical emotion displays (Izard, 1994). It is believed that specific events trigger corresponding bodily responses as reflected in specific emotion expressions that are universal. For example, the thought of an imminent maths test may cause tensioning of face muscles and wide-opened eyes, and this combination may display the emotion 'fear'. On the other hand, appraisal theories – such as the Component Process Model by Scherer (1984) – suggest a combination of cognitive, physiological and motor behaviours that follow the subjective evaluation of an event. This evaluation of the situation controls for novelty, pleasantness or coping resources needed following the onset of an event. And so, the thought of the imminent maths test may be less distressing to the reader than it is to the writer due to different coping resources. Emotion processes can also be driven by social norms which seem appropriate during the experience of an emotional event (*e.g*. Frijda, 1986).

Happiness is the only positive emotion included in the six basic emotions and some researchers have begun to include several other positive emotions such as pleasure or relief in emotion research (*e.g*. Sauter & Scott, 2007). Expressing happiness is part of a very basic and instinct human behaviour and even infants can distinguish between emotions of positive and negative valence (Grossman, Striano, & Friederici, 2007). Happiness can be caused by several factors such as pleasure, excitement, relief or following social approval. In terms of appraisal theories, positive emotions such as happiness are the result of initial stimulus evaluation such as pleasantness which is, for example, followed by the face muscle contraction of pulling lip corners upwards resulting in a smile (Scherer, 2009). Sadness, anger, fear and disgust on the other hand are prevailing negative emotions. Sadness usually is a longer-lasting feeling that expresses suffering as a consequence of a loss. Sadness can be the result of failing to cope with helplessness and grief and can therefore blend with feelings of anger or fear (Ekman & Friesen, 1975/2003).

According to appraisal theories, negative events, which may obstruct reaching a personal goal, are accompanied by bodily changes such as increased heart rate or tensioning of muscles in order to prepare the organism for a possible attack.

This physiological response is associated with an increase in voice frequency and tightening of eye lids, resulting in motor expressions of the felt emotion following the appraisal of the current situation (Scherer, 2009).Very generally, fear – but also anger - processing is thought to be associated with amygdala activation within the human brain (*e.g.* Adolphs, 2002).

Disgust can be a reaction towards physical stimuli such as smells and sounds as well as towards socio-moral types of behaviour that elicit disgust or even anger. In evolutionary terms, disgusted reactions could save lives by protecting the organism from poisonous substances or infections by spitting out spoiled food (Chapman & Anderson, 2012). Disgust perception is commonly associated with activation in the insula (Wicker, Keysers, Plailly, Royet, Gallese, & Rizzolatti, 2001). The last of the six basic emotions is the emotion of surprise which is of neutral nature and usually rapidly followed by either positive or negative emotions. According to discrete accounts of emotion production, surprise is an emotion of short duration, felt after “unexpected or misexpected events” (Ekman & Friesen, 1975/2003, p. 25). According to appraisal theories, motor expressions of positive or negative surprise follow the appraisal of unexpected events with subsequent appraisal of coping resources available to deal with a surprising situation (Scherer & Ellgrin, 2007).

All emotions can vary in intensity and the expression may be influenced by previous experiences, stereotypes as well as personality types (Ekman & Friesen, 1975/2003). Especially appraisal theories suggest individual differences in emotion expression due to underlying differences in appraisal checks and coping resources following an event (Scherer, 2009). More details about specific emotion expressions in either faces or voices and their neural mechanisms are described below.

1.2 Faces

In everyday life, basic and universal emotion cues are most obviously communicated via the visual domain. Faces are very important social stimuli which can communicate information regarding gender (Sung, Gao, & Han, 2010), mental states (Adolphs, 2002), perceived attractiveness (Perrett, Burt, Penton-Voak, Lee, Rowland, & Edwards, 1999) and even dominance levels (Neave et al., 2003) of the expresser.

1.2.1 Vision and Face Perception

As summarised by Goldstein (2001), visual perception starts with objects (or in the present case with faces) that cause light to be reflected from the surface onto the retina inside the human eye. After light has been turned into electrical signals due to chemical reactions in receptive cells in the eye, the signals then travel along the optical nerve to the bilateral lateral geniculate nucleus in the thalamus before being sent to the visual striate cortex in the occipital lobe.

Whilst neurons in the striate cortex are specialised to respond to specific features of the object such as orientation (*e.g*. Bosking, Zhang, Schofield, & Fitzpatrick, 1997), extrastriate areas then allow further in-depth processing of the stimulus such as face expressions (*e.g*. Kanwisher, 1997). Visual information gets then passed on via the ventral or 'what stream', from the visual to the inferior-temporal cortex, which serves object identification. On the other hand, visual information related to location of objects and associated motor response preparation will be sent to the posterior parietal regions via the dorsal or 'where stream' (Goodale & Milner, 1992). Whilst this is a very basic summary of visual perception, the reader is referred to Goldstein (2001) for a more detailed description of visual processing in the human brain.

However, not all visual objects are the same and for humans, faces can play a crucial role during social perception or communication. Even new-born infants pay special attention to human faces as shown by increased eye-tracking behaviour of moving face patterns compared to non-face patterns which suggests inborn abilities for recognising the structure of human faces (Morton & Johnson, 1991). Typical timings of brain responses for processing faces are often observed in ERP[1] studies at around 170ms after stimulus onset which is thought to reflect a face-specific negative component of the event-related potential (*e.g.* Eimer, 2011; Luck, 2005).

Past research from patients with the neurological disorder prosopagnosia, or so-called 'face blindness', has demonstrated separate pathways for object versus face identity recognition within the human brain (Duchaine & Nakayama, 2005). For these individuals, vision and object recognition is intact whilst the recognition of faces is impaired (Bodamer, 1947).The brain has shown a specialised face processing area which consists of three main structures, namely the lateral fusiform gyrus - which is also famously known as the fusiform face area FFA (Kanwisher, McDermott, & Chun, 1997) - as well as the lateral inferior occipital gyrus (occipital face area OFA) and the posterior superior temporal sulcus (STS, Benton et al., 1980). Faces are usually recognised as a whole rather than as a sum of its individual features (Farah, Wilson, Drain, & Tanaka, 1998). Trying to categorise faces that are displayed up-side-down seems to disturb this holistic way of processing faces, resulting in longer processing times based on individual features alone (Rossion, 2009). However, the face-specificity has also been criticised and instead attributed to general expertise recognition (Diamond & Carey, 1986).

1.2.2 Emotional Face Recognition

Humans are generally pretty good at recognising emotions expressed in the face. Kohler and colleagues (2004) for example reported high accuracy rates for categorising four basic emotions in static photos of faces with happy faces being recognised with 91.2% accuracy and sad faces with 84% accuracy.

[1] Event related potential. Average of electrophysiological brain responses following events such as watching emotional faces. Recorded by electroencephalography EEG (Luck, 2005).

ERP studies suggest that emotionality in faces can be recognised as early as100ms after stimulus onset whilst differentiation between specific emotion categories occurred as early as 140 ms after stimulus onset (Batty & Taylor, 2003). Especially the mechanisms for recognising fear is believed to elicit early brain responses: Ashley, Vuilleumier and Swick (2003) reported enhanced negativity for fearful versus happy or neutral facial stimuli at around 200ms after stimulus onset (see also Kiss & Eimer, 2008).

How is the richness of information from faces being processed? In 1986, Bruce and Young first proposed a face perception model that introduced separate functional pathways for speech analysis, identity analysis or expression analysis for extracting specific information from the face. According to this model, we structurally encode the face to extract its features. This information gets stored as face recognition units for future face recognition or it can be used for categorization of emotion expressions and both processes can happen in parallel and functionally separate from each other.

This functional dissociation has been confirmed by data from prosopagnosic patients which suggested that whilst emotion recognition is intact, some individuals had difficulties recognising familiar faces, indicating deficits for identity recognition rather than general face perception deficits (Tranel, Damasio, & Damasion, 1988). In contrast, prosopo-affective agnosia describes impaired face emotion recognition in patients whilst face identity recognition remains intact (Kurucz & Feldmar, 1979). In 2002, Haxby, Hoffman and Gobbini reported separate neural pathways for processing either identity or emotion expression from faces. As illustrated in Figure 1.1, they found that cells in the lateral fusiform gyrus seemed to be specialised for identity recognition whilst cells in the superior temporal sulcus (STS) responded mainly to changes in face expressions and eye gaze.

However, more recent evidence suggests that the route of completely separate pathways for face recognition and face expression may not be as clear-cut. For example, Calder and Young (2005) propose that principal component analysis reveals overlapping visual pixels of images displaying faces that can be used to either decode face identity or face expressions. In addition, Baseler, Harris, Young and Andrews (2014) reported functional connectivity for face identity and face expressions mechanisms between the posterior superior temporal sulcus (STS) and other face areas such as the fusiform face area (FFA). Of particular interest for the present thesis is also the finding that the STS responds to emotion information not only from faces – but also from other modalities such as voices (Charest et al., 2009; Phillips et al., 1998). Hence, the aim of the current thesis is to investigate whether emotional face recognition – irrespective of the debate whether it is completely separated from face identity recognition or not – is modality specific or whether a general emotion mechanism (possibly within the STS) drives emotion recognition for faces and voices equally.

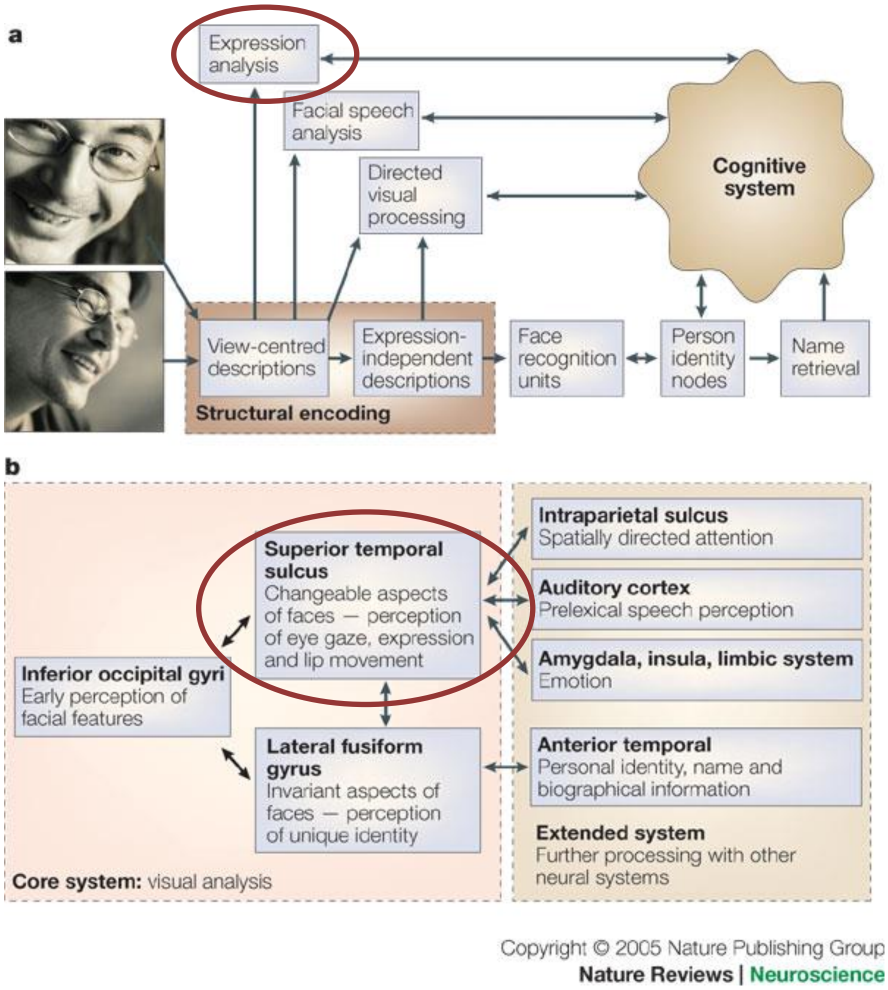


Figure 1.1. Face recognition model. Face identity and face expression recognition seem to be (a) functionally (Bruce & Young, 1986) as well as (b) anatomically segregated (Haxby et al., 2002). In Calder, & Young (2005).

1.2.2.1 Ekman Face Stimuli

How are emotions expressed and recognised in faces? Commonly, facial emotion perception is investigated by manipulating overt changes in facial musculature during face expressions. The Facial Action Coding System (FACS, Ekman & Friesen, 1976) was developed in order to assess facial emotion expression based on the combination of specific muscle activity within the face.
Just like acoustical features within the voice, a combination of action units (physical features based on muscle movement of the face) are believed to give a reliable and universal impression of the underlying emotions displayed (Sauter, Eisner, Calder, & Scott, 2010).

Common action units are summarised in Table 1.1. For example, happiness in faces is expressed by action units of tensing of the lower eyelids, pulling up the lip corners and raising the cheeks (Ekman & Friesen, 1976). Further, wrinkles from the nose down to the outer corner of the mouth as well as wrinkles at the outer corners of the eyes may be visible (Ekman & Friesen, 1975/2003). Sadness is portrayed by raising the inner eye brows and pulling down the lip corners.

Anger is expressed by drawing the eyebrows together and lowering them as well as tensing the lower eyelids and either pressing the lips together or with open mouth. The degree of muscle tension in the face can reflect the intensity of the emotion experienced (Ekman & Friesen, 1975/2003). Fear is demonstrated by raising the eyebrows, eyes are wide open and tense lower eyelids. The lips are usually stretched and tense and the mouth may be open or closed. Interestingly, surprise in the face shares several features with the expression of fear such as the raised eyebrows and wide open eyes. However, surprise shows a lesser degree of tension in the face compared to fear and whilst the lips are stretched with tension in fear, surprise often is expressed with a dropped and relaxed jaw (Ekman & Friesen, 1975/2003). Lastly, disgust is usually expressed by raising the upper lip with the lower lip either raised as well or lowered which results in an open mouth. Further, the nose is wrinkled and the eyebrows are lowered.

Hence, due to its widespread use, the present thesis includes the static black and white Ekman faces from the Pictures of Facial Affect database (Ekman & Friesen, 1975), posed by white Caucasian actors based on the FACS system. Although more recent emotion research has begun to include dynamic (Vieillard & Guidetti, 2009) or spontaneous (Jack, Garrod, & Schyns, 2014) face expression stimuli in order to simulate real-life situations, the Ekman faces are believed to be of a comparable difficulty level to non-verbal affect vocalisations (Hawk et al., 2009) which makes for an ideal basis for comparison.

Table 1.1

List of Action Units (AU) from the Facial Action Coding System (Ekman & Friesen, 1976, p.65)

AU number	Descriptor	Muscular Basis
1.	Inner Brow Raiser	Frontalis, Pars Medialis
2.	Outer Brow Raiser	Frontalis, Pars Lateralis
4.	Brow Lowerer	Depressor Glabellae, Depressor Supercilli; Corrugator
5.	Upper Lid Raiser	Levator Palpebrae Superioris
6.	Cheek Raiser	Orbicularis Oculi, Pars Orbitalis
7.	Lid Tightener	Orbicularis Oculi, Pars Palebralis
9.	Nose Wrinkler	Levator Labii Superioris, Alaeque Nasi
10.	Upper Lip Raiser	Levator Labii Superioris, Caput Infraorbitalis
11.	Nasolabial Fold Deepener	Zygomatic Minor
12.	Lip Corner Puller	Zygomatic Major
13.	Cheek Puffer	Caninus
14.	Dimpler	Buccinator
15.	Lip Corner Depressor	Triangularis
16.	Lower Lip Depressor	Depressor Labii
17.	Chin Raiser	Mentalis
18.	Lip Puckerer	Incisivii Labii Superioris; Incisivii Labii Inferioris
20.	Lip Stretcher	Risorius
22.	Lip Funneler	Orbicularis Oris
23.	Lip Tightener	Orbicularis Oris
24.	Lip Pressor	Orbicularis Oris
25.	Lips Part	Depressor Labii, or Relaxation of Mentalis or Orbicularis Oris
26.	Jaw Drop	Masetter; Temporal and Internal Pterygoid Relaxed
27.	Mouth Stretch	Pterygoids; Digastric
28.	Lip Suck	Orbicularis Oris

Note. AUs can be used in specific combinations to display specific emotions in faces, based on muscular activity. In Rosenberg, 2005.

1.3. Voices

Whilst recognising expressions in faces requires seeing the other person in front of oneself at a not too far distance, voices can be heard from further away and also from all angles without having to turn the head (Scherer, 1995). Hence, communicating emotions via the vocal modality may be especially useful for situations where immediate action is required.

1.3.1. Voice Perception

In basic terms, sound perception is based on the inner ear receiving sound pressure waves from the environment. Those sound waves are then transformed into electrical potentials when vibrations bend hair-cell receptors inside the cochlear which in turn causes those receptors to release neurotransmitters. Next, the electrical signal travels along the auditory nerve to the medial geniculate nucleus in the thalamus before being sent to the primary auditory area in the temporal cortex and further subcortical structures such as the amygdala (Goldstein, 2001). In terms of specialised pathways for auditory perception, comparing two sounds (in the ‘what’-stream) activated the auditory cortex and inferior frontal gyrus whilst locating the origin of two sounds relative to each other (in the ‘where’-stream) activated the posterior temporal cortex, parietal cortex and superior frontal sulcus, suggesting independent processing pathways of auditory information depending on the information extracted from sounds (Alain, Arnott, Hevenor, Graham, & Grady, 2001).

Just like faces are thought to be special kind of visual objects (*e.g.* Duchaine & Nakayama, 2005), voices are thought to be special kind of sounds. The vocal signals we receive and decode are based on acoustical features of the encoder’s voice such as pitch and amplitude and are created by air pressure waves from the lung that moves through the vocal cord and vocal tract. A natural variation in membrane length of vocal folds within the individual’s larynx and its vibrations creates a person’s characteristic pitch whilst the vocal tract above the larynx filters out specific frequencies of the pitch that determines sound production (Banse & Scherer, 1996; Ghazanfar & Rendall, 2008). Variance in size of vocal folds is also the reason for perceived gender differences in pitch with females naturally having smaller vocal folds which in turn produce a higher mean pitch (Juslin & Laukka, 2001). Amplitude in human sounds describes the loudness of the sound (Goldstein, 2001).

Similarly to faces (Kanwisher et al., 1997), voice selective areas, which respond more to voices than to other sounds, have been found and these are believed to be in the superior temporal sulcus (STS) within the auditory cortex (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; von Kriegstein & Giraud, 2004). However, the role of the STS during voice processing does not seem to be exclusive because voice-sensitive areas have also been found in pre-frontal brain areas (Fecteau, Armony, Joanette, & Belin, 2005). Further, it has also been proposed that face or voice-specific neural areas rather reflect general expertise-processing, independent of the object processed (Gauthier et al., 2000). In the same way that prosopagnosic patients show a dissociation between face and general visual object recognition (Duchaine & Nakayama, 2005), phonagnosic patients show dissociations between voice identity recognition and general voice discrimination. Whilst the first is linked to damage in the right parietal lobe, the latter is associated with damage in the bilateral temporal lobe (Van Lancker, Cummings, Kreiman, & Dobkin, 1988). Further, there also seems to be time-locked events in the brain in direct correspondence to human voice processing. Similarly to the N170 component found during face processing (*e.g.* Eimer, 2011), ERP studies have confirmed specific electrical responses between 160 and 200ms after stimulus onset for human voices as compared to non-vocal sounds, called the ‘fronto-temporal positivity to voices’ (FTPV, Charest et al., 2009).

1.3.2. Emotional Voice Recognition

Past research has demonstrated that emotions can be recognised above chance not only from faces – but also from voices. Just like physical features in the face, a combination of auditory features within the voice such as pitch and intensity can give insight into underlying emotional processes of the individual (*e.g.* Sauter, 2010). Across several studies, emotional prosody recognition accuracy seems to be around the 50% mark as observed by Banse and Scherer (1996). They suggested highest recognition accuracy for spoken sentences expressing hot anger (78%) with disgust only being correctly recognised in 15% of cases; this, however, was still higher than making judgements by chance alone. Juslin and Laukka (2001) reported highest decoding accuracy for prosody communicating sadness and fear, followed by anger, happiness and disgust; emotions with higher intensity were recognised with increased accuracy.

Affective information from voices is mainly processed within the right hemisphere such as the right temporal lobe, right inferior prefrontal cortex or right amygdala (Belin, Fecteau, & Bedard. 2004; Morris, Scott, & Dolan, 1999). However, emotional prosody has also been found to be processed bilaterally in the so-called emotional voice areas (EVA) which are located in the superior temporal gyrus close to the primary auditory cortex (Ethofer, Bretscher, Gschwind, Kreifelts, Wildgruber, & Vuilleumier, 2011).

As discussed above, emotional information from faces seems to be processed independently to identity information from faces – although the degree of independence may vary (Calder & Young, 2005). In direct comparison, emotional information from voices also seems to be processed independently from identity information from voices (Belin et al., 2004; see also Belin, Bestelmeyer, Latinus, & Watson, 2011). According to Belin et al.’s voice perception model (2004), after initial low-level analysis of auditory signals, structural encoding of auditory features follows which then leads to either voice identity or voice affect recognition. This is illustrated in Figure 1.2. Separate functional pathways for identity and affect recognition in voices have been demonstrated by phonagnosic patients: A single-case study of developmental phonagnosia has reported impaired recognition mechanisms for famous and new voices which were independent of emotional voice recognition mechanisms (Garrido, Eisner, McGettigan, Steward, Sauter, & Hanley, 2009).
For the present thesis, the pathway of affective processing for both modalities is of special interest. During emotion recognition of prosody, the right posterior STS as well as frontal areas showed increased BOLD[2] responses (Wildgruber et al., 2005)

[2] BOLD: Blood oxygenation level-dependent. Hemodynamic response to increase blood oxygen levels in active brain regions following neural activity (Aguirre, Zarahn, & D’Esposito (1998).

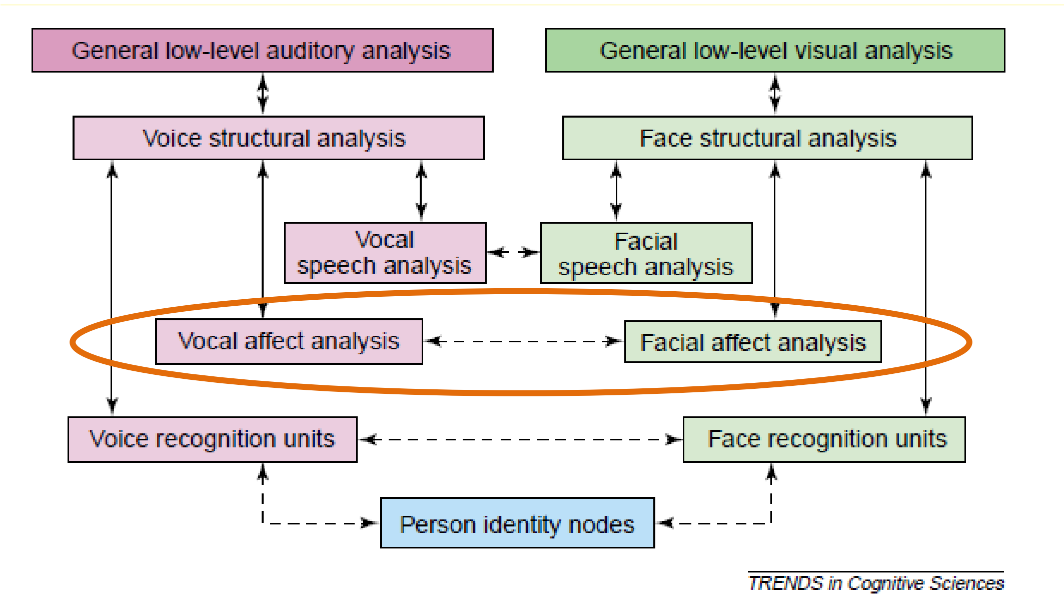


Figure 1.2. Comparable structure between voice (pink) versus face (green) perception. Emotion recognition seems to be independent of identity and speech recognition for both modalities. From Belin et al., 2004

Generally, good verbal knowledge is needed to understand emotion categories and semantic meaning of emotion words can influence general emotion processing (Lindquist, Feldman Barrett, Bliss-Moreau, & Russell, 2006). However, voice perception is still possible in non-verbal individuals such as infants. For example, Grossman, Oberecker, Koch and Friederiki (2010) used near-infrared spectroscopy (NIRS)[3] in order to observe brain development in infants and found adult-like increased oxygenated haemoglobin levels in right temporal regions when listening to words with emotional tone. Those findings suggest that perception of emotional voices has its origin in early infant brain development. Further, emotionality in voices does not always have to be communicated via prosody. As summarised next, non-verbal affect bursts such as laughter have recently begun to be included in emotion research.

1.3.3. Non-Verbal Affect Bursts

An alternative method to observing emotions in speech prosody is to investigate the recognition of non-verbal affect vocalisations which are ‘brief, non-word utterances’ (Simon-Thomas, Keltner, Sauter, Sinicropi-Yao, & Abramson, 2009, p. 838) and contain laughter, crying or an angry

[3] NIRS is a child-friendly functional neuroimaging method which observes changes in hemodynamic responses by investigating optical changes in structures due to different degrees of light absorption following the absence of oxygen in the blood (Zabel & Chute, 2002).

roar. *A*ffect bursts are the shortest possible form of vocal emotion expression which lends itself to the comparison of static face expressions. Vocalisations can be used for cross-cultural research without the limitation of language skills and are thought to represent more natural emotion expressions than prosody (Sauter et al., 2009).

Recognition scores are often higher for emotional affect bursts compared to prosody as expressions are of a more prototypical nature, hence, facilitating emotion recognition (Hawk, van Kleef, Fischer, & van der Schalk, 2009; Simon-Thomas et al., 2009). For example, overall recognition rate for emotion categories was 81.1% with disgust yielding accuracy rates of 93% (Schroeder, 2003) slightly lower recognition rates of 72% for sadness (Simon-Thomas et al, 2009). Whilst anger, disgust, sadness and surprise were especially well recognised, Simon-Thomas and colleagues (2009) concluded that overall, up to 14 out of 22 different emotion categories could be reliably identified from non-verbal affect clues alone. Accuracy levels were above chance with a recognition advantage for emotions expressed by female as compared to male speakers.

By including additional positive emotions expressed non-verbally such as relief or achievement, Sauter and Scott (2007) reported high recognition rates above chance which ranged from 52.7% for contentment to 90.4% for amusement. Even children as young as 5 years old could recognise non-verbal affect vocalisations with overall accuracy of 78.1% which increased to 83.9% in children up to the age of 10 (Sauter et al., 2013). Which features within the voice – other than linguistic content – determine emotion transmission?

Sauter et al. (2012) investigated the relationship of recognition abilities and acoustical features of their own developed dataset of affect bursts created by actors. They concluded that specific combinations of amplitude, pitch and variation of the frequency spectrum could lead to reliable recognition of emotion categories. For example, multiple linear regression analysis revealed that a combination of short stimulus duration, a lower mean pitch and less pitch variation predicted the emotion as belonging to the 'surprise' category.

The nature of non-linguistic affect bursts can be reflexive and – just like for face expressions – they are usually caused by biological processes (so-called 'push effect') or can sometimes be influenced by socio-cultural expectations which are learned and result in intentional exclamations of emotions (so-called 'pull effect'), often containing additional phonetic properties (Scherer, 1994). In terms of neural structures, emotion processing of non-verbal vocalisations is believed to occur in bilateral regions of anterior insula, frontal and temporal regions (Morris, Scott, & Dolan, 1999). Fecteau et al. (2007) additionally reported increased amygdala activation during the processing of positive and negative non-verbal affect bursts.

1.3.3.1. Montreal Affective Voices

When comparing recognition of emotion from affect bursts with the widely used Ekman face set (Ekman & Friesen, 1987), it is crucial to choose vocal stimuli that are of comparable standard in their recordings, speaker ethnicity and emotion categories included. Hence, the present thesis includes the Montreal Affective Voices (MAV) which have been produced by white Caucasian actors (Belin, Fillion-Bilodeau, & Gosselin, 2008). The MAV are solely based on non-verbal variations on the vowel /a/.Recognition rates for the MAV database are generally high with overall recognition accuracy being 68.2%. Sadness received highest rates of 86%, followed by disgust (81%) and anger (87%). Overall, female vocalisations were recognised with higher accuracy than emotions expressed by male actors. As illustrated in Table 1.2, the median pitch across all vocalisations and actors was highest for fear, followed by surprise, sadness, anger and happiness with lowest median pitch for disgusted stimuli. The mean length from stimuli across all 10 actors reached from 385ms for surprised stimuli to 2229s for sad vocalisations (Belin et al., 2008).

Table 1.2

Acoustical features of non-verbal affect vocalisations for all 9 emotions included (Belin et al., 2008).

Category	*f*0 (Hz)				Duration	Power (dB)	
	Minimum	Maximum	Median	*SD*	(msec)	Median	*SD*
Angry	150	413	317	80	924	78	14
Disgusted	108	295	200	58	977	75	12
Fearful	266	642	508	97	603	81	12
Happy	181	421	278	58	1,446	60	14
Neutral	149	184	168	4	1,024	81	6
Painful	134	435	351	87	839	77	13
Pleased	120	261	192	43	1,350	70	13
Sad	185	508	323	73	2,229	63	13
Surprised	228	453	373	69	385	76	13

Note—Values are averaged across the 10 actors.

Overall, previous literature suggests that acoustical features in affect vocalisations reveal sufficient information about the emotion category in question, making the inclusion of non-verbal affect bursts a useful addition to the investigation of emotion recognition across several modalities as well as across cultures and in children.

Neither the MAV nor the Ekman faces rely on linguistic or verbal information which makes them ideal stimuli for the present investigation of emotion recognition in children and across countries as implied by Koeda, Belin, Hama, Masuda, Matsuura and Okubo (2013).

1.4. Role of Modality and Emotion Mechanisms

For person identity recognition, Yovel and Belin (2013) stated recently that – despite very different sensory inputs - there might be unified coding principles across the visual and auditory modalities as based on previous neurological, cognitive and developmental data. Interestingly, during speaker recognition, voice specific areas such as the STS and face-specific areas such as the fusiform face area (FFA) show direct functional connectivity which facilitated identification of familiar persons (Von Kriegstein, Kleinschmidt, Sterzer, & Giraud, 2005). There is little knowledge, however, about whether emotions from faces and emotions from voices are also represented using similar mechanisms; although there is indeed some behavioural evidence for correlations between rating abilities of emotions in faces and voices (*e.g.* Borod et al., 2000). Both voices and faces have been shown to be processed in the STS (Haxby et al., 2002; von Kriegstein & Giraud, 2004), but is there evidence for emotion-specific brain regions that process information independent of the modality?

Although emotion processing may be somewhat subjective and context-dependent (*e.g.* Scherer, 1984), past research has indeed demonstrated biological evidence for recognising the basic emotions in the human brain (*e.g.* Adolphs, 2002; Wicker et al., 2003). There seems to be a general right hemispheric advantage for emotion processing (Blonder, Bowers, & Heilman, 1999; Kesler-West et al., 2001) and emotion recognition in faces as well as in voices is usually associated with specific emotion structures in the brain.

With fear being an emotion representing danger or threat, taking on a locationist[4] perspective of emotion processing (Lindquist, Wagner, Kober, Bliss-Moreau, & Feldman Barrett, 2012), the main subcortical structure commonly associated with fear processing is the amygdala. Lesion studies have consistently shown that although bilateral amygdala damage results in a more general reduced emotion recognition ability of mainly negative emotions (Adolphs, 1999; Blair, Morris, Frith, Perrett, & Dolan, 1999), the recognition of fearful faces was impaired the most (*e.g*. Adolphs, 2002) and can be extended to fearful emotion processing in prosody (Phillips et al., 1998) and even to fearful and angry non-verbal affect-bursts (Scott, Young, Calder, Hellawell, Aggleton, & Johnson, 1997). Fecteau, Belin, Joanette, and Armony (2007) also reported a link between the recognition of non-verbal affect and amygdala function – although this extended to positive and negative emotions equally.

Apart from fear, disgust is another basic emotion that can be mapped to a specific brain location. Calder, Keane, Manes, Antoun, & Young (2000) demonstrated that a patient with lesions to the insula and putamen in the left hemisphere was impaired in recognising disgust from several modalities, including face expressions and verbal as well as non-verbal voices. Further, the connection between somatosensory brain areas and the insula (Stephani, Fernandez-Baca Vaca,

---

[4] Locationist perspective: discrete emotion categories (such as fear) are associated with a discrete brain location (such as amygdala). This is the opposite to a psychological constructionist perspective (see footnote 6) (Lindquist et al., 2012).

Maciunas, Koubeissi, & Lüders, 2010) suggests a link between recognising disgust, feeling sick and activation of the insula.

Orbito-frontal parts of the brain have previously been associated with general emotion recognition abilities across voices and faces (Homak, Rolls, & Wade, 1996) and more specifically with the recognition of angry faces (Blair et al., 1999) and angry voices (Sander et al., 2005). Angry faces also activated the anterior cingulated cortex (Blair et al., 1999) and some studies also reported activation in the right fusiform gyrus in response to angry voices (Johnstone, van Reekum, Oakes, & Davidson, 2006). This indicates that the link between angry stimuli, modality and assigned brain areas is not straight forward. It is possible that frontal parts of the brain are mainly involved in general higher cognition such as the conscious evaluation of emotional stimuli classification rather than specific emotion processing (Nakamura et al., 1999). This may be especially true when stimuli such as affective prosody are ambiguous and need further evaluation (Leitman et al., 2010). The right somatosensory cortex has been associated with a variety of emotion processing within faces (Adolphs, Damasio, Tranel, Cooper, & Damasio, 2000; Pitcher, Garrido, Walsh, & Duchaine, 2008) as well as in voices (Banissy et al., 2010). Lesion studies suggested that impairment of right somatosensory cortices was associated with a deficit in forming neural representations of the observed emotion which in turn disabled simulation and recognition of the emotion expressed by others (Adolphs et al., 2000).

Whilst the current review suggests neural emotion structures that process emotions independently of modality, it also has to be noted that there is some evidence for emotion processing that was dissociated for faces and voices. For example, Bach, Hurlemann and Dolan (2013) reported that for two individuals with bilateral amygdala lesions, fear recognition from prosody was spared although commonly, the amygdala is associated with recognition of fear in faces (Adolphs et al., 1994). Hence, Bach and colleagues (2013) concluded that fear processing in the amygala is modality-dependent. However, it needs to be noted at that point that the locationist approach of emotion brain networks has also been challenged and instead, it has been proposed that more general cognitive networks drive emotion processes. This psychological constructionist approach[5] has been supported by findings from a meta-review by Lindquist et al. (2012); for example, from a density analysis, they reported that voxels within the amygdala did not respond exclusively to the emotion of fear.
Specific brain activations were not only associated with emotions per se but rather with general cognitive processes.

Overall, there is conflicting evidence from neuroimaging studies regarding the role of modality during emotion processing. Hence, the present thesis aims to investigate – mainly behaviourally – how emotions are being processed from two separate modalities in direct comparison by collecting data from adults and children from different North-Western European countries. Since

[5] The psychological constructionist approach assumes that emotions are based on general psychological constructs rather than emotion-specific networks. This is in contrast to locationist approaches (Lindquist et al., 2012).

the present thesis does not include functional neuroimaging methods, a more detailed debate about locationist or psychological constructionist approaches to emotion processing is beyond the focus of this work.

1.5. General Emotion Concepts

Across all modalities, emotions are a highly complex construct that attempts to define a very subjective experience of the human body. Why do we recognise emotions and how can we talk about it in an objective way? After introducing how and where emotions might be processed in the human brain, the following section will describe two common theories regarding the origin of emotion recognition with evidence for innate as well as socially learned emotion recognition processes before familiarising the reader with simulation and perspective taking of others' emotional states.

1.5.1. Innate Versus Learned Effects

Whilst humans have to encode emotions in faces or voices to make their current emotional states apparent to others, decoding the information then allows the observer to recognise emotions expressed by others. In terms of emotion expression, this mechanism is thought to be present from birth onwards in order to promote communication and survival (Darwin, 1872/1965). Hess and Thibault (2009) summarised Darwin's work '*The expression of the emotions in man and animals'* (1872/1965) by saying that emotion expressions evolved as expressive habits over time as a way to symbolize underlying feelings and to promote survival. For example, opening eyes during specific emotion expressions may enhance eye-sight. This habit would be passed on through generations, turning it into an evolved and universal mechanism. Indeed, more recently, Matsumoto and Willingham (2009) suggested that congenitally and non-congenitally blind athletes produced the same emotion expressions in faces as healthy individuals which supports the view of universal and inherited components of emotion expression.

Focusing on emotion recognition rather than expression, years of research have come to the conclusion that the ability to perceive and judge emotions may be included in our genetic make-up (Izard, 1994) but also develops and increases with age throughout childhood (Herba & Phillips, 2004; Gagnon, Gosselin, Hudon-ven der Buhs, Larocque, & Milliard, 2010).
Indeed, several empirical studies have found that infants can already discriminate between different types of emotions expressed in faces. For example, Bornstein and Arterberry (2003) demonstrated that five month-old infants showed preferential looking for new fearful compared to habituated static face expressions of happiness. Those findings suggested that from a very early age onwards, we can distinguish between different facial emotion categories which may be an innate feature aiding survival.

Although even 2-year old children have demonstrated a degree of understanding of the concept ‘emotion’ as evident from ‘pretend-play’, more complex emotion matching and categorizing abilities become evident from the age of 5 onwards (Widen & Russell, 2008).
This goes in hand with findings by Gagnon et al. (2010) who stated that fear and disgust showed higher recognition accuracy in 9 to 10-year old children as compared to 5 to 6-year old children, suggesting a gradual improvement of facial emotion recognition ability with age.
In addition, two earlier studies by McCluskey and Albas (1981) and McCluskey, Albas, Niemi, Cuevas and Ferrer (1975) have shown that just like for face processing, vocal emotion processing in children also improved with age right into adulthood.

Interestingly, on the opposite end of the age-dimension, emotion recognition abilities such as accuracy of labelling facial expressions seems to decrease in older subjects (Isaacowitz et al., 2007) and older adults often demonstrate a positivity-effect with enhanced processing of positive face expressions (Isaacowitz, Wadlinger, Goren, & Wilson, 2006). Whilst the exact interpretation of this finding is beyond the scope of the present work, results demonstrate that emotion recognition abilities change during the course of life and may to some degree depend upon social factors or past experiences and subsequent positive or negative reinforcements (Dollard & Miller, 1941). Especially theories of social learning propose a social account of emotion understanding. Bandura’s Theory of Social Learning (1971) suggested that by observing others, physical as well as emotional behaviour can be copied and learned. In Bandura’s famous Bobo doll experiment (1965), children who observed adults exhibiting aggressive behaviour towards a doll were significantly more likely to also behave aggressively towards the doll compared to children that saw adults acting in a neutral way towards the doll.

Further, the argument of cultural and social rules influencing emotion recognition also contributes to the discussion about universal or learnt emotion processes. Often times, emotions are regarded as universal and inherited mechanisms (Darwin, 1872/1965). For example, it has been proposed that the six basic emotions in faces, which are expressed in prototypically manner, can be expressed and recognised universally across cultures (*e.g*. Ekman & Friesen, 1971; Izard, 1980). The universal recognition of emotions has also been found for emotions displayed within the voice (*e.g.* Bryant & Barrett, 2008). However, more recently, it has also been accepted that social and cultural norms can influence emotion recognition to a degree, similarly to a dialect in an otherwise universal language.
This dialect-theory of emotion recognition (Elfenbein & Ambady, 2002) hints at a learnt contribution of emotion recognition. Overall, there seem to be some innate as well as learned emotion recognition mechanism as evident from developmental and cultural research. By including a developmental as well as cross-cultural perspective on emotion recognition, the present thesis aims to shed light on external factors that may potentially contribute to our emotion recognition abilities.

1.5.2. Mind-Reading and Simulation Approaches

In a philosophical account, our understanding of other's underlying emotions comes from so-called 'mind reading' processes (Goldman & Sripada, 2005). Emotions can be recognised by inferring mental states from others based on one's own knowledge. In terms of facial expressions, one would link the observed emotion to a repertoire of previously known emotion categories in order to classify the observed emotion. One example of mind reading is the Theory of Mind (ToM). According to Baron-Cohen, Leslie and Frith (1985), ToM is a "*mechanism which underlies a crucial aspect of social skills, namely being able to conceive of mental states: that is, knowing that other people know, want, feel, or believe things; in short, having what Premack and Woodruff (1978) termed a 'theory of mind' "* (p. 38). Hence, a degree of ToM and perspective taking is necessary in order to understand others' emotional states and several studies have demonstrated a relationship between impaired emotion recognition and Theory of Mind deficits in autistic children (*e.g.* Baron-Cohen, Leslie, & Frith, 1985; Heerey, Keltner, & Capps, 2003) or schizophrenic patients (*e.g.* Bruene, 2005). Commonly, ToM abilities are associated with activity in the medial prefrontal cortex (Gallagher, Happe, Brunswick, Fletcher, Frith, & Frith, 2000) although most ToM and emotion recognition studies are based on facial emotion recognition rather than on vocal emotion recognition.

On the other hand, simulation theories describe emotion recognition as a process of simulating and replicating others. According to Goldman and Sripada's review (2005), individuals observing others can reproduce the target emotion in corresponding brain areas. This activation can then be related back to how the target individual must feel at a certain moment of time. Consequently, a deficit in emotion production is often paired with a deficit in emotion recognition.
For example, dysfunction of the right somatosensory cortex[6] interrupts simulation of what an actor might feel when expressing emotions in faces (Adolphs et al, 2000; Pitcher et al., 2008) or voices (Banissy et al., 2010) which in turns reduces emotion recognition abilities across modalities. Further, the role of facial feedback also seems to be important during emotion recognition: dampening face expressions following Botulinum Toxin injections in turn reduced subjective experience of emotions as the muscle feedback from faces was limited (Neal & Chartrand, 2011).

Advances in recent neuropsychological research have allowed the investigation of neural mechanisms of simulation based approaches to emotion recognition such as the Mirror Neuron System (MNS). Mirror neurons are individual neurons within the brain, mainly within motor areas and are for example active during observation of other's performances of actions. This in turn gives the individual a deeper understanding of the actions and intentions observed (Fabbri-Destro & Rizzolatti, 2008; Rizzolatti, 2005). More specifically, parts of the premotor cortex and the inferior parietal lobe form a circuit of mirror neurons which fire when observing as well as hearing others

[6] Somatosensory representations help to understand how another person is feeling by internal simulation of facial expressions in brain regions of the right somatosensory cortex (Adolphs et al., 2000).

performing certain motor behaviours by using the hand, foot or mouth (Fabbri-Destro & Rizzolatti, 2008). However, when observing others performing a behaviour, actual execution of the same action is usually inhibited and the activation is only used to get a better insight into what the observed person must feel at that moment of time (Gallese & Goldman, 1998).

The MNS is important for social functioning such as imitation, learning and interpersonal relations and could help to explain why and how humans are able to detect emotions in others in accordance with simulation theories. In 2003, Wicker et al. conducted a study where participants either experienced a feeling of disgust or observed disgusted others' facial expressions. As a result, mirror neurons dealing with experiencing as well as observing the emotion disgust fired in areas of the insula and the anterior cingulated cortex (see also Jabbi, Swart, & Keysers, 2007). Interestingly, electrically stimulating the anterior insula resulted in feeling sensations in the throat and mouth which is associated with experiencing the feeling 'disgust' (Krolak-Salmon et al., 2003). Another study by Calder et al. (2000) demonstrated that brain damage involving the insula resulted in decreased feeling of disgust across several modalities.

Overall, the present thesis assumes that normally developed people are able to take perspectives of others and simulate emotions expressed in faces and voices to some degree. By investigating behavioural ToM abilities during childhood in the present Chapter 3, for the first time, the current thesis could provide a link of children's' mind reading abilities and emotion recognition not only from faces, but also from voices. However, the present thesis does not investigate the localisation of brain activity to support the involvement of specific brain areas or simulation networks during emotion recognition.

1.6. Research Rationale and Outline of Thesis

As demonstrated above, emotion recognition from faces and voices has received a large amount of attention – especially in the past two to three decades. However, the existence of several sets of databases which were especially developed for specific experiments (*e.g.* Morris et al., 1999; Schroeder, 2003; Simon-Thomas et al., 2009) makes comparisons of results between studies difficult. The emotion categories included in databases varies enormously from study to study and expressions can be either posed or spontaneous.

Further, depending on the nationality of presenters, rating scores may be biased by in-group advantages and emotions displayed by people from the own culture may be recognised with higher accuracy than members from another culture (Elfenbein & Ambady, 2002). Also, recording facilities for stimuli will vary depending on resources and laboratories. Consequently, it is important to choose sets of stimuli that have been widely validated and to implement them in a within-subject design to reduce external variance. Hence, the main aim of the present thesis is to investigate human's emotion recognition abilities within the face and voice as two separate, yet interdependent domains, but in a within-subject design. The researcher plans to investigate whether emotion recognition in one modality is superior over another modality – independent of the variable of task difficulty. Or whether there is one general, modality-independent emotion recognition mechanism which may possibly be grounded in shared neural networks.

More specifically, Chapter 2 of the present thesis behaviourally tests emotion recognition in adults and aims to find out just how similar the recognition of the six basic emotions is when presented either in faces or in voices. Investigating both modalities in parallel addresses the question of whether we recognise emotions based on modality-independent constructs such as general appraisal mechanisms (Scherer, 1984) or on distinct perceptual features such as a smile in happiness (Ekman & Friesen, 1975/2003).

Chapter 3 of the present thesis aims to find out how the emotion recognition abilities across two separate modalities develop across middle years of childhood. For this, children at primary school age completed the same emotion task in order to demonstrate whether both modalities develop in parallel or whether recognition from one modality may be acquired earlier in life. It will also investigate whether emotion recognition from two modalities depend on specific emotions and whether young children find it easier to recognise specific emotions in one modality compared to the other modality, which may change as a function of age. Chapter 2 includes a cross-sectional as well as a smaller longitudinal study by re-testing a subset of children after 1.5 years. Additionally, it examines the effect of including child – rather than adult - face stimuli to be rated by children by validating a new child face database (Dalrymple, Gomex, & Duchaine, 2013) for use in children.

Can the results from Chapter 2 and 3 including British adults and children be generalised to another North-Western European country such as Germany.
Chapter 4 includes the cross-cultural comparison of an adult sample in order to investigate cultural differences in emotion recognition in voices and faces in two similar European countries. Further, it includes the cross-cultural comparison of a child samples in order to examine early socialisation effects in emotion recognition.

Lastly, in order to add more depth to the collection of behavioural experiments in the present Chapter 2-4, Chapter 5 provides data from a neurophysiological ERP experiment – which are of preliminary nature at this stage.
The aim here is to investigate whether different modalities such as faces and voices have comparable temporal stages of emotion processing such as a broad but early discrimination between emotional and neutral stimuli and a later but narrow discrimination between individual emotion categories.

Overall, the present thesis aims to examine the role of cognition, child development, culture and display rules as well as temporal brain dynamics in order to answer the question whether emotion recognition is comparable across two separate modalities within one and the same person.

2

Emotion Recognition: Just how similar are voices and faces?

Non-verbal perception of emotions has high evolutionary implications for survival as well as for communication (Darwin, 1965/1872; Hampson, van Anders, & Mullin, 2006). We need to rapidly classify emotions in order to recognise threat, assess social situations, and behave accordingly by, for example, protecting our offspring from enemies. Previous studies have shown that – although emotions are off highly complex nature, - we can perceive and reliably classify the six basic emotions from different modalities, such as faces (*e.g.* Ekman & Friesen, 1978; Herba & Phillips, 2004) and voices (Belin, 2011; Sauter, et al., 2010). However, to the best of the researcher's knowledge, no previous study has specifically investigated via within-subject designs how similar the emotion coding mechanisms for two independent modalities are in direct comparison - unless contrasted to multimodal emotion integration (see Lambrecht, Kreifelts, & Wildgruber (2014) for a summary). In the present Chapter 2, the aim was to examine whether emotions are recognised in comparable manner from faces and voices in British adult participants.

2.1.1. Shared Emotion Coding Mechanisms Across Modalities?

It is possible that perception of emotions presented in different modalities may share the same neural (Calder & Young, 2005) or behavioural mechanisms (Borod et al. 2000). For example, it has been proposed via fMRI multivoxel pattern analysis that activity for five basic emotions such as anger, fear, sadness, happiness and disgust is correlated highly across the face, voice and body in the medial prefrontal cortex and the left superior temporal sulcus (Peelen, Atkinson, & Vuilleumier, 2010). As outlined in the introductory Chapter 1, subcortical emotion structures like the amygdala are commonly associated with facial (*e.g.* Adolphs, 2002) as well as for auditory (Fecteau, Belin, Joanette, & Armony; 2007; Phillips et al., 1998) and even during emotional music recognition (Aube, Angulo-Perkins, Peretz, Concha, & Armony, 2014). Further, ERP studies have suggested that temporal discrimination of emotional versus non-emotional voices occurs as early as 150ms after stimulus onset (Iredale, Rushby, McDonald, Dimoska-Di Marco, & Swift, 2013; Sauter & Eimer, 2009) which is comparable to ERP data from emotion processing in visual domain (Batty & Taylor, 2003; Eimer & Holmes, 2007).

From brain lesion studies it becomes evident that damage to the temporal lobe (Bonora et al., 2011; Dellacherie, Hasboun, Baulac, Belin, & Samson, 2011) or amygdala (Scott, Young, Calder, Hellawell, Aggleton, & Johnson, 1997) is typically associated with impaired facial emotion recognition as well as reduced vocal emotion recognition. In addition, transcranial magnetic stimulation (TMS) of right somatosensory cortex temporarily interrupts representations and correct classifications of facial emotions (Pitcher et al., 2008) as well as non-verbal affect vocalisations (Banissy, Sauter, Ward, Warren, Walsh, & Scott, 2010).

Support for modality-independent emotion processing mechanisms also comes from behavioural data: Studies of individual differences have found significant (though not high) correlations between adults' visual and vocal emotion recognition abilities (Borod et al., 2000; Palermo et al., 2013). Further, emotion recognition across modalities often seemed to be impaired conjunctionally in Autism (Philip et al., 2010), Schizophrenia (Simpson, Pinkham, Kelsven, & Sasson, 2013) and in recently detoxified alcoholics (Kornreich et al., 2013), affecting a core emotion network rather than single modalities. Similarly, boys aged 8 to 16 with callous-unemotional traits[7], have been found to be impaired in recognising fear in others across a range of modalities such as faces and body posture (Munoz, 2009). It is possible then that cognitive appraisal of events such as novely and degree of pleasantness is a substantial component of emotion processing (Ellsworth & Scherer, 2003; Scherer, 1984) - independent of perceptual differences due to mode of presentation.

Overall, it seems plausible that our brain exhibits an emotion processing mechanism involving a general emotion circuit, combining input from several modalities in order to aid evolutionary mechanisms (Calder & Young, 2005). However, referring to the studies mentioned above, heterogeneity of methods and lack of within-subject designs across modalities makes generalizability of outcomes difficult.

2.1.2. Separate Mechanisms

In contrast to the view that the same brain regions support emotion processing across modalities, emotion recognition in the visual and auditory domain have also been independently affected. For example, patients with right hemisphere damage have shown isolated deficits in recognising fear in the face but not in vocal stimuli (Adolphs, Tranel, & Damasio, 2001). In addition, prosody discrimination performance was found to be independent of impaired visual emotion after anterior temporal lobectomy (Milesi et al., 2014) or amygdala damage (Anderson & Phelps, 1998; Bach et al., 2013).

[7] Callous-Unemotional Trait: Psychopathy that describes antisocial patterns in behaviour and emotional impairments due to lack of empathy (Munoz, 2009).

For clinical and pre-clinical patients with the neurogenerative disorder of Huntington's disease, which typically affects the insula and basal ganglia, recognition of disgust seemed to be impaired for face expressions (Sprengelmeyer, Rausch, Eysel, & Przuntek, 1998) but intact in the vocal domain (Sprengelmeyer, Schroeder, Young, & Epplen, 2006), suggesting separate processors for emotion recognition based on modality.

However, on the other hand, a meta-analysis suggested that neurological patients with another form of basal ganglia pathology –Parkinson Disease – have difficulties recognising emotions mainly from voices, but also from faces (Gray & Tickle-Degnen, 2010). This would suggest a common underlying neural network that can affect emotion recognition – independently of modality. Hence, the degree to which underlying mechanisms are shared or dissociable may depend on the specific neural structures involved as well as the specific emotion affected. However, it is important to note that, especially for Parkinson Disease, deficits in emotion recognition across several modalities may be secondary to other executive function deficits such as impaired working memory (Gray & Tickle-Degnen, 2010).

Those findings are not surprising, because at the perceptual level, signals from the voice and from the face are indeed quite different. Within the auditory domain, each of the basic emotions is believed to have a corresponding acoustic profile, containing an emotion- specific combination of pitch, amplitude and spectrum which is produced by movements of the larynx (Sauter et al., 2010). Summarising past neuroimaging studies, Schirmer and Kotz (2006) proposed a hierarchical multi-step working model and suggested that initial acoustic analysis of vocal stimuli is processed within the first 100ms in the auditory cortex. Consequently, information is sent to the right superior temporal sulcus for emotional evaluation and transmitted to frontal parts of the brain as well as the right inferior frontal gyrus for further cognitive appraisal of emotional significance at around 400ms after stimulus onset.

On the other hand, in the visual domain, emotion recognition is based on configural changes within the face initiated by movements of muscles (Ekman & Friesen, 1978). Different combinations of contracted face muscles over time such as the early wrinkled nose followed by raised upper lip during production of dynamic disgusted face expressions allows for an emotion-specific profile. This emotion profile can be reliably discriminated whilst access to incomplete emotion profiles can cause confusions between similar emotion expressions that share common features (Jack, Garrod, & Schyns, 2014; Schyns, Petro, & Smith, 2009). Hence, emotions like happiness produce large scale features in the face such as an open mouth showing teeth which aid rapid recognition. Encoding probably starts with processing of the eyes before 'zooming out' to a whole face analysis before zooming back in to encode more specific perceptual features (Schyns et al., 2009).

Given the differences in basic perception between faces and voices, it is possible that emotion recognition mechanisms depend on their presentation domain and are based on bottom-up perceptual features rather than modality independent top-down cognitive evaluation. This is, however, in direct contrast with studies which suggest that the same brain regions may be responsible for emotion processing in faces and voices alike. In the present study, this current issue is investigated by directly comparing the behavioural emotion recognition profiles across modalities in a within-subject design. In order to compare two perceptually different modalities, data on valence superiority, gender effects and confusion rates within each modality were collected.

2.1.3. Valence Superiority

Assuming similar neural and cognitive mechanisms across different modalities, rating preferences of emotions may be comparable across modalities. For example, if a negative emotion such as anger is easily recognised in faces, one expects to see the same pattern in voices. Indeed, similarity between modalities at the level of individual emotions has previously been demonstrated. For example, multiple ERP studies have reported fast and subliminal detection of fearful auditory (Sauter & Eimer, 2009) and visual stimuli (*e.g.* Kiss & Eimer, 2008) at around 150 to 200ms post stimulus onset. It is possible then that general evolution-motivated mechanisms guide attention towards specific emotions – independent of source. Such mechanisms could be based on i) biological preparedness and threat avoidance (Seligman, 1971), or ii) social bonding (Baumeister & Leary, 1995) and altruism theories (Hauser, Preston, & Stansfield, 2014), which support the importance of attending to people that radiate positive rather than negative attitudes.

Supporting threat-avoidance theories, often, previous studies have reported that threatening, angry faces popped out during visual search tasks of emotional faces (Hansen & Hanson, 1988). This may be guided by experience and has been used to explain police officers' enhanced abilities in detecting angry faces within a riot crowd (Damjanovic, Pinkham, Clarke, & Phillips, 2014). However, it does not take professionals to detect threatening faces within a crowd. The working memory of the average British undergraduate student also seems to be enhanced for detecting angry as compared to happy faces (Thomas, Jackson, & Raymond, 2014). Most recently, studies have begun to investigate this 'anger-superiority effect' in more detail, though to the best of the researchers' knowledge, exclusively within the visual domain. Electrophysiological studies for example have described N2pc[8] amplitudes which occurred earlier and with larger amplitude for detecting angry versus happy faces (Weymar; Loew, Oehman, & Hamm, 2011) indicating increased attention capturing for evolutionary threatening faces.

However, studies demonstrating anger superiority – possibly motivated by threat avoidance mechanisms - cannot account for studies reporting the opposite with early and unconscious brain activation during ratings of faces expressing positive emotions. This has often been reflected in

[8] N2pc is a ERP component and though to reflect spatial selective attention during visual search of angry faces and correlated with differences in behavioural reaction times of about 50ms (Weymar et al., 2010).

increased N170 components (Batty & Taylor, 2003) or increased accuracy and fast reaction times below one second for happy faces in behavioural experiments (Smith, 2012). Questioning the lack of attention directed at happy stimuli, which may promote attachment and caretaking of offspring (Babchuck, Hames, & Hampson, 1985), Becker, Anderson, Mortensen, Neufeld and Neel (2011) conducted a series of experiments, demonstrating consistent evidence that happy faces were detected efficiently during visual search tasks.

This 'happiness superiority effect', – which has also been found by Juth, Lundquist, Karlsson and Oehman (2005) - reflected behaviourally in enhanced accuracy as well as speed. Similarly, Leppänen, Tenhunen and Hietanen (2003) suggested faster response selection times for happy faces as compared to angry and disgusted faces and refuted the automatic anger detection within a crowd. This has refuelled the discussion of threat avoidance versus attachment mechanisms which can be reflected in the duality of valence (positive or negative) of emotions. Problematically, data from auditory sources is almost non-existing which makes identification of supramodal emotion perception mechanisms difficult. Hence, investigating valence superiority effects across modalities could firstly provide data on anger or happiness superiority effects within the vocal domain and secondly demonstrate the degree of emotion recognition similarities across two separate modalities.

2.1.4. Gender Effects

It has been proposed that facial emotion classification in women may –to some degree - be happening automatically. For example, subliminally presented emotions elicited correct accuracy ratings above chance in women but not in men (Hall & Matsumoto, 2004). Emotion recognition may be influenced by gender with females appearing to use different visual features in the face such as the mouth area than men do (Blais, Fiset, & Gosselin, 2013). Further, Schulte-Ruether, Markowitsch, Fink and Piefke (2007) linked higher levels of empathy in women with greater activation of the mirror neuron system during an fMRI study involving facial emotion judgements. Generally, women seemed to be better at recognising emotions from faces, voices or audio-visual presentations (Collignon, Girard, Gosselin, Saint-Amour, Lepore, & Lassonde, 2010). However, gender differences cannot always be reliably established: for example, women were found to show enhanced emotion recognition in the auditory but not in the visual domain (Lambrecht, Kreifelts, &Wildgruber, 2014). Further, Hawk, van Kleef, Fischer and van der Schalk (2009) did not find any gender differences at all in accuracy rates for rating emotional non-verbal vocalizations.

Summarising past research, the relationship between gender differences and emotion recognition in faces and voices remains somewhat unclear. Summarising past research, the extent to which there are gender differences for emotion recognition in faces and voices remains somewhat unclear. In a meta-analysis, Thompson and Voyer (2014) recently suggested that there may be a female advantage in recognising non-verbally displayed emotions, but the size of the effect depended on interactions between the specific emotions and modality presented. For example, the effect sizes for gender differences in anger were significantly larger than for any other basic emotion. Similarly, the gender difference was largest in combined audio-visual conditions rather than in visual or auditory conditions alone.

One reason to expect gender differences in emotion recognition can be related back to the ‘primary care-taker hypothesis’ (Babchuck et al., 1985), which describes that females ‘will display evolved adaptations that enhance the probability of survival of their offspring’ (Hampson et al., 2006, p. 402). This could be reflected in female’s enhanced emotion recognition to ensure the offspring’s wellbeing. Significant gender effects could demonstrate adaptive child-rearing mechanisms in female participants by correctly classifying other’s emotions displayed in faces and voices. Again, assuming neuronal and cognitive similarities in emotion processing across modalities which are possibly linked to evolutionary concepts of survival and attachment promotion, the present chapter aims to investigate whether gender differences in emotion recognition occur not only in one but within both domains.

2.1.5 Confusion Rates

Apart from valence of stimuli and gender of judges, the degree of confusion amongst the six basic emotions lends itself as third comparisons across different modalities. It is understood that specific emotion pairs such as fear and surprise within facial expressions often are confused due to many shared perceptual features such as raised upper eye lids and raised eye brows (Ekman & Friesen, 1978). The lesser amount of action units that fearful and surprised faces share, the higher is the recognition accuracy (Gosselin & Simard, 1999). Moreover, angry and disgusted faces also share many perceptual features during early stages of emotion transmission (Jack et al., 2014).

However, perceptual features cannot explain possible confusions amongst emotions in the auditory domain. Interestingly, Belin, Fillion-Bilodeau and Gosselin (2008) also found common confusions between fearful and surprised affective vocalisations. They suggested that very short length and highest average pitch ($f0$) in both types of emotional stimuli could lead to confusions. In addition, Sauter et al. (2010) found that for emotional vocalisations, anger and disgust as well as disgust and anger often got confused. All three emotions share “low spectral variations and higher spectral centre of gravity” (p. 2262).

2.1.6. The Present Study

Up to this day, no study has specifically investigated how similar emotion recognition profiles are across two separate modalities. To compare across the visual and auditory modalities, the similarity of emotion rating patterns across voices and faces was investigated via valence superiority and gender effects, as well as the confusability rates across different emotions. To the researcher’s best knowledge, the present study is the first to investigate correlations amongst confusion rates for faces as well as for voices in a within-subject design in order to investigate similarities in processing emotions across modalities. High correlations of rating profiles across modalities would indicate processing similarities that are independent of perceptual features of each individual modality. Assuming underlying similarities in cognitive and neural mechanisms as well as emotion recognition which promotes survival or attachment, signals should not only be communicated by face, but also by vocal, non-verbal features within the voice. Finding comparable behavioural profiles across the visual and auditory modalities could suggest shared emotion coding systems.

The present study predicts the following:

Hypothesis 1: The present study predicted to see either anger or happiness valence superiority - replicated across both modalities.

Hypothesis 2: Patterns of gender effects of emotion processing will be comparable across both modalities. It is predicted that female participants will rate emotional stimuli more efficiently than male participants due to parent-offspring attachment formation and offspring survival.

Hypothesis 3: Rating profiles across modalities – as investigated by confusion matrixes – are correlated and suggest similarities in emotion rating patterns.

2.2. Methods

2.2.1. Participants

Participants were 54 British adults recruited through the participant pool at Brunel University and social networks. Data was excluded for nine participants – all undergraduate university students - whose non-verbal IQ scores were lower than two SDs below the overall participants' average IQ scores of 48.78, which seemed to be due to lack of motivation rather than genuine IQ deficits. The final sample consisted of 45 British adults (15 male, 30 female), aged between 18 and 61, (*M* = 30.47; *SD* = 16.81), with average non-verbal IQ scores (*M* = 48.78; *SD* = 4.41) which falls into GRADE III '*intellectually average'* according to British norms from the 1992 standardisation (Raven, Raven, & Court, 2000, page 81). Participants presented different educational backgrounds: GSCE level and below (*N* = 9), A levels (*N* = 6), Undergraduate (*N* = 24) and Postgraduate level (*N* = 6). All participants reported normal or corrected-to-normal vision and hearing.

The study was approved by the Ethics Committee at the Department of Psychology (01/03/2012) Brunel University and all participants gave informed consent to participate (see Appendix).

2.2.2. Research Material

Visual & Auditory Stimuli

Face expressions for happiness, sadness, anger, fear, surprise and disgust were displayed by two male (identity JJ and EM) and two female (identity C and SW) actors from the Ekman Pictures of Facial Affect series (Ekman, 1975). The Ekman faces were created according to the Facial Action Coding System (FACS) as developed by Ekman and Friesen (1976) where specific facial muscles and their position within the human body had been identified which were used for expressing certain voluntary facial emotions. Each of the 24 pictures was in grey-scale and was randomly presented in the middle of a computer screen. Non-verbal voice samples for the six emotions were selected from the Montreal Affective Voices (Belin et al., 2008) and displayed by two male (identity 42 and 55) and two female (identity 45 and 53) French-Canadian actors. Vocalisations included laughter or moans based on the vowel /ɑ/ expressing the 6 basic of happiness, sadness, fear, disgust, anger and surprise in terms of non-verbal affect bursts such as laughter or moans.

Each stimulus was rated on each of the six basic emotions (2 x 2 x 6 x 6 auditory stimuli and 2 x 2 x 6 x 6 visual stimuli) on a 7-point Likert-scale scale ranging from '*not happy/sad/… at all = 1'*, to '*extremely happy/sad/… = 7'*. No definition of emotions was provided in order to avoid biasing the responses into any direction.

Computer Program

The task has been programmed on the PsychoPy software in Python (Peirce, 2007), measuring reaction times (RTs) as well as intensity perception. Stimuli were presented randomly within alternating face and voice condition blocks (on Acer ASPIRE 5735). Voices were presented via closed-back-on-ear headphones (Sennheiser HD 202) for the duration of the stimulus, ranging from 378ms (*M* surprise) to 2039ms (*M* disgust) whilst faces were presented until button click.

Intelligence Assessment & Demographics

The paper based Raven's Standard Progressive Matrices (SPM) test for adults was included in order to assess 'fluid IQ' and problem solving abilities, independent of language skills (Raven et al., 1998, *p*. CPM29). The adult version consists of 60 items and is presented in black and white. Studies across different countries have reported high test-retest reliability around .90 for short term reliability and around.60 for long term reliability of up to four years (Raven et al., 2000). In addition, this current study collected demographic details such as age, gender, ethnicity, first language, country of birth, culture they belong to and educational status of participants (see above).

2.2.3. Procedure

Before starting the study, all participants signed an informed consent sheet and provided some basic demographic data (Appendix A). Participants received a short introduction and could then work through the emotion judgement and IQ task with a maximum time allocation of 25 minutes for each task. Finally, all participants received debriefing and psychology undergraduate students received credits as part of their course requirement.

2.2.4. Data Analysis

For each participant and each stimulus, choice reaction times (RTs) as well as intensity ratings were recorded. RT outliers exceeding 50s were removed and replaced with the mean (*N* trial = 1). Higher hit intensity ratings as well as lower mix-up intensity ratings reflect correct inclusion/exclusion of the target emotion.

For the first analysis (Analysis of Target Emotions), only stimulus scores were included which matched the corresponding emotion question (target emotions) and two separate mixed measures ANCOVAs were conducted, one for RTs and one intensity ratings as the dependent variables. Between-subject variables were gender of participants (male or female).The within-subject variables were modality (face or voice), emotion category (happy, sad, angry, fear, surprise and disgust) and gender of the actor (male or female). Age was included as a covariate in order to control for a significant age difference between female (M = 26.43, SE = 15.62) and male participants (M = 38.53, SD = 16.67), as demonstrated by an independent t-test, $t(43) = 2.4$, $p < .05$.

The age difference between male and female participants was the result of recruiting young Psychology undergraduates as participants which commonly consist of more female than male students.

For the second analysis (Analysis of Confusion), the responses using all emotion labels for each type of stimulus were used, separated by modality in order to create full rating profiles for each emotion. Due to the design of the study, it was possible to collect the mean-intensity ratings for each type of emotional label given to each of the six emotions. Correlation matrixes were computed for mean intensity ratings of all possible combinations of emotion pairs - resulting in 36 conditions. Representational Similarity Analysis (Kriegeskorte et al., 2008) was employed in order to compare how similar the representation and therefore how high the correlation of emotion responses are within as well as across modalities.

2.3. Results

2.3.1. Analysis 1: Target Emotions

Means (*M*) and Standard Error (*SE*) for modality and emotions for adults for Intensity and Choice Reaction Times are tabulated in Table 2.1.

Table 2.1

Means and Standard Error for modality and emotions for adults for intensity and Choice reaction times

	Emotion	**Mean Intensity *(SE)***			**Mean Choice RTs (*SE*)**		
		Female Adults	**Male Adults**	**All Adults**	**Female Adults**	**Male Adults**	**All Adults**
		N = 30	*N* = 15	*N* = 45	*N* = 30	*N* = 15	*N* = 45
		M *= 26.43 (16.62)	*M* = 38.53 (16.67)	*M* = 30.47 (16.81)	*M* = 26.43 (16.62)	*M* = 38.53 (16.67)	*M* = 30.47 (16.81)
Face	*Happy*	6.54 (.10)	6.62 (.14)	6.58 (.08)	3.31 (.13)	3.63 (.19)	3.47 (.11)
	Sad	5.35 (.19)	5.54 (.28)	5.34 (.16)	4.28 (.22)	4.29 (.32)	4.28 (.19)
	Angry	5.41 (.17)	4.79 (.25)	5.10 (.15)	3.76 (.21)	4.34 (.30)	4.05 (.18)
	Fear	5.87 (.18)	5.44 (.26)	5.66 (.16)	4.04 (.25)	4.47 (.36)	4.25 (.21)
	Surprise	6.40 (.12)	6.31 (.17)	6.35 (.10)	3.33 (.24)	4.63 (.34)	3.98 (.20)
	Disgust	5.82 (.19)	5.13 (.28)	5.47 (.16)	3.57 (.28)	5.04 (.41)	4.31 (.24)
	All	5.90 (.10)	5.60 (.150)	5.75 (.09)	3.82 (.09)	4.40 (.23)	4.06 (.13)
Voice	*Happy*	6.12 (.16)	6.16 (.23)	6.14 (.13)	3.66 (.15)	4.39 (.22)	4.03 (.13)
	Sad	6.23 (.15)	5.85 (.22)	6.04 (.13)	4.16 (.16)	4.53 (.23)	4.35 (.14)
	Angry	5.07 (.20)	4.52 (.29)	4.81 (.17)	4.02 (.18)	4.38 (.26)	4.20 (.15)

		Mean Intensity *(SE)*			Mean Choice RTs		
		Female Adults	**Male Adults**	**All Adults**	**Female Adults**	**Male Adults**	**All Adults**
	Fear	5.22 (.21)	5.10 (.30)	5.16 (.18)	3.80 (.17)	4.24 (.25)	4.02 (.15)
	Surprise	5.27 (.20)	5.14 (.29)	5.21 (.17)	3.45 (.11)	3.49 (.16)	3.47 (.10)
	Disgust	5.55 (.17)	5.58 (.25)	5.56 (.15)	3.82 (.13)	4.33 (.19)	4.08 (.11)
	All	5.58 (.11)	5.40 (.16)	5.49 (.10)	3.82 (.09)	4.23 (.13)	4.02 (.08)
Overall	*Happy*	6.33 (.10)	6.39 (.14)	6.36 (.08)	3.48 (.11)	4.01 (.16)	3.75 (.10)
	Sad	5.79 (.13)	5.59 (.19)	5.69 (.11)	4.22 (.14)	4.41 (.21)	4.32 (.12)
	Angry	5.24 (.15)	4.67 (.22)	4.95 (.13)	3.89 (.16)	4.36 (.23)	4.12 (.14)
	Fear	5.55 (.17)	5.27 (.25)	5.41 (.15)	3.92 (.16)	4.36 (.23)	4.14 (.14)
	Surprise	5.84 (.12)	5.72 (.18)	5.78 (.11)	3.39 (.13)	4.06 (.19)	3.72 (.11)
	Disgust	5.68 (.15)	5.35 (.21)	5.52 (.12)	3.70 (.15)	4.68 (.22)	4.19 (.13)
All Female stimuli		5.85 (.09)	5.71 (.13)	5.78 (.08)	3.74 (.11)	4.29 (.16)	4.02 (.10)
All Male Stimuli		5.62 (.11)	5.29 (.15)	5.46 (.09)	3.79 (.12)	4.34 (.17)	4.06 (.10)
Grand Mean**		5.74 (.09)	5.5 (.13)	5.62 (.08)	3.77 (.10)	4.31 (.15)	4.04 (.09)

Note. * = Mean age ** = Average scores across all participants and conditions.

a) Intensity Ratings

Results from intensity ratings suggested that whilst intensity ratings across modalities did not differ, the category of emotion, gender of the stimulus and age of participants influenced intensity perception and intensity ratings also depended on the interaction of emotion and modality as well as on emotion and gender of the stimulus. Huynh-Feldt corrections were used for factors that violated the assumption of sphericity.

A repeated measure ANCOVA for intensity ratings with gender of participants (2) as between-subject factor and modality (2), emotion (6) and gender stimulus (2) as within-subject factors whilst controlling for age of participants as covariate revealed that there was a significant main effect of emotion, $F(4.77,\ 200.25) = 10.15$, $p < .001$, $\eta^2 = .2$ and gender of the stimulus, $F(1, 42) = 12.85$, $p = .001$, $\eta^2 = .23$. Female stimuli receiving higher intensity ratings than male stimuli. In addition, there was a significant main effect for the covariate age of participants, $F(1, 42) = 4.23$, $p = .05$, $\eta^2 = .09$ and as age increased, intensity ratings decreased ($r = -.38$, $p = .01$).

To follow up on the main effect of emotions, Sidak post-hoc for multiple comparisons revealed that happy stimuli received highest ratings which were significantly higher than for all other emotions, followed by surprise, sadness, disgust which did not differ significantly from each other and then followed by fear which was significantly lower than all emotions but sadness and disgust. Anger received lowest ratings which were significantly lower than for all other stimuli displayed, all at $p < .001$. The main effect of modality as well as of gender of participants was not significant ($p > .05$).

Further, the interaction effect between (i) emotion and modality, $F(4.56, 191.65) = 4.7$, $p = .001$, $\eta^2 = .1$, as well as between (ii) emotion and gender of the stimulus, $F(5, 210) = 3.54$, $p = .004$, $\eta^2 = .08$, was significant (for *M* and *SE* see Table 2.1).

i: Happy stimuli received highest ratings in the face condition and happiness and sadness equally received highest ratings in the voice condition. For the face condition, this was followed by ratings for surprise which were different to all other emotions, followed by fear and disgust which did not differ significantly from each other. For the voice condition, happiness and sadness were followed by disgust which was different from all other emotions. Anger received lowest ratings in the face and voice conditions, which were significantly lower than all other emotions but sadness (in the face condition) or fear (in the voice condition), all at $p < .05$. Additionally, although no overall significant main effect for modality, for emotions of happiness, fear and surprise, faces were being rated with higher intensity than voices whilst for sadness, the opposite was true and voices were being rated with higher intensity than faces. For disgust and anger, the modalities did not differ in their intensity ratings.

ii: The interaction between emotions and gender stimulus revealed that the main effect for gender of the stimulus with higher intensity ratings for female than for male stimuli was true for emotions displaying anger, fear and disgust, $p < .05$. Further, for male stimuli, happiness was rated higher than all other emotions with anger and fear receiving lowest ratings. A similar pattern was found for female stimuli with happiness receiving ratings which were higher than for all other emotions and with anger receiving lowest intensity ratings (p < .05). This is illustrated in Figure 2.1.

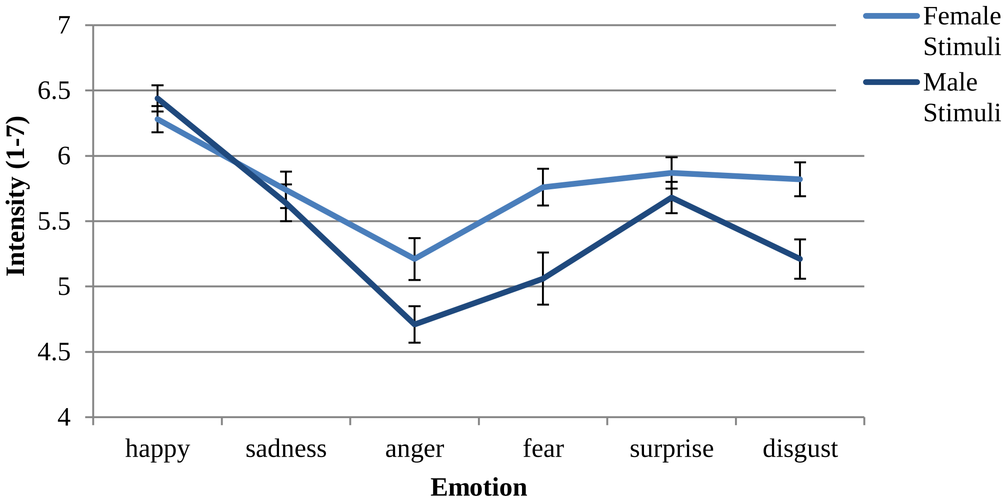


Figure 2.1. Means and Standard Error for intensity ratings for emotion and gender stimulus, p < .05. N = 45.

b) Choice Reaction Times (RTs)

Results from RTs suggest that speed for making decisions depended on the emotion category, on age as well as gender of participants and also on the interaction of modality, emotions, gender of participants and gender of the stimuli. Huynh-Feldt corrections were used for factors that violated the assumption of sphericity. A repeated measure ANCOVA for RTs with gender of participants (2) as between-subject factor and modality (2), emotion (6) and gender stimulus (2) as within-subject factors whilst controlling for age of participants as covariate revealed that there was a significant main effect of emotion, $F(4.92, 206.74) = 2.85$, $p = .02$, $\eta^2 = .06$), age of participants, $F(1, 42) = 4.91$, $p = .03$, $\eta^2 = .11$ and of gender of participants, $F(1, 42) = 8.64$, $p = .005$, $\eta^2 = .17$. To follow up on the main effect of emotions, Sidak post-hoc for multiple comparisons revealed that surprised and happy stimuli had fastest RTs which did not differ significantly from each other, followed by anger, fear, disgust and sadness which all did not differ from each other significantly ($p > .05$). Further, age correlated positively with speed and as age increased, speed of decision making also increased, $r = .43$, $p = .003$. Lastly, female participants made significantly faster decisions ($M = 3.77$, $SE = .1$) compared to male participants ($M = 4.31$, $SE = .15$). However, the main effects for modality and gender of stimuli were not significant ($p > .05$).

Further, analysis revealed that the three-way interaction between modality, emotions and gender of participants was significant, $F(5, 210) = 2.95$, $p = .014$, $\eta^2 = .07$. Whilst the present study did not find overall significant differences in choice RTs for modality, faces were rated faster than voices for happy stimuli – for male as well as for female participants whilst for surprise, there was an opposite modality effect. Only ratings made by male participants showed faster ratings for surprised voices than for surprised faces. This modality effect for surprised stimuli was not visible in female participants. Furthermore, the significant main effect for gender of participant with females rating stimuli faster than male participants seemed to be true for surprise and disgust in the face condition and for happiness and disgust in the voice condition, $p < .05$. This is illustrated in Figure 2.2.

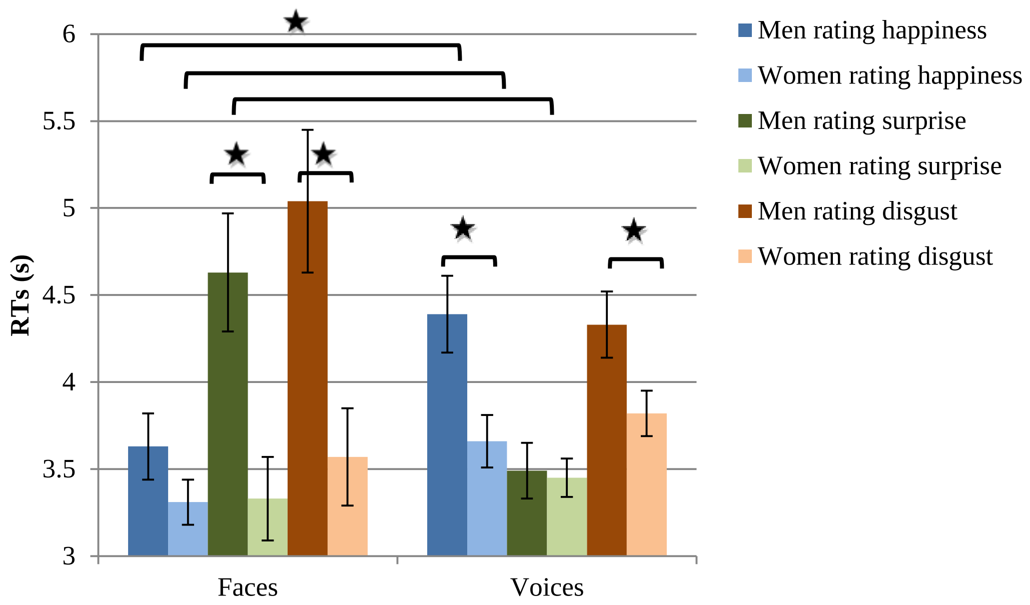


Figure 2.2. Means and Standard Error for RTs. Significant three-way interaction effect between modality, emotions and gender of participants, $*p < .05$. $N = 45$

Lastly, the three-way interaction between modality, emotion category and gender of participants also interacted with a fourth variable gender of stimuli, $F(5, 210) = 2.21$, $p = .05$, $\eta^2 = .05$. Men showed a face superiority effect for happy and surprised faces over voices only if they were displayed by other males. Further, for faces, the faster RTs for women as compared to men for surprised faces only seemed to be true when displayed by male actors whilst for disgusted faces, the RT effect was true for stimuli displayed by actors of both genders. Additionally, faster RTs for women than for men was also true for angry faces when displayed by female actors. Men also rated disgusted faces faster when they were displayed by women as compared to men.

For voices, faster choice RTs for women over men for happy voices was only true when stimuli were displayed by male actors. Men rated angry voices faster when they were displayed by other men as compared to women, all at $p < .05$. This is illustrated in Figure 2.3.

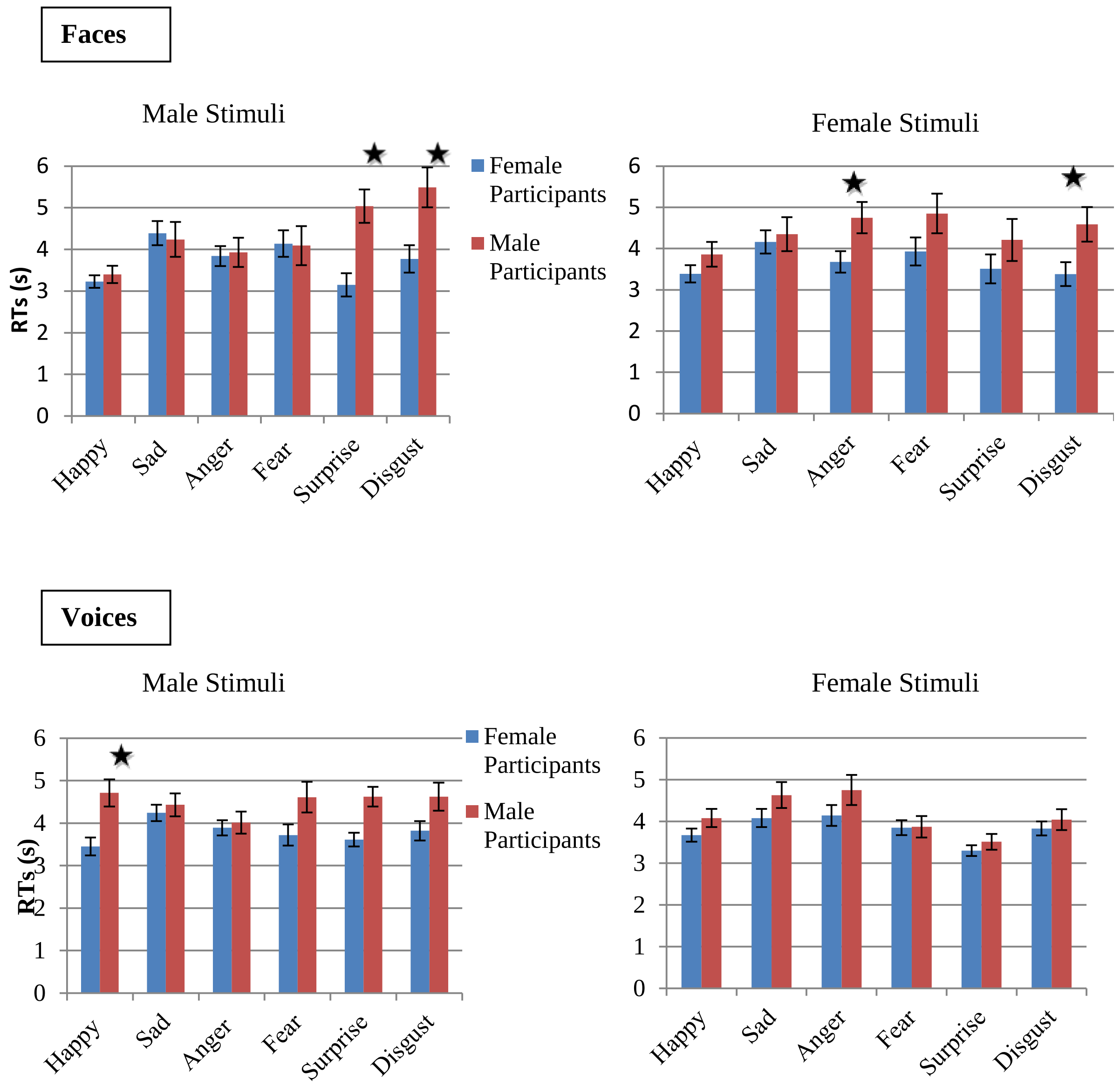


Figure 2.3. Means and Standard Errors for RTs. Significant four-way interaction effect between modality, emotions, gender of participants and gender of stimulus, $*p < .05$. $N = 45$

2.3.2. Analysis 2: Confusion Data

For each individual participant, the whole spectrum of possible emotion combination (36 pairs) was analysed. Confusion data which display the mean intensity ratings for each fo the 36 possible combinations across all participants are displayed in Figure 2.4. For voices and faces, the confusion matrices illustrate some similarities in rating profiles of individual emotions across two separate modalities. As can be seen, for example surprise and fear are often confused within the visual modality – but also within the auditory modality. Happiness on the other hand is very rarely misinterpreted as another emotion – in voices as well as in faces.

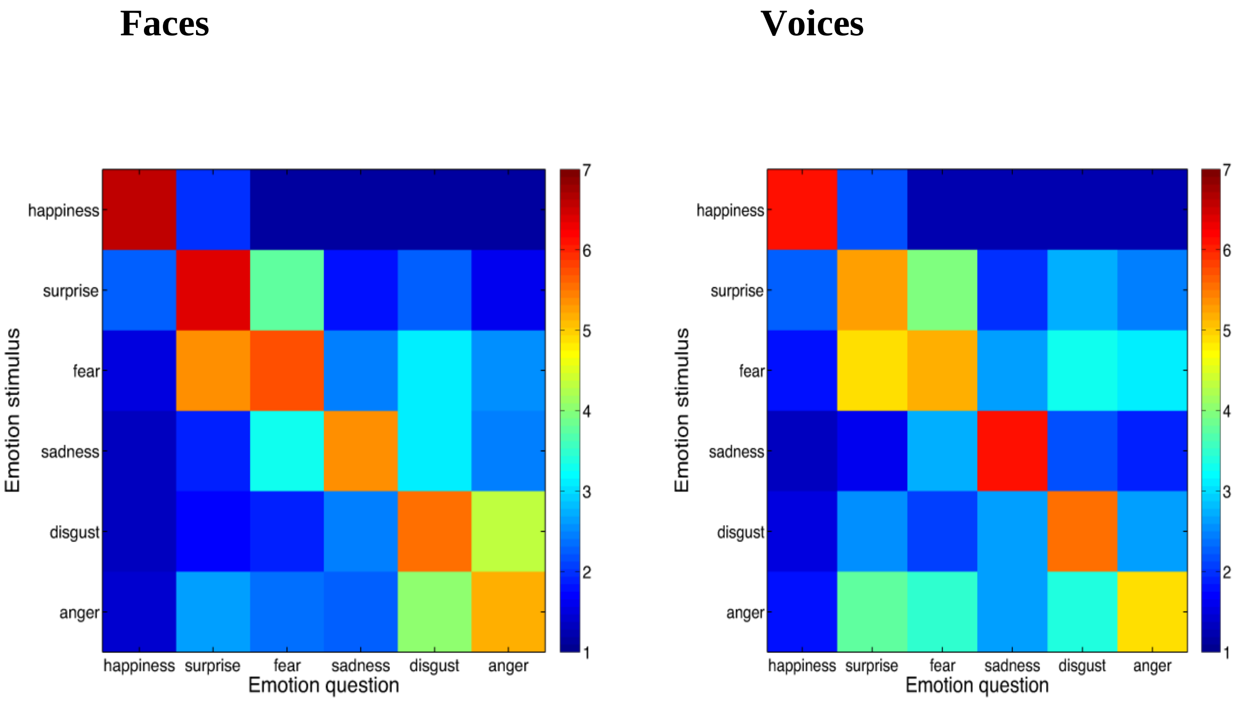


Figure 2.4. Confusion matrices for voices and faces with mean intensity ratings from 45 participants.

The confusions between emotion pairs across two modalities are also illustrated in Figure 2.5, which shows the specific emotion rating profile for each of the six basic emotions.

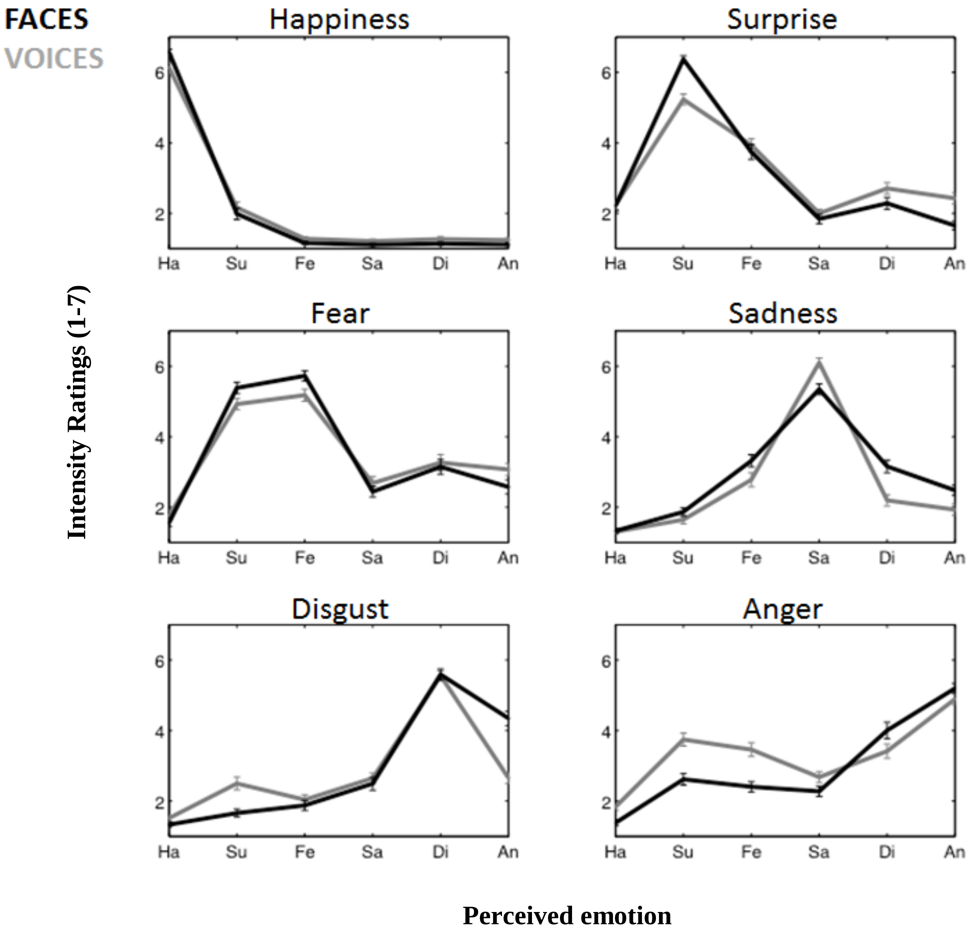


Figure 2.5. Complete emotion rating profiles for each of the six emotions, split by modality. *N* = 45. Ha = happiness, su = surprise, fe = fear, sa = sadness, di = disgust, an = anger.

In order to calculate the similarity of correlations between auditory and visual rating profiles per emotion and participant, the face and the voice conditions were each split into two separate datasets (yielding four datasets in total). The first two presentation blocks of the experiment (F1 and V1) formed the first pair whilst the third and fourth presentation blocks of the experiment (F2 and V2) formed the second pair. Hence, each dataset contained the same number of stimuli per emotion label and emotion category. Thus, four datasets for each participant were obtained: two datasets with average intensity ratings for each label and each emotion for faces (F1 and F2), and two datasets with average intensity ratings for each label and each emotion for voices (V1 and V2). To examine similarity of emotional response profiles, the correlations between response profiles within same modality (F1 *versus* F2 and V1 *versus* V2) and across modalities (F1 *versus* V1) were computed.

Hence, there were four independent datasets that included mean intensity ratings for the full emotion profile for faces (dataset F1 & F2) and the same for voices (dataset V1 & V2). Then, for each individual participant, correlation analysis was performed between response profiles within the same modality (F1 *versus* F2 and V1 *versus* V2) and across modalities (F1 *versus* V1).

The mean correlation across all participants between both face datasets (F1 *versus* F2) was high, r = .78 (SD = 0.10) and emotion ratings from half of the face stimuli correlated with the rating profiles in the other half of the face stimuli and explained about 60% of variance.

For voices, similar results can be found with mean correlations of r = .71 (SD = 0.12) across participants between both voice sets (V1 *versus* V2), explaining about 50% of variance in rating profiles. This suggests consistency in emotion labels associated with stimuli across two sets of stimuli within each modality. Interestingly, not only within but also across modalities, the rating profiles also moderately correlated (F1 *versus* V1), r = .68 (SD = 0.14). About 46% of the variance in the rating profiles for faces can be predicted by the rating profiles for voices (and vice-versa).

In order to investigate whether the mean correlations for rating two subsets of the same modalities differ from correlations for rating two separate modalities, a repeated-measures ANOVA with condition (3 levels) as within-subject variable was conducted. Since correlations are usually not normally distributed, correlation scores were z-transformed before entered into the ANOVA. The repeated-measures ANOVA revealed a significant main effect for condition, $F(2, 88) = 19.11$, $p < .001$. Post-hoc pairwise comparisons revealed that on average, correlations for rating profiles from F1 versus F2 were significantly higher than for correlations for rating V1 versus V2 and higher than correlations for ratings across modalities F1 versus V1. Interestingly, mean correlations for ratings profiles across two separate modalities F1 versus V1 were not significantly different from correlating rating profiles within one modality V1 versus V2 (p > .05). This is illustrated in Figure 2.6.

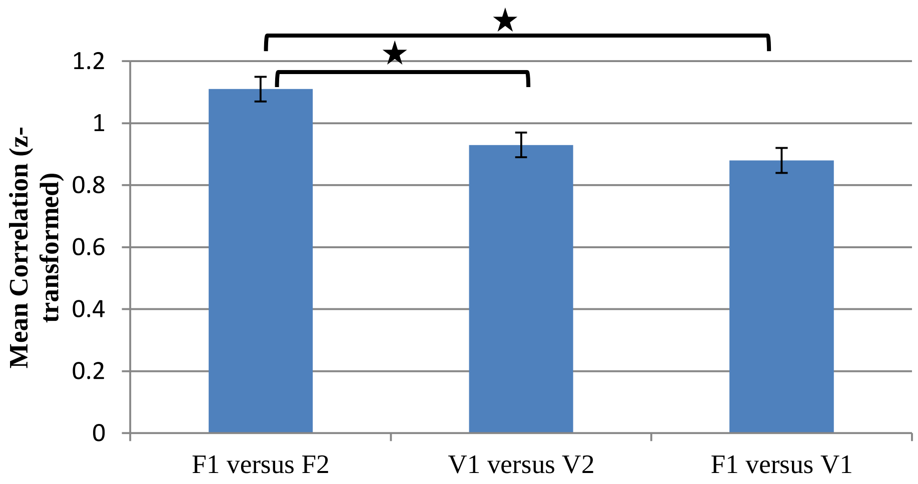


Figure 2.6. Significant mean correlation scores (N = 45) between the complete rating profiles of emotional stimuli either from the same (F1 vs F2 and V1 vs V2) or across different modalities (F1 and V1). With Standard Errors. *p < .05.

2.4. Discussion

The aim of the present study was to establish just how similar the independent modalities of the voice and face are in facilitating recognition of the six basic emotions within the same participant. To the best of the researcher's knowledge, the current study is the first to investigate whether happiness or anger superiority effects are predominantly a feature of visual emotion recognition or whether the same valence effects are also observable in parallel within the auditory domain. Further, the possible existence of shared emotion decoding mechanisms across modalities was investigated by examining gender effects (of actors as well as of participants) within the voice and the face condition. In addition, the present study investigated whether the full rating profiles of emotions, as evident from confusion data – were also correlated between both modalities.

In line with Hypothesis 1, the analysis of target emotions in the present study pointed towards a positive valence superiority effect that was replicated across both modalities. Previous findings regarding the anger superiority - as reported for example by Damjanovic et al. (2013) - could not be replicated. Further, there was evidence for the influence of gender of participants on speed of making emotion decisions – independent of modality which supports Hypothesis 2. Gender of actor/stimuli additionally influenced the perceived intensity of emotions – again, independent of modality. In line with Hypothesis 3, the analysis of confusion data revealed highly correlated confusion matrixes within as well as across modalities. This finding hints towards the possibility of common underlying emotion recognition mechanisms which are independent of modality. This finding may potentially support neurological studies of shared impaired emotion recognition in faces and voices after isolated brain damage (*e.g.* Fecteau et al., 2007) or shared representations of object categories (Shinkareva et al., 2011).

2.4.1 Modality Similarities

In hand with neuroimaging studies which demonstrated activation in shared brain regions during emotion categorisation across a range of modalities (Peelen et al., 2010), data from the present behavioural study emphasizes the idea of modality-independent emotion recognition which may be due to a common coding mechanism. For example, neither RTs nor intensity ratings differed as a function of modality. This has previously also been reported by Hawk and colleagues (2009) who reported lowest accuracy scores for prosody recognition whilst recognition rates for emotional face expressions and non-verbal vocalisations did not differ from each other.

Looking closer at the interaction between individual emotions and modality, data revealed that faces communicated the emotions of happiness, fear and surprise with higher perceived intensity than voices whilst for sadness, the opposite was true and voices communicated sadness with higher perceived intensity than faces. For disgust and anger, participants did not differ in their intensity perception across both modalities. It is plausible that for emotions which convey immediate information relevant for survival such as anger and disgust, it is beneficial to recognise signals with equal intensity across a range of modalities. For emotions that require less immediate action such as happiness, fear or surprise, focusing on the face alone may be sufficient whilst sadness seems to be communicated especially well within the auditory affect bursts. In order to further demonstrate similarities in emotion recognition across voices and faces, the following section will firstly discuss data on valence superiority effects, before moving on to gender and age effects as well correlations of emotion confusions across both modalities.

Overall, current behavioural evidence on emotion recognition similarities across modalities supports findings from brain imaging studies that suggested that emotions from faces (*e.g*. Adolphs, 2002) as well as from voices (*e.g*. Phillips et al., 1998) may be processed within comparable neural circuits. Further, in line with previous behavioural evidence from autism (Phillips et al., 2010) or Schizophrenia (Simpson et al., 2013), the ability to recognise emotions in one modality may indeed be directly linked to the ability to recognise emotions in other modalities, suggesting one emotion core network rather than separate, modality-dependent mechanisms.

Apart from neural evidence, modality-independent processes may, for example, also be rooted in top-down cognitive mechanisms rather than on the bottom-up analysis of modality-specific perceptual features. It is possible that the same semantic representation of emotion categories may be activated across different types of stimulus presentation formats as has previously been shown for objects depicted in pictures or as written words, possibly reflecting a common neural representation (Shinkareva et al., 2011). Further, parallels in emotion recognition across modalities may also originate from perceptual interdependence of physical features during emotion production. For example, Ohala (1980) suggested that the smile in the animal and human face may be a consequence of vocal tract movements to increase the pitch of vocalisations in order to resemble an infantile-like, submissive sound. This pitch variation may protect the sender from an imminent attack and hence, the contraction of the lip corners accompanies changes in the vocal tract resonance. In order to further demonstrate to the reader of the similarity across modalities found in the present study, the following section will discuss valence and gender effects across two independent modalities.

2.4.2 Happiness Superiority in Faces and Voices

This present study is one of the first to have obtained data on valence-superiority effects across non-verbal auditory as well as visual stimuli from the same participant. Contrary to past research on anger superiority effects such as by Damjanovic et al. (2013), the current study did not find a general negative valence superiority effect. In fact, for faces, anger was consistently one of the worst recognised emotions whilst happiness showed fastest RTs and highest intensity ratings. This seems surprising in the light of evolutionary theories since threatening emotions ought to be recognised effectively in order to avoid danger. So what could be a possible explanation for this happiness superiority effect?

In happy faces, a smile takes up a very large proportion of the face. This large perceptual feature within the face may enhance rapid recognition of happiness (Smith, 2011). Indeed, Du and Martinez (2013) suggested only 10-20ms of exposure time for happy faces are needed for correct classification, indicating that happiness in faces is rapidly recognised – even if image resolution was reduced. However, from the findings within the visual domain alone, it is not clear whether happiness superiority effects are only due to perceptual features within the face or may also be due to more modality-independent top-down processes such as attachment promotion or offspring bonding (Babchuck et al., 1985). Interestingly, the current study also provides evidence for happiness superiority within the voice. Again, anger was the emotion recognised with less intensity and longer reaction times throughout. This seemed to be independent of the length of the auditory stimuli since the mean length for angry stimuli used in this study was 970.5ms which was significantly shorter than the average vocal expression of happiness (1267ms).

From a cognitive perspective, anger approach-and-avoidance theories (*e.g*. Marsh, Adams, & Kleck, 2005) suggest that angry stimuli often elicit avoidance rather than approach behaviour. This has been tested by the tendency to manually increase the spatial distance between the written word '*me*' on a computer screen (which represented the participant) in response to a static angry face. It is possible that we prefer to avoid uncomfortable situations and instead pay attention to people displaying positive attitudes - especially when aggression towards the angry person is not an option (Krieglmeyer & Deutsch, 2013). This tendency may be modality-independent and may be based on evolutionary concepts. For example, efficient recognition of positive emotions may enhance social bonding and offspring care-taking (Baumeister & Leary, 1995). Further, happiness superiority could be explained by studies on biological aspects on altruism, suggesting that humans are more willing to help happy people as compared to people in immediate distress which is in accordance with social affiliation models[9] (Hauser, Preston, & Stansfield, 2013). In our everyday life, it may simply be beneficial for us to attend to social partners that have a positive attitude towards us and may pose as a potential mating partner or be sympathetic to helping out in need.

However, it is important to acknowledge that the modality-independent happiness superiority effects found in the present study could simply be a confounding result of comparing one positive emotion (happiness) with one bidirectional (surprise) or four negative (sadness, anger, fear and disgust) emotions. Hence, it may be easier to distinguish one single positive emotion from several negative emotions which could also account for higher confusion rates amongst several negative emotions of the same valence. Further, happy stimuli may be particularly prone to iconic

[9] Social affiliation models suggest the idea to invest in 'cooperative partnerships that benefit survival and success in group-life' (Hauser et al., 2013, p. 2).

representations by actors (Scherer, Clark-Polner, & Mortillaro, 2011). This exact issue of limited numbers of positive emotions typically included in emotion research was addressed by Sauter and Scott (2007) who compared non-verbal vocalisations of several positive emotions such as relief, contentment and amusement. They reported that especially the positive emotion of 'amusement' was recognised with high accuracy rates that exceeded 90%. When conducting the study with children as young as five years old, amusement and relief yielded highest accuracy rates compared to several negative non-verbal affect bursts (Sauter, Panattoni, & Happe, 2013). Further, for face recognition, Juth and colleagues (2005) compared one single positive (happiness) with one single negative (anger) and one neutral emotion, and also found fastest RTs for faces expression happiness. The current issue may not be as black and white but previous studies such as by Sauter et al. (2013) - at least partially - support the notion that emotions of positive valence may enhance emotion recognition, independent of the proportion of positive and negative emotions displayed. The current data suggests that this could be extended to auditory as well as visual stimuli.

2.4.3 External Factors

a) Age and Gender of Participants

Accordingly to the predictions set aside, the present study found significant differences between gender of participants for RTs; this, however, was not true for intensity ratings. Findings may possibly reflect a female advantage in emotion recognition which has previously been linked to greater mirror neuron system activation in females during emotion judgements (Schulthe-Ruether et al., 2008). In the current study, gender differences for speed were true for both modalities. In the face as well as in the voice condition, female participants made faster emotion categorisation decisions than male participants. It is possible that whilst gender may have an impact on the rating speed, intensity perception may be comparable between male and female judges.

In the light of the 'primary care-taker hypothesis' (Babchuck et al., 1985), speed – rather than intensity - of recognising emotions in others may play a very important role in protecting offspring. The current data suggested an importance of disgust because across both modalities, female participants rated disgust faster than male participants. This finding may be related to females' necessity to perceive disgust during offspring rearing. In addition, women rated happiness in voices faster than men, which again, might related to attachment theories and females' evolutionary tasks of forming attachment with offspring and mating partners (Babchuck et al., 1985). This female superiority was not only limited to positive emotions; findings suggest a female care-taking role that includes threat-avoidance as well as attachment formation.

However, a significant difference for gender was not found for intensity ratings – although, again, this applied to both modalities. Unlike result reported by Lambrecht et al. (2014), gender effects did not depend on modality as neither the facial nor the auditory domain showed significant gender effects for intensity ratings. However, it is important to note that the ratio of male and female participants used in the present study was not equally distributed and further research with larger sample sizes needs to be conducted to address this current issue.

Although not specifically worded as hypothesis, the co-variate of age demonstrated that age of participants influenced the way emotions are processed. For intensity ratings, increasing age seemed to decrease emotion intensity perception and for RTs, increasing age resulted in longer choice RTs across participants. This was true for the face as well as for the voice condition. This is a consistent finding in emotion literature with older adults often showing decline in emotion recognition tasks in faces (Williams, Mathersul, Palmer, Gur, Gur, & Gordon, 2009) and in voices (Demenescu, Mathiak, & Mathiak, 2014). Age-related emotion recognition differences could be the result of changes in general cognitive decline, neuronal activity or motivation (Issakowitz & Stanley, 2011). Overall, the same pattern of presence - or absence for that matter – of gender and age affects across both modalities suggests the existence of general supramodal emotion perception mechanisms.

b) Gender Stimuli

Although not specifically predicted, the present study found that emotions expressed by females were generally perceived as more intense than men. This was especially true for emotions of anger, fear and disgust. For disgusted faces, female stimuli were also rated faster than male stimuli – when rated by male participants. Interestingly, as stated above, present findings suggest a connection between being female and either recognising or being expected (by males) to express disgust in faces. This finding possibly indicates females' importance of expressing and perceiving disgust during offspring rearing as stated by the 'primary care-taker hypothesis' (Babchuck et al., 1985). For example, it is equally important for children to learn from the primary caretaker when something is disgusting in order to avoid rotten food as it is for mothers to react rapidly to offspring that shows signs of having consumed poisonous objects.

In addition – although RTs for gender of stimuli was not a significant main effect– men showed faster speed when rating angry voices expressed by other men as compared to women. This finding is interesting because it shows a) that men may be conditioned to rapidly recognise anger in voices, suggesting a possible threat-avoidance or fight preparedness and b) that men might show own-gender bias for recognising anger in other men. It has been stated from face research that male faces tend to be recognised faster than female faces when they express anger (Aguado, García-Gutierrez, & Serrano-Pedraza, 2009; Becker, Kenrick, Neuberg, Blackwell and Smith, 2007). A recent ERP study (Valdés-Conroy, Aguado, Fernández-Cahill, Romero-Ferreiro, & Diéguez-Risco, 2014) also suggested a stronger brain response to male as compared to female angry faces as early as 150ms after stimulus onset. Recently, anger displayed by male actors also elicited stronger startle reflexes in viewers than female faces, demonstrating higher perceived dominance in threatening male faces (Paulus, Musial, & Renn, 2014). This could be a result of sensitivity to stereotypes such as men are expected to express anger (Aguado et al., 2009) or because of male features, such as the relative position of eyebrows within the face, which may naturally appear more angry than female faces (Hess, Adams Jr, Grammer, & Kleck, 2009). Here, the present data suggests that this association between rating anger and male actors may go beyond visual features within the face and could potentially be extended to the vocal domain.

However, the current data only supports the above statement if the judges were also male. This in turn may hint at possible gender in-group effects for rating speed of emotions. For example, past research has shown amygdala lateralization depending on gender of judges with men showing right amygdala activation when judging emotional male faces with the opposite to be true for females (Armony & Sergerie, 2007). However, again, it is to remember that the present sample size of male and female participants was unequal and so results have to be interpreted with caution. Further, present trials per condition were relatively small when split for gender of stimulus and would need to be replicated with increased trial size before generalizing the current findings.

2.4.4 Confusion Data

Results from the confusion matrices indicate that the rating profiles from faces correlated highly with the rating profiles from voices. Independent of the mode and perceptual features of emotions presented, judges associated specific emotions with specific emotion labels which was consistent within, as well as across modalities. For example, in faces, confusions occurred between pairs of emotions such as fear and surprise and between anger and disgust. Confusions between fear and surprise in faces have often been reported due to similar visual features within the face such as wide opened eyes and raised eye brows (Castelli, 2005; Etcoff & Magee, 1992; Matsumoto & Ekman, 1973). In a recent study, Jack et al. (2014) demonstrated that fearful and surprised as well as disgusted and angry dynamic face expressions share the same basic perceptual features during early processing stages and only later do more complex, distinguishable emotion features for six individual emotion categories develop. Anger has also been classified as disgust – although not the other way around - in a face emotion recognition study by Du and Martinez (2011) and the confusion rate increased the lower the picture resolution was.

For voices in the present study, confusions also occurred between emotion pairs such as fear and surprise as well as fear and disgust and anger and disgust, which has partly been supported by Belin et al. (2008) and Sauter et al. (2010). It seems surprising that pairs such as fear and surprise often get confused in faces as well as in voices – independent of perceptual features. The current finding of highly correlated rating profiles as well as confusion matrices across modalities contributes to the current argument of supramodal emotion perception based on shared underlying mechanisms.

In the present thesis, certain confusions between emotions – as evident from the confusion matrices - were more common than others, such as disgust being identified as anger. For facial expressions, disgust and anger share early perceptual signals such as the shared nose wrinkle and lip movements (Jack, Garrod, & Schyns, 2014) which may account for common confusions. From an evolutionary perspective, angry expressions can carry the meaning of aggression and hence, it is important to correctly recognise anger in others in order to avoid threat (Damjanovic, Pinkham, Clarke, & Phillips, 2014; Seligman, 1971).

Error management theory has proposed that it might be more costly to underestimate- rather than overestimate- anger expressed in others (Haselton & Nettle, 2006). Hence, we may be biased towards overestimating emotion expressions as angry (even if they are not) as this has important implications for survival (Holbrook et al., 2014). Consequently, in the present thesis, faces which share certain muscle activation, such as disgust and anger, may both be biased towards 'looking angry', in order to avoid costly underestimation of anger and hence risking attacks. Interestingly, humans are also biased towards judging moving angry face expressions as approaching faster than neutral faces perhaps because overestimating the speed of potentially threatening objects can enhance survival (Brendel, DeLucia, Hecht, Stacy, & Larsen, 2012). Hence, emotion judgements and associated common errors in emotion judgements may be influenced by an evolved bias towards making a less costly error of overestimating rather than underestimating anger in others. Interestingly, in the present thesis, this seemed to be common to both modalities, suggesting a modality-independent bias of emotion judgements.

2.4.5 Limitations & Future Ideas

Although the static Ekman faces (Ekman, 1975) of basic expressions are widely used and validated, recent studies have started to include newly recorded dynamic visual face stimuli which evolve over time in order to create more natural situations within the laboratory.
It would be interesting to see whether the current happiness superiority effect could also be replicated in dynamic faces. It has previously been stated that for dynamic faces, anger – rather than happiness – superiority effects may occur due to anger-typical features within the moving face that may appear later in time and are not instantly accessible in static faces (Ceccarini & Caudek, 2013). Hence, for future studies, it would be interesting to collect dynamic, coloured visual as well as dynamic vocal data from the same poser in order to investigate whether emotions expressed across different modalities correspond when they are expressed by the same person.

Although several studies have identified emotion recognition locations within the human brain that may be shared across modalities (*e.g.* Peelen et al., 2010), less is known about the temporal dynamics of emotional face and voice processing within the same individual. Conducting objective neurophysiological studies as a within-subject design could determine how similar the temporal make-up of emotion transmission across modalities within the brain might be. The idea that visual and vocal domains have very similar patterns for individual emotion processing also raises the option to create intervention or training programs for individuals with emotion recognition deficits. Most previous work focuses on emotion recognition training based on visual input, especially in children with Autism Spectrum Disorder (Baron-Cohen, Golan, & Ashwin, 2009; Williams, Gray, & Tonge, 2012). It could therefore be beneficial to create training programs based on auditory as well as visual input since emotion recognition in both modalities may be based on shared mechanisms.

Lastly, this study included British adults, reportedly belonging to the British culture. Although many studies claim the universal existence of basic emotions across cultures (*e.g*. Ekman & Friesen, 1971, Scherer, Banse, & Wallbott, 2001), one cannot conclude that the current results may be stable across non-Western countries. For example, Jack et al. (2009) repeatedly found that Eastern and Western cultures use different perceptual features of the face (such as the eye region in Eastern groups and extract information from the whole face in Western groups) to decode specific emotions. Since cultural display rules can influence emotional intensity experience which in turn has an effect on emotion intensity recognition (Schneider, Hempel, & Lynch, 2013), it would be beneficial to conduct the same study from a cultural perspective.

2.4.6 Conclusion

Assuming shared cognitive (Borod et al., 2000) as well as neural networks (Peelen et al., 2010), emotions deemed to be important for survival or reproduction within one modality ought to also be recognised with same intensity in a parallel, yet separate modality. Indeed, data from the present behavioural study suggested similar emotion communication mechanisms across modalities based on valence superiority, gender or age effects and confusion rates. The current results could mean emotion recognition is based on general emotion-specific, but modality-independent mechanisms. Hence, the current study may extend Yovel and Belin's hypothesis (2013) of person recognition by adding information about similar cognitive mechanisms during emotion recognition across two separate modalities. Whilst inferences from the present Chapter 2 about a modality-independent emotion mechanism are drawn upon adult participants, it remains unanswered whether children also demonstrate similarities in emotion processing across modalities or whether children rely on one modality more than on another. Consequently, the following Chapter 3 will describe a developmental study of emotion recognition across modalities in primary school children.

3

Children's Emotion Recognition in Human Voice and Face.

Chapter 3 aims to examine findings from the adult data in Chapter 2 within a developmental context including children aged 5 to10. Whilst the adult study in Chapter 2 demonstrated a) similarities in rating emotions displayed in faces and in voices as visible in choice speed and intensity perception, b) superior recognition for happy as compared to angry stimuli replicated across two independent modalities as well as c) highly correlated confusion patterns across independent modalities suggesting supramodal processing, it is of considerable interest how and when these patterns develop across childhood. Investigating emotion perception development across separate modalities in primary school-aged children could help to assess the component of social learning and developmental factors to emotion understanding. This is especially important under conditions where only one modality is accessible such as during phone conversations or extremely noisy play-ground situations. Finding similar patterns to the adult data could indicate supramodal emotion perception mechanisms which are acquired early in childhood.

Overall, the present chapter focuses on the development of emotion perception across two separate modalities in a within-subject design. Individual factors such as children's gender, Theory of Mind (ToM) abilities and non-verbal IQ scores were controlled for in Experiment 3-I. Further, some longitudinal data from a smaller number of 7 and 8-year-old children aims to deliver a deeper understanding of developmental trajectories for emotion perception in Experiment 3-II. In addition, a smaller Experiment 3-III evaluates the appropriateness of using adult face stimuli in a child study by comparing children's emotion recognition abilities for either judging adult or child faces. This methodological comparison also aims to investigate whether children demonstrate in-group age effects during emotion perception in faces of different age groups as otherwise sometimes shown in adult studies (Kuefner, Kassia, Piccozi, & Bricolo, 2008).

3.1. Experiment 3-I

Children's Emotion Recognition in Human Voice and Face:

A Developmental Perspective.

Normal development of emotion perception is a crucial prerequisite for social competence and communication skills in children. For example, abnormally enhanced perception of angry faces in childhood has been linked to later substance abuse in adolescents (Ernst et al., 2010) and caregiver's reports of children's social behaviour has been associated with children's recognition accuracy in face expressions (Parker, Mathis, & Kupersmidt, 2013). It has now been established that the ability to perceive and judge emotions in faces and voices develops and increases with age throughout childhood (Herba & Phillips, 2004; Gagnon, Gosselin, Hudon-ven der Buhs, Larocque, & Milliard, 2010) and even until adolescence (Aguert, Laval, Lacroix, Gil, & Le Bigot, 2013; Thomas, de Bellis, Graham, & LaBar, 2007).

3.1.1.1. Children's Emotion Recognition in Faces

Are young children able to identify emotions expressed in faces? By observing children's looking time at faces with different emotion expressions, it has been claimed that 4 to 6 month old infants can already discriminate between two negative expressions such as fear and anger (Serrano, Iglesias, & Loeches, 1999). According to Pons, Harris and de Rosnay (2004), children undergo three different stages of face emotion recognition. By reflecting on emotional stories told by cartoon characters, it became evident that most 5-year-olds were able to name basic face expressions. Seven-year-old children were able to understand emotions communicating beliefs, desire and hidden emotions which required understanding of other person's mental state. Not before the age of 9 did children pass the final stage of correctly identifying emotions that are mixed in nature or are influenced by morality. Accordingly, although accuracy did not improve after the age of 7, *speed* of recognition increased between the ages of 7 to 11-year old children (De Sonneville, Verschoor, Njiokiktjien, Op het Veld, Toorenaar, & Vranken, 2002), suggesting that – although basic emotion recognition has its onset early in childhood – performance continued to stabilize with age.

Although most developmental studies of emotion recognition have focused on early years of childhood, several studies have demonstrated that emotion recognition develops well into adolescent years which may be especially true for the perception of mixed emotions whose successful interpretation depends upon appraisal of social context (Izard & Harris, 1995) and emotions displayed with varying intensities (Montirosso et al., 2010). Gagnon and colleagues (2010) stated that fear and disgust showed higher recognition accuracy in 9 to10-year-olds as compared to 5 to 6 year-old-children which demonstrated a gradual improvement of emotion recognition. This goes in hand with findings by Thomas and colleagues (2007) who investigated emotion recognition in morphed faces across older children, adolescents and adults and found that recognition of changes in emotion expressions improved well past adolescent years in terms of speed and correct responses. Those changes seemed to be linked to specific emotions as shown by increased sensitivity to changes in angry faces between adolescence and adulthood which happened later than for fearful stimuli. Overall, it seems evident that even basic categorisation of emotions within the face is by no means complete before the time children enter primary school and many developmental changes happen just before they move into secondary school around the age of 10.

3.1.1.2. Children's Emotion Recognition in Voices

Although the majority of emotion recognition studies have focused on facial expressions, some studies have attempted to assess developmental trajectories in vocal emotion recognition which are mainly based on prosody of spoken sentences. It has been proposed that even new-borns (around 34 hours after birth) distinguish between different emotions expressed in the mother's voice. This has been shown by a higher tendency in infants to open eyes when listening to happy compared to sad, angry or neutral voices (Mastropieri & Turkewitz, 1999). Moving on to middle childhood, two earlier studies by McCluskey and Albas (1981) and McCluskey and colleagues (1975) have shown that just like for face processing, vocal emotion recognition in children also seems to continuously improve with age right into adulthood. They demonstrated that Canadian as well as Mexican participants aged 5 years onwards showed increased accuracy in content-free vocal emotion identification until the age of 25. However, McCluskey et al. (1975) made no connection to any external variables such as gender, IQ or Theory of Mind development in relation to emotion processing development which limits the generalizability of results. In a more recent study by Rothman and Nowicki (2004), a nonverbal measure of children's paralanguage (DANVA2-CP) was validated and results indicated that children's intensity ratings of child voices performance increased with age with errors decreasing significantly between the ages of 4 to 12.

Recently, Aguert and colleagues (2013) suggested that emotional prosody recognition of even very basic emotions like happiness and sadness is difficult for children at the age of entering primary school and only becomes more stable when most children move into secondary schools. Apart from linguistic prosody-recognition, one of the first studies to look at children's understanding of emotions presented via non-verbal affect vocalizations such as a cry or moan was conducted by Sauter, Panattoni and Happe (2012). Looking at developmental progress children make in recognising emotions from vocalizations rather than prosody, they suggested that children between the ages of 5 to 10 were able to recognise ten different emotions such as relief, fear or amusement. This result is interesting since the study extends beyond the six basic emotions and also offers the comparison of several emotions of positive valence rather than just focusing on happiness. Sauter et al. (2012) concluded that - similarly to face expression or prosody understanding - recognition abilities for emotional vocalizations also improve with age.

Overall, previous studies have attempted to pin down the time course of emotional development from either the visual or the auditory domain. Although across both channels, infants show early awareness of different emotion categories, advanced emotion classification does not seem to be in place before primary school age and does not cease to develop well into adolescent years. However, comparisons across studies or even across modalities are extremely difficult due to a large variety of stimuli, task instructions and differences in sample selection. Only a limited number of studies have explicitly observed children's emotion recognition across several modalities in a within-subject design. For example, Nelson and Russell (2011) reported that during early childhood (children aged 3 to 5) generally showed high recognition rates for emotion presented in voices, faces, body posture or multimodal conditions but recognition improved with age. The auditory condition showed lowest recognition rates.

It has been proposed that –even for basic emotions like sadness and happiness - children only start to utilise acoustical information from prosody aged 4 to 5 (Quam and Swingley, 2012). In addition, Tonks, Williams, Frampton, Yates and Slater, (2007) investigated the ability to recognise emotions from faces, vocal prosody and eyes during late childhood (children aged 9 to 15) and concluded that, whilst decoding emotions from faces and eyes improved until the age of 11, performance on vocal cues remained stable across age. Across early and late childhood, it seems that children are more equipped to read emotions from faces rather than from voices. However, to the best of the researcher's knowledge, no previous studies have examined emotion recognition across several modalities during middle childhood. Hence, the questions arises whether the social brain develop continuously throughout the years of primary school and whether this depends on modality? Does this also include the emotion recognition in non-verbal affect vocalisations?

3.1.1.3. Neurological Evidence for Changes in Emotion Recognition

Assuming behavioural changes in emotion recognition across childhood, what could account for those age-related changes? Several neuroimaging studies have recently begun to investigate how emotional-related brain structures such as the amygdala change with age. Wierenga and colleagues (2014) conducted a longitudinal fMRI study and concluded that some subcortical structures within the developing brain continue to mature right into adolescent years. They found that volume of the amygdala and hippocampus increased with age. Especially the amygdala is commonly associated with emotion processing such as the recognition of emotions from faces (Morris et al., 1998), non-verbal vocalisations (Fecteau, Belin, Joanette, & Armony) and even music (Gosselin, Peretz, Johnsen, & Adolphs, 2007). Hence, it seems that structures which are important for emotion processing are not fully developed in primary or even secondary school aged children. Similarly, Mills, Lalonde, Clasen, Giedd, and Blakemore (2014) indicated that social brain structures – such as the anterior temporal cortex, medial prefrontal cortex and posterior superior temporal sulcus - continued to develop between the ages of 7 and 30. Data was collected in a longitudinal design with scans conducted roughly every 2 years and investigated an underlying cause for the behavioural age-related changes in emotion recognition abilities reported in the present study.

Apart from structural changes, functional changes during emotion processing across childhood have also been observed. A decrease of amygdala activation during observation of emotional visual scenes has been found to be replaced with an increase of prefrontal activation between the ages of 10 and 24 (Vink et al., 2014). Hence, processing of emotional stimuli in children may be associated with amygdala activation whilst this function seems to be shifted to frontal brain areas in adults. This is supported by developmental studies suggesting that younger children whose frontal lobe hasn't fully matured yet, showed similar performance in matching faces to cartoon situations to patients with frontal lobe damage (Kolb, Wilson, & Taylor, 1999). It has been stated that the connection between limbic structures such as the amygdala and prefrontal areas may consolidate with age as shown in an fMRI study investigating changes in the functional connectivity of the amygdala at rest for participants aged 4 to 23 (Gabard- Durnam et al., 2014). However, according to a review by Herba and Phillips (2004), 'the role of the amygdala in emotion processing throughout childhood and adolescence remains unclear' (p. 1191) and not all studies have shown age related changes in amygdala activation during emotion processing (*e.g.* Baird et al., 1999).

Overall, neuroimaging data from children suggests that brain areas and connections which are commonly associated with emotion processing, mature between childhood and adult years. In line with this, it is predicted that emotion recognition abilities in the present child sample improves with age. However, functional imaging studies linking neuronal responses and emotion processing are typically based on visual processing of emotional stimuli.

As summarised below, it is possible that neuronal structures may influence emotion recognition not only in the visual but also in the auditory domain.

3.1.1.4. Evidence for Similarities Across Modalities

Data from the adult data in Chapter 2 suggested the possibility of underlying coding mechanisms for emotion recognition which may be shared across modalities. Indeed, visual and auditory emotion recognition was simultaneously impaired after temporal lobe (Bonora et al., 2011; Dellacherie et al., 2011) or amygdala damage (Scott et al., 1997), suggesting underlying supramodal emotion networks. Considerably less is known about children's neurological activity during emotion recognition – especially across separate modalities. Is it possible that modality-independent emotion processing is already a feature of the child's brain? It is not clear whether children at primary-school age also show similarities in recognising emotions across modalities as seen from adult data or whether children's judgements rely mainly on external perceptual analyses of stimuli.

For unimodal emotion recognition in children, previous neuroimaging studies have suggested impaired face expression recognition in children aged 6 to 10 presenting with medial temporal lobe epilepsy (Cantalupo et al., 2013). This has also been found for prosody – again, in patients aged 6 to 11 with right temporal lobe epilepsy (Cohen, Prather, Town, & Hynd, 1990). Those findings parallel data from adults with temporal lobe epilepsy which also suggests simultaneous impairment of emotion recognition from faces and voices (Bonora et al., 2011). Above arguments suggest global deficits during right hemisphere or temporal lobe damage which may in turn indicate modality independent emotion mechanisms in similar brain areas by the age of 6 as well as in adults. Further, studies including children with Autism[10] also demonstrated that participants aged 7 to 17 showed less brain activity in the medial prefrontal cortex as compared to healthy children during irony recognition in vocal or in visual conditions (Wang, Lee, Sigman, & Dapretto, 2007). Further, in a behavioural experiment, Munoz (2009) investigated emotion recognition abilities in boys aged 8 to 16 with callous-unemotional traits and suggested impaired recognition of fear not only in faces but also in body posture.

Overall, evidence from neuroimaging as well as behavioural studies suggested that emotion processes may have similar underlying mechanisms across modalities in adults as well as in children. Those mechanisms may already be in place by the age of 5 and could potentially be a result of shared brain areas for emotion processing which mature with age. However, lack of data fails to provide evidence; hence, the present study aims to fill this gap by delivering behavioural data on children's patterns for recognising emotions across modalities.

[10] Autism: Neuropsychiatric disorder which impacts social functioning and empathy understanding, possibly with abnormal amygdala function (Baron-Cohen, Ring, Bullmore, Wheelwright, Ashwin, & Williams, 2000).

3.1.1.5. Individual Factors

Several individual factors such as gender, Theory of Mind (ToM) and Intelligence Quotient (IQ) may play a role in children's ability and development of emotion recognition. This present study aims to control for all of those three factors in order to investigate the possible contribution of each to emotion recognition across two modalities.

a) Gender

For neural as well as behavioural adult data, gender seemed to influence emotion recognition in previous studies. For example, women showed greater event-related potentials in response to passively viewing unpleasant pictures compared to men (Lithari et al., 2010) and females matched semantic content and intonation of emotional stimuli more correctly than their male counterparts (Szymanowski, Kotz, Schroeder, Rotte, & Dengler, 2007). However, this pattern has not always been replicated and several studies could not find significant gender differences for emotion recognition in adults (*e.g.* Hawk et al., 2009). Similarly, results from the adult Chapter 2 did not find any gender effects for intensity ratings for either voices or faces.

For children, the role of gender is equally unclear with some – but by far not all - studies reporting gender differences during emotion decoding for a range of modalities. For example, van Beek and Dubas (2008) demonstrated that girls aged 9 to 15 rated face expressions as more intense than boys. This was especially true for faces displaying anger, suggesting an interaction between gender and specific emotions. The male disadvantage in emotion processing abilities has previously been linked to gender-related neurological structures with the male brain specializing in 'systemizing' rather than 'empathizing' skills, suggesting males preference of analysing systematic patterns rather than specializing in understanding feelings of others (Baron- Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003). However, as mentioned above, gender related effects cause conflicting evidence. For example, for non-verbal affect vocalisations ratings, no gender effects in children have been found (Sauter et al., 2012), suggesting similar perception of emotions in non-verbal vocalisations in children at primary school age. So what role exactly does gender play during emotion recognition in children? This current study aims to address this question by including gender of participants as additional between-subject variable.

b) Theory of Mind

Theory of Mind (ToM) comprises the ability to understand other people's beliefs and desires (Baron-Cohen, Leslie, & Frith, 1985) and a natural consequence of advanced ToM skills is enhanced emotion understanding. For example, a longitudinal study by Ketelaars, van Weerdenburg, Verhoeven, Cuperus and Jansonius (2010) tested 77 children aged 5 for three consecutive years on several ToM aspects such as emotion attribution via story telling.

Findings suggested that children undergo a significant developmental change between the age of 5 to 7, with aspects of ToM such as false belief[11] and emotion understanding improving significantly with age. In addition, Weimer, Sallquist and Bolnick (2012) reported similar findings and demonstrated that children aged 4.5 to 6.5 showed significant improvements in their general emotion understanding - such as emotion recognition from facial expressions - parallel to age and increased understanding of false beliefs.

Indeed, neurological studies support the relationships between emotion recognition and Theory of Mind abilities: adults with traumatic brain injury showed reduced face emotion recognition and the ability to understand feelings of others (Henry et al., 2006) and so did autistic children (for example Frith & Happe, 2002; Heerey, Keltner, & Capps, 2004). However, none of the studies mentioned above have explicitly linked emotion recognition across two separate modalities in relation to age and ToM abilities. Hence, this current study aims to control for children's current ToM level and, in addition, number of participants' siblings was reported since family structure also seemed to be linked with children's ToM abilities (Cutting & Dunn, 1999; Perner, Ruffman, & Leekam, 1994).

c) Non-verbal IQ

It seems obvious that children need basic language capacities in order to verbally describe the emotion in question. Indeed, young children's emotion recognition abilities were positively correlated with language development (Cutting & Dunn, 1999). Since the children included in this study all attend a primary school for normally developed children, sufficient basic language knowledge for emotion categorisation of participants was assumed. The ability to solve novel problems and understand abstract reasoning – as tested by non-verbal IQ tests such as Coloured Progressive Matrices (Raven, Raven, & Court, 1998) – may be particularly important to emotion categorisation tasks. Albanese, de Stasio, Di Chiacchio, Fiorilli, and Pons (2010) tested children aged 3 to 10 and concluded that whilst age correlated positively with general emotion understanding, non-verbal intelligence was also related to emotion understanding; however, this was only true for higher levels of emotion understanding tasks such as rating mixed emotions. It remains unclear whether a link between non-verbal intelligence also exists for rating emotions expressed in faces and voices.

[11] 'False-belief understanding is defined as the ability to understand that others can have an inaccurate understanding (a false belief) of reality.' (Keteelars et al., 2010, p. 87).

3.1.1.6. The Present Study

To the best of the researcher's knowledge, this current study is the first to assess directly how similar two different modalities are in communicating emotions within primary school aged children. The aim is to investigate the time-course of emotion recognition within voices and faces in British children aged 5 to 10 in a within-subject design. It aims to investigate whether emotion develop continuously across childhood and across both modalities. Including an adult control group will help to put children's developing recognition abilities into perspective of the fully developed adult brain. Assuming similar emotion recognition patterns across modalities in adults (see Chapter 2), the present study investigates whether children also show signs of modality-independent emotion processing. Results from the present study can help to identify how children during primary and secondary schools years understand emotions within others – especially when only one modality is available such as during phone conversations. Further, the present study is the first to compare different modalities during middle childhood by using non-verbal vocalisations rather than prosody stimuli.

The present study predicts the following:

Hypothesis 1: Emotion recognition abilities improve continuously with age. This should apply equally to both faces as well as non-verbal vocalisations.

Hypothesis 2: Children will be able to recognise emotions from faces better than emotions from voices although – like for the adult data set - patterns will be similar across modalities.

Hypothesis 3: Similarities across modalities should be reflected by high correlations for confusion matrixes across modalities. Adults may show higher correlations amongst modalities than children.

Hypothesis 4: If individual factors such as gender, IQ and ToM correlate with children's emotion recognition abilities in one modality (more specifically: within the faces), they should equally correlate with emotion recognition abilities in the other modality (more specifically: within the voice).

3.1.2. Methods

3.1.2.1. Participants

Fifty-seven British children (23 male, 34 female) were recruited; however, three children had to be excluded from the study due to English not being the first language and reporting not to belong to the British culture. Adult participants were part of a larger adult sample (*N* = 45) from a separate study in Chapter 2; for the purpose of forming a control group for the present study, only data from younger adults (adults aged 25 or younger) were selected.

The final sample consisted of 54 British children (23 male, 31 female) aged 5 to 10 (*M* = 7.52, *SD* = 1.72) and 27 younger adults in the control group (10 male, 17 female, *M* age = 18.59). Following Sauter, Panattoni and Happe (2013), the children were divided into two age groups: one younger group including children aged 5 to 7 and one older age group including children aged 8 to 10. This included 26 children in Group I (12 male, 14 female, *M* age = 5.96) and 28 children in Group II (11 male, 17 female, *M* age = 8.96). Participants were recruited from a local primary school and Brunel University in West London and were from mixed socioeconomic backgrounds. Children had age-appropriate non-verbal IQ skills (Raven et al., 1998, p. CPM55) within both age groups (*M* = 23.19, *SE* = 5.69 and *M* = 31.46, *SE* = 3.42 respectively). All participants reported normal or corrected-to-normal hearing and vision and English was the first language.

The study was approved by the Department of Psychology, Brunel University, PsyREC Committee (17/10/2012) and informed consent as well as a child-friendly version of the consent form was filled in both by participants and parents prior to the study. All participants were debriefed after the study. Forms are detached in Appendix B.

3.1.2.2. Research Material

Visual & Auditory Stimuli

As for Chapter 2, facial expressions for happiness, sadness, anger, fear, surprise and disgust were individually displayed from the Ekman Pictures of Facial Affect series (Ekman, 1975). However, in contrast to four actors in Chapter 2, for the present Chapter 3 only stimuli from two actors were included: one male (identity JJ) and one female actor (identity C). Again, the Montreal Affective Voice (MAV) samples (Belin et al., 2008) were used as the auditory stimuli and included two actors only: one male (identity 42) and one female (identity 53). Non-verbal stimuli can be used in young children as well as cross culturally to investigate the problem of auditory emotion perception (Hawk et al., 2009).

Like for the adult Chapter 2, each stimulus was presented 6 times at random in order to be questioned on each of the 6 basic emotions resulting in 2 x 6 x 6 visual and 2 x 6 x 6 auditory stimuli for the child condition.

Computer Program

PsychoPy software in Python for psychological experiments by the University of Nottingham (Peirce, 2007) was used to run the experiment. After a short introduction and practice session, the stimuli were presented in a block design with two *voice* trials and two *face* trials presented each in two alternate runs. Within each category, the stimuli were presented at random for counterbalancing. After listening to the sound or seeing the face, each stimulus was rated on a 7-point Likert-scale scale ranging from '*not happy/sad/… at all = 1'*, to '*extremely happy/sad/… = 7'* on each.

For more details on emotional stimuli see Chapter 2.

Nonverbal IQ assessment.

The child-friendly paper based Coloured Progressive Matrices (CPM) test as well as the adult Standard Progressive Matrices (SPM), which had been validated across different nations, was used as a non-verbal IQ test in order to assess non-verbal learning and problem solving abilities such as 'abstract reasoning by analogy and pattern completion' (Raven et al., 1998, p. CPM29). The child version (CPM) consisted of 36 coloured items and had been chosen since it had been tested and validated for the use in children aged 6 years onwards as well as across different cultures such as Germany (Raven et al., 1998). Li et al. (1988) reported test-retest reliability of.95 in Chinese children after 10 days (in Raven et al., 1998). In the CPM, participants had to complete patterns by picking one of four possible choices which ought to match a logical continuous sequence of patterns and/or colours. The difficulty increased over time and finding the correct pattern built on the ability to solve previous problems. For more details about the adult version, see Chapter 2.

Theory of Mind Task and Demographics.

For the child sample, a combination of four Theory of Mind (ToM) stories assessing first, second and third order of ToM as well as False Belief understanding was used. The stories were taken from a study by Liddle and Nettle (2006), see Appendix B. There were two questions for each ToM level and only if children answered both questions correctly, they were classified as having acquired the ToM level in question. In addition, there was a basic Perspective Taking task for 5 and 6-year old children. This included the understanding that a cuddly toy - which was hiding under the table – was not aware of the true content of a box holding chocolate coins (adapted from O'Brien et al., 2011). Tasks ensured that all children had at least a ToM level one as basic requirement for taking part in the present study. Lastly, demographic details of all participants such as age, gender, ethnicity, first language, country of birth, cultural background, number of siblings and educational status (of parents) were collected. See Appendix B.

3.1.2.3. Procedure

Prior to the study, all participants as well as parents of child participants signed an informed consent sheet and provided some basic demographic data described above. This study tested up to two participants at the same time – except from young children aged 5 and 6 where a one-to-one session was always necessary. Participants did not receive a clear definition for each emotion in order to represent an unbiased emotion judgement and participants were not told how often each emotion was displayed. However, it was ensured that all children could verbally explain what the six basic emotions are in order to check their general emotion understanding of the emotions used in the study. In addition, children completed a training session in order to get used to the handling of the computer mouse as well as to grasp an understanding of the Likert scale concept. Participants worked through the stimuli in their own time with an average of 10 minutes to complete each task. No feedback about accuracy or speed of responses was provided. After finishing the two main tasks, children answered four short Theory of Mind questions and 5 and 6-year-olds took part in a playful Theory of Mind task. All children received a small present such as stationary and the school received book vouchers as a 'Thank You' for taking part whilst undergraduate Psychology students received credit points as a course requirement.

3.1.2.4. Statistical Analysis

For the present study, intensity ratings, accuracy rates as well as choice RTs for emotion ratings were collected in three separate mixed measures ANOVAS. Stimuli could either match the corresponding question ('target analysis') or could not match the corresponding question ('confusion analysis'). Between-subject variables were age-group of participants (5 to 7 year-olds, 8 to 10 year-olds and young adults) and gender of participants (male or female). The within-subject variables were modality (face or voice), emotion category (happy, sad, angry, fear, surprise and disgust) and gender of the actor (male or female).

Accuracy Scoring

In order to demonstrate that higher intensity ratings partially reflect better recognition accuracy, accuracy ratings were included in the current analysis. Across all possible emotion combinations that resulted from multiple emotion scales, it was of interest which emotion-pair received the highest intensity ratings. If the congruent target-emotion received highest intensity ratings *i.e.* happy stimuli received highest intensity ratings on the happiness scale), participant received a score of 1 for correct recognition. If the highest intensity rating occurred for an emotion that was not the target emotion (*i.e.* happy stimuli received highest intensity ratings on the angry scale), participant received a score of 0 for incorrect recognition.

Similarly, if the highest intensity rating was shared between several emotions and not classified as one specific emotion, participants also received a score of 0. For each emotion, accuracy scores where then averaged across trials and participants (for more information on accuracy scoring see Kornreich et al., 2012).

RTs

If choice RTs for items exceeded more than 20 seconds, they were classified as outliers and believed to be an error. Outliers were deleted and replaced with the new group mean for that particular variable. This was true for 10 trials. Huynh-Feldt corrections were used for factors that violated the assumption of sphericity. Participants were not aware of RT measurements, indicating that longer RTs may reflect unbiased decision making processes.

Multiple Regression

For the child sample, additionally, separate multiple regression analysis were run for mean scores from either the face or the voice condition with predictors of age, gender, non-verbal IQ (CPM) scores, Theory of Mind levels and number of siblings. Variables were either continuous or dummy-coded into two categories.

Confusion Matrices

For the Analysis of Confusion, confusion matrixes were computed in order to illustrate the average rating profiles across participants for each modality by analysing the mean intensity ratings for all possible combinations of emotion pairs (36 conditions). Further, the correlations of response profiles across modalities for each individual participant were compared across the three age-groups in order to investigate whether the strength of correlation of face and voice rating profiles changes as a function of age.

For more details on statistical analyses see Chapter 2.

3.1.3. Results

3.1.3.1. Analysis 1: Target Emotions

The following Table 3.1 displays Means (*M*) and Standard Error (*SE*) for intensity, accuracy and RT measures in the child sample.

Table 3.1

Means and Standard Errors for modality and emotions for children for intensity ratings, accuracy scores and choice reaction times

	Emotion	**Means Intensity (*SE*)**				**Mean Accuracy (*SE*)**				**Mean Choice RTs (*SE*)**			
		Young Children	**Old Children**	**Young Adults**	**All**	**Young Children**	**Old Children**	**Young Adults**	**All**	**Young Children**	**Old Children**	**Young Adults**	**All**
		N = 26	*N* = 28	*N* = 27	*N* = 81	*N* = 26	*N* = 28	*N* = 27	*N* = 81	*N* = 26	*N* = 28	*N* = 27	*N* = 81
		M *= 5.96	*M* = 8.96	*M* = 18.6	*M* = 11.2	*M* = 5.96	*M* = 8.96	*M* = 18.6	*M* = 11.2	*M* = 5.96	*M* = 8.96	*M* = 18.6	*M* =11.2
		(0.50)	(0.88)	(1.12)	(5.47)	(0.51)	(0.88)	(1.12)	(5.47)	(0.51)	(0.88)	(1.12)	(5.47)
Face	*Happy*	6.38	5.92	6.58	6.29	0.92	0.86	1.00	0.93	4.76	5.06	2.89	4.24
		(.21)	(.20)	(.21)	(.12)	(.04)	(.04)	(.04)	(.02)	(.35)	(.35)	(.36)	(.20)
	Sad	5.56	5.94	6.10	5.87	0.79	0.77	0.82	0.79	5.79	4.62	3.15	4.52
		(.26)	(.25)	(.26)	(.15)	(.05)	(.05)	(.05)	(.03)	(.29)	(.29)	(.30)	(.17)
	Angry	5.14	4.87	4.23	4.75	0.40	0.39	0.68	0.49	5.83	4.36	3.44	4.54
		(.31)	(.31)	(.31)	(.18)	(.06)	(.06)	(.06)	(.03)	(.39)	(.39)	(.39)	(.22)
	Fear	5.30	4.91	6.11	5.44	0.40	0.45	0.66	0.51	5.92	3.91	2.91	4.24
		(.33)	(.33)	(.34)	(.19)	(.07)	(.07)	(.07)	(.04)	(.31)	(.30)	(.31)	(.18)
	Surprise	5.59	5.09	6.59	6.03	0.40	0.50	0.84	0.58	5.21	4.40	3.02	4.21
		(.23)	(.22)	(.23)	(.13)	(.07)	(.06)	(.07)	(.04)	(.41)	(.30)	(.41)	(.23)

		Means Intensity (*SE*)				Mean Accuracy (*SE*)				Mean Choice RTs (*SE*)			
		Young Children	**Old Children**	**Young Adults**	**All**	**Young Children**	**Old Children**	**Young Adults**	**All**	**Young Children**	**Old Children**	**Young Adults**	**All**
	Disgust	4.04	4.39	5.79	4.74	0.14	0.16	0.62	0.30	5.37	4.84	3.28	4.50
		(.37)	(.36)	(.37)	(.21)	(.05)	(.05)	(.05)	(.03)	(.30)	(.30)	(.31)	(.18)
	All	5.33	5.32	5.90	5.52	0.51	0.52	0.77	0.60	5.48	4.53	3.11	4.38
		(.16)	(.16)	(.16)	(.09)	(.03)	(.02)	(.03)	(.01)	(.20)	(.19)	(.20)	(.13)
Voice	*Happy*	5.35	5.79	6.40	5.85	0.58	0.75	0.91	0.75	6.45	4.29	3.43	4.72
		(.29)	(.28)	(.29)	(.16)	(.06)	(.06)	(.06)	(.04)	(.35)	(.34)	(.35)	(.20)
	Sad	5.25	5.38	6.09	5.57	0.29	0.52	0.87	0.56	7.27	6.00	3.92	5.71
		(.28)	(.28)	(.28)	(.16)	(.06)	(.06)	(.06)	(.04)	(.44)	(.44)	(.45)	(.26)
	Angry	3.58	3.26	4.08	3.64	0.27	0.23	0.64	0.38	5.97	5.22	3.95	5.05
		(.27)	(.26)	(.27)	(.15)	(.06)	(.05)	(.06)	(.03)	(.37)	(.36)	(.38)	(.21)
	Fear	4.78	4.95	5.59	5.11	0.19	0.38	0.50	0.36	6.28	6.06	3.63	5.32
		(.33)	(.33)	(.34)	(.19)	(.06)	(.06)	(.06)	(.04)	(.43)	(.42)	(.43)	(.25)
	Surprise	4.35	4.75	5.30	4.80	0.15	0.27	0.53	0.31	6.49	4.41	3.09	4.66
		(.32)	(.32)	(.33)	(.19)	(.06)	(.06)	(.06)	(.03)	(.38)	(.38)	(.39)	(.22)
	Disgust	4.49	4.63	4.97	4.70	0.29	0.43	0.77	0.50	6.70	5.41	3.87	5.32
		(.32)	(.31)	(.32)	(.18)	(.06)	(.06)	(.06)	(.03)	(.40)	(.40)	(.41)	(.23)
	All	4.63	4.79	5.41	4.94	.030	0.43	0.70	0.48	6.52	5.22	3.65	5.13
		(.16)	(.16)	(.16)	(.09)	(.03)	(.03)	(.03)	(.02)	(.26)	(.26)	(.08)	(.15)
Overall	*Happy*	5.87	5.85	6.49	6.07	0.75	0.80	0.96	0.84	5.60	4.67	3.16	4.48
		(.19)	(.18)	(.19)	(.11)	(.04)	(.04)	(.04)	(.02)	(.26)	(.26)	(.27)	(.15)

	Means Intensity (*SE*)				Mean Accuracy (*SE*)				Mean Choice RTs (*SE*)			
	Young Children	**Old Children**	**Young Adults**	**All**	**Young Children**	**Old Children**	**Young Adults**	**All**	**Young Children**	**Old Children**	**Young Adults**	**All**
Sad	5.40	5.66	6.09	5.72	0.54	0.64	0.84	0.67	6.53	5.29	3.63	5.12
	(.21)	(.21)	(.21)	(.12)	(.05)	(.04)	(.05)	(.03)	(.30)	(.29)	(.27)	(.17)
Angry	4.36	4.07	4.16	4.19	0.34	0.31	0.66	0.44	5.09	4.79	3.70	4.79
	(.23)	(.23)	(.24)	(.13)	(.04)	(.04)	(.04)	(.02)	(.33)	(.32)	(.33)	(.19)
Fear	5.04	4.93	5.85	5.27	0.30	0.41	0.58	0.43	6.10	4.98	3.27	4.78
	(.28)	(.27)	(.28)	(.16)	(.05)	(.05)	(.05)	(.03)	(.27)	(.27)	(.28)	(.16)
Surprise	4.97	5.33	5.95	5.41	0.28	0.38	0.68	0.45	5.85	4.41	3.05	4.44
	(.22)	(.22)	(.23)	(.13)	(.05)	(.05)	(.05)	(.03)	(.32)	(.32)	(.32)	(.19)
Disgust	4.26	4.51	5.38	4.72	0.21	0.30	0.69	0.40	6.03	5.12	3.57	4.91
	(.28)	(.28)	(.29)	(.16)	(.04)	(.04)	(.04)	(.02)	(.29)	(.28)	(.29)	(.17)
All Female	4.76	5.09	5.65	5.17	-	-	-	-	5.97	4.86	3.41	4.76
stimuli	(.18)	(.18)	(.19)	(.11)					(.24)	(.23)	(.24)	(.13)
All Male	5.21	5.03	5.56	5.30	-	-	-	-	6.03	4.90	3.35	4.75
Stimuli	(.15)	(.14)	(.15)	(.08)					(.22)	(.21)	(.22)	(.14)
Grand	4.98	5.06	5.65	5.23	0.40	0.48	0.74	0.54	6.00	4.88	3.38	4.76
Means**	(.15)	(.14)	(.15)	(.08)	(.02)	(.02)	(.02)	(.01)	(.31)	(.18)	(.07)	(.12)

Note. * = Mean Age in years. ** = Average scores across all participants and conditions.

a) Intensity Ratings

Results from intensity ratings suggested that emotional intensity perception depended on modality, emotion category and age of participants. A repeated measures ANOVA of modality (2) x emotion (6) x gender stimulus (2) x age group (3) x gender of participants (2) revealed a significant main effect for modality, $F(1, 75) = 51.21$, $p < .001$, $\eta^2 = .41$, emotion, $F(5, 375) = 32.95$, $p < .001$, $\eta^2 = .31$ and for age group, $F(2, 75) = 6.35$, $p < .01$, $\eta^2 = .15$. Emotions in faces were rated with higher intensity than voices and Sidak post-hoc test for multiple comparisons revealed that participants rated happiness with highest intensity ratings which differed from all emotions, then followed by sadness which also different from all emotions, followed by ratings for surprise and fear which did not differ from each other and disgust, with anger receiving lowest overall intensity ratings, all at $p < .001$. Furthermore, whilst the two child groups did not differ from each other in their average intensity ratings, adults yielded higher average intensity ratings than either of the child groups. This pattern was true for both modalities, $p < .05$. There was no significant main effect for average intensity ratings for gender of the stimulus or gender of participants ($p > .05$). For *M* and *SE* see Table 3.1.

There was also a significant interaction effect between (i) modality and emotion, $F(5, 375) = 5.49$, $p < .001$, $\eta^2 = .07$ and between (ii) modality and gender of the stimuli, $F(1, 75) = 11.9$, $p < .01$, $\eta^2 = .14$. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate the interactions.

i: The significant modality effect with higher intensity ratings for faces compared to voices was true for happy, angry and surprised stimuli whilst the other three emotions of sadness, fear and disgust did not differ as a function of modality. For both modalities, happiness reached highest average intensity ratings and anger lowest average intensity ratings, which did not differ from disgust in the voice condition, $p < .05$.

ii: Male faces received ratings which were significantly higher than for female faces, suggesting rating differences in faces as a function of stimulus gender. Gender of the stimulus did not influence emotion intensity perception in the vocal condition, $p < .05$. In addition, male faces received highest intensity ratings which were significantly higher than for male voices, showing that the face over voice advantage was significant for male stimuli. For female stimuli, intensity perception did not vary as a function of modality. This is illustrated in Figure 3.1.

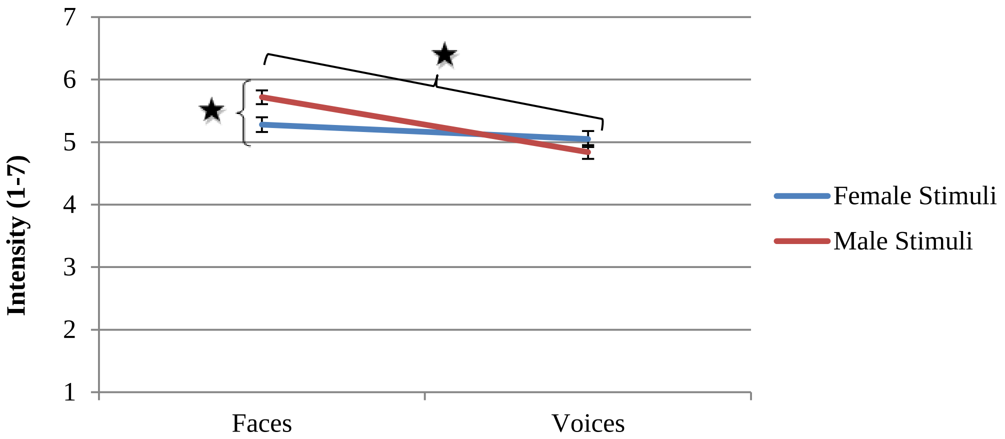


Figure 3.1. Means and Standard Errors for intensity ratings. Significant interaction between modality and gender of stimulus for all participants, $N = 81$, $*p < .05$.

There was also a significant three-way interaction between (i) modality, emotion and gender of the stimulus, $F(5, 375) = 12.19$, $p < .001$, $\eta^2 = .14$ and between (ii) modality, emotion and age group, $F(10, 375) = 2.36$, $p < .01$, $\eta^2 = .06$. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate the three-way interactions.

i: The gender effect in faces with male faces being perceived as more intense than female faces was true for angry stimuli. In addition, the gender stimulus effect in the voice condition with female voices being perceived as more intense than male voices was true for stimuli displaying anger and disgust, $p < .05$. For surprise, female voices were perceived as less intense than male voices. This is illustrated in Figure 3.2.

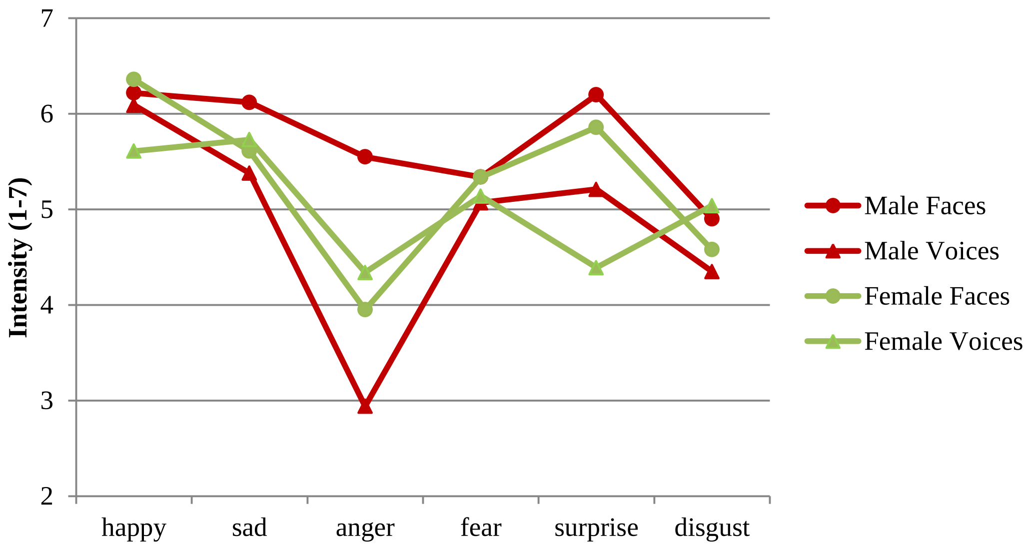


Figure 3.2. Means and Standard Errors for intensity ratings. Interaction between modality and gender of stimulus, dependent on emotion category for average intensity ratings, $p < .05$.

ii: The youngest age-group containing children aged 5 to 7 showed significant modality effects for intensity ratings with faces being perceived as more intense than voices for stimuli displaying happiness, anger and surprise. The older child group showed the same modality pattern for anger and surprise, $p < .05$ whilst the adult control group showed modality effects for stimuli displaying surprise and disgust, $p < .05$. In addition, all three-age groups showed lowest overall intensity ratings for anger in the voice condition and lowest ratings for anger or disgust in the face condition. Happiness received highest overall ratings for faces in both child conditions and for voices in the older child and in the adult group, $p < .05$. This is illustrated in Figure 3.3.

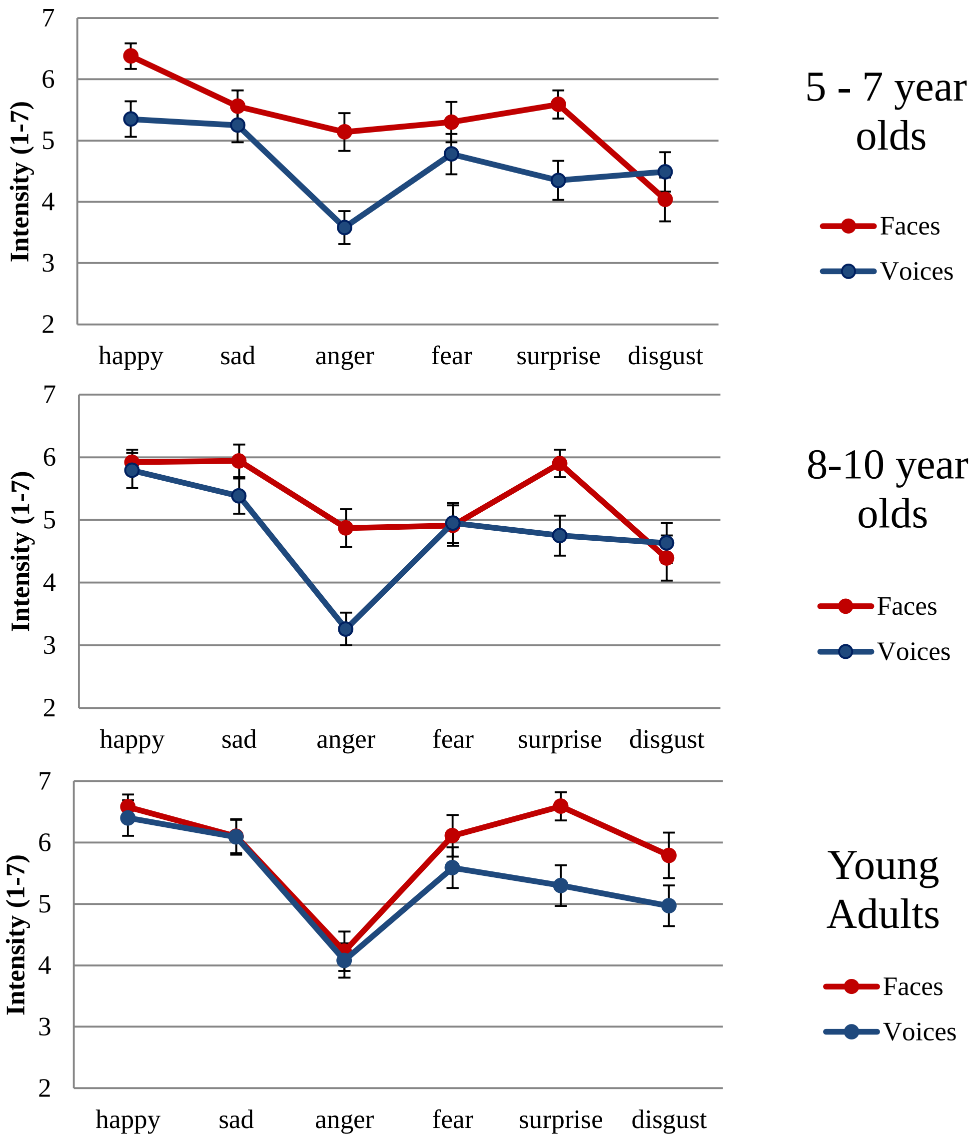


Figure 3.3. Means and Standard Errors for intensity ratings. Interaction between age and modality, dependent on emotion category for intensity ratings.

b) Accuracy Ratings

In order to support results from intensity ratings, accuracy scores suggested that correct recognition rate depended on modality, the specific emotion category as well as age of participants. A repeated measures ANOVA of modality (2) x emotion (6) x age group (3) was conducted. Accuracy ratings were computed in order to demonstrate a link between intensity perception and recognition accuracy. Since the main focus of the current analysis remains on intensity perception and choice reaction times of stimuli, accuracy measures were collapsed across gender of stimuli and gender of participants.

The ANOVA revealed a significant main effect for modality, $F(1, 77) = 51.56$, $p < .001$, $\eta^2 = .4$, for emotion, $F(4.9, 377.55) = 55.99$, $p < .001$, $\eta^2 = .42$, as well as for age-group, $F(2, 77) = 66.89$, $p < .001$, $\eta^2 = .64$. Faces received higher mean accuracy rates than voices. Post-hoc Sidak test for multiple comparisons revealed that happiness received higher accuracy rates than any other emotion, followed by sadness which was lower than happiness but higher than all other emotions, followed by surprise, anger, fear and disgust which did not differ from each other in accuracy rates, all at $p < .05$. Lastly, the youngest age group received lower accuracy ratings than the older child group which in turn received lower accuracy ratings than the adult group, $p < .05$. For *M* and *SE* see Table 3.1.

There was also significant interaction effects between (i) modality and age-group, $F(2, 77) = 6.77$, $p < .005$, $\eta^2 = .15$, between (ii) emotion and age-group, $F(10, 385) = 2.18$, $p < .05$, $\eta^2 = .05$, and between (iii) modality and emotion, $F(4.92, 378.55) = 14.13$, $p < .001$, $\eta^2 = .16$. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate the interactions.

i: The modality effect with higher accuracy scores for faces than for voices was true for all three age-groups, $p < .05$. After splitting for modality, it becomes apparent that for faces, the younger and older child-group did not differ significantly from each other whilst the adults had higher mean accuracy scores than both child-groups, $p < .05$. For voices on the other hand, the younger child group had significantly lower scores than the older child group which in turn had lower scores than the adult group, $p < .05$. This is illustrated in Figure 3. 4.

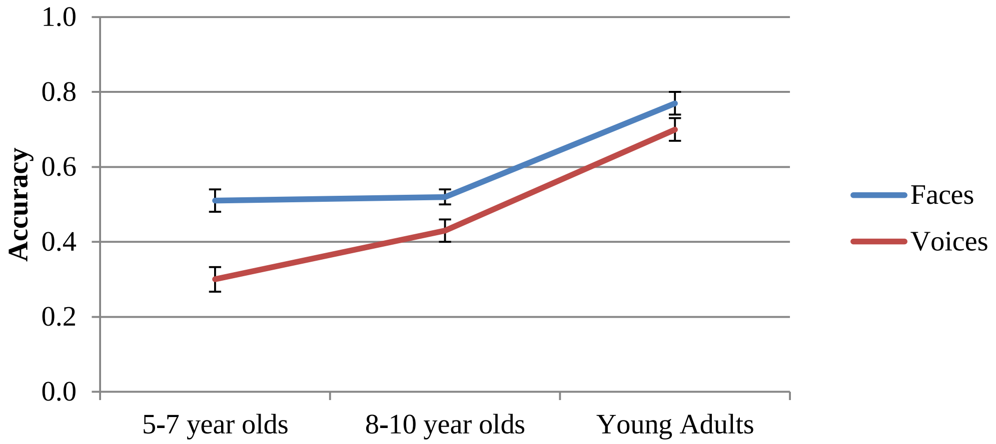


Figure 3.4. Means and Standard Errors for accuracy scores. Significant interaction between modality and age-groups, $p < .05$.

ii: For the youngest child-groups, accuracy scores were highest for happiness which was higher than for all other emotions. This was followed by sadness which had lower mean ratings than happiness but higher ratings than all other emotions. This again was followed by anger, fear and surprise which did not differ from each other. Disgust received lowest accuracy ratings which, however, were not significantly different from anger or surprise. For the older child group, accuracy ratings were highest for happiness, followed by sadness which was significantly different than for all other emotions. This was followed by fear, surprise, anger and disgust. Again, disgust received lowest ratings but this did not differ significantly from anger and surprise. For the adult group, happiness again received highest accuracy ratings, followed by sadness, disgust, surprise, anger and fear received lowest accuracy ratings but did not differ from each other significantly. Bonferroni post-hoc tests confirmed that all differences were significant at $p < .05$. For *M* and *SE* see table 3.1.

iii: The significant modality effect with higher accuracy scores for faces compared to voices was evident in all six emotions. For both modalities, happiness received highest accuracy scores which were higher than for all other emotions. For faces, this was followed by accuracy rates for sadness which differed from all others, then by surprise, fear and anger which did not differ significantly from each other. Disgust in faces received lowest accuracy rates which was lower than for all other face expressions. For voices on the other hand, ratings for happiness were followed by sadness and disgust which did not differ from each other and fear and surprise received lowest accuracy scores which did not differ significantly from each other. Bonferroni post-hoc tests confirmed that all differences were significant at $p < .05$. For *M* and *SE* see table 3.1.

Lastly, there was also a significant three-way interaction between modality, emotion and age-group, $F(10, 385) = 2.51$, $p < .05$, $\eta^2 = .06$. Post-hoc split file ANOVAs revealed that the significant modality effect with higher accuracy rates for faces than for voices within the youngest age group was true for all emotions but for disgust, where ratings were higher for voices than for faces. Anger did not show a significant modality effects. For the older child group, face over voice superiority was true for sadness, anger and surprise. For disgust, again, there was accuracy superiority for voices. For happiness and fear, accuracy ratings did not differ for modality. For the adult group, there was a modality effect for surprise with higher ratings for faces compared to voices, whilst the opposite was true for disgust with higher ratings for voices compared to faces. All other emotions did not show a significant modality effect. This is illustrated in Figure 3.5.

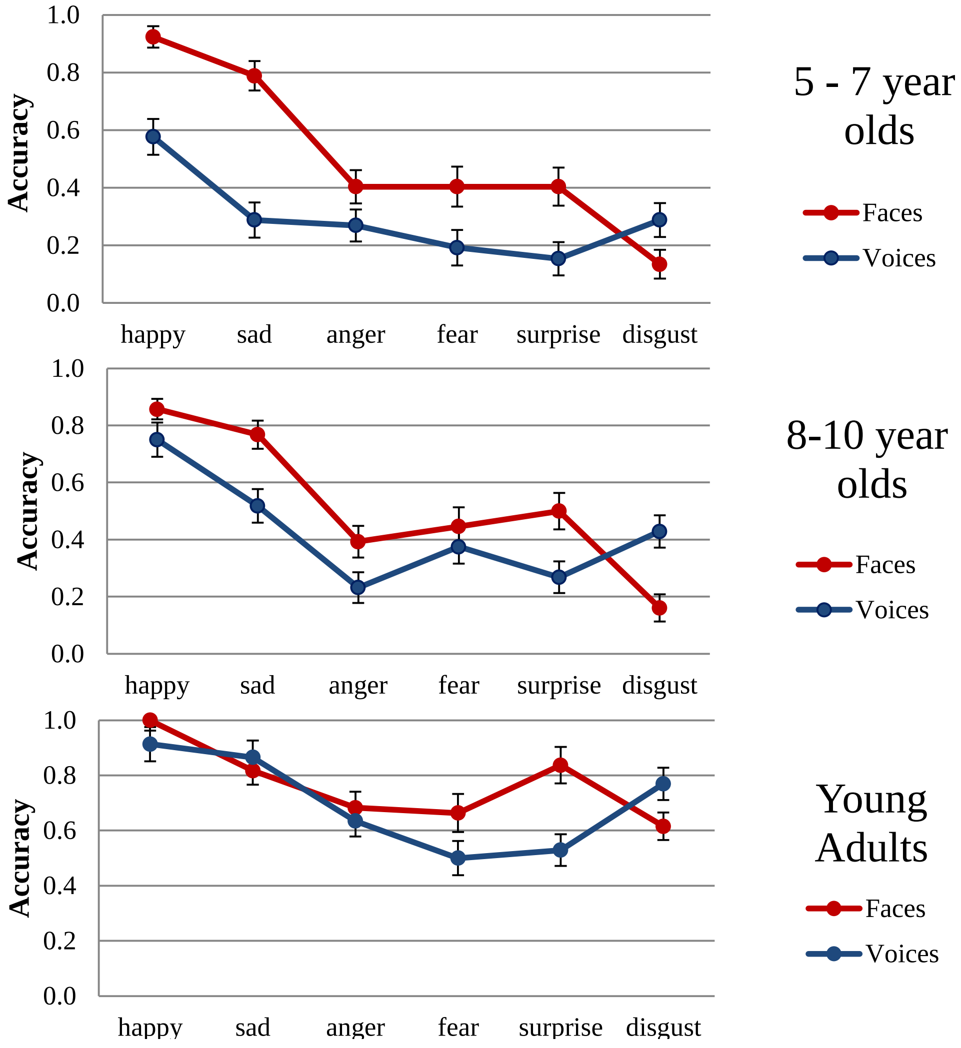


Figure 3.5. Means and Standard Errors for accuracy scores. Interaction between age and modality, dependent on emotion category, $p < .05$.

c) Choice RTs

Results from RTs suggested that speed of rating stimuli depended on modality, emotion category and age group of participants. A repeated measures ANOVA of modality (2) x emotion (6) x gender stimulus (2) x age group (3) x gender of participants (2) revealed a significant main effect for modality, $F(1, 75) = 44.77$, $p < .001$, $\eta^2 = .48$, emotion, $F(5, 375) = 3.88$, $p = .002$, $\eta^2 = .05$ and for age group, $F(2, 75) = 38.64$, $p < .001$, $\eta^2 = .51$. Choice RTs for faces were faster than for voices in all three age groups and Sidak post-hoc test for multiple comparisons revealed that whilst happiness and surprise were rated fastest, this did not differ significantly from anger or fear. This was followed by disgust which was significantly slower than happiness and surprise. Participants took longest to rate sadness which did not differ significantly from anger and disgust, all at $p < .05$. Furthermore, RTs increased linear with age, with younger children being significantly slower than older children who, in turn, were slower than the adult control group. This pattern was true for both modalities, $p < .05$. There was no significant main effect for gender of the stimulus or gender of participants ($p > .05$). For *M* and *SE* see Table 3.1.

There was also a significant interaction effect for RTs between modality and gender of the stimuli, $F(1, 75) = 24.97$, $p <.001$, $\eta^2 = .25$. Sidak post-hoc test for multiple comparisons revealed that participants also rated male faces significantly faster than female faces, suggesting differences in rating speed in faces as a function of stimulus gender. The opposite pattern was true for the voice condition, with female voices receiving faster ratings than male voices, $p < .05$. In addition, – in accordance with IR data - male face stimuli were rated fastest which was significantly faster than male voices. For female stimuli, rating speed did not vary as a function of modality. This is illustrated in Figure 3.6.

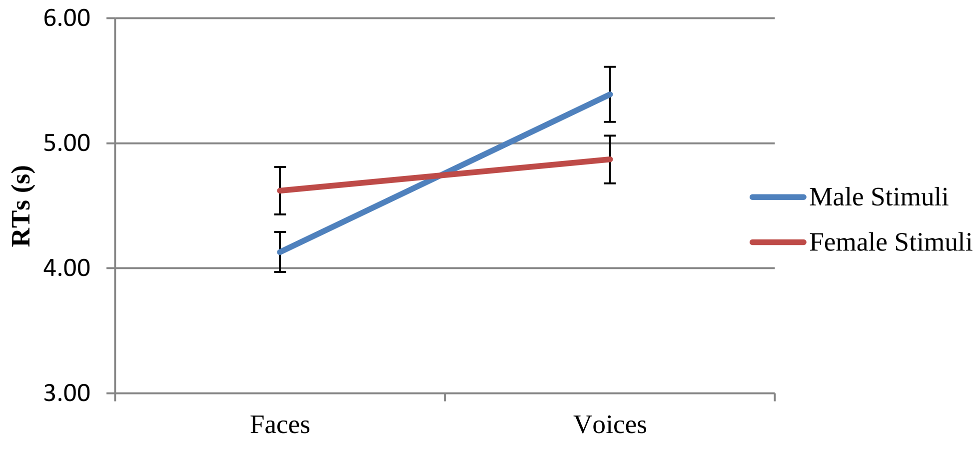


Figure 3.6. Means and Standard Errors for reaction times. Interaction between modality and gender of the stimulus, $p < .05$.

There were also significant three-way interactions between (i) modality, emotion and age-group, $F(10, 375) = 3.09$, $p = .001$, $\eta^2 = .08$, between (ii) age-group, modality and gender of the stimulus, $F(2, 75) = 3.13$, $p =.05$, $\eta^2 = .08$, and between (iii) age-group, modality and gender of participants, $F(2,75) = 4.03$, $p < .022$, $\eta^2 = .1$. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate the three-way interactions.

i: Post-hoc Sidak comparisons revealed that the youngest group containing children aged 5 to 7, showed faster RTs for faces than for voices across emotions displaying happiness, sadness, surprise and disgust. The older child group containing children aged 8 to 10 showed the same modality effect for stimuli displaying sadness, anger and fear, $p < .05$. The adult control group on the other hand did not show any modality effects for any of the emotions displayed ($p > .05$). When split for modality, the face condition revealed significant age-effects between both child-groups for sadness, anger and fear whilst speed for happiness, surprise and disgust did not differ between both child groups. For all emotions in the face but anger, were adults significantly faster than both child groups, all at $p < .05$. For the voice condition, speed for happy, sad surprise and disgust changed significantly between both child groups whilst anger and fear did not vary between both child groups. For all emotions displayed in the voice but happiness, were adults significantly faster than both child groups, $p < .05$. This is illustrated in Figure 3.7.

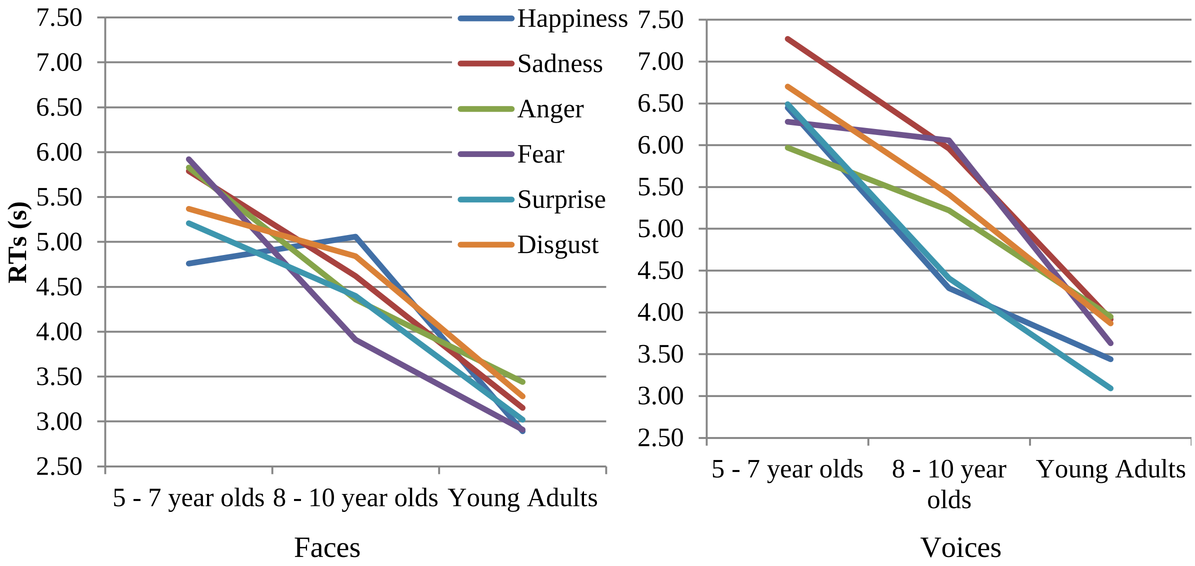


Figure 3.7. Means and Standard Errors for reaction times. Interaction between emotion category and age group for RTs, split for modality, $p < .05$.

ii: The interaction of age-group, modality and gender of the stimuli suggests that children in the youngest age group showed differences in RTs depending on the gender of the stimuli across both modalities. As illustrated in Figure 3.8., the younger children made decisions faster for male than female faces and slower ratings for male compared to female voices, $p < .05$. The older child group did not show gender stimuli effects for the face condition; however, for the voice condition, a significant gender of stimuli effect remained with female voices being rated faster than male voices, $p < .05$. For the adult control group, choice RTs did not differ depending on stimulus gender across both conditions, suggesting similar rating patterns of modalities independent of external factors such as gender of the stimulus ($p > .05$).

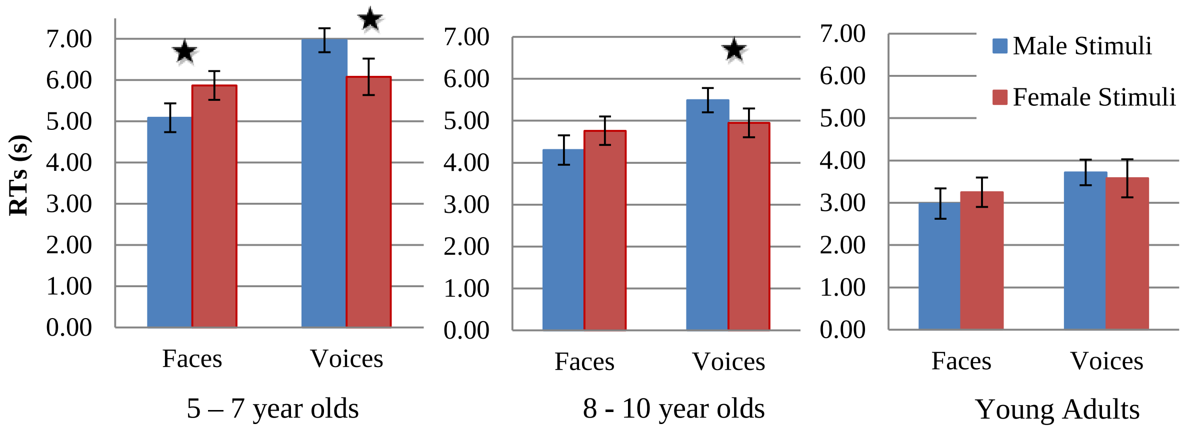
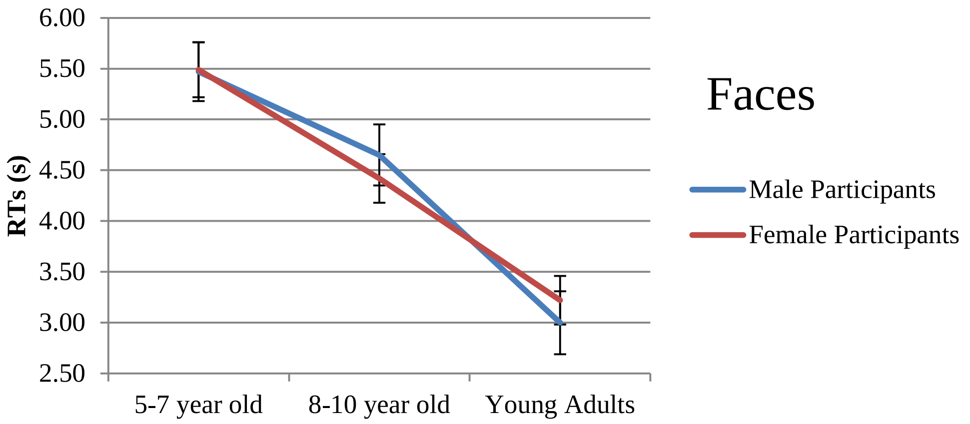


Figure 3.8. Means and Standard Errors for reaction times. Interaction between stimulus gender and modality, split for age. *p < .05.

iii: The interaction of age-group, modality and gender of participants revealed that the linear increase of choice speed - with the younger child-group taking longer than the older child-group which in turn took longer than the adult control group - was true for female participants across both modalities and for male participants within the face condition only. However, for male participants within the voice condition, the younger boys were just was quick as the older boys in rating voices, outperforming younger girls, $p < .05$. Figure 3.9 illustrates that this gender effect in the voice condition was eliminated in both older age-groups ($p > .05$).



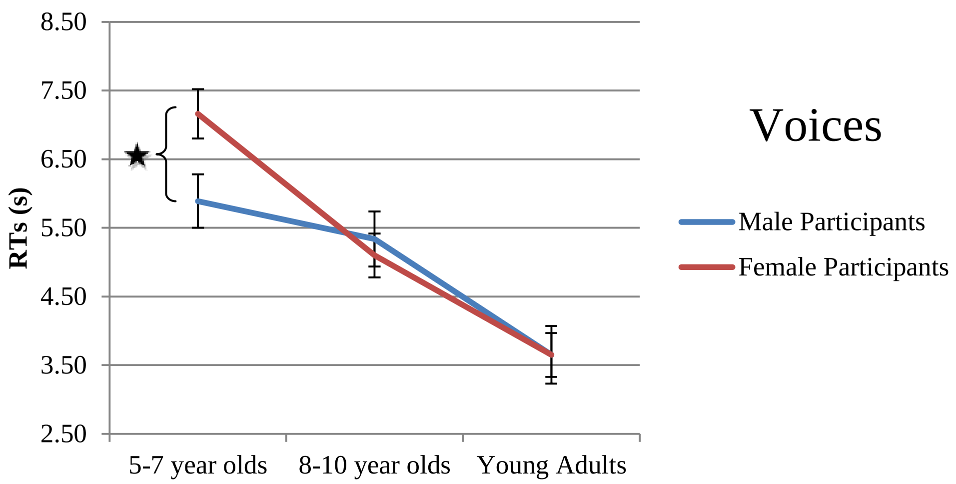


Figure 3.9. Means and Standard Errors for reaction times. Interaction between age and gender of participant, split for both modalities. *p < .05

3.1.3.2. Multiple Regression

For the child sample (*N* = 54), multiple regression analysis was used to test whether additional factors such as age, gender, non-verbal IQ, level of Theory of Mind or number of siblings influenced emotion recognition in voices and faces. Six separate regression analyses were conducted: two for intensity ratings for both modalities, two for accuracy scores for both modalities and two for RTs for both modalities. The following Table 3.2 summarises results and statistics for the multiple regression model. Across all six regression analyses, unsurprisingly, age correlated positively with the non-verbal IQ scores (r = .71, $p < .001$), with Theory of Mind Level 3 (r = .6. $p < .001$) and non-verbal IQ also correlated positively with having Theory of Mind Level 3(r = .49, $p < .001$).

a) Intensity Ratings

For face intensity ratings, all five predictors together only explained 18% of the variance and the overall model was not successful in predicting the emotional intensity ratings of faces (R^2 = .18, $F(8, 53)$ = 1.22, $p > .05$). However, two factors correlated with intensity perception of faces: Having two siblings as compared to none correlated negatively with the intensity perception of faces and having three or more siblings compared to none correlated positively with facial intensity ratings. The other variables did not contribute to the overall multiple regression model ($p > .05$).

For voice intensity ratings, all five predictors only explained 19% of the variance and the overall model was not successful in predicting the emotional intensity ratings of faces, $F(8, 53)$ = 1.34, R^2 = .19, $p > .05$). None of the individual factors correlated with the ratings of intensity in children ($p > .05$).

b) Accuracy Scores

For face accuracy scores, the overall multiple regression model was not successful in explaining the variance [$F(8, 53) = 1.22$, $R^2 = .18$, $p > .05$]. However, non-verbal IQ (CPM correlated positively with accuracy rates for rating face expressions. The other variables did not contribute to the overall multiple regression model, $p > .05$.

For voice accuracy scores, the overall multiple regression model predicted 37% of the variance. The overall model was successful in predicting accuracy rates for rating emotional voices, $F(8, 53) = 3.32$, $R^2 = .37$, $p < .005$. Age, non-verbal IQ (CPM), having 2 as compare to no siblings as well as having a ToM level 3 compared to level 1 all correlated positively with accuracy scores in the voice condition.

c) Reaction Times

For face RTs, all five factors explained 34% of the variance and the overall model was significant in predicting the choice RTs for rating faces , $F(8, 53) = 2.87$, $R^2 = .34$, $p = .01$. As expected, for children, age correlated negatively with rating speed as did having a Theory of Mind Level 3 as compared to Level 1. Having one sibling as compared to none correlated positively with speed for rating faces.

For voice RTs, all five factors explained 35% of the variance and the overall model was significant in predicting the choice RTs for rating voices, $F(8, 53) = 3.01$, $R^2 = .35$, $p < .01$. As expected, age correlated negatively with children's speed of rating voices and non-verbal IQ also correlated negatively with voice RTs. Further, having Theory of Mind Level 3 compared to Level 1 also correlated negatively with speed for rating voices. Finally, having one as compared to no siblings correlated positively with voice RTs.

Table 3.2.

Means, Standard Errors and Statistics for correlations of age, gender, non-verbal IQ, number of siblings and level of ToM with intensity ratings, accuracy scores and RTs for faces and voices.

		Age	Gender	Nonverbal IQ (CPM)	Siblings 1 vs 0	Siblings 2 vs 0	Siblings 3+ vs 0	ToM 2 vs 1	ToM 3 vs 1
Mean		7.52	1.57	27.48	.57	.20	.13	.15	.48
SE		1.72	0.50	6.21	0.50	0.41	0.34	0.36	0.50
INT Faces	Corr	0.02	0.03	-0.03	0.10	**-0.32****	0.24*	-0.10	-0.03
	B	0.08	0.22	-0.01	0.02	-0.57	0.56	-0.40	-0.30
	β	0.18	0.13	-0.06	0.01	-0.28	0.23	-0.17	-0.18
INT Voices	Corr	0.20	0.09	0.03	0.06	-0.10	0.20	-0.03	0.04
	B	0.25	0.24	-0.05	0.64	0.39	1.11	-0.15	-0.17
	β	0.50	0.14	-0.37	0.38	0.19	0.45	-0.06	-0.10
ACC Faces	Corr	0.14	0.21	**0.26***	-0.19	0.16	0.12	0.11	-0.04
	B	0.01	0.06	0.01	-0.08	-0.03	-0.01	-0.01	-0.08
	β	0.06	0.22	0.32	-0.29	-0.09	-0.03	-0.04	-0.29
ACC Voices	Corr	**0.44****	0.22	**0.50*****	-0.07	**0.28***	0.07	0.08	**0.29***
	B	0.02	0.03	0.01	0.13	0.19	0.14	0.04	0.04
	β	0.21	0.21	0.19	0.40	0.48	0.28	0.09	0.11
RT Faces	Corr	**-0.40*****	-0.07	-0.19	**0.23***	-0.11	-0.14	-0.15	**-0.37*****
	B	-0.31	-0.23	0.07	0.29	0.01	-0.09	-1.22	-0.98
	β	-0.43	-0.08	0.35	0.12	0.01	-0.02	-0.35	-0.40
RT Voices	Corr	**-0.48*****	0.11	**-0.31****	**0.30***	-0.14	-0.18	-0.18	**-0.27***
	B	-0.48	0.31	0.02	0.92	0.37	0.26	-1.38	-0.42
	β	-0.47	0.09	0.08	0.26	0.09	0.05	-0.42	-0.12

Note. INT = intensity, ACC = accuracy, RT = reaction times, Corr = correlation, * $p < .05$ ** $p < .01$ *** $p < .001$

3.1.3.3. Analysis 2: Confusion data

In order to investigate how the full emotion rating profiles of faces and voices change with age, the present chapter presents the confusions data from voices and faces. As illustrated in Figure 3.10, the mean intensity ratings for each of the possible 36 emotion-pairs seem to change with age and also the overall rating profiles for voices and faces seem to become more similar with age. Confusions such as between surprise and fear as evident in higher intensity ratings seem to exist in all age-groups and also across both modalities.

In order to statistically investigate whether the confusion matrices across two modalities become more similar with age or whether even the youngest children show high correlations between voice and face ratings, the full rating profiles for all displayed emotions in voices were correlated with the rating profiles for faces – individually for each participant. The individual correlation scores for each participant were then entered into a one-way ANOVA. Since correlations usually don't show normal distributions, the scores have been z-transformed to fulfil the basic assumptions of ANOVA.

A one-way ANOVA with age-group (3) revealed a significant main effect for age, $F(2, 79) = 33.31$, $p < .001$. Post-hoc Bonferroni correction for age-groups reported that adults had a significantly higher mean correlation for voices and faces compared to both child groups. As illustrated in Figure 3.11., the two child groups did not differ significantly in their mean correlation for voice and face ratings ($p > .05$). However, there was a trend and the mean correlation was higher for older compared to younger children.

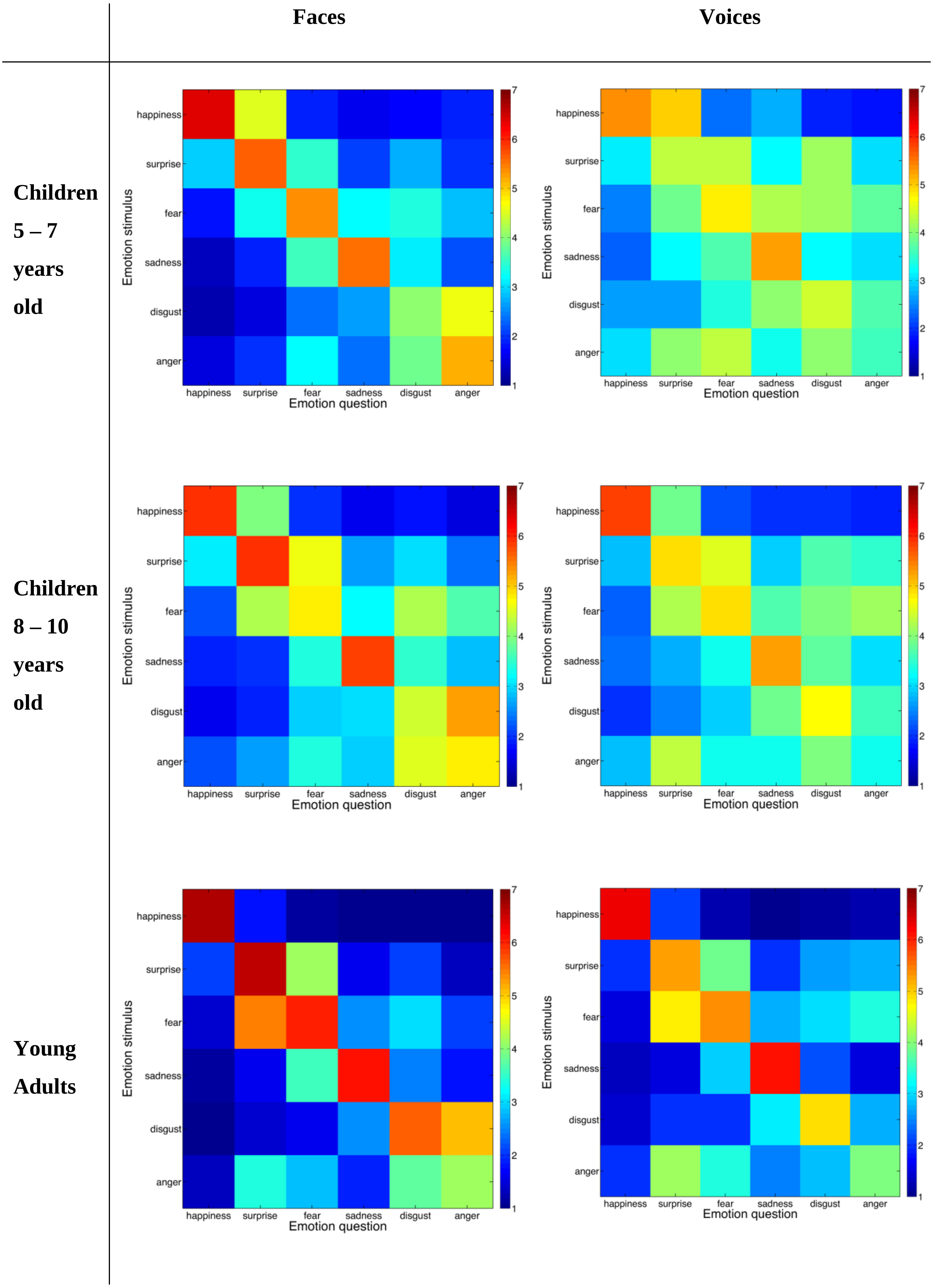


Figure 3.10. Illustration of mean intensity ratings (1-7) for each of the 36 possible emotion combinations, split for modality and age-groups.

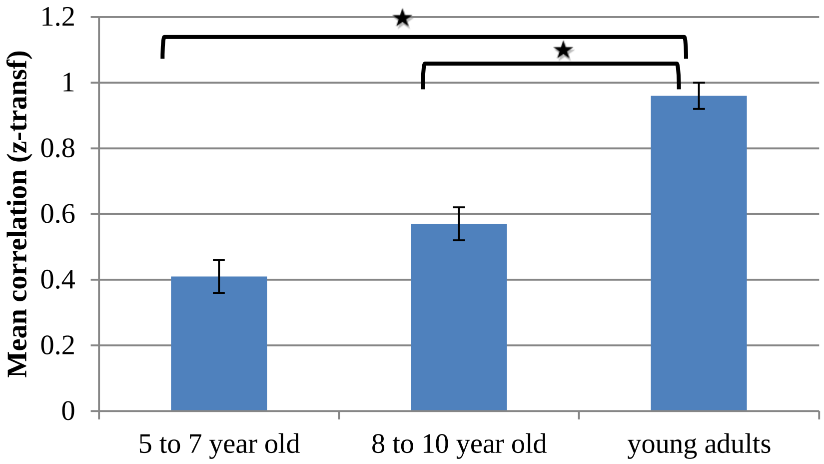


Figure 3.11. Mean score (z-transformed) for all individual face-voice correlations for each participant within each age-group, **p* < .05.

3.1.4. Discussion Experiment 3-I

Since previous studies delivered evidence for continuous emotion recognition development during early (Nelson & Russell, 2011) as well as late childhood (Tonks et al., 2007), the aim of the present experiment 3-I was to investigate how emotion recognition across two independent modalities develops in primary school aged children. For the first time, this comparison was based on contrasting static face expressions with non-verbal emotion exclamations. In a cross- sectional design, the developmental pathway of intensity perception, accuracy rates as well as choice reaction times for vocal and visual emotion perception has been observed in children aged 5 to 10 and in an adult control group. In connection to adult data from Chapter 2, it was of interest how similarities in emotion recognition across modalities develop in childhood. Further, additional variables that may contribute to individual variance such as children's non-verbal intelligence or Theory of Mind levels during emotion decoding - across vocal and visual domains – were investigated.

Results from the present study suggest that emotion recognition develops with age and even past the age of 10 as reflected in higher intensity judgements, higher accuracy scores and faster speed for adults as compared to children. This was true for both modalities, supporting Hypothesis 1. In addition, children from the current experiment rated faces with higher intensity, higher accuracy rates and with faster speed than voices, supporting Hypothesis 2 which assumed a face advantage in emotion rating. There were also significant interactions between modality and emotions and only certain emotions showed face superiority effects whilst for other emotions, ratings were independent of mode of presentation and showed similar patterns across modalities. Overall, the similarity in rating profiles as illustrated by confusion matrixes between faces and voices seemed to increase with age, suggesting a developmental effect of modality-independent emotion processing which supports Hypothesis 3. Furthermore, although neither non-verbal IQ nor Theory of Mind influenced intensity ratings of emotions, this was true for both modalities. For accuracy scores, non-verbal IQ was positively associated with accuracy and again, this was true for both modalities. For RTs of emotion ratings, age, having one as compared to no sibling as well as having Theory of Mind Level 3 as compared to Level 1 significantly correlated with emotion rating speed across both modalities which partly supports Hypothesis 4.

3.1.4.1. How does emotion recognition develop with age?

In line with neurological studies that suggested changes in brain structures such as increased frontal lobe activation across childhood and adolescence (*e.g*. Vink et al., 2014), the current study found significant behavioural age affects with older age-groups showing better emotion recognition skills than younger age-groups.

Accuracy ratings continued to improve across all three age-groups whilst for intensity ratings, no significant changes in perception from both types of modalities between the ages of 5 to 10 occurred. However, the adult group perceived emotions as significantly more intense than all children together, suggesting that intensity perception does not develop in a linear trend across childhood but increases sharply in teenage years. This indicates that some further development for emotion perception happens between the ages of 10 and adulthood. Those findings are not unexpected and have been supported individually for the face (De Sonneville et al, 2002; Vicari et al., 2000) and for the voice domain (Aguert et al., 2013).

The current study suggested that age-effects affected both modalities in parallel, indicating that developmental changes in brain structures may not only affect visual emotion recognition but could potentially be extended to the auditory domain. In line with neurological evidence, activity in children's temporal lobe has, for example, been associated with emotion recognition from faces (Cantalupo et al., 2013) as well as from voices (Cohen, Prather, Town, & Hynd, 1990). This could hint at similar underlying emotion recognition trajectories in childhood across two separate modalities. However, due to a lack of neuroimaging data, the current interpretation is of hypothetical nature.

Further, results demonstrated that for younger children, speed was influenced by external cues like gender of stimuli to categorise emotions in the voice as well as in the face condition. Emotion classification speed from the adult control group – as well as from the adult Chapter 2 - on the other hand across both modalities was not prompted by gender of stimuli at all. This goes in hand with evidence suggesting very early ERP signals around 200ms in adults for recognising emotional and non-emotional voices independent of individual characteristics of the speaker such as gender (Paulmann & Kotz, 2008). It is possible that young children's emotion recognition is guided by external features such as gender of actor more than it is for adults.

3.1.4.2. Does the degree of modality similarity change with age?

Although abilities to recognise emotions improved with age for both modalities, the degree of modality similarity also changed with age and depended on additional factors such as specific emotions expressed. Children in the younger child group seemed to perceive faces with higher intensities and also with higher accuracy rates than voices for a wider range of emotions. This discrepancy between face and voice was less evident in the older child group. Adults showed this pattern only for more complex emotions like surprise and disgust; all other emotions had comparable ratings irrespective of whether they were displayed in the visual or in the auditory domain. Findings from the current within-subject study go in hand with previous findings suggesting reliance on visual cues early in life whilst auditory emotion recognition starts to play a role in older children (Nelson & Russell, 2011).

From the confusion analysis, it becomes evident that the degree of similarity of rating profiles - as measured by comparing the individual correlations for each participant from voices and faces – changes as a function of age. The correlations between rating emotions in voices and faces became stronger with age and adults had significantly higher correlations compared to both child-groups. Overall, results indicate that whilst younger children may rely more on faces than on voices, ratings of target as well as of confused emotions become more similar between voices and faces the older children were. For the youngest child-group, is it is not likely that the way children rated emotional faces can reliably predict how children will rate emotions in voices.

3.1.4.3. Time course of specific emotion-recognition within each modality

Whilst there may be a common underlying emotion recognition mechanism as evident in similar recognition patterns across modalities, onsets of recognition of specific emotions seems to be moderated by modality. For faces, children’s decisions about sadness, anger and fear became faster with age, whilst happiness, surprise and disgust did not become faster over age and only demonstrated significant changes in RTs in adulthood. However, speed for rating happiness, surprise and disgust may have started off with a ceiling effect in children’s rating speed. For sadness, anger and fear on the other hand, rating speed caught up with happiness, surprise and disgust in the older child group. Intensity ratings did not change with age. It is possible that fast RTs reflect a genuine rating speed improvement in youngest children, independent of intensity perception. For voices, children’s decisions about happiness, sadness, surprise and disgust became faster with age whilst anger and fear did not become faster with age and only demonstrated changes in RT in adulthood – which was later than for the face condition.

Overall, threatening emotions such as anger and fear had slower RTs than non-threatening emotions and for the face condition, speed for those two emotions during childhood seemed to peak between the ages of 5- 7 and 8-10. For the voice condition on the other hand, the peak happened between the ages of 8-10 and adulthood. Hence, the overall rating pattern of emotions seems to be comparable between both modalities with happiness, surprise and disgust improving earlier than anger and fear; however, the onset for all emotions in the voice conditions seems to be one age-step behind the face condition, suggesting possibly difficulties for rating emotions displayed in the voices as compared to faces in younger children.

In addition, for accuracy rates in faces, the younger and older child-group did not differ which suggests that even younger children were equally equipped to categorise emotions as efficiently as older children. For voices on the other hand, younger children were significantly poorer at correctly recognising emotions compared to older children.

Again, present findings support the idea of earlier visual reliance during emotion recognition whilst auditory emotion recognition develops later in childhood children (Nelson & Russell, 2011). The ability to discriminate between acoustical information representing specific information may develop later in childhood, in line with Quam and Swingley's (2012) findings who suggests that children only start to make use of prosody information during middle childhood.

Although improvement occurred at different time-points depending on modalities, patterns of emotion improvement were comparable across modalities. One likely explanation for this finding is the possibility of the same brain mechanisms being responsible for emotion recognition across different modalities. This interpretation goes in hand with brain imaging data which, for example, suggested decreased medial prefrontal activation that was visible equally during visual and auditory emotion recognition tasks in children with Autism Spectrum Disorder (Wang et al., 2007). However, the story of supramodal emotion recognition in childhood undoubtedly is not as straightforward as this and future neuroimaging studies are needed to investigate the development of emotion areas in children across modalities.

3.1.4.4. Individual factors contributing to emotion recognition

The present study could not find gender effects in children for intensity ratings and speed of categorising emotions. This supports previous research on children's visual emotion recognition by Thomas and colleagues (2007) as well as on auditory emotion recognition (Sauter et al., 2012). Interestingly, the adult Chapter 2 did not find any significant effects of participants' gender on intensity ratings during emotion recognition either. In fact, the only occasion where gender of participants seemed to play a role was for the youngest child group in the voice condition. Whilst younger boys and older boys did not differ in their speed for rating voices, younger girls were significantly slower than boys in the same age range for rating voices. This difference was not evident in the face condition and seemed to have disappeared in the older child group, suggesting speed for rating voices and faces was at similar level - independent of gender of participants - by the time children were 8 years old. However, this interaction for RTs was not replicated for intensity ratings, suggesting slower responses in young girls but not lower intensity perception compared to young boys in the voice condition.

How could this gender difference be explained? Hypothetically, young girls may have found the voice condition more difficult than face condition, resulting in girls taking more time to make their judgements. Alternatively, young boys may just generally be faster in making judgements based on emotional voices than young girls. Interestingly, gender differences for changes in brain volume have been reported in children aged 6 to 17 (de Bellis et al, 2001). For example, de Bellis and colleagues (2001) reported a faster increase in white matter and corpus callosum volume in boys compared to girls between the ages of 6 to 17. This may reflect growth of neurons and myelination to enhance neuronal communication. Underlying anatomical brain development could be the reason for behavioural gender differences in the younger child sample; however, data from the present study only allows making hypothetical interpretations due to the lack of brain imaging data to support this idea.

Whilst results were less clear for intensity ratings of emotional stimuli, for accuracy rates, recognition accuracy for both modalities benefited from non-verbal IQ (CPM) abilities. Accuracy in voices additionally benefited from an advanced Theory of Mind and having two compared to no siblings. For choice speed, having an advanced Theory of Mind Level 3 – which correlated positively with age - contributed significantly to faster choice reaction times. Additionally, non-verbal IQ scores were also related to speed of rating emotions – but only for the vocal condition, suggesting IQ-independent RTs for the face but not for the voice modality. Interestingly, having one as compared to none siblings seemed to slow down the performance for rating speed - which was true for both modalities. At first, this direction of effect seems counterintuitive; however, a negative relationship between recognising fear in faces and number of siblings has also been reported by Felisberti, Chipi and Trueman (2014). One possible interpretation could be that sibling rivalry and jealousy in children (Volling, 2003) results in focusing on one's own emotions in order to gain most attention from parents rather than in trying to understand the feelings of siblings and other children. In contrast, it is possible that children with siblings have learned from the beginning to put effort and consideration (*i.e.* time) into understanding others whilst children with no siblings may make more rapid and less well considered other-judgements as their necessity to understand others is less developed.

Findings from the present study support previous findings by Ketelaars et al. (2010) and Weimer et al. (2012) who linked emotion understanding and ToM development – although only for the speed of rating emotions and not for accuracy or intensity perception. However, for the first time, the present data suggests that the link between emotion understanding and ToM abilities is not limited to visual emotion recognition and extends to the auditory domain as well. The dependence on ToM abilities for choice RTs across two separate modalities during emotion recognition hints at modality-independent cognitive requirements for rating speed of emotions. However, it is important to note that the present study used a very basic paradigm for assessing ToM abilities. In order to draw reliable conclusions about emotion understanding across modalities and ToM, one needs to include a more holistic ToM task as used for example in Ketelaars and colleagues' study (2010).

3.1.4.5. Implications from the current study

Results from the current study suggested that emotion recognition develops continuously across childhood and information from the vocal domain may only be recognised to a higher standard later in childhood. In addition, children seem to be especially equipped to recognise happiness across modalities from a very early age onwards and this pattern seemed to be similar to data from adults and becomes even more adult-like with age. Anger generally was not well recognised. Findings may have implications for children in primary or even secondary schools: Young children seem to be better at seeing than hearing what other people feel. This may be a helpful feature in noisy situations such as playground settings where children might not be able to hear what other peers say, but can see their expressions on faces rapidly and from far away. However, the present finding also indicates that settings where voice information is the only available source such as during radio advertising, young children may benefit from visual input; auditory emotion communication must be targeted at the appropriate age group and may not communicate the intended emotion to children under the age of 8 to 10. Present data suggests that young boys aged 5 to 7 might be more susceptible to vocal emotions than girls whilst the girls' performance catches up with boys' abilities around the age of 8 to 10.

In addition, whilst children with Autism Spectrum Disorder (ASD) and especially Asperger's Syndrome commonly show impaired emotion recognition abilities across faces as well as voices (Linder & Rosen, 2006), findings from the current study suggested that - although both modalities may have slight differences in onset and easiness - modalities showed parallel recognition patterns. This suggests that some intervention programs which were aimed at improving ASD patients' facial emotion recognition abilities such as by Baron-Cohen, Golan and Ashwin (2009) could potentially become extended to auditory emotion training in children as has been done previously in adults (Golan & Baron-Cohen, 2006).

3.1.4.6. Limitations & Future Research

The current study presented data on children's emotion recognition across two independent modalities in a within-subject design. It was found that speed of rating decisions increased in linear manner with age and so one could argue that faster speed may simple be a result of increased exposure to handling computers with age and underlying maturation of general executive functions. Whilst this is a likely interpretation, it is important to note that accuracy rates also increased in a linear manner with age. Further, all young children completed a training session before the main experiment to get used to pressing the buttons on the computer mouse as quickly as possible. The similar pattern across all three age groups with faster speed for happiness as compared to angry emotions suggests that speed in younger children is tied to specific emotion recognition rather than to more general and emotion-independent cognition.

The current study included children aged 5 to 10 which covers the range of primary school children.

However, present data as well as several previous studies (*e.g.* De Sonneville et al., 2009) suggested that developmental changes for speed of emotion recognition occurred even later in childhood and early adolescent years. Hence, it is advisable for future studies to also include older children in order to localise the peak of emotion rating abilities at which they become adult-like.

In addition, the present study delivered behavioural evidence about possible shared mechanisms across modalities which may develop with age but could be influenced by other external factors in childhood. However, to be best of the researcher's knowledge, no studies have investigated children's development of emotion recognition across several modalities from a neuroimaging perspective. It would be interesting to see whether children and adults employ the same brain regions for rating emotions delivered in faces as well as in voices which could hint at supramodal emotion recognition circuits which are either already existing in childhood or develop as a function of age and possibly emotion category.

3.1.4.7. Conclusion

For the first time, the current study provided data for children at primary-school age for emotion recognition across faces and non-verbal vocalisations in direct comparison. By including intensity, accuracy as well as choice reaction times measures from two separate modalities, the present study tested whether potential similarities in emotion recognition across modalities developed with age. It was demonstrated that whilst children were slower than adults and intensity perception as well as accuracy improved with age, this was true for both modalities. Further, there was a face-superiority effect which was evident for all age-groups although rating patterns for face and voice became more similar with age. This was also evident in increasing correlation-strength for full rating profiles of voices and faces which suggests that modality-independent rating of emotions is a feature of development. Whilst general patterns for rating emotions within the voice and the face in children represented adult-like emotion understanding such as highest ratings for happiness across all age and modality groups, this was not consistent. For example, young children's decision speed was influenced by gender of stimulus for both modalities but this effect disappeared with age. Further, young girls were slower in rating emotions than young boys but this was limited to the auditory domain and again, this effect disappeared with age. In addition, individual factors such as levels of ToM or having one sibling influenced the speed of rating faces and voices alike, suggesting general cognitive factors that influence emotion recognition on a supramodal level.

However, the present study provided data from a cross-sectional design; hence, one cannot draw conclusions about age-related changes between 5 to 7 and 8 to 10 year-olds that hold up in a longitudinal design.

By comparing children of different age-groups at only one time point increases the likelihood of measuring external differences such as family settings or classroom structure rather than true emotion understanding development. A longitudinal design could allow drawing conclusions about the relationship of cause and effect (Farrington, 1991), such as age and emotion understanding. Hence, the following Experiment 3-II will address this exact issue by re-testing a smaller sub-sample of 16 children with a time-gap of 18 months between the first and the second session.

3.2. Experiment 3-II

A Longitudinal Perspective on Children's Emotion Recognition in Faces: What Difference do 1.5 Years Make?

Since Experiment 3-I provided cross-sectional data of children's emotion rating abilities development across three different age–groups, the researcher returned to the same primary school 1.5 years later in order to collect some longitudinal face recognition data from a number of previously included participants. Children were in the younger age-group for 5 to 7 year-olds at Time 1 (autumn term 2012) and would have been in the older age-group for 8 to 10 year-olds at Time 2 (spring term 2014). The longitudinal data was collected as part of Experiment 3-III, comparing children's performance of rating emotional adult as compared to emotional child faces. Hence, due to the lack of a non-verbal vocalisation children database, Experiment 3-III only collected data from faces, providing Experiment 3-II with longitudinal data from the face conditions only.

Previous longitudinal data has suggested that emotion recognition abilities are changing with age and develop with experience. For example, Hills (2012) suggested that 6 to 10-year-old children's everyday experiences, such as exposure to school mates of the same age, over a period of three years, influenced the own-age bias for perceiving faces from the own or other age group. In addition, Bracovic, de Heering and Maurer (2012) reported that recognising child faces as compared to adult faces suddenly improved in young children once they had entered primary school which increased contact with other children of the same age. Further, emotion understanding abilities in 3-year-olds was related to the same children's emotion understanding three years later (Brown & Dunn, 1996). Developmental changes in facial emotion recognition may partially be due to gradual development of emotion-related structures within the developing brain between the ages of 7 and adulthood. For example, Mills and colleagues (2014) suggested that brain structures typically associated with mentalizing developed with age: grey matter volume as well as thickness for anterior temporal cortex peaked in early adulthood. In addition, Wierenge and colleagues (2014) demonstrated that the development of emotion related structures like the amygdala or hippocampus increased with age across childhood. However, as may be evident from the current literature review – longitudinal data on age-related changes on emotion recognition in voices is sparse.

3.2.1.1. The Present Study

Comparing two age groups in a longitudinal design seemed of particular interest since results from the previous cross-sectional study (Experiment 3-I) indicated that speed for making judgements increased significantly between both child groups. In addition, for RTs, the younger child-group relied on external cues such as gender of stimuli in both modalities whilst gender of the stimulus in faces did not prompt emotion rating speed for either the older child group or the adult control group. A clear direction on the involvement of external features like gender of stimuli and emotion recognition cannot be given from a cross-sectional study alone.

Consequently, a longitudinal study – as proposed here – may be able to identify developmental trajectories on the use of external features such as gender of stimuli during emotion recognition over a prolonged period of time. In order to exclude the possibility that developmental changes in emotion recognition are due to external or individual factors amongst participants such as differences in family structure or classroom setting, a smaller sub-sample of the original participant group was re-tested.

The present study predicts the following:

Hypothesis 1: Across a time-span of 1.5 years, children's emotion recognition will improve; this should reflect results from a larger sample in Experiment 3-I and may apply especially to RTs of rating emotions.

Hypothesis 2: External factors such as gender of the stimuli may play a critical role at children's emotion recognition at Time 1, but may disappear at Time 2, suggesting that emotion recognition may become more independent of external perceptual features with age as indicated by Experiment 3-I.

3.2.2. Methods

3.2.2.1. Participants

The 16 child participants (6 male, 10 female) in this present experiment were part of the original child sample (Experiment 3-I) and re-tested after 18 months on their ability to rate emotional faces, providing some longitudinal data for rating emotional adult faces. At Time 1 (autumn term 2012), children had a mean age of 7 years 2 months (*SE* = 1.14). At Time 2 (spring term 2014), children had a mean age of 8 years 8 months (*SE* = 1.1). When returning, the two groups differed in CPM non-verbal IQ scores which is a reflection of normal development (Raven et al., 1998, p. CPM55). Children did not receive feedback at time one in order to ensure they wouldn't remember the correct answers. At Time 1, children on average scored 28.13 *(SE* = .17) and at Time 2, 32 (*SE* = .85) out of 36 possible correct answers. Two out of 16 children scored lower at Time 2 than at Time 1.

The study was approved by the Department of Psychology, Brunel University, PsyREC officer (17/10/2012) and informed consent as well as a child-friendly version of the consent form was filled in both by participants and parents prior to the study. All participants were debriefed after the study. Forms are detached in Appendix B.

3.2.2.2. Research Material

Child and Adult Faces

See Experiment 3-I but containing visual face stimuli only.

Computer Program, Intelligence Assessment, Demographics

See Experiment 3-I.

3.2.2.3. Procedure

See Experiment 3-I.

3.2.2.4. Statistical Analysis

For the present study, data from matching target emotions were included. Two separate repeated-measures ANOVAs were conducted for Intensity Ratings and Reaction Times (RTs) comparing data from two time points. The repeated measures were time-point (2), emotion (6) and gender of stimulus (2). Since there were only 16 children included in the longitudinal data, no between-subject variables such as age or gender were included.

3.2.3. Results

a) Intensity Ratings

Due to the small samples size, the variable of gender or age of participants was not included in this analysis. A repeated measures ANOVA of time-point (2) x emotion (6) x gender stimulus (2) revealed a significant main effect for emotion, $F(5,75) = 8.96$, $p < .001$, $\eta^2 = .37$, as well as for gender of the stimulus, $F(1,14) = 7.11$, $p = .018$, $\eta^2 = .32$, but no effect of time points ($p > .05$). Post-hoc LSD comparisons suggested that for the subset of participants used in the longitudinal study ($N = 16$), happiness and surprise had highest intensity ratings whilst anger, fear and disgust received lowest intensity ratings, $p < .05$. Additionally, male stimuli ($M = 5.56$, $SE = 19$) were perceived as more intense than female ($M = 5.4$, $SE = .22$) stimuli.

The study also revealed a significant interaction effect for emotion and gender of the stimulus, $F(5,75) = 2.4$, $p = .045$, $\eta^2 = .14$. The significant main effect for gender of stimulus was only true for faces displaying anger with higher intensity perceptions for angry male faces ($M = 5.19$, $SE = .44$) as compared to angry female faces ($M = 3.59$, $SE = .44$). For all other emotions, gender of stimuli did not influence intensity perception. However, there was no significant main or interaction effect for the variable of time point, suggesting stable intensity ratings over time ($p > .05$). This is illustrated in Figure 3.12.

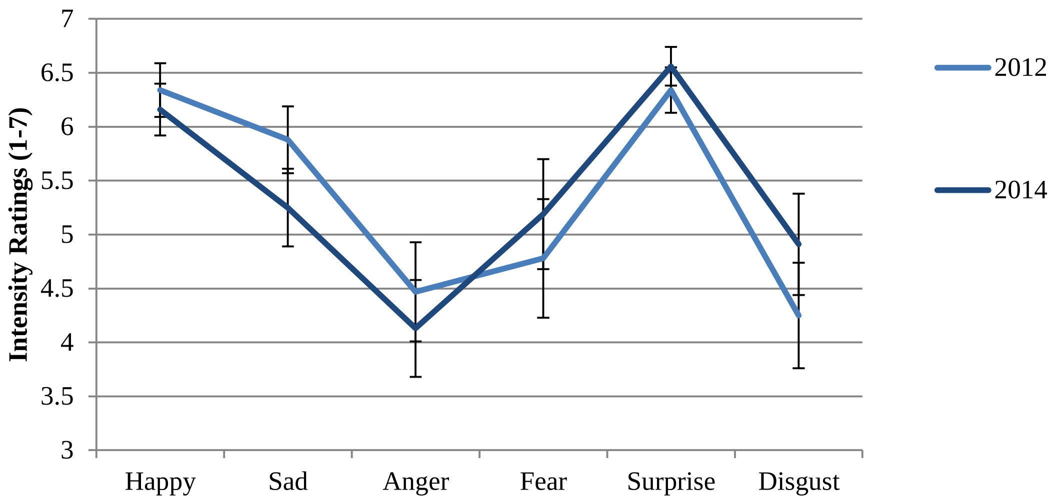


Figure 3.12. Means and Standard Errors for intensity ratings from the same participant group at two different time points (2012 and 2014) for each of the six emotions, $N = 16$. N.s.

b) Reaction Times (RTs)

Individual trials with RTs of 20 seconds and above were classified as outliers and replaced with the mean (*N* trials = 3). Due to the small samples size, the variable of gender or age of participants was not included in this analysis.

A repeated measure ANOVA of time-point (2) x emotion (6) x gender stimulus (2) revealed no significant main effects for either time - point, emotion category or gender of the stimulus ($p >$ .05). Although there was a trend of rating stimuli faster at Time 2 than at Time 1, this was not a significant difference and speed of rating emotions was stable over 1.5 years ($p > .05$). This is illustrated in Figure 3.13.

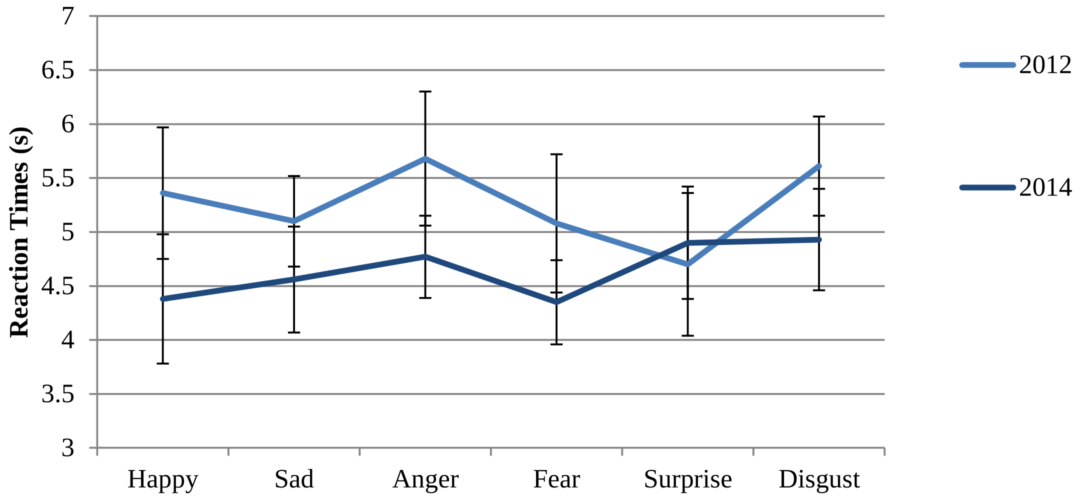


Figure 3.13. Means and Standard Errors for Reaction times from the same participant group at two different time points (2012 and 2014) for each of the six emotions, $N = 16$, n.s.

In fact, the only significant interaction effect for this cohort of participants for RTs was between time point, emotion and gender stimulus, $F(5,75) = 2.51$, $p = .034$, $\eta^2 = .14$. Opposite to predictions, post-hoc LSD comparisons revealed that whilst children at Time 1 did not show any stimulus-gender effects for RTs across the six emotions, at Time 2, children now showed stimulus-gender effects in speed for rating happy and disgusted faces: at Time 2, male faces were rated faster than female faces for happy and disgusted stimuli. In addition, children rated happy stimuli faster at Time 2 than at Time 1, but only if they were expressed by male actors. Further, children rated angry stimuli faster at Time 2 than at Time 1 when they were expressed by female actors, $p < .05$. This is illustrated in Figure 3.14.

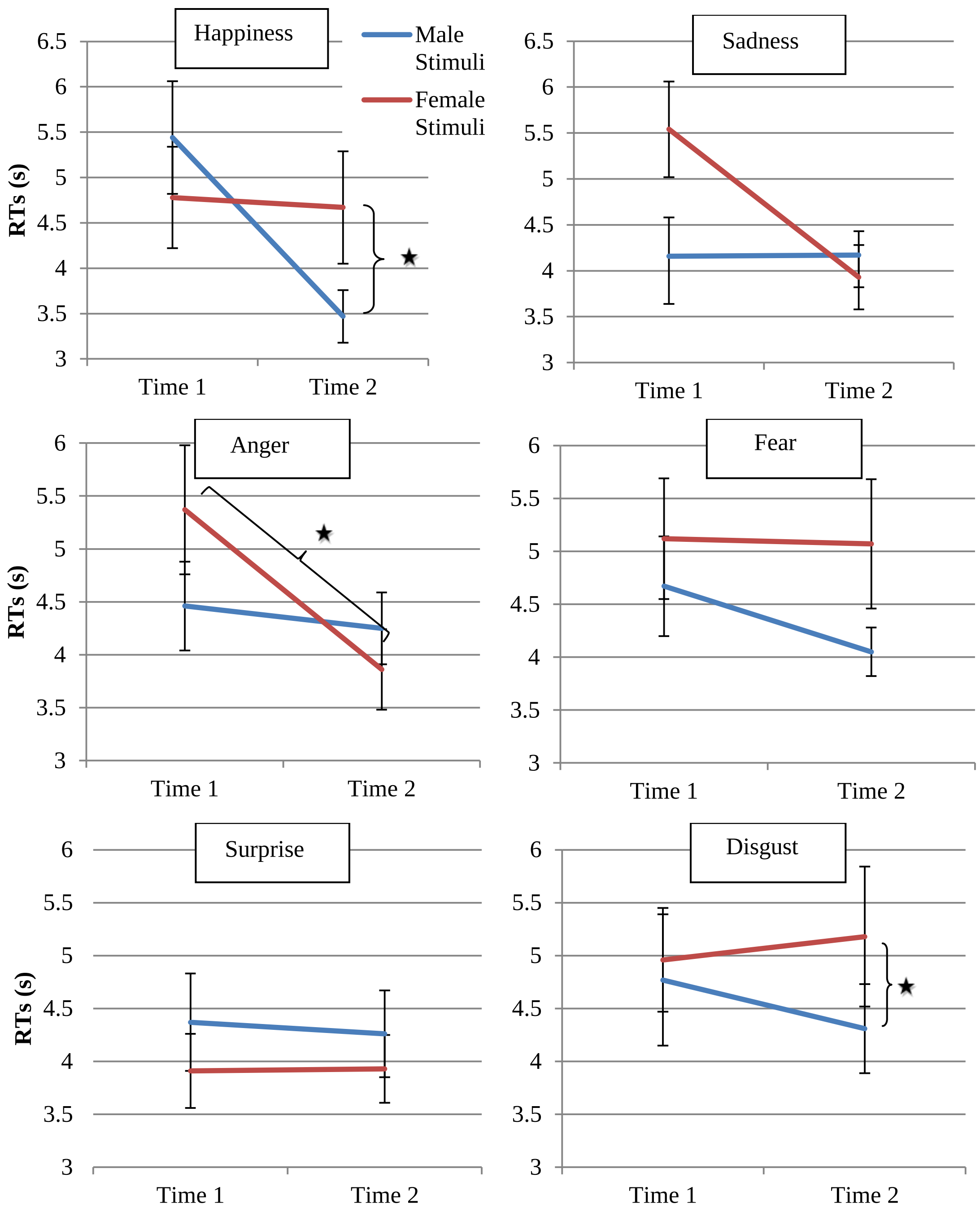


Figure 3.14. Means and Standard Errors for reaction times. Significant interaction effects between time of experiment and gender of stimulus, split by each of the six emotions. *$p < .05$.

3.2.4. Discussion Experiment 3-II

The present study on longitudinal age effects for emotion recognition within the visual domain aimed to replicate developmental effects from the cross-sectional Experiment 3-I within a sub-sample of child participants. The current study reported children's emotion recognition abilities with a mean age of 7 years 2 months at Time 1 and were re-tested 18 months later at a mean age of 8 years 8 months. By providing longitudinal data, the current study aimed to validate findings from the cross-sectional experiment by reducing external variables and pinning down developmental effects of emotion recognition across a time span of 1.5 years. Data suggests that RTs as well as intensity ratings for the sub-sample used in Experiment 3-II were stable between Time 1 and Time 2. Results do not support either of the two hypotheses in relation to developmental changes in RTs or intensity ratings.

3.2.4.1. Age effects

Results for intensity ratings from the longitudinal study somewhat reflect the findings from the cross-sectional Experiment 3-I. Neither the cross-sectional (Experiment 3-I) nor the present longitudinal experiment reported age-related effects for intensity perception between the two age-groups. Data from the adult control sample in Experiment 3-I, however, indicates that intensity ratings increase between the age of 9 to10 and adulthood and important developmental changes may happen during teenage years. Findings from Experiment 3-I and 3-II equally support the idea that intensity perception of emotion in faces may require a higher degree of cognitive flexibility; hence, sensitivity to emotion expression may change in early adulthood which is not captured in the present study. However, the reader is to remember that the actual intensity of faces was not modified for the purpose of the current study.

For RTs data, results from the current experiment did not replicate findings from the cross-sectional Experiment 3-I. Time of data collection did not influence RTs, although there was a trend towards faster RTs at the second data collection time. A relationship between age and speed is typically a common finding in emotion literature and has often been reported (*e.g.* de Sonneville et al., 2009). It is possible that the sample size for Experiment 3-II was not sufficient and including a larger sample of participants may result in a significant difference of speed over time. Another potential explanation is that the children included in the present experiment were younger (M = 8.1) than the average child in age-group II in Experiment 3-I (M = 8.96). It is possible that age-related effects for speed, which resulted in faster RTs for older versus younger children in Experiment 3-I, is not seen in the older child group at Time 2 in Experiment 3-II because the age does not match completely.

However, the interesting fact remains: children did not improve in RTs between the ages of 6.5 and 8 in a longitudinal design, suggesting that speed-related changes may happen in stages which may be longer than the included 1.5 years period.

In support, Mills and colleagues (2014) conducted several MRI scans in children 2 years apart and reported changes in brain structures related to emotion mechanisms such as the anterior temporal cortex at a time gap of 2 years.

3.2.4.2. External Variable (Gender of Stimulus)

For the cross-sectional Experiment 3-I, results suggested heightened intensity perception in children for angry male versus female faces. Interestingly, the same finding was reported in the longitudinal study in the present Experiment 3-II. The male superiority effect in angry faces seemed to be a robust finding and also supports previous literature such as Becker and colleagues (2007) who reported a connection between increased masculinity in a male face and perceived angriness. Becker and colleagues suggested that 'the angry face naturally looks more masculine' (p. 189).

Further, the present Experiment 3-II aimed to demonstrate a link between stimulus gender-related effects and emotion rating speed in younger children which may change over the period 1.5 years. Experiment 3-I reported gender stimulus effect for RTs for faces in younger but not in older children or adults. This possibly indicate that younger children relied on external factors such as gender of stimulus for making faster decisions about emotional meaning whilst older children and adults blocked out external factors and focus on the emotions displayed per se. However, findings from the current experiment led to a rejection of this hypothesis because the sub-sample from the longitudinal study actually demonstrated the opposite effect. The present smaller sub-sample did not exhibit gender stimulus effect for faces at Time 1 but on return after 1.5 years, the same children now showed gender-stimulus effects for rating speed of happy and disgusted faces with male superiority effects.

There are two possible explanations for this finding. It is quite likely that the sub-sample used in the longitudinal data collection is too small and hence does not deliver significant statistical effects. Including a larger sample and observing development over a longer period of time may give more reliable conclusions about emotion recognition across childhood. Alternatively, it is also possible that children between the aged of 6 and 8 – as included in the present longitudinal study - become indeed more sensitive to gender effects in stimuli with age. It is possible that this effect was overshadowed in the cross-sectional Experiment 3-I due to a higher degree of variability and external influences between the two child age-groups. Current results suggest that further investigation of the link between emotion recognition and external factors such as stimulus gender during development is necessary, especially with a larger sample that may include a wider range of age-groups and adult control groups.

3.2.4.3. Limitations & Future Research

One drawback of the current longitudinal study is that the sample size was relatively small and the current study only included data from face – rather than voice - stimuli due to the study design requirements for Experiment 3-III which compared emotion recognition in child versus adult faces. For future investigations, it would be important to include a voice emotion recognition task in order to observe children's emotion recognition developments across two independent modalities over time.

The lack of changes in intensity ratings and RTs whilst rating adult faces in the given child sample over time may be the result of children's stable exposure to adults over a period of 1.5 years. However, when children enter primary school, they are suddenly surrounded by a large number of same- age peers. Young children between the ages of 7 to 10 make more friends than they break up with and relationships only become more stable with age (Berndt & Hoyle, 1985). Whilst adults such as parents or teachers are usually a more or less consistent and stable component in a child's life, young children may encounter many new social situations with same-aged peers. This may include becoming a new member in a football club or attending after-school clubs which requires the ability to adhere to a social group and efficiently read emotions of same-aged peers. In order to validate the usefulness of including adult faces to be rated by children, it would be useful to investigate children's ability to rate emotions expressed in children as well as in adults.

3.2.4.4. Conclusion

Overall, the current longitudinal data was not able to answer questions that resulted from the cross-sectional Experiment 3-I. However, it did replicate findings such as stable intensity perception in middle childhood and also suggested that important developmental changes in emotion perception in faces may happen after the age of 10. One must include a larger sample and possibly observe performance over a longer period of time in order to make a more qualified judgement regarding the development of emotion perception based on external factors such as gender of the stimuli. The following Experiment 3-III aims to investigate the contribution of including adult rather than child faces to the inability to report significant changes in emotion perception across 1.5 years.

3.3. Experiment 3-III

Children's Emotion Recognition in Faces and the Role of Age-Group Biasing

In Experiment 3-I and 3-II, stimuli produced by adult actors were judged by adults and children alike. However, several face recognition studies have reported in-group advantage for identifying familiar faces (Sporer, 2001) and emotional faces (Vogel, Monesson, & Scott, 2012) from the own race as well as from the own age group (Kuefner et al., 2008) as compared to out-group faces. This finding may be a result of increased contact with own-group members which in turn leads to expert recognition of individual in-group faces (Sporer, 2001). For example, adult teaching trainees may demonstrate enhanced recognition of child-faces despite the age difference as they are exposed to children on an every-day basis (Harrison & Hole, 2008). Enhanced ability to recognise others of similar age-groups may be related to enhanced temporal brain dynamics in younger adults for memorizing younger compared to older adult faces (Wiese, Schweinsberger, & Hansen, 2008). However, previous studies usually focused on adults' perception of faces from members of different age-groups. Interestingly, Anastasi and Rhodes (2005) added child participants to the equation and found that 5 to 8-year old children also showed enhanced identity recognition abilities for faces displayed by children of similar age, suggesting in-group age effects for adults and children alike. Furthermore, a longitudinal paradigm demonstrated age-related face identity recognition biases in children which changed over time, depending of own experiences with the particular age group in question (Hills, 2012).

However, only a limited number of studies have investigated adults' as well as children's age-biases towards *emotional* faces of either child or adult faces. For example, Ebner and Johnson (2009) suggested no interaction between the age of faces and emotions in younger and older adult participants. However, in terms of children, friendships amongst individuals are generally more fluid and younger children may change their friendship bonds within very short periods such as three weeks (Cairns, Leung, Buchanan, & Cairns, 1995). Children move in between friendships and are usually exposed to a varying number of peers which may enable them to rapidly recognise emotions from friends of the same age. Adults such as parents or teachers on the other hand usually play a more stable role in a child's life. This pattern will obviously change as a function of age during childhood - until adults are the new own-age in-group members. It is possible that the lack of age-effects for intensity ratings or RTs from the longitudinal Experiment 3-II may be a result of stable encounter of adult faces over a period of 1.5 years.

Possibly, there could have been age- related effects for rating child faces which may have an adaptive function during early childhood. Further, whilst other children are usually at a similar height to the child observer, adult faces are further away from a child face, possibly resulting in different viewing angles which in turn may influence the emotion expression observed by children. Hence, the present experiment aims to investigate children's perception of emotions in child versus adult faces in order to validate the appropriateness of using adult faces to examine children's developmental effects over age.

Apart from the NIMH-ChEFS database (Egger et al., 2011) or the Radboud Faces Database (Langner et al., 2010), the number of available child-face databases are limited. For this exact reason, Dalrymple, Gomex and Duchaine (2013) have developed and validated the Dartmouth Database of Children's Faces with pictures of Caucasian children aged 6 to 16, displaying eight emotions which had been recorded from five different angles. However, to the researcher's knowledge, this database has not previously been evaluated for the use in child participants nor has it been used explicitly to investigate children's own-age bias for recognising emotional face expressions.

3.3.1.1. The Present Study

The present study aimed to investigate these exact issues reported above as it is important to understand potential age-in-group effects in emotion processing in order to improve social interactions. Data that indicates that children may have heightened sensitivity to emotions which are expressed by one age-group but not by another could help explain behaviour observable in playground settings. For example, children might be more likely to copy peers of a similar age as they can identify better with them. Teaching children to understand emotions displayed by other children of a similar age group could encourage prosocial behaviour and may be used in an attempt to explain antisocial acts such as bullying.

Furthermore, in order to ensure that children's lower efficiency in rating emotional faces compared to adult participants –as found in the previous Experiment 3-I - is not simply due to the disadvantage of rating out-group faces, but rather represents a general developmental effect, the present study with a smaller sub-sample was conducted to investigate children's capabilities for either rating emotional adult or emotional child faces. Additionally, this present experiment will produce data that allows investigation of whether children show own-age biases for certain emotions when using the newly developed Dartmouth Database of Children's Faces. Although verbal databases for children speaking emotional sentences have been produced (DANVA2-CP, Nowicky & Duke, 1994), the lack of child databases for non-verbal emotional vocalisations limits the researcher to the investigation of emotional faces only.

The present experiment predicts the following:

Hypothesis 1: Children show a small in-group advantage for rating emotions from children of a similar age as compared to the adult database as a result of increased exposure associated with entering primary school.

Hypothesis 2: Similar patterns of emotion perception across both adult and child data-bases are also expected, demonstrating that children can reliably recognise emotions from faces of different age groups.

3.3.2. Methods

3.3.2.1. Participants

Participants in this present experiment were part of the original child sample (Experiment 3-I) and re-tested after 18 months on their ability to rate emotional faces. The longitudinal data – comparing children's performance on rating emotional adult faces across two time points - was presented in Experiment 3-II. At the same time, the returning 16 children (6 male, 10 female) also performed a rating task of child-faces displaying the six basic emotions. The current Experiment 3-III investigated children's ability to compare between emotional adult faces and new stimuli containing emotional child faces. At Time 2 (spring term 2014), children had a mean age of 8 years 8 months.

The study was approved by the Department of Psychology, Brunel University, PsyREC officer (17/10/2012) and informed consent as well as a child-friendly version of the consent form was filled in both by participants and parents prior to the study. All participants were debriefed after the study. Forms are attached in Appendix B.

3.3.2.2. Research Material

Child and Adult Faces

For details on adult face stimuli, see Experiment 3-I. Here, an additional child face condition was implemented and the condition displayed children's face expressions for happiness, sadness, anger, fear, surprise and disgust from one male (identity 87, 9 years old) and one female (identity 77, 10 years old) Caucasian child actor from the Dartmouth Database of Children's Faces (Dalrymple et al., 2013). Each of the 24 pictures was randomly presented in the middle of a computer screen. Each stimulus was rated on each of the six basic emotions (2 x 6 x 6 adult faces and 2 x 6 x 6 child faces) on a 7-point Likert-scale scale ranging from '*not happy/sad/... at all = 1'*, to '*extremely happy/sad/... = 7'*.

No definition of emotions was provided in order to avoid biasing the responses into any direction; however, all children were verbally tested on their understanding on each of the six basic emotions prior to the task to ensure they were familiar with the emotions displayed. Within each block, order of stimuli was randomised. Across participants, the order of the blocks (child faces or adult faces) was counterbalanced.

Computer Program, Intelligence Assessment, Demographics

See Experiment 3-I.

3.3.2.3. Procedure

See Experiment 3-I.

3.3.2.4. Statistical Analysis

See Experiment 3-I, but within-subject variable was ‘face-type’ (child versus adult face). Since there were only 16 children in the present experiment, participants were not divided into different age-groups but instead analysed as one sample. Data was collapsed across gender. Further, no adult control group was included in the present experiment.

3.3.3. Results

a) Intensity Ratings

A repeated measures ANOVA of face-type (2) x emotion (6) x gender stimulus (2) revealed a significant main effect for emotion, $F(5,70) = 4.35$, $p = .002$, $\eta^2 = .24$ but not for face type ($p > .05$). Post-hoc LSD comparisons revealed that happiness and surprise received highest intensity ratings whilst the other emotions did not differ from each other significantly, $p < .05$.

There was no significant main effect for face-type. Paired t-tests were performed for each of the six emotions between adult and child faces. Paired t-test revealed that for anger, child faces were perceived with higher intensity than adult faces, $t(15) = 2.11$, $p = .05$, whilst for surprise, adult faces were perceived with higher intensity than child faces $t(15) = -2.3$, $p < .05$. This is illustrated in Figure 3.15.

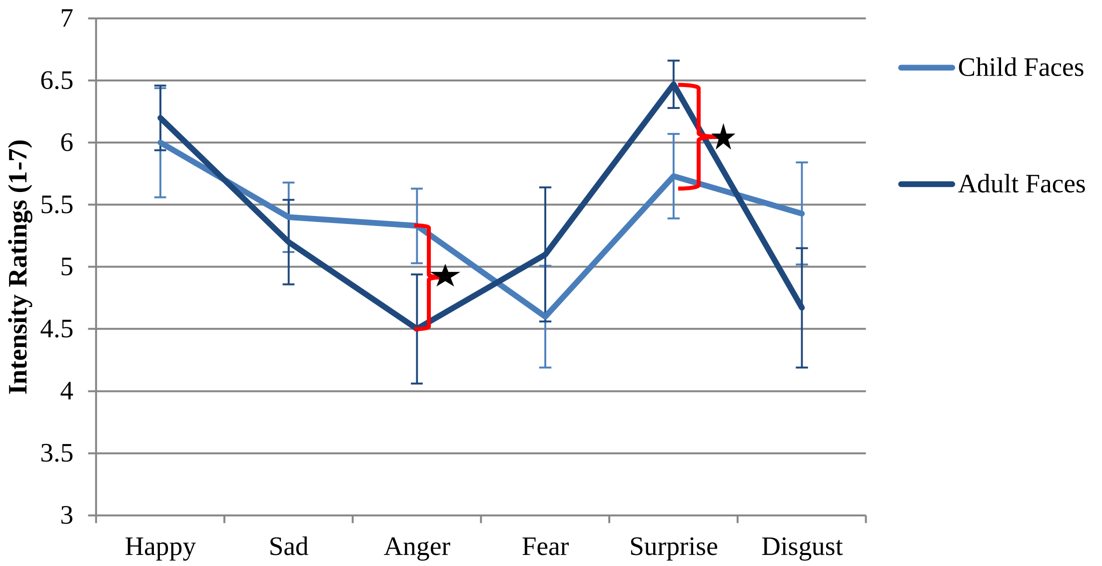


Figure 3.15. Means and Standard Errors for intensity ratings for the child as well as the adult face database for each of the six emotions, $N = 16$, $*p < .05$.

Further, there was a significant interaction effect between emotion and gender of the stimulus, $F(3.31, 28.34) = 7.78$, $\eta^2 = .36$. Post-hoc univariate comparisons revealed that for angry faces, male stimuli ($M = 6.07$, $SE = .27$) were perceived with almost double the intensity than female faces ($M = 3.77$, $SE = .34$) whilst for surprise, the opposite was true and female faces ($M = 6.53$, $SE = .2$) were rated with higher intensity than male faces ($M = 5.67$, $SE = .3$), $p < .05$.

b) Choice RTs

A repeated measures ANOVA of face type (2) x emotion (6) x gender stimulus (2) revealed no significant main or interaction effects for any of the three variables ($p > .05$), suggesting that faces from the child database were perceived with similar RTs than faces from the adult database, independent of all six emotions and gender displayed. This was confirmed by six individual t-tests for each of the six emotion pairs ($p > .05$). For illustration of rating similarity across both face databases, see Figure 3.16.

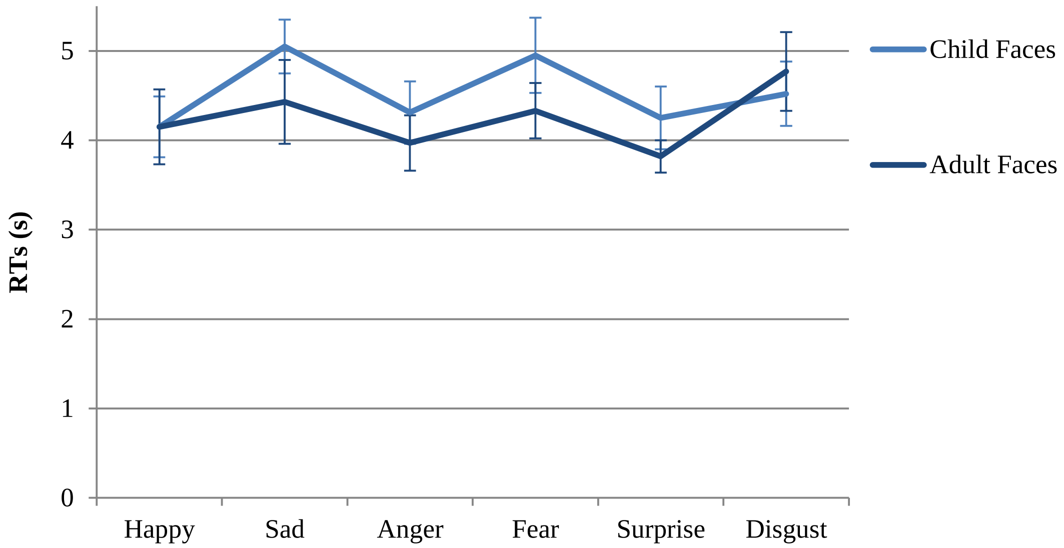


Figure 3.16. Means and Standard Errors for RTs for the child as well as the adult face database for each of the six emotions, $N = 16$, n.s.

3.3.4. Discussion Experiment 3-III

The current Experiment 3-III investigated whether children perceived emotions differently depending on whether faces are expressed by adult or by child actors. Additionally, the present study aimed to deliver pilot data for own-age bias of child participants for rating either adult faces or the newly developed Dartmouth Database of Children's Faces DDCF (Dalrymple et al., 2013). Whilst the present data suggests that adult faces are reliable stimuli to assess children's emotion perception because they did not differ significantly from child faces in RTs and intensity, results also implicate that children did not exhibit own-age bias towards emotional faces of their own age groups. The present findings support Hypothesis 1 partially, suggesting children's in-group age-bias towards higher intensity perception of angry faces when displayed by children as compared to adult actors. However, in contrast to Hypothesis 1, intensity ratings for surprised stimuli seemed to benefit from presentation by adult faces rather than child faces. In addition, results support Hypothesis 2, suggesting that children can recognise emotions from adult and child faces alike in a similar manner, independent of the age-group of actor.

3.3.4.1. Stimulus-Effect

Investigating the potential own age bias effect in child participants rating emotions from the DDCF, results suggested that children aged 7-10 did not rely on age information of the actor when judging emotions. Results could not demonstrate general age-related biases within the DDCF. Fluid and often changing relationships in younger children as described by Cairns et al (1995) did not seem to cause enhanced rating abilities of emotional faces in same-age peers – in direct contrast to research from face identity recognition in children (*e.g.* Anastasi & Rhodes, 2005). In addition, the angle of viewing emotional expressions that children are used to when talking to either children of the same height or taller adults did not seem to influence children's recognition abilities of emotional faces.

The present results are in line with Ebner and Johnson's (2009) findings who suggested that the age of adult participants did not interact with the age of the actors displaying emotional faces. It is possible that the lack of own-age bias in adults during the rating of emotional faces can be extended to children as well. However, when looking at the data in more detail, the present study found that for faces displaying surprise, adult faces were perceived as more intense than child faces. This possibly suggests that children had a higher degree of certainty when judging surprise in adult as compared to child faces. The current findings could indicate that children are wired to perceive surprise with higher intensity in adults than in children. This may be an important evolutionary feature for survival as it may be beneficial to be sensitive to the caregiver's facial expression in response to the onset of sudden events.

In turn, emotions expressing anger resulted in children's in-group age effect with heightened sensitivity towards angry children as compared to angry adults. Since angry stimuli displayed in Experiment 3-I received lowest intensity ratings across all child groups, it is possible that this is – at least partially – due to the fact that stimuli were displayed by adult actors. Had the researcher used child stimuli instead of adult stimuli, anger may not have yielded lowest scored for intensity ratings. In addition, children's in-group age bias towards angry child faces may be adaptive in nature. Children that have entered primary schools and have to acclimatize to new social groups and friendships that may change often may benefit from recognising anger in peers rapidly. It is possible that the average middle-class child is less often exposed to aggressive anger displayed by adults as compared to anger displayed by children in playground situations who have yet to learn how to regulate their feelings. Hence, recognising anger in peers or similar age-groups could help to avoid bullying situation or getting into arguments.

3.3.4.2. Limitations & Future Research

Infants' understanding of visual and auditory congruency of emotional stimuli seems to underlie the concept of own-race group biasing (Vogel et al., 2012) but little is known about the age in-group biasing of emotional voice processing in primary school children. The researcher is not aware of a child database for non-verbal affect vocalisations which would be needed in order to investigate children's own- age-group bias across modalities. Investigating age-group biases across two separate modalities during emotion perception in children could contribute to a deeper understanding of supramodal mechanisms facilitating emotion understanding. Furthermore, such a vocal child database could help to validate the use of adult voice samples as used within the previous child Experiment 3-I.

Since this current study was designed as a small pilot study for assessing children's own-age bias in emotional faces within a newly developed child face database, including a larger sample size could also allow controlling for gender effects on own-gender bias in face emotion recognition in children. For adults, this has been found to exist for general face recognition memory with females being better at recognising female faces, independent of attentional load (Loven, Herlitz, & Rehnman, 2011). It would be interesting to see whether female children also demonstrate enhanced processing mechanisms for emotional faces of the own-gender group.

In addition, it would be interesting to collect data from children rating child stimuli over a longer period of time which can then be compared to the longitudinal study in Experiment 3-II. The lack of developmental changes for RTs and intensity ratings for rating adult stimuli over a period of 18 months may be due to the stability of exposure to adults during childhood.

Ever changing exposure to different social networks during primary school may elicit more enhanced developmental effects for rating child faces in a longitudinal context which may not be visible for adult faces.

3.3.4.3. Conclusion

In conclusion, data from Experiment 3-III suggested that overall, children were equally able to rate child as well as adult faces and did not show age-related in group effects for stimuli displayed by actors of a similar age. This suggests that the stimuli used in Experiment 3-I are reliable and findings are not due to age-effects of adult faces. However, looking into data in more detail, this may depend on specific emotions and especially anger was perceived by children as more intense when displayed by children as compared to adults. In addition, to the best of the researcher's knowledge, the present study is the first to evaluate the newly developed DDCF (Dalrymple et al., 2013) for the use with child participants. Current results indicate that this particular database elicited proficient ratings in children which are comparable to adult stimuli. However, the lack of a non-verbal child data base on affect bursts limited this current study to the investigation of visual stimuli.

3.5.General Discussion

The main aim of this Chapter 3 was to investigate findings from the Chapter 2 within a developmental context in order to examine children's emotion recognition of two separate, yet parallel modalities. Data from adults in Chapter 2 showed similar patterns of emotion recognition across voices and faces but it is not clear whether this –potentially supramodal – emotion processing mechanisms already exists in similar manner in children. It is possible that the modality-independent emotion processing develops with age, experience and maturation of certain brain structures. Chapter 3 aimed to answer the questions of how similar children's emotion perception is to adults (*i.e*. how performance changes as a function of age) and whether children's emotions recognition shows parallel patterns in separate modalities. More specifically, the researcher tested the above issue in a cross-sectional as well as in a smaller longitudinal study and also evaluated the usefulness of adult stimuli to be judged by children.

Overall, results from Chapter 3 suggested that children showed superior emotion recognition for visual compared to auditory emotion recognition. Speed for rating emotional voices decreased at a later onset than for faces and accuracy for rating voices also improved later than for faces, depending on the specific emotions displayed. However, interestingly, children showed parallel patterns for emotion recognition across modalities with individual factors such as levels of Theory of Mind not only affecting speed for rating emotional faces bur also for emotional voices. Similarly, gender of participants did not influence emotion recognition in either voices or faces. This was also replicated the findings for intensity ratings in the adult data. Further, superior recognition for happy emotions was evident in both modalities, in all age groups, in the longitudinal as well as cross-sectional study and independent of the age of actors portraying the stimuli. Hence, certain patterns of children's emotion recognition seemed to be independent of the mode of presentation and reflected adult-like performance, maturing with age.

In addition, the present data found that children can rate emotional adult faces without presenting out-group age disadvantages. Results indicated similar patterns for rating individual emotions and did not show significant differences for either RTs or intensity ratings of adult or child faces across different emotions. Hence, the fact that children decode emotions less efficient than adults - is not due to children rating out-group faces from a different age group but rather exists due to more general developmental effects – at least within the visual condition. Across all three experiments, Experiment 3-I, 3-II and 3-III, in Chapter 3, three common themes emerged. Those are the superior emotion recognition from faces as compared to voices, enhanced recognition of happy stimuli as compared to threatening stimuli and age-related effects across modalities and will be discussed in the following section.

3.5.1. Happiness Superiority

Across modalities, age groups, longitudinal or cross-sectional designs and adult versus child faces, Chapter 3 consistently found superior emotion recognition for happy stimuli which was also replicated findings from the adult Chapter 2. Enhanced effects for happy faces may be a result of obvious configural features within the face such as showing teeth as well as open mouth during a smile (Schyns, et al., 2009). However, finding happiness superiority effect across independent modalities suggests that visual, perceptual features are not sufficient in explaining happiness superiority. Rather, modality independent cognitive processes may modulate emotion recognition which may already be deployed within a child's brain and may be robust to age.

In spite of this, it is important to note that the current Chapter 3 only examined one positive emotion (happiness) as compared to multiple negative emotions (such as anger, fear, sadness and disgust). Hence, critics may argue that advanced recognition of happy stimuli is due to lack of similar emotions. However, Sauter, et al. (2013) tested children's non-verbal vocalisation recognition of several positive as well as negative emotions and findings suggested that children still recognised positive emotions like amusement and relief better than negative non-verbal affect bursts. Hence, the current finding of happiness superiority effects may replicate a general enhancement for positive emotion recognition which may be independent of visual features within the voice or – at least – be accompanied by equally distinct acoustic features within the voice. For a more detailed discussion see Chapter 2.

Further, the categorical classification of the six basic emotions may be problematic as it is possible that emotion classification is based on the dimensions of valence and arousal rather than specific categories. It is possible that especially younger children base their emotion perception on valence-based judgements (Widen & Russell, 2010). The dimensional Circumplex Model of Affect (Posner, Pussell, & Peterson, 2005) suggests that different combinations of valence and arousal ratings which result in physiological responses describe emotions better than separate responses for each of the basic emotions. For example, 'calm' would be classified as low in arousal and high in pleasantness. Hence, the happiness superiority effect in children could be a result of the lack of sensitivity to similar negative emotions as children may base their judgements on a continuum such as 'good versus bad' rather than on individual categories.

3.5.2. Age Effects

The present Chapter 3 provided evidence for continuous changes in emotion recognition abilities across two separate modalities which were evident across childhood and continued beyond the ages of 10. This was especially true for speed and accuracy but was also somewhat related to the specific emotion category in questions. However, data from the longitudinal study did not support changes over a period of 18 months and it is possible that a) the child sample was not large enough, b) the period of 18 months was too short and did not capture a crucial stage of improvement or c) children did not change in their recognition abilities of adult stimuli but may have shown changes over time in recognising child stimuli as a result of increased exposure to other children.

In contrast, data from Chapter 3 did suggest that young children were less stable than older children and adults in their emotion perception and often relied on additional external factors such as gender of stimulus to classify the emotion in questions. Further, the similarity of emotion recognition across two separate modalities seemed to be evident in children but matured with age. Although basic emotion recognition is thought to be universal and even innate (Darwin, 1872) and emotion imitation is already evident in new-born infants (Meltzoff & Moore, 1989) recent studies – including the present – have delivered ample evidence for the continuous improvement of emotion recognition over time. This could be due to maturation of brain mechanisms (*e.g.* Vink et al., 2014) or advances in cognitive development (Herba & Phillips, 2004).

Further, it is possible that young children are learning to make eye contact during conversations as a cultural norm of politeness and social skills in Western cultures which could promote more observant behaviours of emotion expressions in others as a function of age. The discussion about nature versus nurture for emotion recognition abilities continuous but present results suggest that nurture and social development certainly influences the sensitivity of emotion perception with age. Social learning theories support this notion, suggesting that social behaviour can be learned by observing others (Bandura, 1965). This in turn supports the idea that children improve their emotion recognition abilities during their years of primary school which results in increased exposure of social situations. Further, Denham, Mitchell-Copeland, Strandberg, Auerbach and Blair (1997) suggested that children’s emotion competence and emotion expression abilities was directly linked to their parent’s emotional expressiveness, indicating environmental learning effects on emotion recognition rather than recognition abilities at adult standard from very early age on.

3.5.3. Implications

Results from the current study indicate that children’s emotion recognition abilities across two separate modalities improved during their years of primary school, *i.e.* between the ages of 5 to10. Hence, the present data fills gap in existing emotion research in children during either early childhood (Nelson & Russell, 2011) or during late childhood (Tonks et al., 2007). More precisely, children between the ages of 5 to 7 might be able to recognise certain emotions like happiness in faces whilst this develops a little later for voices.

In general, children continued to show improvement in speed and intensity perception as well as accuracy over childhood and even 10-year old children were not as competent as adults in categorising emotions. This finding has to be taken into consideration for parenting or teaching guidelines as it might help to understand miscommunications between adults and children. Especially in the light of rising figures for children under the Child Protection Register which reached 50,732 registered children in the UK in March 2013 (NSPCC, 2013) or rising numbers of care applications (CafCass, 2009), parents that are unable to cope may benefit from education about their children's emotion understanding abilities. It may shed light on young children's inability to efficiently classify negative emotions – especially when expressed in the voice – suggesting that carers as well as teachers need to communicate emotions in very obvious and prototypical ways.

In addition, the notion that younger children may be better at seeing rather than hearing what other people feel, might be beneficial for noisy playground situations where children might not be able to hear what other children say, but can see their expressions on faces from far away. In turn, results indicate that emotional intention in radio advertising as well as audio books or puppet shows such as 'Punch and Judy', which is based on vocal story telling rather than changing face expressions, might only be understood by children from the ages of 8 to 10 onwards. However, this obviously depends on the individual and may be influenced by factors such as gender of child or number of siblings.

In terms of research implications, findings from Chapter 3, which suggested similar emotion recognition patterns between two independent modalities that are already to some extend visible within children, provide the basis for enhanced inclusion of auditory emotion recognition tests in research. This may apply especially to non-verbal emotion vocalisations. Further, it may advance investigations of links between emotion recognition and future social behaviour as well as for autism research and the development of emotion recognition training programs which have previously been based on visual emotion recognition only (*e.g.* Baron-Cohen et al., 2009).

3.5.4. Limitations & Future

The current child chapter focused on one aspect of emotion understanding: namely the basic emotion recognition of faces and voices. It did, however, not include further general emotion comprehension factors such as external cause attribution, mixed emotions or socially influenced emotion recognition such as recognition of shame or pride which may all show age-related effects during development (Albanese et al., 2010). Future studies on child development ought to control for several additional emotion understanding factors and their relationship to emotions expressed across several independent modalities.

In addition, the current chapter investigated emotion recognition in British children who were either born or raised in England and were exposed to the British culture. However, although the basic emotions in faces as well as in voices are believed to exist and be recognised universally (*e.g.* Bryant & Barrett, 2008; Ekman & Friesen, 1971; Izard, 1980), emotion recognition and perception may differ as a function of culture and certain display rules. For example, culture may influence intensity perception of emotions for Mexican and Canadian children (McCluskey & Albas, 1981) as well as accuracy of emotion categorisation (Markham & Wang, 1996) for Chinese and Australian children as use and expression of specific emotions may be encouraged or discouraged in some but not in other cultures. Further, certain cultural values which are displayed within the family setting such as power distance or the way parents express emotions influences children's socialisation and emotion recognition (Halberstadt & Lozada, 2011; Eisenberg, Cumberland, & Spinrad, 2009).

Further, different schooling systems as well as differences in child-care facilities depending on countries may influence parent's contribution to emotional development in their offspring and children's social networks. Hence, the present data is limited to interpretations within a British context and might not be replicable in other cultures. Since Chapter 3 used non-verbal auditory as well as visual stimuli, a cross-cultural adaptation of the paradigm is possible in order to investigate children's emotion recognition development across different cultures – independent of language. Data for this exact issue will be presented in a cross-cultural Chapter 4.

3.5.5. Conclusion

Overall, across three experiments in Chapter 3, results demonstrated that children during primary school had similar emotion recognition patterns to adults such as enhanced sensitivity to happy as compared to threatening stimuli. However, performance stabilised with age - which was especially true for the auditory condition - and speed improved significantly between the age of 10 and adulthood. Whilst faces generally communicated emotions better than non-verbal vocalisations, patterns for individual emotions or contributing factors such as participants' age paralleled performance across modalities, suggesting shared emotion recognition mechanisms which may already exist in childhood. However, modality-similarities matured and became stronger with age, suggesting that modality-independent emotion processing may be a result of development. Findings are independent of the use of child or adult faces as children did not exhibit in-group age bias for stimuli expressed by actors from the same age-group. However, in order to extend data from Chapter 2 and 3 to other cultures, a cross-cultural comparison of children's emotion recognition is of interest and will be conducted in the following Chapter 4.

4

Cross-Cultural Comparison of Emotion Recognition in British and Germans: Not so Similar After All?

When we see someone smiling, our natural assumption is that this person must be happy; when we hear someone groan, we probably conclude that that person is angry or frustrated. Results from the previous Chapters 2 and 3 have suggested that we are able to categorise basic emotions presented in the voice and the face – and this ability improved throughout childhood. However, an important question remains: Do people from different cultures vary in the way they perceive emotions when watching or listening to other people?

The effect of cultural variation within emotion recognition has been studied extensively within either the facial (Ekman & Friesen, 1971) or the auditory domain (Bryant & Barrett, 2008). However, to the best of the researcher's knowledge, no study has specifically investigated whether the recognition of emotions - not only in one, but in two separate modalities – is influenced by cultural context. Hence, the present study will expand previous literature by including within-subject comparisons of emotion recognition in faces as well as in voices. This will be set within a cross-cultural as well as a developmental research context. It will introduce more specific ideas regarding variation in basic emotion recognition across two modalities in German and British participants by firstly analysing cross-cultural data from an adult sample in Experiment 3-I and then analysing cross-cultural data from a child sample in Experiment 3-II.

4.1. Experiment 4-I

Emotion Recognition Across Cultures and Modalities

Does everyone recognise emotions in the same way? How much influence do cultural norms have on the experience and expression of emotion? The last decades have attempted to answer those questions by investigating emotion recognition across different cultures (*e.g*. Ekman & Friesen, 1971). Early research has proposed that emotions are universally recognised (*e.g.* Izard, 1980). There seem to be certain factors within basic emotions which allow different cultures to agree about the emotion displayed by others. In order to exclude the confounding factor of shared access to Western media, literate as well as illiterate cultures showed that facial (Ekman & Friesen, 1971; Izard, 1980) as well as vocal emotions (Bryant & Barrett, 2008) could be recognised in a similar manner – independent of culture or media influence. This finding has been interpreted as evidence that emotion recognition and therefore expression seemed indeed to be universal and accessible from birth onwards. The universal communication of basic emotions may be rooted in evolution as it aids identification of dangerous situations and would therefore be a result of universal biological heritage (Darwin, 1872/1965; Russell, 1994).

On the other hand, however, the cultural universality of emotion recognition has also been contested. In 2009, Jack, Blais, Scheepers, Schyns and Caldara collected behavioural and eye-movement data and proposed that Eastern groups relied mainly on the eye region to decode emotional face expressions whereas Western groups focused on the whole face. This explains how emotions such as fear and disgust which share the feature of wide open eyes can cause recognition problems within the East-Asian group. Similarly for the auditory domain, recent data has brought fresh air into the discussion of the universality of emotion recognition. Gendron, Roberson, van der Vyver and Barrett (2014) for example have proposed that emotional non-verbal vocalizations may not be universally recognised. A remote cultural group from parts of Namibia did not use the same emotion terms as American participants during a label task of the emotion sounds. In contrast, the Himba culture in Namibia categorise emotions mainly on valence and arousal ratings and tended to describe the action (*e.g*. 'scream') rather than the emotion (*e.g.* 'fear'). From the above, the debate about the universality of emotion recognition within the auditory and visual domain across two very different cultures is still ongoing. It is, however, possible, that certain basic elements of emotion recognition are universal whilst culture adds some variation to emotion perception as proposed by interactionist theorists.

4.1.1.1. Interactionist account of emotion recognition across cultures

In 1971, Ekman and Friesen reported that accuracy of facial emotion recognition in illiterate cultures such as New Guinea was comparable to recognition abilities of Western, literate cultures – although there was some slight cultural variation in distinction between certain emotion pairs. By reviewing 97 studies, Elfenbein and Ambady (2002) concluded that certain emotions were indeed recognised universally but recognition was often influenced by society and cultural boundaries. Happiness and anger were the most accurately recognized facial expressions across several cultures; however, the degree of understanding those emotions varied slightly depending on external factors such as the similarity between the person that expressed the emotion and the person that perceived the emotion. Generally, emotions were better recognised when they were expressed by members of the same culture as compared to another culture.

For vocal emotion recognition, Sauter, Eisner, Ekman and Scott (2009) demonstrated overall agreement in identifying basic non-verbal vocal emotion expressions between British and remote Namibian participants. However, recent data has called into question the universality of emotion recognition. Gendron, Roberson, van der Vyver and Barrett (2014a) for example have proposed that emotional non-verbal vocalizations may not be universally recognised when focusing on discrete emotions alone. The Himba, a remote cultural group from parts of Namibia did not use the same emotion terms as American participants during a label task of the emotion sounds. In contrast, Himba emotion categorisation seemed to be mainly based on valence and arousal ratings and tended to describe the action (*e.g.* 'scream') rather than the emotion (*e.g.* 'fear'). Similarly, for face expressions of emotions, Himba participants did not show the anticipated universal pattern of emotion categorization and this was especially true when no conceptual context was provided such as forced-choice labelling (Gendron, Roberson, van der Vyver, & Barrett, 2014b).

In response to this, Sauter et al. (2015) replied that this cross-cultural difference in emotion recognition may be due to the inclusion of culture-specific positive emotion vocalisations which are generally less well recognised across cultures (*e.g.* triumph) rather than simply valence-based judgements on part of the Namibian sample. However, this discussion also shows that a 'Western way' of setting up the experiment (such as cross-checking that remote cultures understand the 'Western way' of correctly completing the task) as done by Sauter et al. (2015) provides conceptual context itself to participants. This in turn could influence the universality of categorical emotion recognition negatively in remote cultures and only free-labelling tasks can truly investigate cultural variation in emotion recognition (Gendron, Roberson, & Barrett, 2015). Hence, the discussion whether different cultures perceive emotions universally based on discrete categories or on dimensional valence and arousal ratings is ongoing and may be going back to the appropriate use of experimental design to answer the experimental question.

Overall, however, the majority of recent cross-cultural studies have now consistently reported a subtle degree of variation on otherwise universally recognised emotions across cultures. In the light of those findings, Elfenbein and colleagues postulated the ‘Dialect Theory of Emotion Recognition’ which acknowledges the role of cultural specificity on emotion recognition (Elfenbein, 2013; Elfenbein & Ambady, 2003). This theory proposes the idea that general emotion recognition abilities of basic emotions are controlled by innate and biological processes. However, this universal process would then vary across cultures like a dialect within an otherwise universal language. Hence, social learning and different social norms could therefore contribute to emotion expression and create an in-group advantage for recognising emotions expressed by some member of one’s own culture. Elfenbein, Beaupre, Levesque and Hess (2007) presented evidence for actual culture-specific differences in the use of face muscles to create expressions of emotions: Canadian as well as African French-speakers posed emotional face expressions and those expressions were then compared for their differences in typical muscle activations. Whilst an universal understanding and expression of the basic emotions was clear, emotions such as happiness, anger or surprise showed ‘dialects’ in their expression and members of different cultures used different face muscles to communicate the intended emotion. Those dialects can account for in-group advantages of recognising emotions better when they are expressed by members of one’s own culture as compared to a foreign culture (Elfenbein & Ambady, 2002).

Very recently, this particular in-group advantage for emotion recognition has also been extended to the vocal domain by Laukka, Neiberg and Elfenbein (2014). English-speakers from Australia, India, Kenya, Singapore and the United States expressed emotions verbally by speaking a neutral sentence with different emotional prosody. Acoustic features used to express the intended emotion were then compared across cultures. Similarly to facial expressions, many acoustical features were shared across participants. However, in accordance with the Dialect Theory, classification analyses revealed an ingroup-advantage for matching vocal stimuli produced by speakers of the same versus other culture. Therefore, it is of interest how culture-specific values and norms contribute to differences in emotion expression as well as emotion recognition between members from two different Western nations.

4.1.1.2. How do display rules influence recognition of emotions?

The term ‘display rules’ of emotions (Buck, 1984) attempts to describe the process of social learning within any culture which creates social norms and influences social behaviour and in turn the display of emotional information. If certain emotional behaviours are doomed to be inappropriate, members of a culture will learn to either mask or even ignore unfavourable emotion displays. In contrast to ‘dampening’, ‘amplification’ of emotion defines the opposite and describes a situation where emotions are expressed or decoded more intensely than the actual experience (Matsumoto, Kasri, & Kooken, 1999). Often times, people from collectivistic nations - such as Asian nations - display and perceive emotions in a less intense way than their individualistic counter parts - such as Americans - do due to the existence of certain social norms such as emotional restraint (Friesen, 1972; Matsumoto et al., 2008). To illustrate this, Soto, Leveson and Ebling (2005) for example could not find any physiological differences such as changes in finger temperature or skin conductance levels across Chinese-Americans and Mexican-Americans during acoustic startle conditions. Interestingly, however, Chinese-American participants reported experiencing less emotional feelings of 'being startled' than Mexican-Americans which suggests culture-related 'dampening' of subjective emotion experience.

How exactly are emotion perception, emotion expression and emotion recognition linked in the light of display rules such as emotion suppression? When suppressing one's own emotional face expressions – which may be seen as appropriate in certain cultures – emotion recognition sensitivity can be impaired as a result. For example, Schneider, Hempel and Lynch (2013) demonstrated that preventing emotional face expressions whilst watching emotional faces – and therefore hindering facial mimicry – slowed down emotion recognition and also negatively affected sensitivity to the expressed emotion. This phenomenon has further been supported by limiting the mimicry of facial expressions of emotions such as by biting on a pen (Oberman, Winkielman, & Ramachandran, 2007) or after Botulinum Toxin injections of participants (Neal & Chartrand, 2011) which resulted in lower emotion recognition accuracy. Those findings have demonstrated that changes in facial emotion expression are indeed linked to emotion recognition abilities. It is believed that by reducing afferent facial feedback from contractions of face muscles to the brain, subjective emotion perception is also dampened as a result (Neal & Chartrand, 2011).

In the present experiment, participants were residents of two nations –Germany and Great Britain – with a high degree of shared cultural heritage. Consequently, it is expected that overall recognition abilities of facial as well as vocal emotion expression may not necessarily differ as a function of culture: Chances are high that in both nations, a happy face is indeed labelled as happy. However, certain subtle differences in display and decoding rules and small but important differences in cultural and social behaviour might still give cause for variation in emotion detection. A more detailed account of the assumption of typically British and German display rules which may affect emotion recognition are given below.

4.1.1.3. Why compare Great Britain and Germany?

Traditionally, emotion recognition has been investigated between different ethnic groups such as individualistic or collectivistic nations, based on very different display rules (Ekman, 1987; Matsumoto et al., 2008). As defined by Triandis (1990), collectivist nations such as East Asia represent harmony and are taught to value the need of the group over the need of the individual. Individualistic western countries such as the United States of America on the other hand value autonomy and independency over social unity. Up to this date, it remains unclear whether emotional judgement differs between two similar cultures as is often the case for cultures with obvious differences in social and cultural behaviour.

Germany and Great Britain form a particular interesting comparison because of their increasingly strong goods and service trading bond. In 2013, Germany was UK's largest import partner with 13.7% of total UK imports for 2013 and Germany was UK's second largest export trading partner after the United States contributing to 9.8% of total UK exports (Office for National Statistics, March 2014). For Germany, the UK was the third most important export partner in 2013, exporting goods and services in the excess of 75bn € and in 6^{th} place for import trade (Statistisches Bundesamt, 2014). Further, in 2012, after India, Poland, Pakistan and Republic of Ireland, Germany was the 5^{th} most common country of birth from oversea residents currently living in the United Kingdom which amounts to about 300000 UK residents (Office for National Statistics, August 2013). German companies provide work for about 330000 people in Great Britain (British Chamber of Commerce in Germany e.V., April 2014).

From those statistics, it is apparent that Great Britain and Germany have one of the strongest bonds within Europe which requires a high degree of understanding of the opposite partner for example during business negotiations. This in turn requires the awareness of potential cultural differences in display rules between countries and emotion perception in face-to-face meetings as well as in situations where only vocal information is available, for example, during telephone conversations. Indeed, being able to accurately recognise facial emotions expressed by others increases business negotiation success: Elfenbein, Foo, White, Tan and Aik (2007) demonstrated that during a negotiation task sellers with high personal emotion recognition abilities were more successful in completing business transactions than individuals with lower emotion scores.

4.1.1.4. How similar are Germans and the British?

Why should German and British participants show differences in emotion recognition? Hofstede (2001) has developed five cultural dimensions to compare up to 76 different countries by investigating differences in power distance, individualism, masculinity, uncertainty avoidance and pragmatism. In 2010, this has been extended to include the dimension of indulgence (Hofstede et al., 2010). Hofstede's data suggests that Germans scored lower on individualism than British which suggests more society oriented behaviour in Germans. This other-direction behaviour might result in a willingness to restrain their emotional expression to avoid drawing attention to oneself (Schneider, Hempel, & Lynch, 2013). British participants on the other hand may be less concerned about their emotional expressivity and the impact it has on others, resulting in higher emotion expression and hence increased sensitivity to intensity perception of expressive emotions such as happiness and disgust. Similarly, indulgence scores are higher for British than for German citizens, suggesting that British participants are more likely to be led by impulses.

Most behavioural evidence regarding cultural variation between European countries comes from business and leadership studies. In 2002, Brodbeck et al. investigated desired leadership characteristics within 22 European countries and over 6000 participants rated common leadership attributes on a 7-point Likert Scale which the participants valued to be desirable for a successful leadership in business. In regards to the Anglo and Germanic cluster[12], differences in cultural values seemed to account for over 60% of overall variance in leadership styles.

Within Europe, it seemed that features like being administrative and organised was highly valued in Germanic cultures. Further, ´integrity` which included being honest and trustworthy as well as ‘autonomy’ was a more important perceived trait in leaders for Germanic nations than for Anglo countries. On the other hand, Anglo countries valued traits such as being generous and compassionate as high in business leaders, rating warmth and interpersonal relationships (Brodbeck et al., 2000). However, although Brodbeck and colleagues’ study was conducted in 2000, data stems from a large international study (GLOBE project, Hanges et al., 1999) with original data collected between 1995 and 1997. Hence, some of the findings regarding personal attitudes may be outdated by now and reflect views of a previous generation of business managers.

In addition to leadership studies, the second strand of research which claims to reveal information about cultural differences of behaviour and norms between countries is the field of communication and language. If communication styles vary over time and generations, one can assume that underlying cultural norms and attitudes change as well (Wierzbicka, 1998). For example, a change in the use of formal titles may indicate underlying changes in social distance. Hence, lexical hints within the language of a society can reveal information about cultural values within that society. Grieve (2009) qualitatively investigated structure and content of workplace telephone conversations and reported that Germans were more likely to overtly show disappointment, tell the truth and apologized more often compared to Australians. In addition, the German language was more direct than English because German speakers used more infinitives and imperatives, showed less verbal concern about others and believed in fixed rules. Germans demonstrated a higher degree of social distance by including the formal you (´Sie`) as well as the informal you (´Du`) within their vocabulary (House, 2006) which clarifies hierarchy between two individuals (Wouters, 2011). It has been attempted to explain this social distance with events in German history or as a result of strict educational systems which focus on education rather than on social acts (House, 2007).

[12] European countries can be clustered according to their cultural values or leadership styles. The anglo cluster contains Ireland and United Kingdom, whilst Germany, Austria and Switzerland form their own Germanic cluster (Brodbeck et al., 2002).

Overall, data from business and conversation-style studies have suggested that there may indeed be differences in social norms and therefore display rules between two very similar individualistic countries such as Germany and Great Britain. Assuming a link between one's own emotional expression and emotion perception (Schneider et al., 2013), the question is whether those subtle differences in display rules between German and British nationals also rub off on emotion recognition abilities. Following previous findings presented above, the current study works on the assumption that Germans value a more reserved but direct style as compared to a more emotional and less hierarchical British style. However, it is to note that the present study did neither investigate actual physical differences in emotion expression between cultures nor differences in social norms.

4.1.1.5 The Present Study

The present study approached the question of universal emotion recognition from a novel perspective by comparing two similar, Western and individualistic countries. To the best of the researcher's knowledge, this is the first attempt to investigate cross-cultural effects on emotion recognition not only within one – but also within two separate modalities such as the face and voice. Hence, the aim of the present study will be the investigation of emotion recognition abilities in the voice and face – based on intensity as well as choice reaction time data in German and British adults. Whilst general emotion categorisation concepts may be comparable between British and German people (according to the universality hypothesis), investigating differences in intensity perception will allow the investigation of more subtle differences in emotion recognition (according to the dialect hypothesis).

Further, since most of the past research has been conducted either in the field of facial (*e.g.* Matsumoto et al., 2008) or vocal (Bryant & Barrett, 2008) emotion recognition, there is limited evidence regarding the interplay between emotion recognition across separate modalities and cultural aspects such as display rules. Extending previous research, the current project not only included facial but also vocal emotion recognition data in a within-subject design. Overall, possible differences between British and German participants could a) demonstrate whether cultural factors have equal influence on emotion recognition across modalities, suggesting a modality-independent vulnerability to cultural influences on emotion recognition and b) demonstrate that even similar countries with strong business and trading bonds ought to pay attention to cultural differences in emotion perception in order to avoid communication misunderstandings.

Overall, the experimental hypotheses of the present study are as follows:

Hypothesis 1: Assuming general universality of basic emotion recognition across two individualistic countries, the current study predicts that both cultures have comparable patterns of emotion recognition.

Hypothesis 2: Due to the observed differences in expected social norms between German and British participants in studies reported above, within the comparable patterns predicted by Hypothesis1, emotion intensity perception may deviate and Germans will perceive emotions as less intense than British participants.

Hypothesis 3: Assuming common underlying emotion coding mechanisms across voices and faces as stated in Chapter 2, potential cultural differences between German and British participants should not only influence facial but will also affect vocal emotion processing abilities.

4.1.2. Methods

4.1.2.1. Participants

Adult participants were 45 British adults (15 male, 30 female), aged between 18 and 61, (*M* = 30.8; *SD* = 16.81) from the adult Chapter 2 (see Chapter 2 for more details) as well as 32 German adults (15 male, 17 female), aged between 18 and 61 (*M* = 35.03, *SD* = 13.6). German as well as British participants were matched for age and non-verbal IQ which both did not differ significantly between groups (all $p > .05$). German adults were recruited via social networks and from a local horse-riding club. All participants reported normal or corrected-to-normal hearing and vision and English or German was the first language, respectively. Participants were tested in their native language. All material was translated by the experimenter (who is a native German speaker living in England) and, for verification, translated back by a native English-speaker living in Germany. A list of translated words is attached in Appendix A.

In addition, all adult participants from the behavioural main study (ERS 1) as well as all adult participants from the larger online questionnaire study (ERS 2) completed an online or paper-version of the Emotional Reactivity Scale (Nock et al., 2008). Again, all participants were tested in their native language and the scale was translated from English into German by the experimenter (who is a native German speaker living in England) and, for verification, translated back by a native English-speaker living in Germany. The German version of the ERS scale is attached in Appendix A.

The study was approved by the Department of Psychology, Brunel University, PsyREC officer (01/03/2012 & 17/10/2012) and informed consent was filled in both by participants and parents prior to the study. All participants were debriefed after the study. Forms are attached in Appendix A.

In addition, a separate online study was run to assess whether participants from different cultures actually differ in their emotional reactivity. For this, data was collected from a larger online sample of British and German adults (ERS 2). For the Emotional Reactivity Scale questionnaire (ERS, Nock, Wedig, Holmberg, & Hooley, 2008), 104 British adults (15 male, 89 female), aged 18 to 62 (*M* = 23.49, *SD* = 10.07) as well as 111 German adults (41 male, 70 female) aged 18 to 56 (*M* = 26.65, *SD* = 7.14) completed the online questionnaire. Participants were recruited via Brunel University's participation pool and through social networks. Informed consent was given online and debriefing was sent out via emails. The study was approved by the Department of Psychology, Brunel University, PsyREC officer (08/11/2012). Again, all participants were tested in their native language and the scale was translated from English into German by the experimenter (who is a native German speaker living in England) and translated back by a native English-speaker living in Germany. The German version of the ERS scale is attached in Appendix A.

4.1.2.2. Stimuli

See Chapter 2. The stimuli were chosen according to their appropriateness for the use in cross-cultural research. Both visual and auditory stimuli are created by white-Caucasian actors in order to avoid potential out-group biasing (Russell, 1994) and do not contain any verbal or language content.

In addition, all adult participants from the behavioural main study (ERS 1) as well as all adult participants from the larger online questionnaire study (ERS 2) completed an online or paper-version of the Emotional Reactivity Scale (Nock et al., 2008). The ERS score is based on a 21-item self-reported questionnaire on a 5-point Likert scale and measures subjective experience of emotions such as intensity or persistence in response to emotional situations. The ERS questionnaire includes items such as ‘When I experience emotions, I feel them very strongly/intensely’ and shows internal consistency of .94 (Nock et al., 2008). For a copy, see Appendix A.

4.1.2.3. Procedure

See Chapter 2.

4.1.2.4. Statistical Analysis

For the main analysis of intensity ratings and reaction times, a mixed measures ANOVA was conducted for matching target emotions. Between-subject variables were country (Great Britain or Germany) and gender of participants (male or female). The within-subject variables were modality (face or voice) and emotion category (happy, sad, angry, fear, surprise or disgust). For the current analysis, gender of stimuli was not included in order to limit a loss of power due to multiple comparisons.

For the Emotional Reactivity Scale online questionnaire, independent t-tests were conducted with country (Germany or Great Britain) as between-subject variable and a continuous ERS score (sum of all answers) as within-subject variable.

4.1.3. Results

This section will present results from the current cross-cultural chapter which includes results from the Emotional Reactivity Scale (ERS 1 & 2), intensity ratings as well as choice reaction times. For *M* and *SE* see Table 4.1. To ensure that British and German adults were matched for age and non-verbal IQ scores, two independent t-tests were performed. The tests were not significant for either variables which suggests that both groups had comparable age and non-verbal IQ scores ($p > .05$).

a) Emotional Reactivity Scale (ERS 1 & 2)

ERS 1: To control whether both cultures significantly differed in the degree of emotional reactivity and sensitivity to emotional events, participants completed the 21-item Emotional Reactivity Scale Questionnaire (Nock et al., 2008). An independent t-test revealed no significant difference in emotional reactivity scores between German (M = 27.41, SD = 9.61) and British (M = 25.56, SD = 16.15) participants [$t(65) = -.37$, $p > .05$]. However, only 22 Germans and 48 English adults completed the ERS which suggests an unequal and relative small sample size.

ERS 2: In order to investigate how representative the above result was, a larger-scale online survey with British (N = 104) and German (N = 111) adults was conducted. Due to a technical error, only answers for 17 out of 21 question items were recorded; however, internal consistency was measured by Cronbach alpha and was still very high, with .93 for the British sample and .85 for the German sample. With a score calculated from 17 answers, an independent t-test revealed no significant difference between British (M = 24.78, SD = 13.54) and German (M = 26.05, SD = 10.42) scores for emotional reactivity. Hence, there were no significant difference in the way participants perceived their own emotional reactivity across Germany and England.

b) Intensity Ratings

The following Table 4.1 displays Means and Standard Errors for intensity ratings for British and German adults.

Table 4.1

Means (M) and Standard Errors (SE) for intensity ratings for modality and emotions for British and German adult participants

		Mean Intensity (*SE*)						
		British			**German**			
		Female	Male	All	Female	Male	All	**Total**
		N = 30 *M* = 26.72 (15.82)	*N* = 15 *M* = 35.53 (16.67)	*N* = 45 *M* = 30.47 (16.81)	*N* = 17 *M* = 34.78 (13.67)	*N* = 15 *M* = 35.33 (13.98)	*N* = 32 *M* = 35.03 (13.60)	*N* = 77 *M* = 32.58 (15.62)
Face	*Happy*	6.54 (.10)	6.62 (.14)	6.58 (.08)	6.17 (.19)	6.30 (.20)	6.23 (.14)	6.39 (.09)
	Sad	5.35 (.19)	5.54 (.28)	5.34 (.16)	5.65 (.27)	4.97 (.29)	5.31 (.20)	5.32 (.13)
	Angry	5.41 (.17)	4.79 (.25)	5.10 (.15)	6.13 (.19)	5.48 (.21)	5.80 (.14)	5.48 (.10)
	Fear	5.87 (.18)	5.44 (.26)	5.66 (.16)	5.65 (.28)	5.17 (.30)	5.41 (.20)	5.53 (.14)
	Surprise	6.40 (.12)	6.31 (.17)	6.35 (.10)	6.32 (.17)	6.10 (.18)	6.21 (.12)	6.27 (.08)
	Disgust	5.82 (.19)	5.13 (.28)	5.47 (.16)	4.68 (.27)	4.35 (.30)	4.52 (.20)	5.00 (.14)
	All	5.90 (.10)	5.60 (.15)	5.75 (.09)	5.77 (.15)	5.39 (.12)	5.58 (.11)	5.66 (.07)
Voice	*Happy*	6.12 (.16)	6.16 (.23)	6.14 (.13)	5.74 (.23)	5.40 (.26)	5.57 (.17)	5.84 (.12)
	Sad	6.23 (.15)	5.85 (.22)	6.04 (.13)	6.54 (.21)	5.65 (.23)	6.10 (.16)	6.04 (.11)
	Angry	5.07 (.20)	4.52 (.29)	4.81 (.17)	4.81 (.23)	4.32 (.25)	4.56 (.17)	4.70 (.11)
	Fear	5.22 (.21)	5.10 (.30)	5.16 (.18)	4.60 (.28)	4.15 (.31)	4.37 (.21)	4.75 (.14)
	Surprise	5.27 (.20)	5.14 (.29)	5.21 (.17)	5.18 (.26)	4.68 (.28)	4.93 (.19)	5.06 (.13)
	Disgust	5.55 (.17)	5.58 (.25)	5.56 (.15)	6.17 (.23)	5.28 (.25)	5.73 (.17)	5.65 (.12)
	All	5.58 (.11)	5.40 (.16)	5.49 (.10)	5.51 (.16)	4.91 (.18)	5.21 (.12)	5.34 (.08)

		British			German			
		Female	Male	All	Female	Male	All	**Total**
Overall	*Happy*	6.33 (.10)	6.39 (.14)	6.36 (.08)	5.95 (.17)	5.85 (.19)	5.90 (.13)	6.11 (.09)
	Sad	5.79 (.13)	5.59 (.19)	5.69 (.11)	6.1 (.2)	5.31 (.22)	5.70 (.15)	5.68 (.10)
	Angry	5.24 (.15)	4.67 (.22)	4.95 (.13)	5.47 (.17)	4.90 (.19)	5.18 (.13)	5.09 (.09)
	Fear	5.55 (.17)	5.27 (.25)	5.41 (.15)	5.13 (.25)	4.66 (.27)	4.89 (.18)	5.14 (.12)
	Surprise	5.84 (.12)	5.72 (.18)	5.78 (.11)	5.75 (.17)	5.39 (.18)	5.57 (.13)	5.67 (.08)
	Disgust	5.68 (.15)	5.35 (.21)	5.52 (.12)	5.42 (.02)	4.82 (.22)	5.12 (.15)	5.32 (.10)
Grand Means		5.74 (.09)	5.50 (.13)	5.62 (.08)	5.64 (.14)	5.15 (.15)	5.40 (.10)	5.50 (.07)

Note. * = Mean Age in years. ** = Average scores across all participants and conditions.

Results from intensity ratings across all German and British adults suggested that emotional intensity perception did not differ between countries in general but rather depended on specific emotions and modality. A mixed-measures ANOVA of modality (2) x emotion (6) x country (2) x gender of participants (2) revealed a significant main effect for modality, $F(1, 73) = 23.49$, $p < .001$, $\eta^2 = .24$, gender, $F(1, 73) = 9.52$, $p = .003$, $\eta^2 = .12$ and for emotion, $F(4.76, 347.46) = 26.94$, $p < .001$, $\eta^2 = .27$. Faces were rated with higher intensity than voices and female participants associated stimuli with higher intensity ratings than male participants. Sidak post-hoc test for multiple comparisons revealed that participants rated happiness with highest intensity ratings which was higher than for all other emotions, then followed by sadness and surprise and fear which did not differ from each other. Anger received lowest overall intensity ratings; however ratings were not significantly different to fear and disgust, all at $p < .001$ (for *M* and *SE* see Table 4.1). There was no significant main effect for country ($p > .05$).

There was also a significant interaction effect between (i) country and emotion, $F(5, 365) = 3.39$, $p = .005$, $\eta^2 = .04$ and between (ii) modality and emotion, $F(4.62, 336.98) = 41.29$, $p <.001$, $\eta^2 = .36$. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate the interactions.

i: Although there was no overall main effect for country, for happy, fearful and disgusted stimuli, British participants rated emotions with higher intensity than German participants. In addition, the same pattern of happiness superiority was visible in both countries. In Great Britain, happiness was rated higher than all other emotions and anger received lowest ratings. For Germany, happiness was rated highest which did not differ significantly from sadness and fear received lowest intensity ratings which was not significantly different from anger and disgust ($p > .05$). This is illustrated in Figure 4.1.

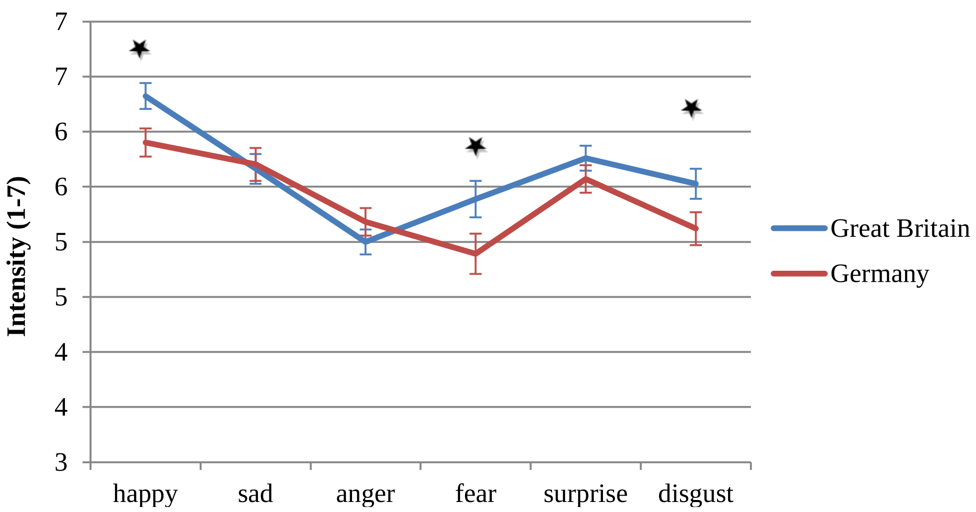


Figure 4.1. Means and Standard Errors for intensity ratings. Significant interaction effect between country and emotion, *$p < .05$. $N = 77$.

ii: There was a significant modality effect visible for all emotions. However, higher ratings for faces as compared to voices were only true for stimuli displaying happiness, anger, fear and surprise. For sadness and disgust, the opposite was true and voices received higher intensity ratings than faces, $p < .05$. In addition, for faces, happiness received highest ratings which did not differ significantly from surprise, followed by fear, anger and sadness. Disgusted faces received lowest intensity ratings which were lower than all other emotions displayed in the face. For the voice on the other hand, sadness received highest ratings which was not significantly different from happiness. This was followed by ratings for disgust which did not differ from happiness ratings, followed by fear and anger. Interestingly, surprised voices received lowest intensity ratings which was the opposite pattern compared to surprised faces.

There was also a significant three-way interaction between country, modality and emotion [$F(5, 365) = 7.62$, $p < .001$, $\eta^2 = .09$]. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate the three-way interaction. Although no overall country effect, both cultures differed in their intensity perception for happy, angry, fearful and disgusted stimuli – depending on the modality. British participants rated happy and fearful voices as well as disgusted faces as more intense than Germans. Germans on the other hand, perceived angry faces as more intense than British participants, $p < .05$. This is illustrated in Figure 4.2.

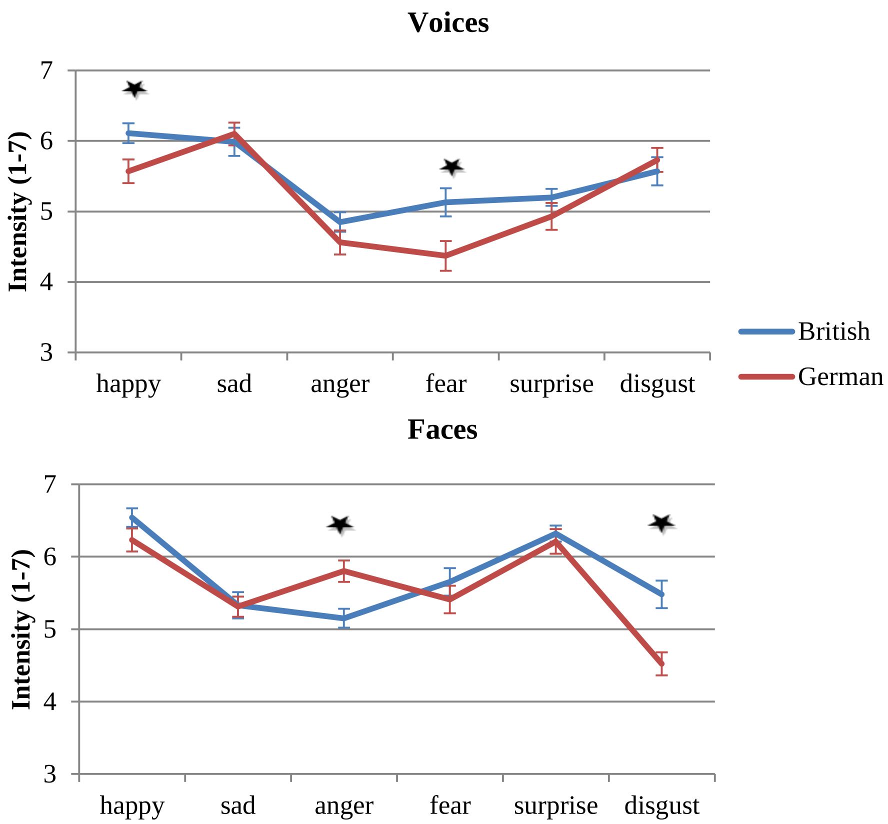


Figure 4.2. Means and Standard Errors for intensity ratings. Significant three-way interaction between country, modality and emotion, $*p < .05$. $N = 77$.

Further, for happy, fear and surprise, both countries rated faces with higher intensity than voices, and the opposite for sadness: voices with higher intensity than faces. Interestingly, the nations differed in their ratings for anger and disgust as only Germans demonstrated a modality effect with higher intensity perception for angry faces over angry voices and disgusted voices over disgusted faces. British judges rated voices and faces with similar intensity for the emotions of anger and disgust.

The following Table 4.2 summarises mean reaction times (RTs) and Standard Errors (SE) for adults across Germany and Great Britain.

Table 4.2

Means and Standard Errors for reaction times for modality and emotions for British and German adult participants

		Mean Choice RTs (*SE*)						
		British			**German**			
		Female	Male	All	Female	Male	All	**Total**
		N = 30 *M* = 26.72 (15.82)	*N* = 15 *M* = 35.53 (16.67)	*N* = 45 *M* = 30.47 (16.81)	*N* = 17 *M* = 34.78 (13.67)	*N* = 15 *M* = 35.33 (13.98)	*N* = 32 *M* = 35.03 (13.60)	*N* = 77 *M* = 32.58 (15.62)
Face	*Happy*	3.31 (.13)	3.63 (.19)	3.47 (.11)	4.16 (.32)	3.62 (.35)	3.89 (.24)	3.70 (.16)
	Sad	4.28 (.22)	4.29 (.32)	4.28 (.19)	4.52 (.36)	4.54 (.38)	4.53 (.26)	4.44 (.17)
	Angry	3.76 (.21)	4.34 (.30)	4.05 (.18)	3.69 (.30)	3.96 (.32)	3.82 (.22)	3.93 (.15)
	Fear	4.04 (.25)	4.47 (.36)	4.25 (.21)	4.90 (.38)	4.47 (.41)	4.69 (.28)	4.47 (.19)
	Surprise	3.33 (.24)	4.63 (.34)	3.98 (.20)	3.38 (.29)	3.43 (.30)	3.40 (.21)	3.71 (.14)
	Disgust	3.57 (.28)	5.04 (.41)	4.31 (.24)	4.15 (.35)	3.74 (.37)	3.94 (.25)	4.13 (.17)
	All	3.82 (.09)	4.40 (.23)	4.06 (.13)	4.13 (.25)	3.96 (.26)	4.05 (.12)	4.06 (.12)
Voice	*Happy*	3.66 (.15)	4.39 (.22)	4.03 (.13)	4.48 (.67)	6.70 (.71)	5.59 (.49)	4.82 (.33)
	Sad	4.16 (.16)	4.53 (.23)	4.35 (.14)	4.55 (.23)	4.70 (.25)	4.62 (.17)	4.51 (.11)
	Angry	4.02 (.18)	4.38 (.26)	4.20 (.15)	4.53 (.24)	4.77 (.26)	4.65 (.18)	4.42 (.12)
	Fear	3.80 (.17)	4.24 (.25)	4.02 (.15)	4.43 (.27)	4.99 (.28)	4.71 (.20)	4.37 (.13)
	Surprise	3.45 (.11)	3.49 (.16)	3.41 (.10)	4.44 (.34)	4.23 (.36)	4.34 (.25)	3.91 (.17)
	Disgust	3.82 (.13)	4.33 (.19)	4.08 (.11)	4.19 (.21)	4.26 (.22)	4.22 (.15)	4.13 (.10)
	All	3.82 (.09)	4.23 (.13)	4.02 (.08)	4.44 (.18)	4.94 (.19)	4.69 (.13)	4.36 (.09)

		British			German			
		Female	Male	All	Female	Male	All	**Total**
Overall	*Happy*	3.48 (.11)	4.01 (.16)	3.75 (.10)	4.32 (.37)	5.16 (.39)	4.74 (.27)	4.26 (.18)
	Sad	4.22 (.14)	4.41 (.21)	4.32 (.12)	4.53 (.23)	4.62 (.25)	4.58 (.17)	4.47 (.12)
	Angry	3.89 (.16)	4.36 (.23)	4.12 (.14)	4.11 (.21)	4.36 (.23)	4.24 (.16)	4.18 (.10)
	Fear	3.92 (.16)	4.36 (.23)	4.14 (.14)	4.67 (.24)	4.73 (.26)	4.70 (.18)	4.42 (.12)
	Surprise	3.39 (.13)	4.06 (.19)	3.72 (.11)	3.91 (.24)	3.83 (.25)	3.87 (.17)	3.81 (.12)
	Disgust	3.70 (.15)	4.68 (.22)	4.19 (.13)	4.17 (.22)	3.99 (.24)	4.08 (.16)	4.13 (.11)
Grand Means		3.77 (.10)	4.31 (.15)	4.04 (.09)	4.28 (.17)	4.05 (.18)	4.37 (.13)	4.21 (.09)

Note. * = Mean Age in years. ** = Average scores across all participants and conditions.

c) **Choice RTs**

Speed of making decisions about judgements for emotional stimuli depended on the interaction of modality, specific emotions and culture. A mixed measures ANOVA with modality (2) x emotion (6) x country (2) x gender of participants (2) revealed a significant main effect for modality, $F(1, 72) = 5.85$, $p = .02$, $\eta^2 = .08$, for emotion, $F(3.54, 530.95) = 5.48$ $p < .001$, $\eta^2 = .07$, and for gender, $F(1, 72) = 6.37$, $p = .01$, $\eta^2 = .08$.

Post hoc comparisons revealed that across all participants, decisions were made faster for faces as compared to voices and male participants ($M = 4.43$, $SE = .13$) took longer to categorise emotions than female participants ($M = 4$, $SE = .11$). For emotion category, surprised stimuli were rated fastest which did not differ significantly from happy stimuli. Happy in turn was not significantly different to all other emotions displayed. Sadness and fear were rated slowest, all at $p < .05$. There was no significant main effect for country, indicating that Germans and British participants took equally long to rate emotional stimuli. For *M* and *SE* see Table 4.2.

There was also a significant interaction effect between (i) modality and country, $F(1, 72) = 7.72$, $p = .01$, $\eta^2 = .1$, between (ii) emotion and country, $F(5, 360) = 3.51$, $p = .004$, $\eta^2 = .05$, and between (iii) modality and emotion, $F(3.24, 232.98) = 4.45$, $p = .991$, $\eta^2 = .06$. Post hoc Sidak comparisons were applied to investigate the interactions in more detail.

i: Although there was no overall main effect for country, the significant interaction between country and modality suggested that for voices, British participants were faster than German participants. For faces, both countries showed similar speed for rating stimuli. In addition, the significant main effect of modality with faster ratings for faces as compared to voices was only true for the German sample. For the British sample, voices and faces were rated with similar speed. These results are illustrated in Figure 4.3.

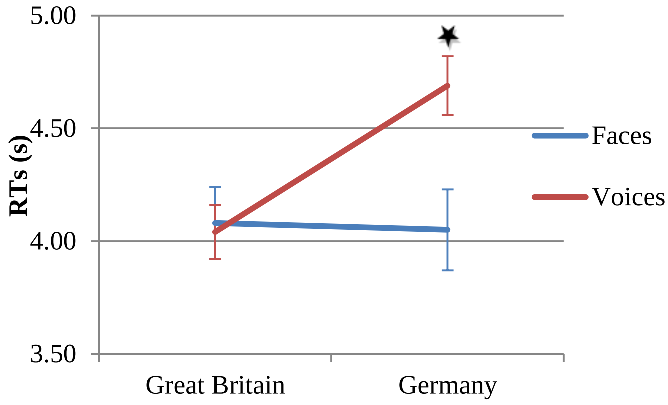


Figure 4.3. Means and Standard Errors for significant interaction effect, $*p < .05$. $N = 77$.

ii: Again, although there was no overall main effect for country, results indicated that for happiness and fear, British participants categorised emotion stimuli faster than the German sample. For sadness, anger, surprise and disgust, countries did not differ in their speed for rating emotions. This is illustrated in Figure 4.4.

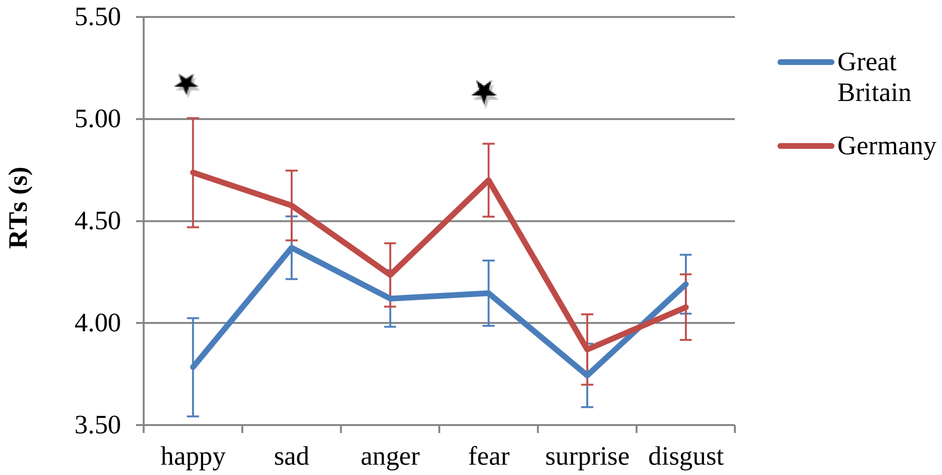


Figure 4.4. Means and Standard Errors for significant interaction effect between country and emotion category of stimuli. **p* < .05, *N* = 77.

iii: Although there was a significant main effect for modality, this depended on the emotion displayed and was only true for happiness and anger. Across all participants, happy and angry faces were rated with faster speed than their vocal counterparts.

There was also a significant three-way interaction between (i) country, modality and gender, $F(1,72) = 5.43$, $p = .02$, $\eta^2 = .07$, and between (ii) modality, emotion and gender, $F(5, 360) = 3.68$, $p = .003$, $\eta^2 = .05$. Post hoc Sidak for multiple comparisons gives a more detailed insight into those interactions.

i: The interaction of decision making speed between country, modality and gender of participants suggested that there was a country effect in voices which was true for both male and female participants. Both gender had faster reaction times for rating voices when they belonged to the British as compared to the German sample although women outperformed men in both countries. For faces, there was no significant country effect and this was true for male as well as female participants, suggesting that German and British nationals rated faces with similar speed - independent of gender. In addition, for British participants, neither men nor women demonstrated a significant modality effect, suggesting that for Britain, gender did not affect rating speed of modality. Germans on the other side demonstrated a significant difference for modality with faces being rated faster than voices. However, this was only true for German men as German women rated voices and faces with similar speed. For faces, German men and women demonstrated same speed but whilst women were equally fast in rating voices, men's speed decreased significantly when rating voices (these results are illustrated in Figure 4.5).

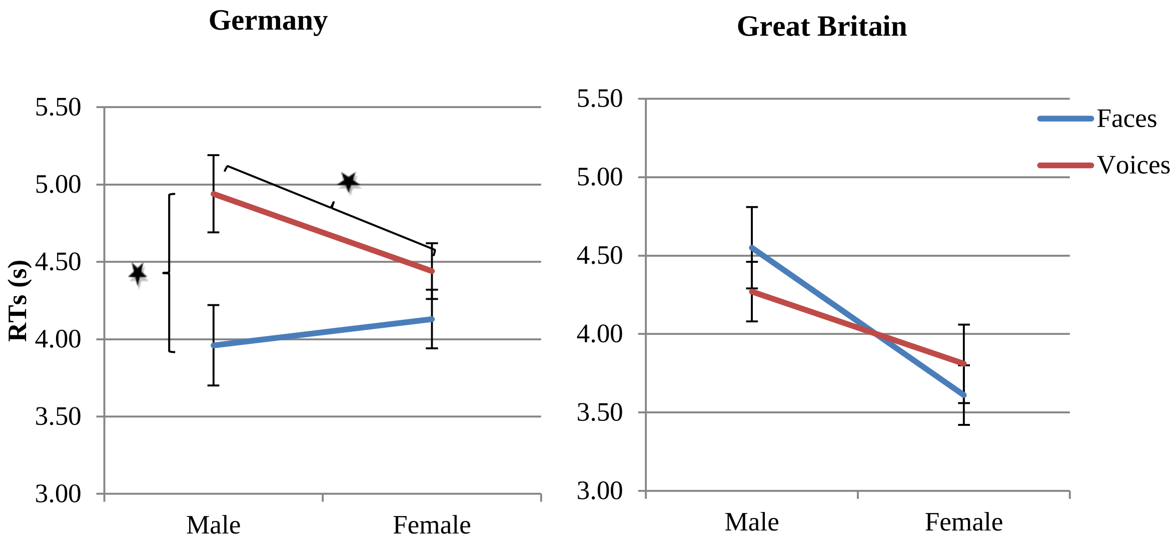


Figure 4.5. Means and Standard Errors for significant three-way interaction between country, modality and gender, $p < .05$. $N = 77$.

iii: Although results suggest an overall main effect for gender as well as for modality and emotion, those three factors interacted with each other. Whilst men generally were slower in making decisions about emotion categorisation, in faces, this was only true for surprise. In the voice condition, this was only significant for happy and fearful voices. For all other emotions across modalities, rating speed did not differ as a function of gender. Further, the significant modality effect in happy and angry stimuli also was dependent on the gender of participants. Men took longer to rate happiness if it was displayed in voices as compared to faces whilst women took longer to rate anger - and additionally surprise - if it was displayed in voices as compared to faces. For happiness, sadness, fear and disgust, women had equally fast decision times across modalities and for sadness, anger, fear, surprise and disgust, men also showed equal rating speed across modalities.

4.1.4. Discussion Experiment 4-I

The purpose of the present experiment was to investigate the contribution of culture on emotion recognition in British and German adults. Whilst Chapter 2 and 3 have found behavioural support for emotion processing which was comparable across voices and faces and which continued to develop with age, Experiment 4-I from this Chapter 4 investigated cross-modal emotion recognition within a cultural context across two similar Western nations. Overall, current findings support Hypothesis 1 of universality of emotion recognition because there was no significant main difference in rating speed as well as intensity perception between German and British adults. There were, however, subtle differences in intensity perception within the adult sample: some emotions were rated as more intense by the British compared to the German sample which was in accordance with Hypothesis 2. Accordingly to Hypothesis 3, in terms of intensity ratings, both modalities seem to be equally affected by culture whilst for reaction times, the culture effect was only seen in voices.

4.1.4.1. Do German and British adults recognise emotions in similar manner?

The present Experiment 4-I did not report a significant main effect of culture for either measures of choice reaction times or intensity ratings during emotion recognition, suggesting that in general, emotion recognition was comparable between British and German adults. This present finding is, however, expected and in accordance with Hypothesis 1 of the current study as it supports the theory of universal emotion recognition across cultures (Ekman & Friesen, 1971). Since Germany and Great Britain are both individualistic countries in North-Western Europe, it is probable that overall emotion recognition does not vary as a function of cultural differences in the present sample due to political, social and geographical proximity.

However, it is of considerable interest whether there are some more subtle differences in emotion recognition between two comparable cultures as suggested by the Dialect Theory (Elfenbein & Ambady, 2002). As hypothesised, data from the present study indeed suggests a relationship between emotion recognition and culture as evident in variation in emotion intensity perception, depending on the specific emotion or modality displayed. For happiness, fear and disgust, British participants perceived the displayed emotion as more intense than their German counterparts. British participants also took less time than Germans to rate happiness and fear, whilst the other emotions did not differ in processing speed as a function of culture and suggest culture-independent emotion processing. This did not depend on the modality of the presented emotion. In addition, participants from both countries rated happiness with highest and anger with lowest intensity ratings which suggests a similar emotion rating pattern independent of culture. This pattern also replicates findings from Chapter 2.

4.1.4.2. Are display rules really the source of cultural emotion recognition differences?

To the best of the researcher's knowledge, there is no existing questionnaire that measures emotional processing abilities as a function of differences in underlying display rules. In an attempt to investigate whether participants actually differed in their emotional temperament and emotion regulation across two cultures, the Emotional Reactivity Scale (ERS, Nock et al., 2008) was used to investigate this exact issue. However, in neither the participants from the main ERS 1 study (N = 77) nor from a larger online sample ERS 2 (N = 215) did the present study find significant differences based on culture, suggesting that participants experienced emotional everyday situations on a similar emotional level. Assuming that display rules can indeed affect emotion expression (Matsumoto et al., 2008) and in turn emotion recognition (Schneider et al., 2013), the present finding may indicate that there were no major differences in underlying cultural norms that significantly influence emotional experience between German and British participants in accordance with the universality hypothesis (Ekman & Friesen, 1971). It is possible that the ERS is not sensitive enough to measure subtle differences in display rules such as emotional reactivity as suggested by the present finding of differences in emotional rating patterns of certain emotions between German and British participants.

Alternatively, however, it is possible that the ERS may not be an adequate test for capturing differences in emotion recognition based on underlying display rules. Self-reported emotional reactivity may not be a reliable indicator of differences in underlying display rules which have an effect on emotion expression and perception. In summary, the present findings of the ERS could not explain the subtle variation of emotion recognition across two countries that were found in the main analysis of emotion recognition across cultures.

4.1.4.3. Cultural differences: what's the role of modality on emotion recognition?

Assuming the German appreciation of reservedness and emotional restrained as suggested by leadership (Brodbeck, 2000) and communication studies (Wouters, 2011), it is possible that the link between masking emotion expressions and decreased emotion perception skills may not only apply to the facial domain as suggested by Schneider et al. (2013) but also the vocal domain. Indeed, for intensity ratings, in accordance with Hypothesis 3, the present results suggested that both modalities may be prone to culture effects: Emotions displayed in faces as well as in voices were rated with varying intensities depending on the culture of participants. However, this pattern depended on specific emotions displayed and different emotions were affected in different modalities whilst the direction of the country effect also varied. For voices, happiness and fear received higher ratings by British than by German participants.

In the visual domain, disgust was also rated higher by British than by Germans. Interestingly, the opposite pattern was true for anger which received higher ratings from German compared to British adults. This indicates a particular German sensitivity to anger expressed in faces which is contrary to the expectation of reduced emotion intensity perception in the German sample. For the remaining emotions in both modalities, there was no variation in emotion recognition across two cultures.

For rating speed of categorising the perceived emotions, however, modality played a significant role between both participating countries. The current data demonstrated that there was a culture-effect on rating speed for voices. However, contrary to Hypothesis 3, this was not true for faces, suggesting that culture does not affect emotion recognition speed equally across separate modalities. Influence of culture and display rules may depend on the mode of presentation and voices may be more vulnerable to cultural variation whilst rating speed for faces seemed to be constant across countries.

4.1.4.4. Modality similarities: What's the role of culture and gender on emotion recognition?

Germany and Great Britain showed different patterns of modality reliance: Germans were faster in rating faces and particularly slow at rating emotional voices whilst British participants showed no significant overall modality effect for rating speed as already demonstrated in Chapter 2. Interestingly, a meta-analysis by Elfenbein and Ambady (2002) also suggested enhanced emotional face over voice processing across cultures, supporting present data for the German sample. In addition, the modality effect within the German sample was only true for male participants whilst all other participants (that is British men and women as well as German women) demonstrated modality-independent emotion processing speed. In fact, German men seemed to be particularly poor at recognising emotions expressed in the voices which probably drove the significant modality effect within the German sample. Hence, the similarity in processing speed across two modalities may to some degree not only depend upon culture but also on gender of participants.

In terms of intensity perception, the current data replicated modality effects for happiness, sadness, fear and surprise across both countries. This suggests a similarity in the pattern of emotion intensity perception across two modalities for two countries in parallel. However, German adults also showed additional modality effects for anger and disgust so there were significant modality effects for all six emotions in the German sample. Overall, for speed, the degree of modality similarity varied depending on the culture and participants' gender whilst intensity perception of both modalities seemed to be comparable across German and British participants.

Across all British and German participants, females perceived emotions as more intense and also categorised emotions faster than their male counterparts.

Interestingly, whilst emotion recognition in general seemed to be inferior when communicated by voices as compared to faces, this effect was enhanced for men when the voice expressed ‘happiness’ and for women when the voice expressed ‘anger’. Often, in face expression studies, ‘happiness’ is associated with a female face and ‘anger’ associated with a male face (Hess, Adams Jr, Grammar, & Kleck, 2009). For the present auditory data, the vulnerability of recognising emotions within the voice also seems to be increased by rating gender-stereotypical emotions which were opposite to the gender of the judge. This indicated that male participants are poor at identifying the ‘female’ expression of happiness in the voice and female participants are poor at identifying ‘male’ expression of anger in the voice, possibly lending support to Hess et al.’s (2009) finding from face expression data.

4.1.4.5 Conclusion

As summarized above, adult data from Experiment 4-I suggested that whilst there were universal patterns of emotion recognition across German and British participants such as happiness superiority in both countries, there was also evidence for subtle differences in intensity perception. British participants tended to perceive some emotions with higher intensities than German participants. In order to investigate how these cultural differences affect emotion recognition in childhood, the following Experiment 4-II will investigate British and German children’s abilities to recognise emotion from faces and voices.

4.2. Experiment 4-II

Children's' Emotion Recognition Across Cultures and Modalities

Following results from the adult Experiment 4-I, the interesting question emerges how those cultural differences in emotion recognition are being transferred across generations and how children learn to apply cultural norms in emotion understanding via socialization processes. For example, differences in primary school entry age (Eurydice National Foundation for Educational Research, 2013) between British and German children as well as in social norms discussed in Experiment 4-I show the importance of investing more into comparative research between similar North-western European countries.

As results from the child Chapter 3 indicated, emotion recognition in faces as well as in voices develop over the course of childhood in accordance with previous studies (*e.g.* De Sonneville, Verschoor, Njiokiktjien, Op het Veld, Toorenaar, & Vranken, 2002). However, little research has been conducted regarding cultural influences on emotion processing development in children – especially across more than one modality. McCluskey and Alba (1981) suggested that *vocal* emotion recognition abilities also increased with age in Mexican as well as Canadian subjects. Further, they also noted that Mexican children perceived emotions in speech more intensely than Canadian children did. In 1996, Markham and Wang investigated developmental differences in *facial* emotion processing between Chinese and Australian children aged 4 to 8 and also reported increased accuracy depending on age. However, they also suggested that children across both cultures differed in their emotion perception with Chinese children recognising facial emotions more accurately than Australian children. According to the Dialect Theory (Elfenbein, 2013; Elfenbein & Ambady, 2003), display rules, which encourage or discourage the expression of specific emotions within one culture, may be established early in life (McCluskey & Albas, 1981) due to parental socialisation.

What are the mechanisms behind cultural differences in emotion recognition not only in adults but also within children? Halberstadt and Lozada (2011) reviewed how culture impacts parents' socialization processes of infants' emotional development. They stated that cultural factors such as power distance within a family, the way a child learns, collectivistic versus individualistic values within a family and the value of emotional expression between family members all influences children's socialization processes and in turn their emotional development. The way children grow up within a family setting will influence future interpretations of emotional events and factors such as parent's beliefs or parent's own social behaviour affects children's emotion recognition skills (Castro, Halberstadt, Lozada, & Craig, 2014).

Hence, family expressiveness may be a product of affective display rules which in turn influences children’s emotion understanding. It is believed that emotional expressiveness in infants as young as three months old is related to the mother’s emotional behaviour which suggests very early socialisation of emotions (Malatesta & Haviland, 1982).

Further, the understanding of display rules and emotion regulation seems to increase with age throughout primary school (Jones, Abbey, & Cumberland, 1998) in line with general cognitive development. Lastly, apart from parenting style, media may also be driven by cultural norms and vice versa: American children books included more powerful and negative emotions like anger in their stories compared to Turkish or Romanian children books (Wege, Gonzales, Friedlmeier, Mihalca, Goodrich, & Corapci, 2014), confronting children with cultural emotion expectations from an early age onwards.

4.2.1.1. The Present Study

Following the above evidence, it seems that culture plays a role in children’s emotion development, although, again, most studies concerning emotional development in varying cultures are based on significant differences in display rules between two very different countries (*e.g.* Markham & Wang, 1996). No previous study has investigated the difference in children’s emotion perception across two similar cultures such as Germany and Great Britain. Hence, the aim here is to fill this gap by providing a cross-cultural comparison of emotion recognition development across German and British children aged 5 to 10. It is predicted that due to socialisation processes, German and British children will also exhibit subtle differences in emotion perception across modalities that correspond to the equivalent adult sample (Hypothesis 1).

4.2.2. Method

4.2.2.1. Participants

Child participants were 54 British children (23 male, 31 female), aged 5 to 10 (*M* = 7.52, *SD* = 1.72) from the child Chapter 3 (see Chapter 3 for more details) as well as 44 German children (15 male, 29 female), aged 5 to 10 (*M* = 7.36, *SD* = 1.63). Since there were 98 children all together, children were divided into three age groups: one younger group including children aged 5 and 6, one older age group including children aged 7 and 8 and one oldest child group with children aged 9 and 10. This included 32 children in age-group 1 (12 male, 20 female, *M* age = 5.47, *SD* = .51), 37 children in age-group 2 (16 male, 21 female, *M* age = 7.59, *SD* = .6) and 29 children in age-group 3 (10 male, 19 female, *M* age = 9.45, *SD* = .57). All three age groups were matched for age and non-verbal IQ between both countries resulting in non-significant differences between German and British children for any of the three age groups ($p > .05$). German children were recruited from two local primary schools and one local nursery. All participants reported normal or corrected-to-normal hearing and vision and English was the first language.

The study was approved by the Department of Psychology, Brunel University, PsyREC officer (17/10/2012) and informed consent as well as a child-friendly version of the consent form was filled in both by participants and parents prior to the study. All participants were debriefed after the study. Forms are detached in Appendix B.

4.2.2.2. Stimuli

See Chapter 3. It is important to note that children completed 2 (actors) x 6 (face expressions) x 6 (voice expressions) trials per modality to avoid loss of concentration, while adults completed 4 (actors) x 6 (face expressions) x 6 (voice expressions) trials per modality

4.2.2.3. Procedure

See Chapter 3.

4.2.2.4. Statistical Analysis

For children, between-subject variables were country (Great Britain or Germany) and age-group of participants (age-group 1, age-group 2 or age-group 3). The within-subject variables were modality (face or voice) and emotion category (happy, sad, angry, fear, surprise or disgust). For the current analysis, neither the variable of gender of participants nor gender of stimuli were included in order to limit a loss of power due to multiple comparisons.

4.2.3. Results

The following Table 4.3 displays Means and Standard Errors for measure of intensity and choice reaction times for British and German children.

Table 4.3

Means and Standard Errors for intensity ratings and reaction times for modality and emotions for British and German child participants

	Emotion	**Mean Intensity (*SE*)**			**Mean Choice RTs (*SE*)**		
		British Children	**German Children**	**All Children**	**British Children**	**German Children**	**All Children**
		N = 54	*N* = 44	*N* = 98	*N* = 54	*N* = 44	*N* = 98
		*M** = 7.52 (1.70)	*M* = 7.36 (1.63)	*M* = 7.45 (1.68)	*M* = 5.96 (0.51)	*M* = 18.6 (1.12)	*M* = 18.6 (1.12)
Face	*Happy*	6.14 (.18)	6.02 (.20)	6.08 (.13)	4.63 (.24)	4.82 (.27)	4.72 (.18)
	Sad	5.70 (.22)	5.13 (.25)	5.42 (.17)	5.03 (.19)	5.33 (.22)	5.18 (.15)
	Angry	4.98 (.21)	5.24 (.23)	5.11 (.16)	4.59 (.21)	5.16 (.23)	4.87 (.16)
	Fear	5.07 (.26)	4.66 (.29)	4.86 (.19)	4.49 (.21)	5.75 (.24)	5.12 (.16)
	Surprise	5.78 (.22)	4.41 (.25)	5.09 (.17)	4.17 (.20)	5.48 (.22)	4.83 (.15)
	Disgust	4.29 (.26)	4.72 (.29)	4.50 (.20)	4.69 (.22)	5.72 (.25)	5.21 (.17)
	All	5.33 (.11)	5.03 (.13)	5.18 (.09)	4.60 (.13)	5.53 (.16)	4.99 (.09)
Voice	*Happy*	5.61 (.22)	5.55 (.25)	5.57 (.17)	5.09 (.21)	5.66 (.24)	5.38 (.16)
	Sad	5.26 (.22)	5.33 (.24)	5.30 (.16)	6.15 (.26)	6.25 (.29)	6.20 (.20)
	Angry	3.43 (.23)	3.65 (.25)	3.54 (.17)	5.44 (.26)	6.05 (.29)	5.75 (.19)
	Fear	4.84 (.26)	4.79 (.29)	4.81 (.20)	5.63 (.26)	6.10 (.29)	5.86 (.20)
	Surprise	4.65 (.24)	3.88 (.27)	4.27 (.18)	5.01 (.21)	5.57 (.24)	5.29 (.16)
	Disgust	4.66 (.27)	4.95 (.30)	4.81 (.20)	5.87 (.27)	6.34 (.30)	6.11 (.20)

		Mean Intensity (*SE*)			Mean Choice RTs (*SE*)		
		British Children	**German Children**	**All Children**	**British Children**	**German Children**	**All Children**
	All	4.74 (.13)	4.69 (.22)	4.72 (.09)	5.38 (.14)	6.00 (.18)	5.76 (.12)
Overall	*Happy*	5.87 (.16)	5.78 (.18)	5.83 (.12)	4.86 (.17)	5.24 (.19)	5.05 (.13)
	Sad	5.38 (.18)	5.23 (.20)	5.36 (.13)	5.59 (.18)	5.79 (.20)	5.69 (.14)
	Angry	4.20 (.18)	4.44 (.20)	4.32 (.13)	5.02 (.20)	5.60 (.22)	5.31 (.15)
	Fear	4.95 (.20)	4.72 (.23)	4.84 (.15)	5.06 (.18)	5.92 (.20)	5.49 (.14)
	Surprise	5.22 (.19)	4.15 (.21)	4.68 (.14)	4.59 (.16)	5.52 (.18)	5.06 (.12)
	Disgust	4.48 (.19)	4.83 (.21)	4.65 (.14)	5.28 (.19)	6.03 (.22)	5.66 (.15)
Grand Mean**		5.03 (.11)	4.86 (.12)	4.95 (.08)	5.07 (.13)	5.69 (.14)	5.38 (.10)

Note. * = Mean Age in years. ** = Average scores across all participants and conditions.

a) Intensity Ratings

Results from intensity ratings from all children across Germany and Great Britain suggest that emotional intensity perception depended on a combination of emotions and modalities presented as well as the age-group and culture of participants. A mixed measure ANOVA with modality (2) x emotion (6) x country (2) x age-group (3) suggested a significant main effect for modality, $F(1, 92) = 33.82$, $p < .001$, $\eta^2 = .27$, for emotion, $F(5, 460) = 20.34$, $p < .001$, $\eta^2 = .18$, and for age-group, $F(2, 92) = 3.81$, $p = .03$, $\eta^2 = .08$. Faces were rated with higher intensity than voices and Sidak post-hoc test for multiple comparisons revealed that children associated happiness with highest intensity ratings which was higher than for all other emotions, then sadness, which was significantly different from all other emotions, and followed by fear, surprise and disgust. Anger received lowest intensity ratings which were significantly lower than for all other emotions, all at $p < .05$. Further, the middle age group with children aged 7 and 8 rated emotional stimuli with lowest intensity whilst the youngest (5 and 6-year old) and the oldest child group (9 and 10-year old) did not differ from each other in their intensity perception. For *M* and *SE* see Table 4.3.
There was no significant main effect for country of participants, $p < .05$

There was, however, a significant interaction effect between intensity ratings for (i) country and emotion, $F(5, 460) = 4.34$, $p = .001$, $\eta^2 = .18$, and between (ii) modality and emotion, $F(5, 460) = 9.12$, $p < .001$, $\eta^2 = .09$. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate the interactions.

i: Although there was no overall main effect for country, there was one exception and for surprise stimuli, British children perceived them with higher intensity than German children. For all other emotions, ratings did not significantly vary as a function of culture. Further, for both countries, happiness and sadness received highest ratings. For British, disgust and anger received lowest ratings whilst for the Germans, anger and surprise received lowest ratings. This is illustrated in Figure 4.6.

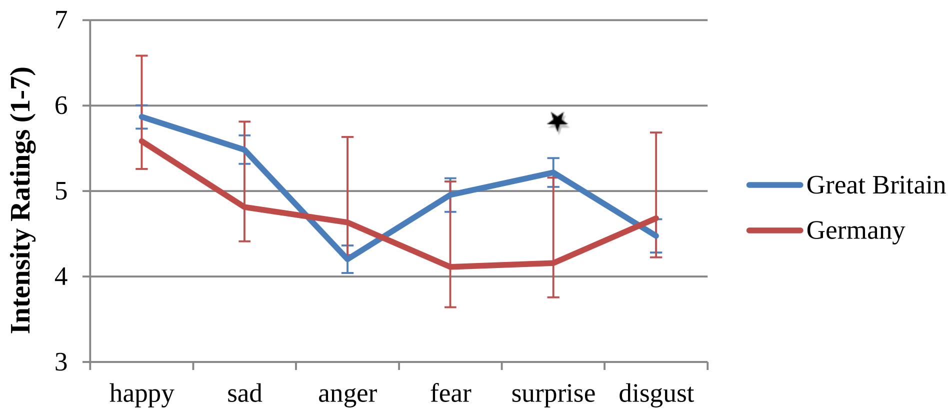


Figure 4.6. Means and Standard Errors for significant interaction effect between country and emotions within the child sample, *$p < .05$.

ii: Although there was a significant main effect for modality, this was only true for some but not for all emotions displayed. For happiness, anger and surprise, faces received higher intensity ratings than for voices, whilst for sadness, fear and disgust, voices and faces were perceived with similar intensity across all children. Further, for both modalities, happiness received highest ratings followed by sadness. For faces, disgust received lowest ratings whilst for voices, anger received lowest ratings.

There was also a significant three-way interaction effect between country, emotion and age, $F(10, 460) = 3.19$, $p = .001$, $\eta^2 = .09$. Post-hoc Sidak test for multiple comparisons was applied in order to further investigate this interaction. In the youngest child group, happiness, anger and disgust was perceived with higher intensity in the German as compared to the British child sample whilst surprise was rated higher by British compared to German children.

In the second child group, the German ratings decreased significantly for happiness, anger and disgust compared to the youngest child group. For happiness and surprise, British children rated emotions now as more intense than German children. In the oldest child group, the British intensity score for disgust has increased significantly compared to the middle child group and for Germans, ratings for surprise have also increased whilst ratings for anger decreased further. Overall, the oldest children from the British and German sample have now caught up with each other and had equal intensity perception scores for all six emotions. This is illustrated in Figure 4.7.

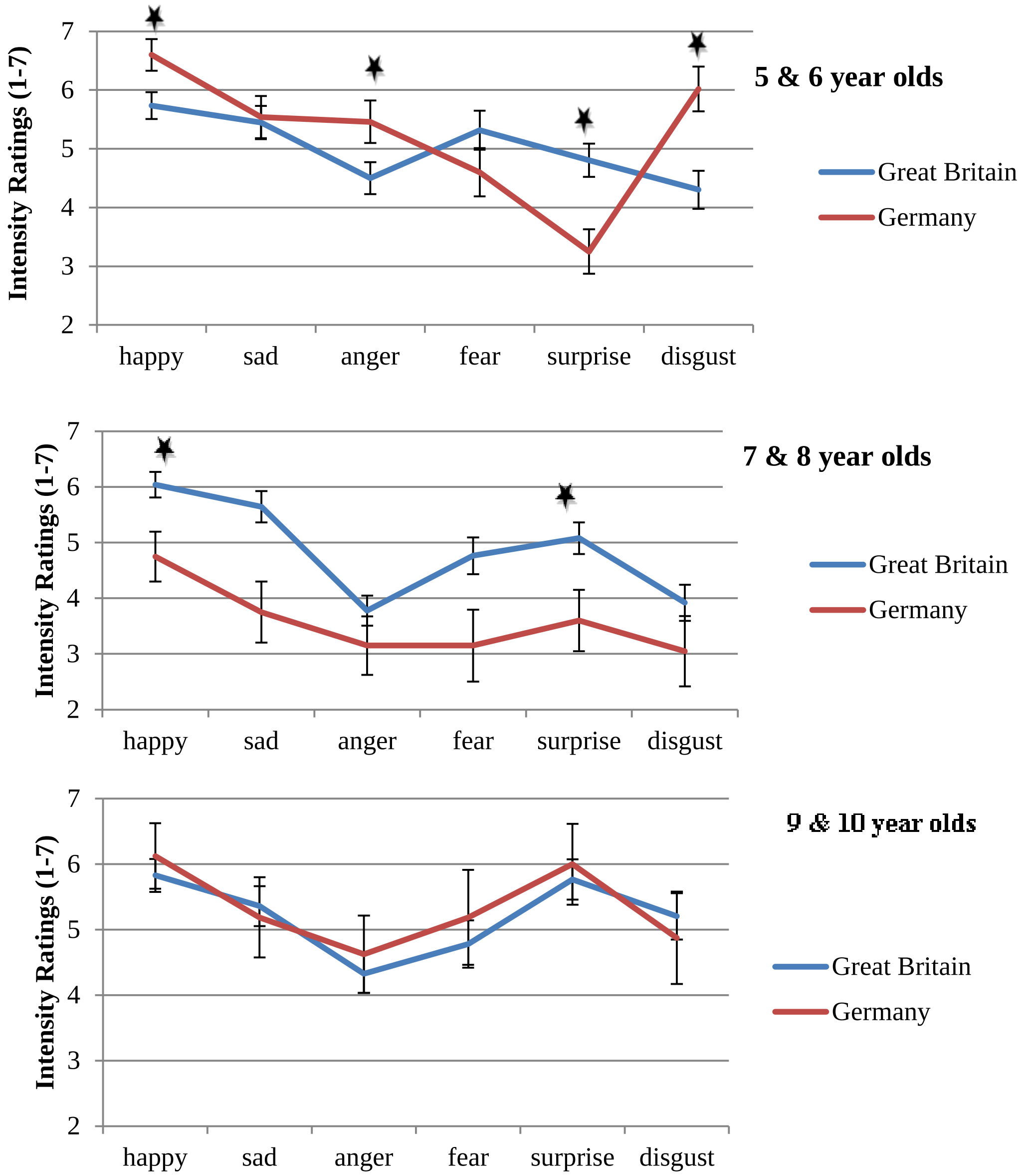


Figure 4.7. Means and Standard Errors for significant three-way interaction between age-group, country and emotions, $*p < .05$, $N = 98$.

b) Choice RTs

Results from RTs from all children across Germany and Great Britain suggested that speed for making decisions about emotional category membership depended on the country and age-group of participants as well as on the modality and emotions presented. A mixed measure ANOVA with modality (2) x emotion (6) x country (2) x age-group (3) suggested a significant main effect for modality, $F(1, 92) = 61.99$, $p < .001$, $\eta^2 = .4$, for emotion, $F(5, 460) = 7.4$, $p < .001$, $\eta^2 = .15$, for country, $F(1, 92) = 10.47$, $p = .002$, $\eta^2 = .1$, and for age-group, $F(2, 92) = 36.82$, $p < .001$, $\eta^2 = .45$.

Faces were rated faster than voices and Sidak post-hoc test for multiple comparisons revealed that judgements regarding happy and surprise stimuli were made fastest and did not differ significantly from anger. Sadness, fear and disgust were rated slower and anger and fear did not differ from each other significantly, all at $p < .05$. Further, British children rated emotional stimuli with faster speed than German children and speed increased linear with age: the youngest child group made slowest judgements, followed by the middle and by the oldest child group. The difference between all three age groups was statistically significant (for *M* and *SE* see Table 4.3).

There was also a significant interaction effect between modality and age, $F(2, 92) = 8.39$, $p < .001$, $\eta^2 = .15$. Post hoc Sidak test confirmed that across all children, there was a significant effect of modality in all three age-groups. Further, speed for rating faces decreased continually between all three age-groups and the youngest child group was slower than the middle child group which was in turn slower than the oldest child group. For voices, however, the youngest child group was slower than both older child groups but the two oldest child groups did not differ from each other for speed of rating voices. Whilst speed for rating faces continued to decrease across all age groups, speed for rating voices did not change between the age of 7 to 8 and 9 to10. This is illustrated in Figure 4.8.

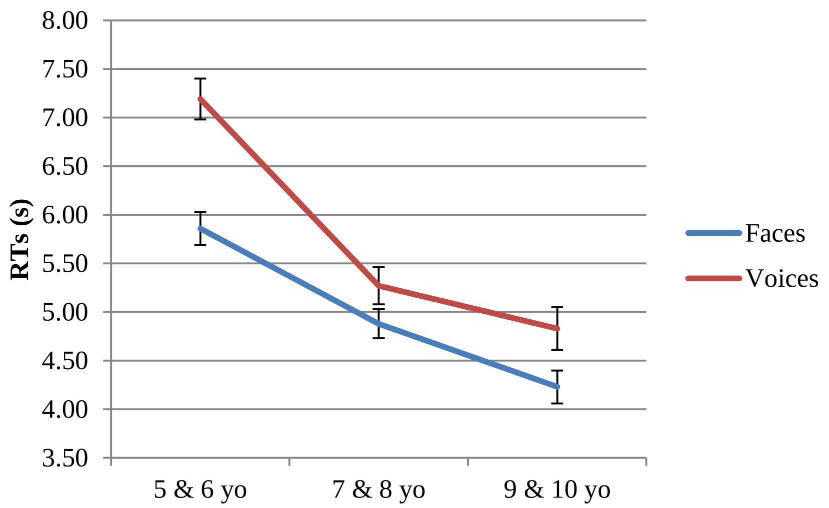


Figure 4.8. Means and Standard Errors for significant interaction between age-group and modality, $p < .05$, $N = 98$.

Further, there was a significant three-way interaction between modality, emotion and age-group, $F(10, 460) = 2.2$, $p = .02$, $\eta^2 = .05$. In the youngest child group, there was a significant modality effect for all six emotions displayed with voices being judged at a slower pace than faces. For the middle year group, all emotions displayed in the voice have now significantly faster choice RTs compared to the youngest child group and ratings for faces displaying sadness, anger, fear and disgust have also accelerated.

This means that in the middle child group, only sadness and fear show a significant modality effect whilst the other emotions of happiness, anger, surprise and disgust have equal choice reaction times. In the oldest child group, ratings happy and surprise in faces have significantly accelerated compared to the middle child group and for voices, sadness and fear also continued to be rated with faster RTs. This consequently results in no significant modality RT effects – apart from the emotion of disgust – in the oldest child groups, suggesting that the speed of rating voices has caught up with the faster speed of rating faces as a function of age in five out of six emotions. This is illustrated in Figure 4.9.

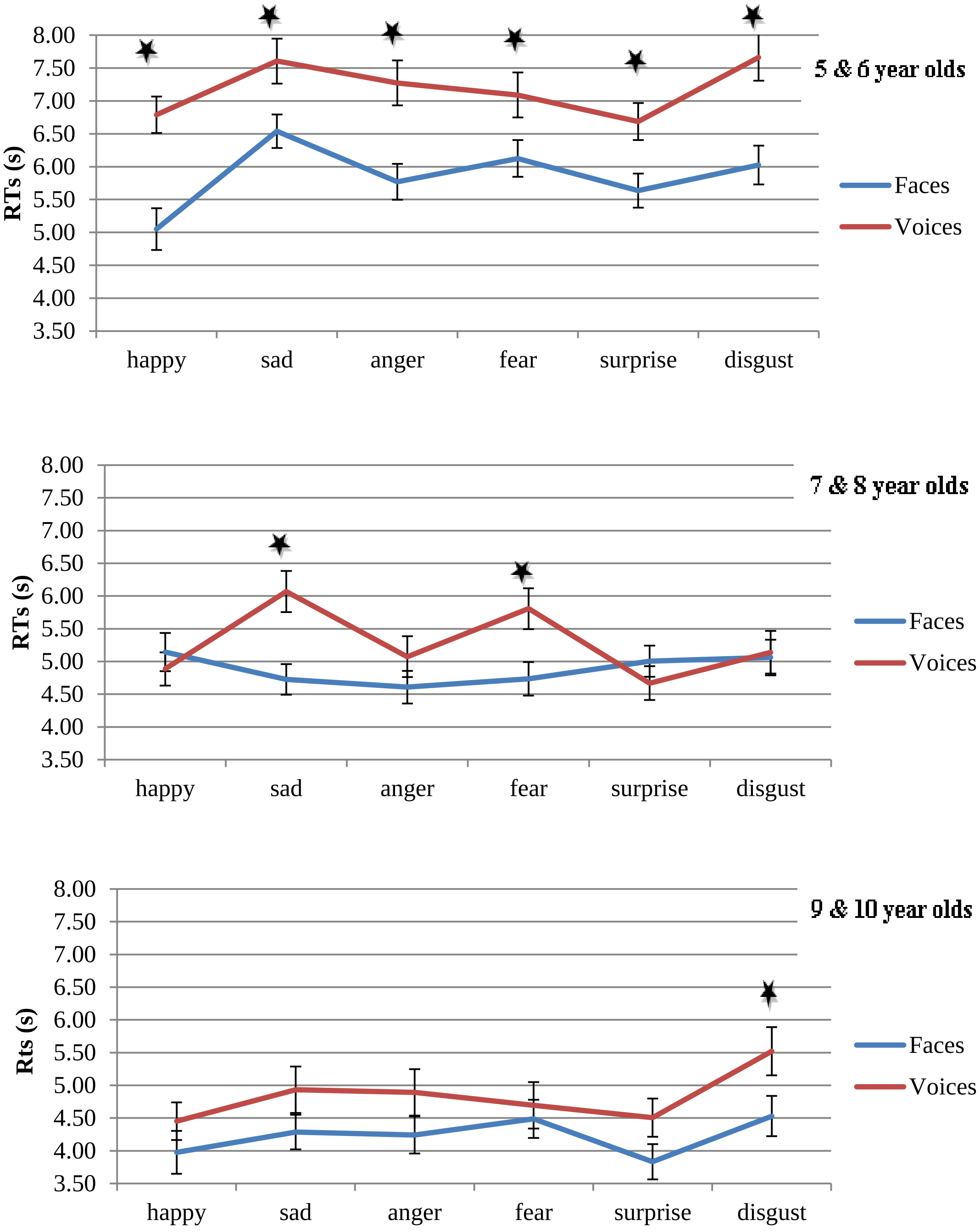


Figure 4.9. Means and Standard Errors for significant three-way interaction between age-group, emotions and modality, *$p < .05$. $N = 98$.

4.2.4. Discussion Experiment 4-II

The purpose of the present Experiment 4-II was to investigate whether potential cultural differences in emotion recognition are already seen within a group of children aged 5 to 10. If one was to follow Castro et al.'s findings (2014) of a link between family values and children's emotion recognition abilities, it is to expect that culture effects in adults (as demonstrated in Experiment 4-I) will introduce comparable culture effects within the child sample due to early socialisation processes. The present results suggested that culture only affected speed but not intensity ratings in children: British children rated emotions faster than German children. This finding was unexpected since there was no significant country effect for rating speed within the adult sample of the current study.

Further, the current data did not find significant country effects on emotion intensity perception in the child sample – contrary to rating speed. The similarity in the perceived emotion intensity between German and British children suggests that children were not prone to significant culture effects. However, 'surprise' was the only exception and British children rated surprised stimuli with higher intensity than German children – this effect is the same direction as seen for the adult group which higher intensity ratings from British compared to German adults. Consequently, the lack of significant effects on intensity perception between British and German children conflicts with previous findings by McCluskey and Albas (1981) who suggested early cultural effects in children from Mexico and Canada. It is possible that cultural values between German and British families are very similar and differences are too subtle in order to significantly interfere with children's emotion intensity perception.

When looking at the different age-groups within the child-sample in more detail, it becomes evident that there were, however, subtle differences in emotion recognition between cultures. Children across Germany and Great Britain differed in their intensity perception depending on the emotion displayed and the age of the judges. For the youngest child group including children aged 5 to 6, for four out of six emotions, ratings differed as a function of culture. For happiness, anger and disgust, young German children associated stimuli with higher intensity ratings whilst for surprise, young British children rated stimuli with higher intensity than German children. For the middle group with children aged 7 and 8, the difference between cultures was reduced to two out of six emotions whilst the oldest child-group with 9 and 10-year old children did not show culture effects for any of the six emotions displayed. Overall, whilst youngest children showed some variation for emotion intensity perception between countries, this variability decreased with age and ratings became more stable across both samples.

Data from the present experiment supported findings from the child Chapter 3 on the role of modality and emotion recognition during childhood: modality differences diminished with age. Across all children, for the youngest age-group, faces were perceived with higher intensity than voices for all six basic emotions – although the pattern for each modality was similar with faster ratings for happiness and slower ratings for negative emotions in both modalities. For the following age-group including children aged 7 and 8, this face superiority was only visible in two out of six emotions, namely sadness and fear. For all other emotions, speed for voices has decreased and caught up with rating speed for faces. For 9 and 10-year old children, faces and voices were rated with the exact same speed for all six emotions, indicating modality-independent processing speed in the oldest child-group. This pattern for rating speed did not interact with country of participants. For further discussion and interpretations of this finding, the reader is referred to Chapter 3.

Further, children rated happiness, anger and surprise with higher intensity when they were displayed in faces compared to voices. Again, this was independent of the culture of participants. Interestingly, this exact same pattern was visible in adults suggesting that children demonstrate some adult-like emotion processing although at a less sophisticated level such as with slower speed. For sadness, fear and disgust, children did not show perceptual differences in intensity between voices and faces, suggesting comparable processing from the auditory and facial domain whilst adults exhibited voice superiority effects for rating sadness and disgust.

4.5. General Discussion

Overall, the present Experiment 4-I and 4-II in Chapter 4 are the first to behaviourally investigate whether there are differences in basic emotion recognition between two similar Western-European countries across two modalities. Current result demonstrated that whilst general recognition of the six basic emotions showed universal and country-independent processing pattern, at a more subtle level, culture seemed to influence intensity perception as well as choice speed. British adults in Experiment 4-I perceived half of the displayed emotions as more intense than German adults did; although, this varied depending on whether emotions were expressed visually or auditory. Germans seemed to react specifically to angry faces and perceived them as more intense than British adults. This seemed to be unrelated to the self-reported emotional temperament in participants as measured here by the Emotional Reactivity Scale (Nock et al., 2008). Whilst speed for rating emotions did not differ in the British sample as a function of modality, for Germans, voices took particularly longer to be rated, suggesting that different cultures may have a modality preference during emotion processing.

Within the child Experiment 4-II, for five out of six emotions, children had comparable intensity rating across both cultures and only for surprise, British children perceived stimuli more intense than German children. This pattern suggests that children are less susceptible to cultural effects and socialisation of social norms between German and British children did not seem to influence emotion recognition. This, however, was influenced by age and younger children showed more cultural variation than older children. Similarly to findings from Chapter 3, modality effects for rating speed seemed to be more pronounced in younger children and diminished with increasing age.

The direction of emotion recognition differences with higher intensity ratings for British compared to German adults supports the notion of the ‘reserved German’ as has been proposed by previous business leadership (*e.g*. Brodbeck, 2000) and communication (*e.g*. Wouters, 2011) studies: Germans tend to value hierarchy and directness rather than warmth during interpersonal relationships. As suggested by Schneider et al. (2013), display rules such as suppressed emotion expression may indeed be linked to emotion recognition abilities. Cultural differences such as a lower score on individualism in Germans (Hofstede, 2001) could indeed be a source of differences in emotion recognition across cultures. Note, however, that the present study did not specifically measure the existence of display rules or actual differences in face expressions (as investigated by Elfenbein et al., 2007) between both participating countries.

The finding of reduced emotion intensity perception in Germans that applies not only to faces but also to voices (with the exception of anger) is a striking and novel finding, which, however, may depend on specific emotions displayed. In accordance with facial feedback theories (*e.g.* Neal & Chartrand, 2011) it has previously been found that vocal emotion expression can also influence emotion experience.

For example, listening to and reproducing positive emotional sounds magnified the experience of positive feelings (Hatfield, Hsee, Costello, Schalekamp Weisman, & Denney, 1995). Consequently, it would be interesting to see whether suppressing vocal mimicry – as opposed to just replicating vocal sounds – could also reduce auditory emotion recognition as shown from facial feedback theories (Oberman et al., 2007).

4.5.1. Development of culture effects across age: When do display rules kick-in?

Whilst children in the youngest group did show rating differences between cultures, it is possible that this variation is the result of noise due to additional factors such as differences in schooling systems (see below) or generally large individual differences in cognitive abilities such as Theory of Mind abilities (Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991). With increasing age and equal school attendance for children across both countries, this variation decreased and children across two separate cultures seemed to catch up in their emotion recognition skills. By the time children reached the end of primary school, emotion recognition capacities in British and German children were at a comparable level and each of the six emotions were being perceived without significant differences. Hence, there were no cultural differences within older children.

When looking at the oldest child-group (9 and 10-year old children) across cultures in the present chapter, it becomes evident that children demonstrated adult-like emotion intensity perception with almost identical intensity ratings for all six emotions. Following the current data, one possible interpretation is that cultural display rules do not contribute significantly to emotion recognition in children. In support, explicit understanding of display rules and emotion regulation develops with age throughout childhood (Jones et al., 1998). It is possible that the subtle culture effect in emotion recognition between German and British children only kicks in during teenager years in line with an awareness of cultural display rules. Indeed, even for undergraduate college students, expressiveness of social environment such as families continued to affect non-verbal communication skills in adolescence (Halberstadt, 1986). This finding suggests a long-lasting relationship between display rules and emotion development.

So where does the cultural variation for the youngest child-group in the present chapter come from? One explanation could be differences in schooling systems: Whilst 5-year olds in Great Britain are already in primary school, in Germany they typically still visit community-run nurseries until the age of 6 (Eurydice National Foundation for Educational Research, 2013). The 5-year old German children tested in the present study attended nursery on a half-day basis and therefore, it is possible that the youngest children in the German sample – which were the slowest out of all groups - are driving this significant difference in rating speed between countries. In Germany, 80% of all nurseries only provide half-day care (Burger, 2010). Consequently, many toddlers spend a large proportion at home with the primary caretaker.

Indeed, within a British sample, attending preschool has been found to have a positive impact on later cognitive and social development (EPPA project, Sylva & Britain, 2003) but long-term effects of attending preschool on later school performance within a German sample were not evident (Burger, 2010; Spiess, Büchel, & Wagner, 2002). This suggests that there could be a link between British system of early preschool attendance and faster rating speed for emotions.

Further, because this study was based on a computer-task, it is possible that British children spend more time with media such as play consoles whereas Germans value free play time and outdoor play as evident from the existence of over 1500 'Waldkindergarten' (forest nurseries) that provide child-care based in the outdoors rather than inside buildings (Bundesverband der Natur – und Waldkindergaerten in Deutschland e.V. 2014). Young German children may simple be not as familiar with handling computers as young children in Britain.

4.5.2. Implications

The present data suggested for the first time that even across two similar Western-European countries, there may be small differences in the way emotions from voices and faces are perceived. Because in the present chapter, Germans perceived certain emotions as less intense than their British counterparts, it is possible that this pattern reflects the German reservedness. For cross-cultural interactions, this suggests that emotional content during phone as well as face-to-face interaction may be interpreted in different ways, depending on the cultural background of business partners. As suggested by Elfenbein et al. (2007), the link between emotion recognition abilities and successful business negotiations explicitly demonstrated the importance of investigating differences in emotion recognition skills between trading partners such as Germany and Great Britain.

The effectiveness of affective advertisements such as production of consumer guilt can also be influenced by cultural norms – and the perceived guilt in turn influences purchase intention (Kim & Johnson, 2012). For example, individualistic nations such as the United States value adverts that target individual benefit whilst collectivistic cultures value products that enhance benefit for a larger group (Han & Shavitt, 1994).

Here, the researcher proposes that the perceived intensity of emotions communicated – which is guided by cultural norms - via adverts such as the radio or billboard could influence clients purchasing or donation behaviour. For example, the same picture of a child in distress may elicit higher perception of sadness in one compared to another culture; this may increase guilt perception and in turn donation intentions. In the present study, British participants perceived happiness, fear and disgust as more intense than German samples. Those emotions are prime example for communicating intentions in advertising: understanding of happiness in others may be targeted in humorous adverts or to demonstrate positive benefits of purchasing items.

Emotions of disgust can be associated with adverts in health-settings such as ‘NHS SmokeFree’ (Department of Health, 2014) or ‘Think!’ road traffic campaigns (Department for Transport, 2014) and fearful expressions can be used to transmit topics of violence and abuse as often used by charity appeals. Additionally, due to the modality effect seen in the German sample, purchasing or donation behaviour may benefit from displaying affective advertisement in the visual domain such as billboards as compared to radio – this may be especially true for intentions of communicating anger.

Results from the present Experiments 4-I & 4-II in Chapter 4 are interpreted on the assumption that there are cultural differences from previous leadership and communication studies due to underlying display rules. The researcher relied on cultural differences between Germany and Great Britain such as individualism as demonstrated by Hofstede’s cultural values survey (2001). However, it is beyond the scope of the current study to measure actual differences in display rules and social norms. There may be many other variables – apart from display rules and social norms – that may differ from country to country - which could cause variation in emotion recognition such as family size (Morand, 1999) or time spent in front of screen-based media (Uhls et al., 2014). Those factors were not controlled for in the present study and so the present study reported subtle cultural differences in emotion perception but cannot prove the importance of display rules as cause for variation. Hence, future studies would include measures such as the human value survey, summarising attitudes towards values such as authority, tradition such as acceptance of norms or security (Schwartz, 1992) across Germany and Great Britain.

4.5.3. Conclusion

Overall, Chapter 4 demonstrated that cultural differences may influence emotion recognition in adults. This pattern was not as clear within the child sample and it is possible that socialisation processes of emotions only kick in after the age of 10. Overall, the present data seems to support the dialect model of cultural emotion recognition (Elfenbein & Ambady, 2003) with universal features that are influenced by cultural variation.

So far, this thesis has tackled the issue of emotion recognition across modalities from a behavioural perspective. In order to investigate whether there may also be common underlying similarities in the temporal processing of emotions across modalities, the following Chapter 5 will present neurophysiological ERP data on the exact issue.

5

An ERP study of emotion processing within faces and voices: How similar are modalities?

Having demonstrated similar patterns of emotion recognition mechanisms across modalities in the previous behavioural chapters of this thesis, the question emerged whether the underlying temporal hierarchy of neural processes may also be comparable across two separate modalities. In the past, there has been conflicting evidence regarding the degree of similarities across modalities due to a large variation of methodologies deployed. Further, the lack of within-subject designs limits the comparability of emotion recognition across modalities. Hence, the main objective of this present EEG (electroencephalogram) study is to collect some data from event-related potentials (ERPs) on the temporal features of emotion processing across two separate and independent modalities in a within-subject design. By comparing typical emotion processing stages for faces and non-verbal vocalisations, the researcher wants to examine whether both modalities have a comparable pattern of emotion processing in terms of timing of neural mechanisms.

5.1.1. Faces

5.1.1.2. Early processing: 70 – 200ms

Early visual perception is believed to happen from 50ms after stimulus onset onwards at occipital and posterior sites and reflects visual processing within the primary and extrastriate visual cortex as well as parts of the fusiform gyrus (Di Russo, Martinez, Sereno, Pitzalis, & Hillyard, 2002). During this early visual processing, it is not thought that the brain pays much attention to the emotional content of stimuli (*e.g*. Krolak-Salmon et al., 2001) and common onset of emotion recognition seems to be between 200-400ms after stimulus onset (see below). On the other hand, Eimer and Holmes (2002) showed that fearful as compared to neutral face expressions were processed as early as 120ms after onset in frontocentral sites with increased positivity for fearful faces. Batty and Taylor (2003) have also reported emotion- related waveforms that were evident as early as 90ms after onset of emotional faces – this, however, did not differ for specific emotions in question.

A negative peak at around 170ms (N170) after stimulus onset has also previously been associated with emotion processing.

This lateral negative peak at occipito-temporal sites between 140 and 200ms is commonly larger for faces compared to other visually evoked responses and may have a counterpart with positive polarity at more central sites (Bentin, Allison, Puce, Perez, & McCarthy, 1996; Eimer, 2011). Kanwisher et al. (1997) originally reported a brain area called the fusiform face area (FFA) that specializes in processing faces as compared to objects. The N170 activity is thought to reflect this face-processing activity in lateral occipito-temporal regions such as the occipital and fusiform face area as well as superior temporal sulcus (Eimer, 2011).

The role of the N170 component in regards to emotion processing in faces has not yet clearly been demonstrated. Several studies have reported that early face processing in the N170 time-range is independent of emotional content. For example, the N170 did not show emotion-specific responses and ERP amplitudes as well as latencies were comparable for different emotional as well as neutral faces in two studies conducted by Eimer and Holmes (2007) as well as Eimer, Holmes and McGlone (2003).

On the other hand, studies have demonstrated that early face processing can also be affected by emotional expressions. For example, Blau, Maurer, Tottenham and McCandliss (2007) reported an enhanced deflection at 170ms after stimulus onset for fearful compared to neutral faces. Marinkovic and Halgren (1998) also found emotion effects in faces as early as 170ms after stimulus onset at temporal sites for positive versus neutral faces. Batty and Taylor (2003) demonstrated that N170 waveforms even differed depending on the very specific emotion displayed. For example, at 140ms after stimulus onset, there was a later effect for fearful and disgusted faces as compared to positive emotions such as happy or surprised faces.

In support, Pizzagalli, Lehmann, Hendrick, Regard, Pascual-Marqui and Davidson (2001) used tomographic source localisation to report early activity within the fusiform gyrus – which has previously been associated with face encoding mechanisms (Bruce & Young, 1986) – at around 160ms post stimulus onset during the processing of liked versus disliked faces. Further, recent neuroimaging evidence also suggested shared activity in the posterior STS during the processing of face expression as well as face identity (Baseler, Harris, Young, & Andrews, 2014). This suggests that the STS may also modulate the perception of emotion to a certain degree. However, it is important to note that there are also studies that argue that the earliest signs of emotion processing are only visible after 200ms (*e.g*. Krolak-Salmon et al., 2001; Marinkovic, & Halgren, 1998) and hence separated in time from general face perception.

5.1.1.2. Late processing: 300-600ms

Neural processes which reflect processing for individual emotions may only reliably be found during processing stages that include higher cognition such as situation appraisal processes. Indeed, previous studies have reported emotion-specific effects at even later latencies, starting at around 400ms.

For example, Krolak-Salmon et al. (2001) reported different ERP waveforms for happy and fearful versus disgusted faces at around 550 to 750ms after stimulus onset in right posterior-temporal brain regions. Additionally, later waveforms for disgust seemed to be emotion-specific and in this case appeared between 700 and 950ms in more frontal regions. According to Krolak-Salmon et al. (2001), deep subcortical emotion structures such as the amygdala or basal ganglia may be responsible for a more wide-spread activity distribution on the scalp and even feed back to the extrastriate cortex in a top-down manner (see also Sato et al., 2001).

However, it is debatable how reliable those apparent emotion-specific ERPs really are. Eimer et al. (2003) for example did not find any evidence at all for emotion-specific ERP waveforms. Instead, it was suggested that - although all six basic emotions demonstrated typical ERP waveforms with an early frontocentral (120-180ms after stimulus onset) and a later posterior deflection (250-1000ms after stimulus onset), this pattern of waveform was comparable across the six basic emotions. In contrast to emotion-specific ERP waveforms reported by studies cited above, results from Eimer et al. (2003) hinted towards a more emotion-general emotion processing mechanisms, independent of the presentation of specific emotions.

Following previous – partly conflicting – results, the debate remains about the exact timing of emotion recognition onset. Some studies have reported early differentiation which may even occur before the face-specific N170 response (Batty & Taylor, 2003) whilst others report late emotion processing that follows structural encoding in a hierarchical manner (Sato et al., 2001). Secondly, it is not clear when time-critical processes within the human brain start to categorise between distinct emotion-categories. Assuming modality-independent emotion mechanisms as suggested by previous behavioural chapters, it is of considerate interest to investigate whether a second, independent modality such as non-verbal affect bursts also shows distinct temporal stages for emotion processing that may be comparable to neural patterns within the face.

5.1.2. Voices

5.1.2.1. Early processing: 70 – 200ms

The auditory signal rapidly travels from the ear to the thalamus and the primary auditory cortex (Goldstein, 2001). Early auditory processing happens within the first 50ms and reflects activity within medial geniculate nucleus and primary auditory cortex (Luck, 2005). Initial acoustical analysis of sounds is often reportedly found at fronto-central sites from 50ms onwards. In order to specific an ERP time-window that is specific to the auditory processing of human voices, Charest and colleagues (2009) have compared the temporal dynamics of processing human voices, bird voices and environmental sounds.

They reported an increased positivity at frontal sites which was typically accompanied by a posterior negativity that peaked at occipital sites at 164ms after stimulus onset and which was specific to human voices. This early voice-sensitive peak has been referred to as fronto-temporal positivity to voices (FTPV) and its latency can be compared to the face-sensitive N170 component which typically occurs in a similar time range (Charest et al., 2009). The FTVP is thought to reflect temporal voice areas within the right anterior superior temporal sulcus (STS, Charest et al., 2009) which may reflect the auditory 'what'-pathway that connects the anterior superior temporal gyrus with the orbitofrontal cortex (Rauschecker & Tian, 2000).

Similarly to the N170 component during face processing, the important question emerges whether components in the time-window that is thought to process human voices is also modulated by emotional content. In support of pre-emotional processing of voices during the first 100-200ms, Chronaki et al. (2012) found a negative deflection between 90 and 180ms specific to human voices in 6 to 11 year old children which did not differ as a function of emotional prosody.

However, there is also evidence which suggests earlier processing of emotional information in human voices. Iredale, Rusby, Mcdonald, Di Marco and Swift (2013) reported that the early voice processing stage may in fact already be modulated by emotional content: Waveforms differed for emotional versus neutral voices in parietal areas as early as 100ms after stimulus onset. At this early stage, however, there was no evidence for differentiation between different classes of emotions. Similarly, Sauter and Eimer (2009) also suggested early processing of fear, achievement and disgust versus neutral affect vocalisations which were reflected in an enhanced positivity at fronto-central electrodes between 150 and 300ms. Overall, as pointed out by Eimer and Holmes (2002), this early processing of emotional voices from 150ms onwards is comparable with data from emotion face processing that shows early emotional versus neutral differentiation at around 150ms after stimulus onset.

5.1.2.2. Late processing: 300 – 600ms

Following previous findings, it does not seem likely that differentiation between specific emotion categories such as happiness or anger happens before 200-330ms after stimulus onset. A study that investigated non-verbal affect vocalisations as compared to semantically or verbal loaded stimuli suggested emotion differentiation between distressed or joyful exclamations from 300ms onwards (Bostanov & Kotchoubey, 2004). Similarly, in 6 to 11 year old children, Chronaki et al. (2012) found a larger effect between 380 and 500ms to angry versus happy or neutral voices.

After establishing whether a voice is of emotional significance within the temporal voice area at around 150 - 200ms, the acoustical signal is then being passed on to more frontal areas like the prefrontal cortex.

This information transmission might reflect travel along the auditory 'what pathway' to prefrontal areas for higher cognitive analysis such as context interpretation (Rauschecker & Tian, 2000). Due to conscious analysis of the emotion stimulus, it may now possible to actively discriminate between several emotional states. Indeed, this statement has been supported by Iredale and colleagues (2013) who reported an enhanced negative deflection for happy versus angry voices at frontal sites between 400 and 650ms which possibly indicated reallocation of cognitive resources.

Following the review on auditory emotion processing, the exact timing onset of emotion processing within the human brain remains unidentified. Previous studies have suggested early emotion-processing that coincided with the general human-voice processing response (Eimer & Holmes, 2007) whilst other reported latencies up until 300ms after stimulus onset crucial for emotion processing (Paulmann & Kotz, 2008). Further, individual emotion discrimination did not seem to happen before 300-400ms post stimulus onset (Iredale et al., 2013).

5.1.3. Hierarchical Emotion Processing

Following past research, it seems possible that emotion processing happens in a hierarchical manner: from sensory processing to broad emotion discrimination finishing with narrow emotion-category distinguishing. For voices, this has previously been summarised by Schirmer and Kotz (2006), who have created a working model of temporal emotion processing stages for prosody based on previous research (for illustration purposes see Figure 5.1 from Schirmer & Kotz, 2006).

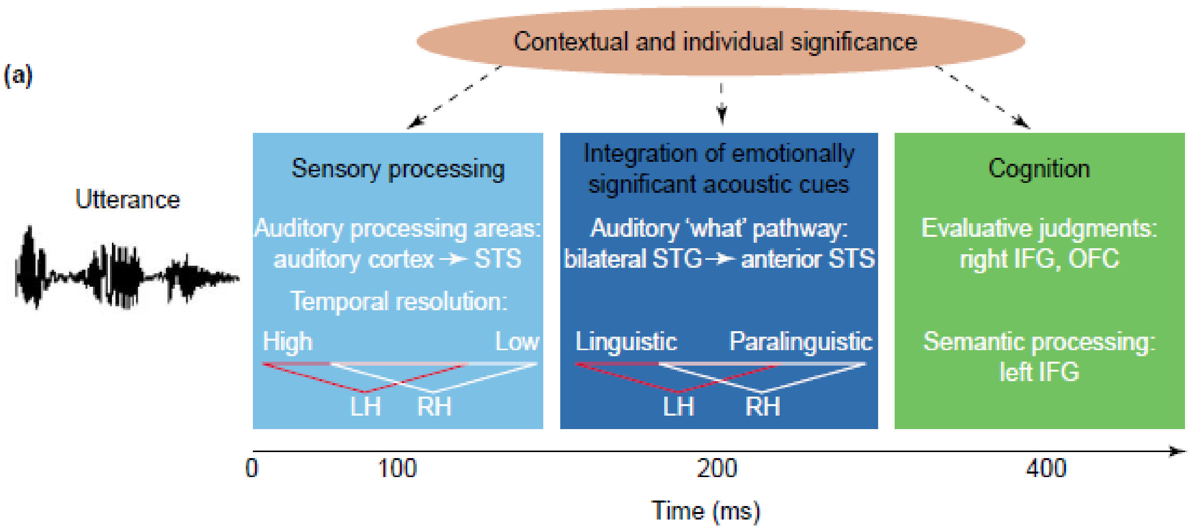


Figure 5.1. Hierarchical three-stage working model of emotion recognition for prosody (Schirmer & Kotz, 2006)

According to Schirmer and Kotz's (2006) three-stage model, the first voice processing stage peaks at around 100ms and defines the analysis of acoustical features such as sound intensity during initial sensory processing of voices as reported by Engelien, Schulz, Ross, Arolt and Pantev (2000). This stage would not be modulated by emotional content. Next, early emotion effects visible during the discrimination of emotional (*e.g*. angry) versus neutral prosody reflect emotion significance evaluation within the superior temporal sulcus (STS, Grandjean et al., 2005).The final stage then describes discrimination between individual emotion categories such as anger and happiness at around 400ms which includes higher cognition and conscious semantic processing of prosody in more frontal areas. For example, the inferior frontal gyrus (IFG) in the frontal lobe has previously been found to be active during the processing of semantic information (Schirmer, Zysset, Kotz, & von Cramon, 2004).

In support, Iredale et al. (2013) have recently confirmed Schirmer and Kotz's (2006) hierarchical emotion processing model in voices. Their findings from an ERP study suggested early emotional versus neutral discrimination of prosody in parietal regions, followed by more specific discrimination of happy and angry prosody and an activation shift to more frontal brain regions during the final stage of semantic processing. However, in contrast to Schirmer and Kotz (2006) – Iredale et al. (2013) reported emotion versus neutral prosody discrimination that was already visible during the first processing stage at around 100ms after stimulus onset. Note, however, that the auditory stimuli used in the current study are not based on semantic or verbal content.

Overall then, it seems likely that for voices, there may be three distinct stages of emotion processing: initial sensory processing which may are may not communicate emotional intent, a subsequent stage specific for the perception of emotion and a final stage of emotional evaluation that allows the differentiation between different emotion categories. Do those three primary stages of emotion processing also apply to emotions processed in non-verbal affect bursts and in faces? Some of the past research certainly suggests hierarchical processing of emotion across modalities: For example, Batty and Taylor (2003) reported earliest emotion effects in face expressions within the first 100ms, followed by specific emotion discrimination after 140ms after stimulus onset and a third stage which was defined as late frontal activity after 330ms. This late stage seemed to be biased towards processing negative emotions expressed in faces. Assuming an emotion-general mechanism across modalities supports the idea of a hierarchical three-stage model of emotion processing that may be independent of modality. However, to the best of the researcher's knowledge, this claim of hierarchical emotion processing across several modalities has not previously been tested in a within-subject design.

Indeed, no one has previously investigated in a within-subject ERP design whether emotional stimuli derived from different sensory modalities show comparable temporal patterns of processing. Very recently, Bayer and Schacht (2014) attempted to investigate underlying event-related potentials in response to rating emotional faces, pictures and written words.

To the best of the researcher's knowledge, this is one of the first accounts to compare three different types of emotion stimuli in a within-subject design during an ERP study – although all within the visual modality. Promisingly, results suggested that all stimulus types showed comparable emotion-typical activity in early posterior negativity component (EPN) as well as in the late positive component (LPC). However, whilst words showed superior processing for happy stimuli as evident in larger amplitudes for positive words, there were larger amplitudes for angry faces and pictures eliciting anger which suggests stimulus-specific processing differences at a lower processing level. Further, by distinguishing intact faces and words from distorted faces or pseudowords (Schacht & Sommer, 2009), increased posterior negativity at occipito-temporal sites was evident for positive versus negative and neutral words between 388 and 438ms as well as for positive versus negative and neutral faces between 128 and 172ms.

Hence, those previous results are promising as they suggest that different stimulus types within the visual modality have comparable stages of emotional processing; however, depending on the nature of stimulus, this may occur at different latencies. Consequently, although auditory and visual stimuli differ significantly in their basic perceptual features, it is of interest whether emotion signals within voices and faces may be processed on a supra-modal level with comparable emotion-processing stages.

5.1.4. The Present Study

As evident from past research, several processing similarities between emotions presented in faces and in voices have been established. For example, not much emotion activity seems to be noted within the first 100ms after stimulus onset in either modality. Secondly, both the face and the voice-sensitive stages N170 and FTVP may – or may not – already be modulated by emotional content of the stimulus. Lastly, both modalities have been shown to process emotions in a hierarchical manner, starting with broad discriminations between emotional and neutral stimuli and narrowing down to individual emotion category distinction at around300ms after stimulus inset. This may involve a shift from sensory and perceptual to more cognitive emotion processing.

The main research aim of this present study is to investigate in a within-subject design whether the pattern of neural activity is comparable across two separate modalities. To keep variance between modalities limited, all stimuli were based on non-verbal emotion communication. In line with behavioural findings from the previous chapters in this thesis, shared cognitive mechanisms should result in comparable emotion processing patterns across early and late time-windows, independent of the modality of presentation.

Further, according to Schirmer and Kotz's (2006) model of hierarchical emotion processing in voices, it is of interest whether both modalities show a hierarchical processing of emotion information, starting with broad processing such as emotional versus neutral discrimination and proceeding to more narrow emotion processing such as the discrimination between happy and angry stimuli.

In order to determine the onset of emotion processing abilities in faces as well as in voices, three time-windows have been determined for the present investigation. The first stage included a pre-emotion period, ranging from 70 to 130ms which is thought to reflect sensory processing in the P1 component (Luck, 2005). Further, in order to investigate whether the voice- as well as face-specific components N170 and FTVP are moderated by emotional content, the second stage was defined between 130 – 200ms. Lastly, in order to investigate the onset of emotion-specific discrimination such as between happy and angry stimuli, the late processing stage was defined between 300 and 600ms post stimulus onset.

The current within-subject study predicts the following:

Hypothesis 1: In line with supra-modal emotion networks (see behavioural results from previous chapters), emotional faces as well as emotional non-verbal vocalisations will show comparable temporal stages of emotion processing across all three time windows.

Hypothesis 2: According to the hierarchical model of prosody processing (Schirmer & Kotz, 2006), earliest emotion processing will not occur within the first 100ms after stimulus onset and then begin with broad discrimination between emotional and neutral stimuli. Later cognitive stages after 300ms after stimulus onset will then show narrower emotion discrimination between angry and happy stimuli. Again, this pattern will be visible within both modalities.

5.2. EEG Methods & Task

5.2.1. Participants

The participants included in the present study were 13 healthy male adults and were either British or have lived in the UK for at least the previous four years to ensure good use of the English language. After controlling for outliers for each of the variables, two participants with extremely noisy data such as constant eye movements were removed from the analysis, resulting in a final sample of 11 male participants (*M* age = 28.9, *SD* = 7.04) with an age ranging from 22 to 47 years. Ten out of 11 participants were white-Caucasian and one was Black-African. Participants were recruited from Brunel University or online social networks and all were in higher education. Ten participants were right handed and one was ambidextrous. Participants had normal or corrected-to-normal eyesight and hearing. Participants with corrected vision were asked to wear their prescription glasses rather than contact lenses in order to reduce excessive blinking due to dry eyes as a result of wearing contact lenses.

The study was approved by the Department of Psychology, Brunel University, PsyREC Committee (07/05/2014) and informed consent was provided before and debriefing after the study (see Appendix C).

5.2.2. Task & Material

The task was programmed on the PsychoPy software in Python (Peirce, 2007), and visual stimuli were chosen from the Ekman Pictures of Facial Affect series (Ekman, 1976); for more details on stimuli see Chapter 1. Five male and five female actors (identity: A, C, EM, JJ, MF, NR, PE, PF, SW & WF) displaying emotional (50% angry and 50% happy expressions) or neutral face expression based on the FACS system were chosen. Of the emotional faces, 50% were displayed with open mouth and visible teeth in order to control for salient visual features. Any background and external features of the face stimuli such as earrings or clothing were removed with Adobe Photoshop Elements 9. For illustration of faces see Appendix C.

Non-verbal auditory stimuli were chosen from the Montreal Affective Voices set (Belin et al., 2008) with five male and five female actors (identity: 6, 42, 45, 46, 53, 55, 58, 59, 60, 61) producing emotional (50% angry and 50% happy expressions) and neutral non-verbal vocalisations. Again, for more details on stimuli see Chapter 1. Sound stimuli were cut after 800ms or continuously repeated and merged until they reached a length of 800ms. Sound stimuli were modified to fade in and out to avoid sudden onsets and loudness of each voice was normalised to a comparable level.

The current study consisted of two separate blocks split for either faces or voices and each block contained two sub-blocks where participants were told to either perform a gender or an emotion discrimination task in order to keep up attention levels. So all together, there were four experimental blocks but for subsequent analysis, ERPs were averaged across both gender and emotion tasks to increase number of trials. Each block was then repeated 3 times, resulting in 120 stimuli for each of the four presentation blocks. Each block was preceded by a short practise session. Order of blocks, tasks and stimuli within each block were manually randomized as well as counterbalanced amongst all participants. Each of the four experimental blocks contained either 50% emotional (divided into 25% angry and 25% happy stimuli) and 50% non-emotional stimuli, resulting in 40 stimuli in each block (20 neutral, 10 angry and 10 happy). Participants were told that emotional stimuli consisted of either angry or happy stimuli but were instructed to only discriminate between emotional and neutral stimuli. This is illustrated in Figure 5.2.

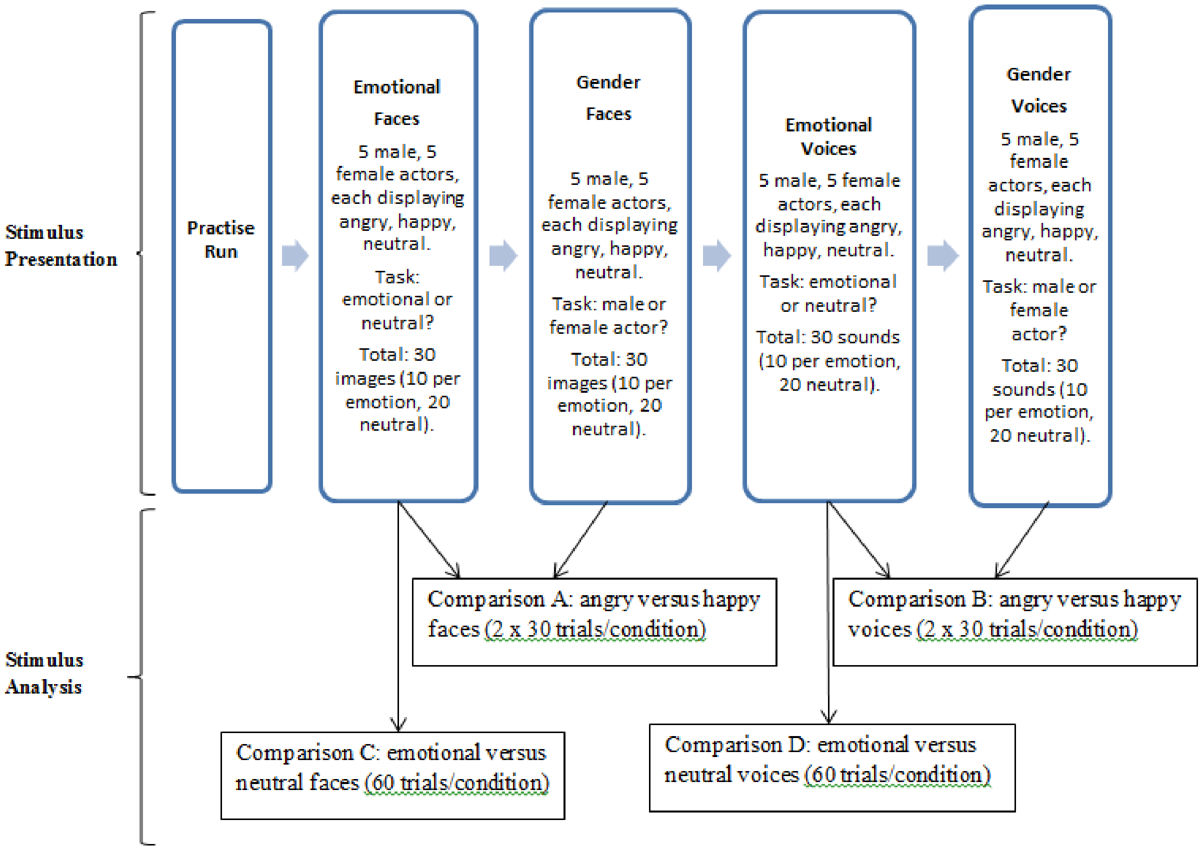


Figure 5.2. Presentation blocks and subsequent conditions for analysis of stimuli. There were four experimental blocks with 2 modality x 2 task, resulting in 8 conditions. Each block was repeated 3 times.

Before each stimulus there was a fixation cross for 500ms. The stimulus then appeared for 800ms, followed by an inter-stimulus interval of 2500ms in which the participants were asked to make a discrimination choice by pressing one out of two buttons.

This button-press happened only after the stimulus has disappeared in order to minimize interference with emotional processing during stimulus perception.

5.2.3. Procedure

Participants completed the task in a quiet, electrical shielded room with natural daylight. They sat on a comfortable chair, 80-85cm away viewing distance from the screen (refreshing rate 60Hz) inside a Faraday cage to minimise electrical noise. Before the start of the study, all participants signed the informed consent sheet, received detailed instructions and completed a practice session after which there was time for questions. Participants were informed about artefact problems and encouraged not to blink too often and sit as still as possible during stimulus presentation. Between each block, participants had a short break before continuing with the next condition. After completion of the experiment, there was time for questions and debriefing.

5.2.4. ERP Acquisition, Analysis and Data Reduction

The EEG (electroencephalogram) signals were recorded from 64 electrodes (FP1, FPz, FP2, AF3, AF4, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCz, FC2, FC4, FC6, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO5, PO3, POz, PO4, PO6, PO8, CB1, O1, Oz, O2, CB2) which were arranged on the scalp according to the 10-20 system. Cz formed the average reference during ERP recording. To improve electrical conduction, a conductive gel was used (Compumedics Neuromedical Supplies, Quik Gel).

To measure vertical eye movements (VEOG), two electrodes were placed just above and below the left eye (about 1 cm). To measure horizontal eye movements (HEOG), two electrodes were placed on the outer canthi of the right as well as the left eye. Electrical impedance was kept below 10KΩ. To analyse the signal, Scan 4.4 acquisition and analysis software (Compumedics Neuroscan Ltd.) were used to amplify (x 1000) and band-pass filter (0.1 – 100Hz) the EEG waves. 'Bad' electrodes were identified by visually observing the EEG data. This only applied to one participant, where the electrode F8 had to be excluded as it was not recording correctly.

Offline, the signal was band-pass filtered (0.1hz-30hz, 12 db/octave) and electrodes were re-referenced to the average of both linked mastoids. A spatial filter paradigm (Scan 4.2.) was applied to remove eye blink artefacts by Principal Component analysis. The EEG signal was baseline corrected (entire sweep) and peaks exceeding ±150µV were automatically excluded in the subsequent analysis. Event-codes identified stimulus-locked responses which were then averaged across all sweeps.

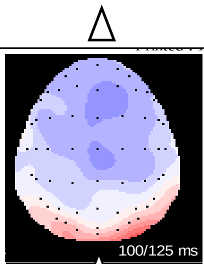
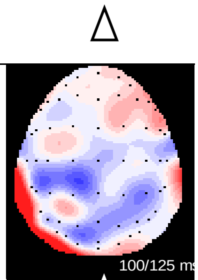
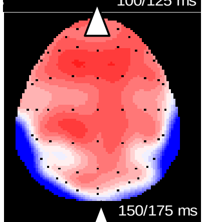
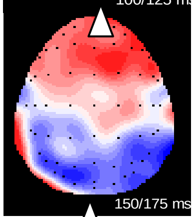
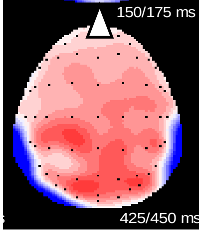
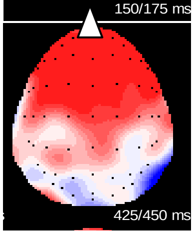
Epochs were created from the cleaned data by including 100ms of pre-stimulus interval and 900ms of stimulus processing which was locked to stimulus onset at 0s. For each participant, ERPs were averaged across 8 conditions (*i.e.* angry/happy face or voice, emotional/neutral face or voice). Averaged ERPs for each individual and each condition were then baseline corrected to pre-stimulus interval of -100ms. For each average ERP per participant and condition, amplitudes and latencies were detected via a peak detection algorithm in Scan 4.3. This was conducted across all electrodes and for three specific time-windows: 70-130ms for early sensory processing, 130-200ms for specific face or voice processing, and 300-600ms for late effects during emotion-specific processing.

For the present analysis of ERPs, separate ANOVAs were conducted for auditory and for facial stimuli due to their perceptual differences in basic perception. Further, separate ANOVAs were conducted for latency and amplitude measures. Although stimuli were presented in four blocks (2 modality x 2 task), for analysis, the data was subsequently split into 8 conditions. For each of the two modalities, there were separate analyses for (1) emotional versus neutral stimuli derived from the explicit emotion-task and secondly for (2) angry versus happy stimuli which were collapsed across gender and emotion-tasks to increase the number of trials. This way, number of trials were kept constant across tasks (N = 60) as is illustrated in Figure 35. Running separate analyses for emotional versus neutral and happy versus angry stimuli allowed observing hierarchical processing stages within the human mind from more general to narrower emotion concepts. The current study did not analyse ERP responses in relation to gender discrimination nor did it specifically investigate the contrast of direct versus indirect emotion ratings because at this point, the present investigation is of preliminary nature only.

<u>Regions of interest (ROI)</u>: Regional averaging by clustering several electrodes in a smaller amount of regions reduces the levels of comparisons and degrees of freedom (Dien & Santuzzi, 2005). Visual inspection of topographical maps for all conditions and each time-window revealed ROIs for the subsequent analyses (see Table 5.1.). After initial screening for hemispheric lateralisation, no significant effects were found (p > .05). Therefore, lateralisation was not included as additional variable. Hence, in the present analysis, several electrodes were averaged together in order to observe activity in three main regions of the brain: Frontal (FP1, FP2, AF3, AF4, F3, F1, F2, F4), Central (FC3, FC1, FC2, FC4, C3, C1, C2, C4, CP3, CP1, CP2, CP4) and Posterior (P3, P1, P2, P4, PO3, PO4, O1, O2) ROIs within the 10-20 system. There was one exception: for faces during the second time-window, four individual posterior electrodes P7, P8, PO7 and PO8 - rather than one averaged region - were selected. Those electrodes are commonly used to investigate the face-typical N170 response in posterior regions (*e.g.* Batty & Taylor, 2003; Smith, 2012). This is illustrated in Figure 5.3.

Table 5.1

Identification of regions of interest (ROI) for each modality and time window

	Faces		**Voices**	
70 – 130 ms	Frontal region: negativity Posterior region: positivity		Frontal region: positivity Posterior region: negativity	
130 – 200 ms	P7, P8, PO7, PO8: negativity		Frontal Region: positivity Posterior region: negativity	
300 – 600 ms	Posterior region: positivity		Frontal Region: positivity Central Region: positivity	

Note. Topographic maps show difference waves between emotional and neutral stimuli.

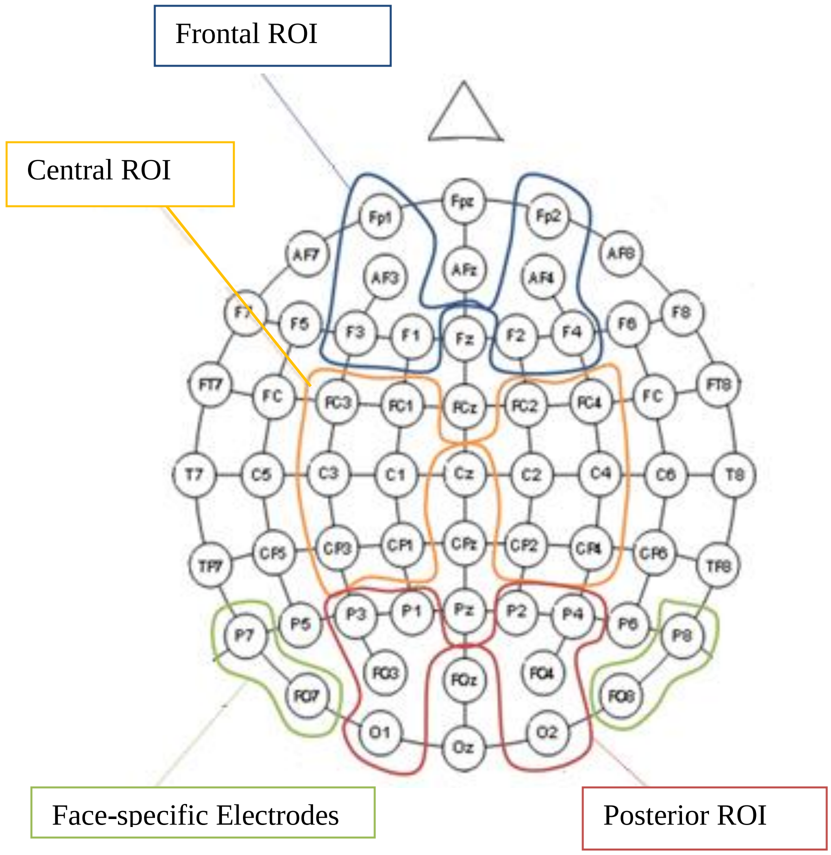


Figure 5.3. Electrodes used to create average for frontal (blue), central (yellow) and posterior (red) regions of interest (ROI). Green electrodes used as individual electrodes for the second time-window in the face analysis.

5.3. Results

The present section will report findings from the ERP study on emotion-effects across two separate modalities across three individual time-windows. It will firstly summarise behavioural findings before moving on to describe ERP patterns which are split according to task: emotional versus neutral and angry versus happy. Within each subsection, data will be reported for all three time-windows and for both modalities. Mean amplitudes and latencies for each time-window and condition are presented in Table 5.2.

5.3.1. Behavioural Analysis

A modality (2) x emotionality (2) repeated measures ANOVA revealed that the number of correct guesses of whether a stimulus was emotional ($M = 26.86$, $SE = .83$) or not ($M = 28.14$, $SE = 1.34$) did not vary across different modalities ($p > .05$). Behavioural reaction times were not analysed as participants were instructed to wait with the button press until the stimulus has disappeared. However, due to a recording error, only behavioural data from seven participants was recorded which significantly limits the generalizability of results. Since the behavioural task was at a very basic level and was only included to ensure that participants made conscious categorizations, the behavioural results will not be further discussed.

5.3.2. ERP Analysis

Table 5.2 displays ERP latency and amplitude means and Standard Errors for all three time-windows and across both modalities.

5.3.2.1. Emotional versus Neutral

a) **Early effects:**

70 – 130ms: For each pre-defined ROIs and its polarity, dependent t-tests were conducted with task (emotional vs neutral) as within-subject variable. For **faces**, although visual inspection of topographic maps suggested posterior positivity, neither amplitude nor latency measures showed any significant differences between categorising emotional or neutral faces. This suggests that for faces, statistically, the earliest processing stage did not carry any emotion information. This is illustrated in Figure 5.4 (electrode P8).

Table 5.2

Means and Standard Error for amplitude and latency measures for each time-window and conditions

	Faces Emotion	Voices Emotion	Faces Neutral	Voices Neutral	Faces Happy	Voices Happy	Faces Angry	Voices Angry
70-130ms Amplitude	1.48 (0.69)	-0.06 (0.23)	1.94 (0.65)	0.57 (0.32)	1.91 (0.70)	0.62 (0.31)	1.86 (0.62)	0.10 (0.24)
70-130ms Latency	98.53 (4.63)	101.65 (2.53)	99.14 (4.68)	97.04 (4.71)	101.03 (4.85)	101.34 (2.42)	97.90 (3.99)	96.45 (4.15)
130-200ms Amplitude	-1.72 (0.74)	-2.23 (0.44)	-1.97 (0.78)	-2.03 (0.53)	-1.54 (0.96)	-1.43 (0.34)	-1.70 (0.87)	-1.16 (0.35)
130-200ms Latency	146.78 (4.53)	143.14 (2.57)	151.96 (5.99)	151.39 (5.64)	147.28 (5.06)	150.21 (4.06)	146.02 (4.47)	146.71 (2.92)
300-600ms Amplitude	8.03 (1.41)	3.30 (1.54)	7.03 (1.47)	1.78 (0.79)	7.74 (1.48)	2.68 (0.49)	7.46 (1.19)	3.51 (0.48)
300-600ms Latency	462.84 (22.89)	442.02 (15.15)	416.29 (26.68)	414.42 (12.72)	470.01 (16.19)	457.29 (18.80)	463.15 (21.50)	418.96 (17.23)

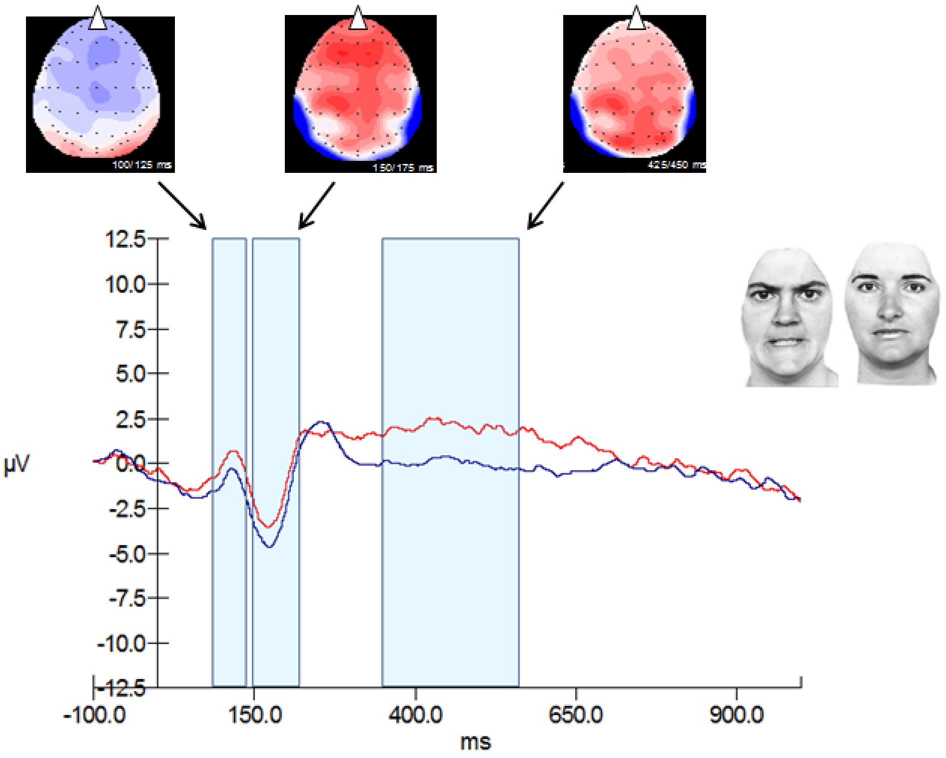


Figure 5.4. Faces: Emotional (red) versus neutral (blue) at posterior ROI (P8). Non-significant effects in mean amplitude and mean latencies at the first positive peak, second negative deflection and late stage. N.s.

For **voices** on the other hand, in the frontal ROI, the mean amplitude differed at the first positive peak, $t(10) = 4.4$, $p = .001$. Emotional voices ($M = .69$, $SD = 1.14$) had a significantly higher mean deflection than neutral voices ($M = -1.93$, $SD = 2.17$). This early frontal positivity for emotional voices occurred at a mean latency of 98.93ms. For posterior ROI, mean amplitudes for the first negative deflection did not differ between emotional and neutral voices; however, there was a significant difference for mean latencies in the posterior regions, $t(10) = -6$, $p < .001$, and neutral voices ($M = 74.4$, $SD = 11.65$) were processed with an earlier mean latency than emotional voices ($M = 108.39$, $SD = 19.93$). This is illustrated in Figure 5.5 (electrode FP2).

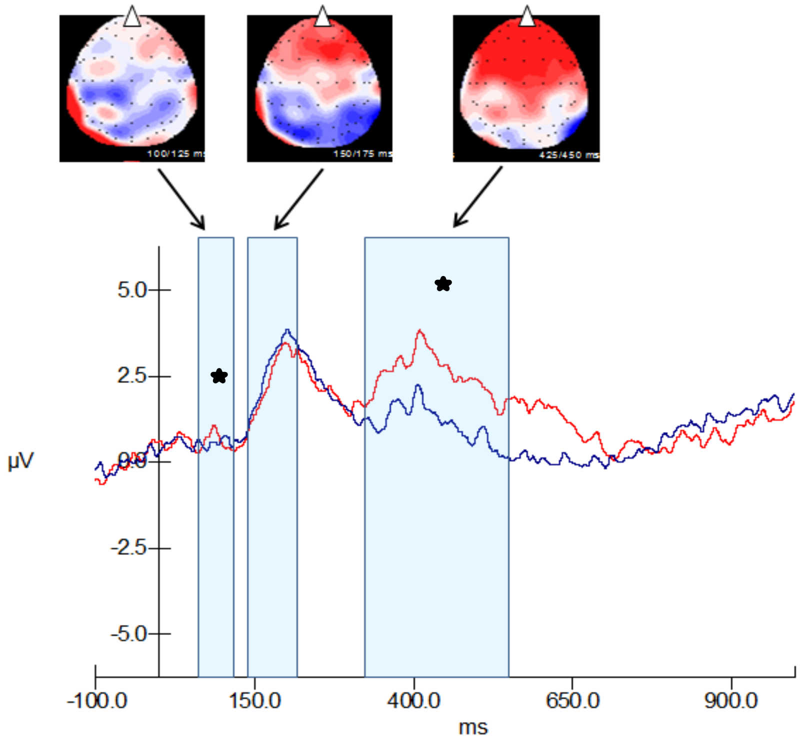


Figure 5.5. Voices: Emotional (red) versus neutral (blue) at frontal ROI (FP2). Significant effects in mean amplitude at the first positive peak and late stage for emotional versus neutral voices ($p < .05$) but not in the second stage. * = $p < .05$

130 - 200ms: A repeated-measures ANOVA was conducted with electrodes (P7, P8, PO7, PO8) x task (emotional versus neutral) as within-subject factors. For posterior negativity in **faces**, results revealed a significant main effect for electrode, $F(3, 30) = 4.36$, $p < .05$, pes = .3 with the electrode P8 ($M = -6.58$, $SE = 1.41$) in the right posterior area showing greatest negativity at a latency of 169.92ms. However, after Sidak post-hoc comparisons, only the difference between P8 and PO7 remained close-to-significant at $p = .058$. There was no significant main effect for task, $p > .05$. Latency for emotional and neutral faces did not differ at posterior sites at the given time-window.

For **voices**, for each pre-defined ROIs and their polarities, dependent t-tests were conducted with task (emotional versus neutral) as within-subject variable. Neither amplitude nor latency measures at the posterior or the anterior ROI differed for emotional versus neutral stimuli ($p > .05$). This suggests that for voices, statistically, early posterior negativity as well as frontal positivity did not differ for emotional versus neutral voices between 130 and 200ms. This is illustrated in Figure 5.5.

b) Late effects:

300 – 600ms: At each pre-defined ROI, dependent t-tests were conducted with task (emotional vs neutral) as within-subject variable. Data from **faces** did not reveal any significant differences in mean amplitude or latency for processing emotional and neutral stimuli in posterior ROI (p > .05). This finding suggests that brainwaves for rating emotions in faces were not different to rating neutral faces during later stages of processing.

For **voices,** two ROIs were identified that both showed positive deflections; hence, a repeated-measures ANOVA was conducted with area (2: frontal versus central) and task (2: emotional versus neutral) as within-subject factors. Results revealed a significant main effect for task, $F(1, 10) = 5.23$, $p < .05$, pes = .34. Emotional voices (M = 3.48, SE = .66) had significantly higher mean amplitudes than neutral voices (M = 1.67, SE = 1.09) across fronto-central sites. There was no main or interaction effect for amplitude differences in area (p > .05). For measures of latency, there was a significant main effect for area, $F(1, 10) = 12.48$, $p < .05$, pes = .56 and positive deflections in the frontal ROI (M = 387.76, SD = 12.81) occurred at significantly earlier latencies than in the central ROI (M = 436.11, SE = 13.11). There was no significant main effect for task latency (p > .05).

In summary, faces did not show any significant differences in processing neutral versus emotional stimuli during any of the given time-windows apart from a subtle right hemisphere advantage. Voice ERPs on the other hand differed for emotional and neutral stimuli for the first time at around 100ms after stimulus onset as reflected in an increased early frontal positivity. For later stages of processing, voices showed increased fronto-central positivity for processing emotional compared to neutral voices which was earlier for frontal than for central ROIs (this is illustrated in Figure 5.5).

600-900ms: Since none of the included time-windows showed any emotion-effects at all during the discrimination of emotional versus neutral **face expression**, it was necessary to ensure that no emotion effects occurred after the first 600ms. Hence, a fourth time-window was included (600 and 900ms), specifically for faces, to observe the possibility of a very late positive component within the posterior region that would have otherwise been missed. Related t-tests showed that even for the very late time-window, the posterior positivity for discriminating emotional from neutral faces did not show any significant effects for either amplitude or latency (p > .05).

5.3.2.2. Emotion Differentiation (angry versus happy)

a) **Early effects:**

70 – 130ms: For each pre-defined ROI and its polarity, dependent t-tests were conducted with task (angry versus happy) as within-subject variable. For **faces**, neither positivity in posterior ROI nor negativity in frontal ROI showed significant differences in mean amplitude or latency measures for categorising happy or angry stimuli.

Similarly, for **voices**, neither frontal positivity nor posterior negativity differed significantly in mean amplitude or latency measures for categorising happy or angry stimuli ($p > .05$).

130 – 200ms: For the present time-window, a repeated-measures ANOVA was conducted with electrodes (P7, P8, PO7, PO8) x task (angry versus happy) as within-subject factors. For **faces**, there was a significant main effect for electrode over posterior regions, $F(3, 30) = 3.93$, $p < .05$, pes = .28. Electrode P8 ($M = -6.61$, $SE = 1.3$) in the right hemisphere showed a greatest negativity at a latency of 166.09ms. However, after applying post-hoc Sidak comparisons, none of the differences between pairs of electrodes remained significant. Further, there was no significant main effect for task. Equally, mean latency was not affected by condition ($p > .05$). This is illustrated in Figure 5.6.

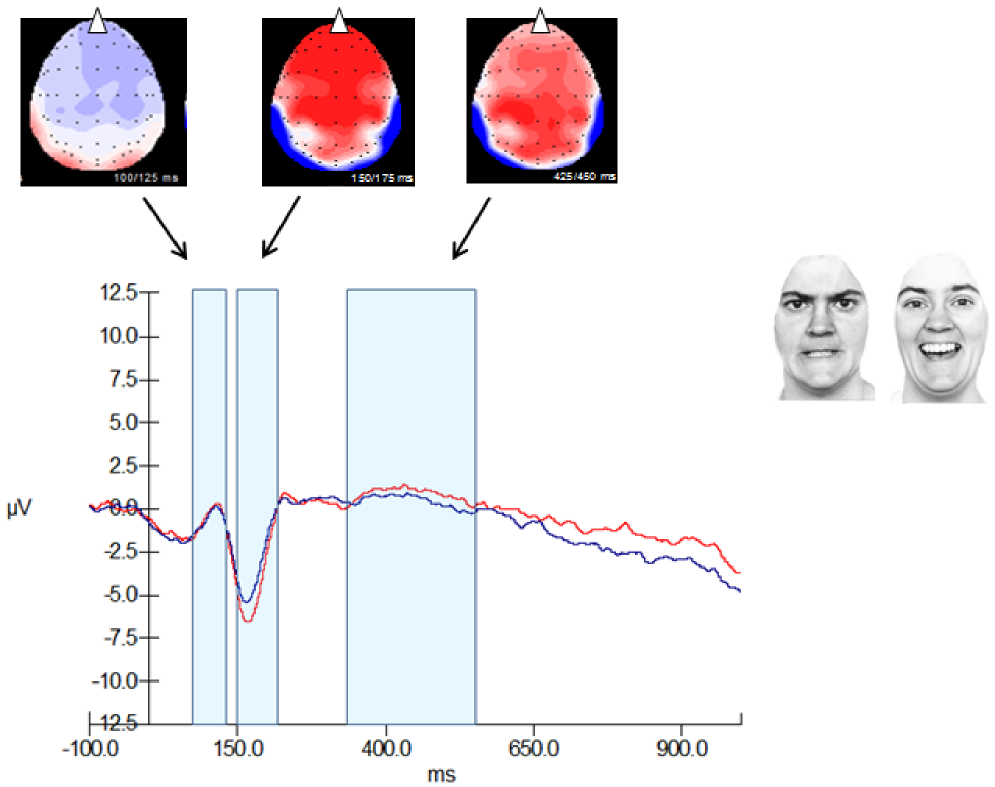


Figure 5.6. Faces: Angry (red) versus happy (blue) at posterior ROI (P8). No significant effects in mean amplitude and mean latencies at the first positive peak, second negative deflection and late stage. N.s.

For each of the pre-defined ROIs and their polarities in **voices**, dependent t-tests were conducted with task (angry versus happy) as within-subject variable. Posterior negativity did not differ in mean amplitude between happy and angry stimuli in the given time-window. Latency showed a close-to-significant difference, $t(1, 10) = 4.27$, $p = .066$ with a relatively high effect size (pes = .3). Negative deflections over the posterior ROI for angry voices ($M = 155.9$, $SD = 3.47$) occurred earlier than for happy voices ($M = 164$, $SD = 3.32$). Anterior positivity did not differ for amplitude as well as latency measures for happy versus angry voices ($p > .05$). This is illustrated in Figure 5.7 (electrode FP2).

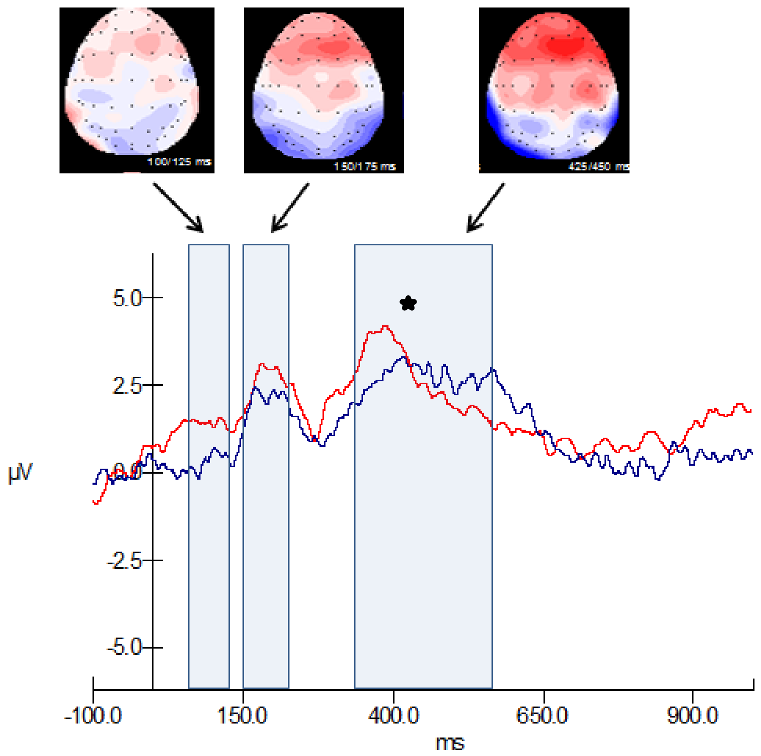


Figure 5.7. Voices: Angry (red) versus happy (blue) at frontal ROI (FP2). No significant emotion effects in the first stage (n.s.), close-to-significant latency differences in the second stage ($p = .06$) and significant latency effects for the third stage (*$p < .05$).

b) Late effects:

300 – 600ms: For the pre-defined ROI, dependent t-tests were conducted with task (angry versus happy) as within-subject variable. For **faces**, there were no significant differences in waveforms over the posterior ROI for processing angry or happy stimuli within the given time-window ($p > .05$),.

For **voices** within the given time-range, the current study has identified two ROIs that both showed positive deflections; hence, a repeated measures ANOVA was conducted with area (frontal versus central) and task (angry versus happy) as within-subject factors.

For amplitude measures in the given time-window, there were no significant main effects for either task or area ($p > .05$). For latency, there was a significant main effect for task, $F(1, 10) = 9.03$, $p < .05$, pes = .47 with angry voices ($M = 394.13$, $SE = 13.52$) showing significantly earlier latencies of positive deflections than happy voices ($M = 433.67$, $SE = 17.71$) over fronto-central sites. Further, there was a significant main effect for area, $F(1, 10) = 6.23$, $p < .05$, pes = .38 and frontal areas ($M = 396.4$, $SE = 13.93$) showed significantly earlier positive peaks during the given time-window than central regions ($M = 431.71$, $SE = 17.73$). This is illustrated in Figure 5.7.

In summary, during the first time-window, neither faces nor voices showed different waveforms for rating happy or angry stimuli - in fact, faces did not show any significant emotion-differentiation effects within any given time-window. Between 130 and 200ms, angry voices showed an earlier deflection than happy voices over posterior sites although this was only marginally significant ($p = .066$). In the final processing stage of voices, this anger-superiority effect became significant over fronto-central sites.

5.4. Discussion

The present ERP study investigated for the first time in a within-subject design whether faces as well as non-verbal vocalisations showed comparable neural patterns of activity during emotion processing. Further, it was of interest whether emotions were processed in distinct, hierarchical stages. This was predicted to start with a pre-emotional processing stage before moving on the broad emotion versus neutral discrimination and ending with narrow discrimination between angry and happy stimuli.

Data from the present ERP study indicated that for faces, there were no early emotion effects visible across three distinct time-windows of processing. Contrary to Hypothesis 1, comparable emotion stages between voices and faces could not be consistently found. For voices, participants showed emotion effects within the first 100ms and - according to Hypothesis 2 - this processing seemed to occur in hierarchical manner. Broad discrimination between emotional and neutral voices happened before the enhanced processing of angry compared to happy voices in later stages. Overall, present data identified specific emotion processing stages in the voice condition; however, those stages could not be replicated in the face condition.

5.4.1. Sensory Processing Stage (70-130ms)

5.4.1.1. Faces

The first time-window was included in order to investigate whether early sensory processing of faces and voices can be modulated by emotion content. The results of the current study suggested that for faces, the first processing stage between 70 and 130ms was not significantly influenced by the emotional load of displayed faces. The waveforms for neutral and emotional as well as for angry and happy faces were comparable. Contrary to the present results, very early emotion effects during visual perception have previously been suggested. For example, Batty and Taylor (2003) suggested automatic processing of emotional versus neutral faces in posterior brain regions as early as 90ms after stimulus onset. Indeed, visual stimuli that have been associated with affective meaning following electric shocks yielded better results in a visual search task and also increased brain activity in the primary visual cortex V1 (Padmala & Pessoa, 2008). This association between affective meaning of faces and early visual perception could not be replicated in the present study.

However, the current findings are in line with previous accounts of pre-emotional processing during sensory face perception. For example, Krolak-Salmon et al. (2001) found earliest discrimination between emotional versus neutral faces at 250 to 550ms after stimulus onset.

This effect was mainly visible in occipital brain areas. In agreement with this, emotional versus neutral faces also showed a larger occipital-temporal negativity as late as 200 to 400ms in a study by Marinkovic and Halgren (1998).

The current ERP data on emotional face processing supports the first pre-emotional stage of the hierarchical three-stage processing model in voices by Schirmer and Kotz (2006). According to this model, early voice processing within the first 100ms based on perceptual features is independent of emotional content as reported by Engelien et al. (2000). Consequently, it is possible that the first stage of Schirmer and Kotz's (2006) voice model can be extended to emotion processing in faces, too.

5.4.1.2. Voices

Data from the present study found very early emotion modulation of the ERP signal during early sensory processing stages: Current results demonstrated enhanced broad discrimination for emotional versus neutral non-verbal affect vocalisations around 100ms after stimulus onset in frontal brain regions. Further, in line with Hypothesis 2 of a hierarchical emotion process model, there was no specific emotion-category discrimination such as between happy and angry vocalisations.

In line with the first stage of Schirmer and Kotz's model (2006) and ERP face data reported above, initial processing of sensory information happens irrespectively of emotional content within the first 100ms after stimulus onset. This statement, however, has been challenged by the current voice data as well as by Iredale et al. (2001) who reported emotion effects in prosody between 50 and 100ms. In support, during an fMRI study, non-verbal affect vocalizations such as laughing did not only activate the amygdala but also directly enhanced auditory cortex activation (Sander & Scheich, 2005). Since Schirmer and Kotz (2006) based their voice processing model on prosody rather than on non-verbal affect bursts, it is possible that the affect modulation on non-verbal vocalisation may be happening earlier due to prototypical presentation of emotion stimuli.

The present brain activity pattern of frontal positivity and posterior negativity during voice perception can typically be found at around 160ms and previously been termed as 'fronto-temporal positivity to voices' (FTPV) by Charest et al. (2009). The current data suggests firstly that the FTPV response may already occur at around 100ms after stimulus onset in non-verbal affect vocalisations. Indeed, Rogier, Roux, Belin, Bonnet-Brilhault and Bruneau (2010) reported a typical FTPV response at fronto-temporal sited in children as early as 60ms after stimulus onset. Secondly, the current data further suggests that the FTPV may be modulated by emotional content of voices. In accordance with this, Kryklywy et al. (2013) conducted a functional neuroimaging study and concluded that the auditory 'what' pathway was also influenced by emotionality of voices.

Alternative accounts of the present frontal activation also exist: For example, it has been proposed that humans may exhibit a 'prefrontal auditory domain' that also shows activation during emotion processing (Fecteau, Armony, Joanette, & Belin, 2005). Following present results, it is possible that this prefrontal auditory region processes emotional significance of non-verbal vocalisations as early as 100ms after stimulus onset.

Overall, the earliest time-window reported an early posterior negativity (EPN) and frontal positivity in voices, but not in faces. This EPN could possible reflect early voice-specific responses in the auditory 'what' pathway' that are modulated by emotional significance as early as 100ms post stimulus onset.

5.4.2. Modality-specific processing stage (130 – 200ms)

5.4.2.1. Faces

For the second time-window, topographic maps from the current ERP data indicated a frontal positivity and occipito-parietal negativity around 170ms after stimulus onset as commonly found during the face-specific N170 component. However, this pattern was not statistically different for processing emotional and neutral faces. Current ERP responses suggest that faces may be encoded independently of emotional significance at around 170ms. The current study is not the only one to suggest emotion-independent face processing responses at the N170 component: For example, Eimer and Holmes (2002) also reported N170 waveforms that were unaffected by the emotionality of faces. According to Bruce and Young's face perception model (1986), it is possible that the extraction of facial features used to recognise face expressions may to some extend be functionally segregated from extracting facial features for general face perception.

The current analysis also found a subtle right hemisphere advantage during emotional face processing at a latency of 170ms after stimulus onset. Although this findings was only close to being significant (p = .06), it is in line with the common assumption of emotion processing within the right hemisphere (*e.g.* Blonder, Bowers, & Heilman, 1991). Hemispheric differences after rating face expressions indicate that some emotion-typical processing may be happening within the right posterior region in the given time-window. This is likely to reflect enhanced activity in the right STS and FFA during processing of face expressions (Ganel et al., 2005).

On a different note, it is, however, possible that emotion-effects reported for the N170 component depend upon the commonly used reference electrode as noted by Rellecke, Sommer and Schacht (2013): When using an average reference, studies such as by Batty and Taylor (2003) or Blau et al. (2007) reported emotion effects for the N170 whilst for studies using the linked mastoid as average reference (*e.g.* Eimer & Holmes, 2002) often failed to find emotion effects in the N170.

Since the present study used a linked mastoid reference, it is possible that emotion effects at the temporal-occipital N170 component are not visible as effects seem to be smaller when reference and recording sites are close to each other (Rellecke et al., 2013).

5.4.2.2. Voices

Although the topographical ERP map suggested a frontal positivity and posterior negativity at around 170ms which may reflect the FTPV response, the present study did not find significant emotion effects during the discrimination of neutral and emotional affect vocalisations.
This finding for the voice condition is in line with findings from the face condition: neither found significant effects of rating emotional versus neutral faces during the second time-window. However, for emotion-specific discrimination between happy and angry stimuli, activity in the second time-window reveals close-to-significant effects (p = .066) with surprisingly early posterior negativity. This occurred earlier for angry compared to happy voices at around 155ms. In line with Hypothesis 2, the current pattern demonstrates a hierarchical emotion mechanism as the early time-window demonstrated broad emotional versus neutral voice discrimination, followed by narrower anger versus happiness voice discrimination in the subsequent time-window.

It is possible that this negativity for angry voices coincides with the FTPV response at round 160ms after stimulus onset. Interestingly, the regions within the STS are also commonly associated with the specific processing of angry voices (Grandjean et al., 2005; Sander et al., 2005). Note, however, that the current study did not find significant frontal positivity effects in the given time-window which is commonly thought to accompany the FTPV response.

The current proposal that threatening emotions such as anger seemed to be processed earlier than happy voices has previously been supported by Sauter and Eimer (2009) who also found enhanced processing of fearful non-verbal vocalisations at around 150ms after stimulus onset. This negativity bias may have evolutionary reasons: Earlier latencies could reflect faster processing and in order to enhance survival and minimize threat, it is important to rapidly detect the angry person (Hansen & Hanson, 1988).

Overall, neither faces nor voices showed discrimination effects between emotional and neutral stimuli during the second time-window. However, voice data indicated an anger-superiority effect at around 170ms in posterior regions which possibly suggests early STS activation for emotion-specific vocalisations. For faces a small right hemisphere advantage was noted in the same time-window and region, suggesting some degree of emotion processing within the right hemisphere. Although this did not survive statistical post-hoc comparisons, by including a larger sample in future studies, this effect from the present study may eventually become significant.

5.4.3. Late processing stage (300-600ms)

5.4.3.1. Faces

The present ERP data did not find any statistically significant differences in waveforms for emotional versus neutral faces during the third time-window.

To ensure no late activity was missed, the time-window got extended to the first 900ms after stimulus onset, but again, no significant brain activity related to the specific processing of emotions was recorded over posterior regions. According to the majority of past ERP studies investigating emotion recognition in faces, emotion effects are consistently reported to occur within the first 200-300ms of processing (*e.g*. Eimer & Holmes, 2007; Sato et al., 2001).

Commonly, within the first 400ms, a late posterior negativity or frontal positivity for emotional compared to neutral faces is found, which possibly reflects projections from the amygdala back to the visual cortex (Sato et al., 2001).

One possible explanation for the lack of emotion effects in faces could be that the used stimuli were not salient enough to elicit a significant ERP signal. Especially the ERP signal around 300 and 450ms seems to be higher for high-arousing compared to low-arousing faces (Rozenkrants & Polich, 2008). Similarly, between 400 and 800ms, the late positive potential (LPP) was found to be enhanced for images with a high compared to low degree of arousal (Leite, Carvalho, Galdo-Alvarez, Alves, Sampaio, & Goncalves, 2012). Hence, although not statistically tested, it is possible that differences in arousal levels between neutral and emotional stimuli in the present study are too subtle and did not get picked up in the ERP signal.

Secondly, the current study did not find different waveforms for angry compared to happy faces during the final time-window. The lack of emotion-specific effects could, again, be attributed to the possibility of similar degrees of arousal in happy and angry faces. However, lack of emotion-specific effects in the late time-window is not uncommon: for example Eimer and Homes (2007) reported global emotion effects that lasted up to 1 second but were unaffected by the specific emotion category. According to Eimer and Holmes (2007), this lack of emotion-specific effects could be evidence for one general emotion core network of face processing. For example, the amygdala has been associated with not only with fear processing (Morris et al., 1998) but also equally with happiness and anger processing (Fitzgerald, Angstadt, Jelsone, Nathan, & Phan, 2006). However, it is to note that whilst Eimer and Holmes (2007) compared brain waves for six individual basic emotions, the current study only compared one positive with one negative emotion category and due to the current study design, these emotions were judged implicitly rather than explicitly.

Further, it is possible that the current design of the study additionally limited emotion effects between angry and happy stimuli by balancing the number of angry and happy stimuli that expressed emotions with open mouths and teeth visible.

Previously, it has been proposed that happiness superiority effects in faces may be the confounding result of global visual features such as displaying white teeth (Schyns et al., 2009). Hence, it is possible that for the current study, visual features were comparable across both emotions which in turn results in comparable ERP waveforms.

5.4.3.2. Voices

During the later stage of emotion processing, a late positive component over fronto-central sites was found for the discrimination of emotional versus neutral as well as for angry versus happy vocalisations. The late stage of processing in frontal brain regions may reflect conscious evaluation of emotions (Eimer & Holmes). It has for example been stated that the amygdala is functionally connected with the medial prefrontal cortex (Kim, Gee, Loucks, Davis, & Whalen, 2010) which may couple bottom-up emotion experience and top-down emotion appraisal (Whalen et al., 2013).

Secondly, for voices within the third time-window, there were separate waveforms for angry and happy voices that differed in their mean latencies but not in mean amplitudes. This late positivity seemed to apply especially to frontal brain areas. According to Hypothesis 2, this hierarchical processing of emotions supports Schirmer and Kotz's (2006) stage model which suggests broad emotion discrimination before narrow emotion-specific processing in the third stage. The present findings are supported by Iredale et al. (2013) who suggested that at the third processing stage between 400 and 650ms, sentences spoken with angry or happy prosody elicited different brain waves at frontal sites. As for the second time-window, angry vocalisations had earlier mean latencies compared to happy vocalisations, but at the third stage, this comparison was statistically more significant. Again, brain activity for emotion-specific processed seemed to move from a posterior negativity in the second time-window to a frontal positivity in the third time-window.

Interestingly, damage to frontal lobe regions has also been associated with impaired emotion recognition in faces as well as in affect vocalisations (Hornak, Rolls, & Wade, 1996). Although not all participants were impaired across both modalities, Hornak et al.'s (1996) data supports the current findings that emotion recognition abilities may indeed be linked to pre-frontal activity. Alternatively, it is possible that the fronto-central activation seen in the current study reflects processes in deeper brain structures such as the amygdala which is commonly associated with emotion processing. For example, the amygdala is often activated when listening to non-verbal exclamations of laughter and crying (Sander & Scheich, 2005).

Overall, faces did not show any emotion-effects for either emotional versus neutral nor for angry versus happy discrimination during late stages of emotion processing. For voices, emotional versus neutral as well as happy versus angry comparisons yielded significant differences in fronto-central regions. This suggests that angry vocalisations had earlier mean latencies than happy vocalisations at 300ms after stimulus onset.

It is possible that in the present within-subject design, participants perceived vocalisations are more arousing that static faces which results in the null-effect seen for the face condition.

5.4.4. Implications, Limitations and Future Research

Results from the present ERP study suggested that Schirmer and Kotz's (2006) hierarchical model of prosody processing may also extend to emotional non-verbal vocalisations. However, the distinct emotion processing stages seem to be occurring earlier for vocalisations compared to prosody which may be due to prototypical representation of emotions in vocalisations. Further, the present study did not find pre-emotional stages of sensory processing, which suggest very rapid and automatic processing of emotional vocalisations.

In order to increase statistical significance in the present ERP study, it is necessary to increase the number of participants. Further, the current data may have contained some low levels of noise as it was conducted during hot summer days and sweat may influence electrical skin conductance and consequently decrease statistical power (Kappenman & Luck, 2010). Additionally, it was not possible to relate brain activity to behavioural findings due to a technical error. Hence, it is not known whether the ERP signal corresponds to accuracy rates of emotion categorisation across modalities. Further, the current ERP study was conducted with male participants only; hence, the results can neither be generalised across gender now can it explore potential gender differences during emotion recognition in voices and faces. However, since the behavioural chapters in this thesis did not find any main effects of gender, this was not the main focus of this ERP study.

Possible implications may arise from the current study: If applied to the real world, it is possible that voices communicate emotional intend earlier than face expressions. In an evolutionary context, it is possible that recognising emotional voices rather than emotional faces is important for threat avoidance in situations where vision is limited such as in darkness. This reflects verbal feedback from participants after the ERP study who found the voices as 'more exciting' than the faces which in turn possibly motivated attention and emotion processing.

5.4.5. Conclusion

In conclusion, the present ERP study investigated in a within-subject design how similar emotion processes across two separate modalities are. Whilst the current data did not suggest significant emotion categorisation effects across any of the three time-windows for faces, voices showed distinct and hierarchical stages of processing with early discrimination of emotional versus neutral vocalisations and subsequent discrimination of angry versus happy vocalisations.

In terms of modality similarities, the current study could not confirm equal emotion-processing stages across faces and voices. However, there were also some similarities – at least to a degree - between emotion processing across faces and voices: for example, during the structural face and voice-decoding response such as the N170 and FTPV at around 150-170ms, neither of the modalities showed waveforms that were influenced by the emotionality of stimuli.

6

General Discussion & Conclusions

How do we recognise emotions from individual modalities such as faces or voices? The aim of the present thesis was to investigate the above question in a multi-method manner and to investigate how this ability develops with age and how it varies depending on cultural factors. Implementing a within-subject design to rate emotions across modalities counteracts the many different measures and stimuli previously used in emotion recognition research which make comparisons of results difficult (Bayer & Schacht, 2014). For the present investigation, cognitive, developmental, longitudinal, cross-cultural and neuroimaging studies as well as an online-survey were conducted in order to investigate from several viewpoints whether emotion recognition was comparable across faces and voices.

Chapter 2 suggested that emotions presented in one modality can be recognised in comparable manner with emotions presented in a second, separate modality. Consequently, there is a possibility that emotions are categorised more on the appraisal of the events itself (Scherer, 1984) rather than on perceptual features such as the typical smile in happy faces (Ekman & Friesen, 2003). It has been hypothesised that there may be one general emotion core network which acts independently of modality (Peelen et al., 2010) and the current behavioural data supports this statement.

Chapter 3 suggested that this modality-independent emotion processing may be a feature of development because as age increased, recognition across modalities became more similar. Across both modalities, emotion recognition improved with age. Findings from the child Chapter 3 additionally validated the newly developed child face database DDCF (Dalrymple, Gomez, & Duchaine, 2013) for the use in children which has not been tested before. The present longitudinal study indicated, however, that emotion processing development in faces may happen in stages which cover periods longer than 1.5 years before changes are visible. Overall, children's facial emotion recognition was not influenced by in-group age effects.

The originality of the current investigation into emotion recognition has further been demonstrated by comparing German and British individuals in their emotion recognition. Chapter 4, showed universal signs of basic emotion recognition which were coloured by subtle differences depending on culture, in line with the Dialect theory (Elfenbein & Ambady, 2003). It is possible that underlying differences in display rules such as the German reservedness (*e.g.* Wouters, 2011) may be linked to dampened intensity perception in the German compared to the British sample.

Children on the other hand showed less vulnerability to culture effects during emotion recognition; hence, it is possible that whilst general developmental aspects drive emotion recognition during childhood, only in adulthood does culture drive emotion recognition.

Lastly, the present ERP Chapter 5 aimed to investigate neurophysiological features of emotion processing across two separate modalities. Hierarchical emotion processing stages were found on the voice condition with emotion processing visible as early as 100ms after stimulus onset. However, no emotion effects were visible for the face condition. Overall, the results of the current ERP study remain preliminary and further testing needs to be conducted with a larger sample size.

6.1. Role of Modality

For humans, it is a necessity to be able to recognise emotions when we only have access to one modality such as in darkness or in noisy situations. Across all three behavioural chapters, the present thesis concludes that participants rated emotions from separate modalities with more similarities than dissimilarities. The similarity across modalities became especially evident during the analysis of confusion matrixes in Chapter 2 and 3. For each participant, rating patterns for each emotion were highly correlated across modalities and this correlation became stronger with age whilst younger children showed a face-preference.

Similar pattern of emotion perception across modalities questions the reliance on perceptual features during emotion classification. Instead, the researcher proposes the idea that there may be similar underlying emotion mechanisms which are shared across modalities and guided by cognitive appraisal such as novelty and pleasantness (Scherer, 1984). This may be despite the great perceptual difference in sensory inputs from faces and voices. An emotion-general core network that acts independently of modality has previously been proposed by Peelen et al. (2010). In addition, neuronal structures such as the amygdala also seem to process emotions independently of their perceptual features (Adolphs, 2002; Phillips et al., 1998). The present findings suggest that not only person-identity from voices and faces (Yovel & Belin, 2013), but also emotion recognition from voices and faces may share common underlying coding mechanisms. Additionally, there may also be a perceptual interdependence of physical features during emotion production. For example, muscular activation of the mouth and tongue, which reflect functionally adaptive behaviours such as vomiting, not only creates the typical face expression of disgust but also activates musculature in the upper vocal tract, creating typical vocal expressions of disgust (Scherer, 1994). This may explain the similarity of rating patterns of emotions across two separate and independent modalities.

This modality-independent rating pattern may, however, be a feature of adult emotion recognition whilst children during years at primary school relied more on facial than on auditory emotion information. Further, cultural differences may affect the strength of modality similarities. The similarity across modalities was weaker within the German adult sample as there was a face reliance in emotion recognition, especially for German men. Children on the other hand did not seem to differ in their degree of modality similarities across cultures and culture-independent early face reliance with later modality-independent processing was evident in both countries.

What could be the benefit of modality similarities during emotion recognition? In real life, emotions in faces or voices may not always be expressed in isolation. In other words, it is very common that emotions are expressed multimodally and simultaneously across modalities. In order to make sense of our environment and everyday life, we need to successfully integrate emotion cues from several independent modalities such as faces or voices. Multimodal emotion recognition is believed to be superior to unimodal emotion recognition from single modalities (Ethofer et al., 2006) and attending to one modality may change the perception of another modality in situation where visual and auditory inputs don't match (Collignon et al., 2008).

The benefit of a possible shared emotion mechanism across faces and voices could be the successful integration of modalities in shared brain regions such as the STS (Kreifelts, Ethofer, Shiozawa, Grodd, & Wildgruber, 2009). Indeed, research into congruent and incongruent emotion cues across modalities hints at a strong link between processing emotion signals from faces and voices that cannot simply be switched off – in line with the idea of a shared emotion mechanism. For example, De Gelder and Vroomen (2000) suggested that the recognition of emotions expressed in faces was influenced by the simultaneously presented emotion expressed in voices and vice versa. Purposely attempting to focus on one channel alone did not break this mandatory link between processing of emotions from faces and voices simultaneously. Similarly, Ethofer et al. (2006) suggested the bidirectional link between two separate modalities during congruent emotion perception. So called cross-modal effects were visible when fearful faces were presented with a simultaneous fearful voice, enhancing the emotion perception of fear by integrating congruent emotion signals from two separate modalities. This enhanced emotion perception of fear across two modalities was also reflected in an enhanced hemodynamic response in the emotion structure amygdala during an fMRI study.

In terms of incongruent emotion perception, the recognition of ambiguous face or voice expressions was influenced by the simultaneous presentation of bodily expressions of emotions (van den Stock, Righart, & de Gelder, 2007). Collignon et al. (2008) suggested that for incongruent emotion pairs, visual emotion signals commonly overrode auditory emotion signals – although this was less common if the visual signal was unreliable and noisy. This carry-over effect of incongruent emotions across modalities suggests a link between emotion signals from faces and voices that influence the bidirectional perception. This in turn is consistent with the idea of a shared emotion network, independent of modality.

The superior temporal sulcus is believed to be involved in dynamic information from faces such as gaze and face expressions (Bruce & Young, 1986; Haxby et al., 2000). Interestingly, parts of the STS are also believed to include temporal voice areas (Charest et al., 2009) which are active during processing of auditory voice information such as prosody. This shared use of neural structures during emotion recognition across modalities may enable the integration of information from several modalities. Indeed, the STS is commonly associated with multimodal emotion recognition: For example, Hagan, Woods, Johnson, Calder, Green and Young (2009) suggested increased activity in posterior regions of the STS during the combined processing of static faces and non-verbal emotion vocalizations. Kreifelts, Ethofer, Shiozawa, Grodd and Wildgruber (2009) reported a functional segregation of emotion processing in the STS with the trunk showing voice-selective activation whilst the ascending branch showed face sensitive activation. Grippingly, in the middle regions of the STS that spatially overlapped between face – and voice-selective regions, activity during audio-visual emotion recognition was recorded.

Hence, it is possible that the similarities of emotion recognition across separate modalities - as demonstrated by the current behavioural results – smooth the way for the integration of emotions from several modalities in the STS. Alternatively, every-day exposure to multimodal emotion expressions may have strengthened the associations of emotion representations from individual modalities so that the recognition from one isolated modality is associated with the recognition from another isolated modality. This could also possibly explain how the degree of modality similarity developed throughout childhood (see Chapter 3) as the associations between emotions expressed in individual modalities become stronger with every-day experience of multimodal emotion expressions.

6.2. Nature of Emotions

One main finding of the current thesis was that, although emotion recognition may be anchored in our genetic make-up, this ability improved with age, in line with research by de Sonneville et al. (2002) and others. The present thesis attempted to shed some light on how much variance is accounted for by nature or nurture effects during emotion recognition by including developmental as well as cross-cultural data. As presented in the following paragraphs, the present thesis has delivered support for – at least to some degree – universal emotion recognition abilities across age and culture, but also specific emotion recognition abilities that varied according to age and culture.

6.2.1. Early Emotion Recognition Mechanisms

Do humans have an innate core system of emotion recognition? One perspective of emotion recognition is the view of emotions as a 'natural kind' (Izard, 2007; Panksepp, 2007) which suggests that emotions are of biological and categorical origin that is enrooted in every mammalian brain. Universal biological responses to arousing events can then be interpreted as emotions (James, 1884; Lange, 1885). In support of this statement, the current thesis demonstrated firstly that children showed similar emotion recognition patterns to adults and secondly that British and German participants showed similarities in basic emotion recognition despite belonging to separate cultures.

As suggested by Darwin (1872) and Izard (1994), humans may be born with an innate emotion expression as well as recognition capacity as part of evolution and successful threat avoidance mechanisms. This can also be seen in 5 month-olds' ability to discriminate between different types of emotions expressed in faces (*e.g.* Bornstein &Arterberry, 2003). In the developmental Chapter 3, the similar rating patterns of emotions across two separate modalities were found for children and adults alike. Those patterns also seem to be largely independent of gender. Findings suggest that certain basic aspects of emotion recognition are independent of external effects and emotion recognition may exist from an early age onwards. However, it is important to note that the present thesis did not include children under the age of 5 and hence the researcher cannot draw conclusions about innate abilities of emotion recognition that are evident from birth. In addition, children's responses were independent of the age of actors suggesting that children did not prefer to rate emotions from members of the same age-group as suggested by Bracovic et al. (2012). Those arguments are in favour of the idea that we may have a general emotion system that recognises prototypical emotions in others, independently of external factors such as age of actors.

Does culture contribute to the way we experience and recognise emotions in faces and voices? Previously, Izard (1980) stated that basic emotions are recognised universally across literate as well as illiterate cultures. Current results from the adult sample in Chapter 4 partly support the idea that there are indeed certain aspects of emotion recognition that may be universal. For example, Chapter 4 reported scores from a relatively large-scale questionnaire on self-reported emotional reactivity that were comparable between British and German participants (N = 215). Looking at emotion recognition scores, there were no main effects of culture for either the adult or the child sample. Further, the current results for adults suggested that happy stimuli were rated preferentially compared to angry stimuli and this was found to be true for both countries. For children, general emotion recognition performance improved across both countries and older children showed almost identical emotion recognition scores for all six emotions between Germany and Great Britain. This finding is interesting as it suggests that emotion recognition processes in older children may be guided by culture-independent developmental factors such as brain maturation or general school attendance. Hypothetically, it is possible that we humans have an innate emotion core-system that is universal but at the same time also universally programmed to develop in certain predetermined stages, independent of external factors such as culture and society.

6.2.2. Learnt Emotion Recognition Mechanisms

On the other hand, it is possible that emotion recognition is additionally influenced by external factors and may change throughout a lifetime. For example, older adults often show a positivity bias in emotion recognition (Isaacowitz et al., 2006), which, according to learning theories, may be based on past experiences and learnt positive or negative reinforcements (Dollard & Miller, 1941). Results from the present thesis support the idea of emotion recognition abilities that may be influenced by developmental as well as by cultural aspects. It is commonly believed that emotion recognition abilities mature with age throughout childhood (Herba & Phillips, 2004) and even until adulthood (Thomas, de Bellis, Graham, & LaBar; 2007). This has for example been related to the development of social brain structures such as the anterior temporal cortex throughout childhood (Mills et al., 2014). In line with appraisal theories (*e.g.* Scherer, 1984), it is possible that cognitive appraisal of events such as novelty and pleasantness changes with age.

Current results from the developmental Chapter 3 support the idea of age-related effects on emotion processing. For example, modality ratings became more similar with age. This finding is intriguing as it poses the idea that younger children may base emotion recognition on perceptual differences across modalities whilst older children may recognise emotions independently of the modality and perceptual features. Up until adulthood, RTs as well as intensity ratings also increased with age; this was true for both modalities. Again, this is in support of appraisal theories (*e.g*. Frijda et al., 1989; Scherer, 1984) as this suggests that general cognitive evaluation of events takes place in adulthood which is independent of modality.

In line with one assumption of the current thesis, which stated that emotion recognition may be based on inferring mental states of others, children's Theory of Mind abilities also increased with age which in turn had a positive impact on speed for rating emotions in faces as well as in voices. Further, children rated child faces with higher intensity than adult faces when they expressed anger. Interestingly, the recognition of angry faces may be context-dependent: Children may be more programmed to recognise anger in others when the emotion is expressed by members of their own age-group. Lastly, in support of importance of age on emotion-recognition, results from the adult Chapter 2 implied that emotion intensity ratings as well as rating speed were negatively affected by increasing age. Overall, findings from the present thesis have outlined that external factors such as age may indeed influence the pathway of emotion development. The age-variable could be replaced with several other interacting factors such as cognitive development – such as Theory of Mind abilities.

Although basic emotions are believed to be recognised universally (Izard, 1980), it is now commonly accepted that cultural norms and display rules (Buck, 1984) may influence emotion recognition. The present results of Chapter 4 are in line with the Dialect theory (Elfenbein & Ambady, 2003) which suggests that otherwise universal emotion recognition processes may be coloured by culture-specific factors. For example, results from Chapter 4 suggested that British children rated emotions faster than German children whilst intensity ratings, did not differ between countries. Since the youngest children in Germany visited a nursery whereas the youngest children in Britain already visited primary schools, it is possible that external social factors such as school attendance influence emotion recognition processes. Further, although Chapter 4 did not find main effects for culture within the adult sample, British adults nonetheless rated 50% of emotion categories as more intense than German adults. The exception was for angry faces which were rated with higher intensity by the German than the British samples. Depending on cultural background, it is possible that subjective appraisal of events may have been altered which accounts for variation in emotion recognition across cultures (Scherer, 1984). However, it is possible that those social norms and display rules only take effect in adulthood because older children showed almost identical emotion intensity perception for all six emotion categories across both countries.

6.3. Appropriateness of Design & Stimulus Choice

A natural draw-back of past literature on emotion recognition across modalities is the vast amount of different measures, stimuli and samples used (Bayer & Schacht, 2014). Hence, the current within-subject design allowed for direct comparison of emotion recognition across faces and voices. Further, by rating each emotion on multiple emotion-scales, forced-choice answers were avoided which in turn reduced methodological artefacts, especially when comparing across cultures (Russell, 1994). The choice of stimuli from the present thesis proved to be appropriate to investigate emotion recognition across children and adults from different cultures. For example, all actors as well as participants were White-Caucasian in order to reduce the possibility of out-group biasing during emotion recognition (Russell, 1994). The Montreal Affective Voices (Belin et al., 2008) are based on non-verbal features of the voice which facilitated emotion research across different cultures and children which may not be as verbally developed as adults.

Although the Ekman POFA stimuli date back to 1975 and only contain black and white photographs, it is one of the most widely used and validated emotion face databases (Belin et al., 2008).

The present developmental study included static face expressions as well as dynamic, non-verbal affect bursts. Previous research suggested comparable task-difficulty for non-verbal affect vocalisations and face expressions (Hawk et al., 2009) whilst emotions recognised from speech prosody showed higher error rates (see also Sauter, Panattoni, & Happe, 2012). Present findings agreed with this and it seems possible that for static faces as well as for non-verbal affect bursts, emotions were portrayed in iconic and prototypical ways which are relatively easy to recognise across both modalities. This impression has been supported by Simon-Thomas et al. (2009) who reported higher recognition rates for prototypical compared to standard emotion expressions in vocal affect bursts. However, it is to note that whilst the current data supports the notion of participants' comparable emotion recognition abilities for vocal affect bursts and static face expressions, this may not apply to more general vocal emotion recognition paradigms such as prosody.

Further, in order to investigate whether the current results could also be extended to dynamic face expressions, future studies could include dynamic 3D emotional face stimuli as developed by Girges, Spencer, & O'Brien (in print). So far, inferences from the present thesis are limited to judgements from categorical and static emotion expressions only. Stimuli used in the present thesis are also acted rather than spontaneous which allows a higher control in the experimental setting. However, acted emotion expressions may not be as realistic as spontaneously elicited emotions as their prototypical and stylistic representation of emotions may inflate recognition rates compared to prosody (Hawk et al., 2009; Wallbott & Scherer, 1986). Hence, the stimulus choice may have some influence on generalizability of results although Sauter (2014) suggested that affect bursts in general, whether acted or spontaneous, seem to be well recognised and appropriate for emotion research.

The present study included six basic emotions as stated by Ekman and Friesen (1973). Whilst this covers a good range of emotion expression, it has been suggested that more than six emotions need to be studied in order to ensure that judgements aren't simply based on discrimination processes (Banse & Scherer, 1996). This issue also applies to the happiness superiority effect consistently found in the present behavioural chapters. Because the present thesis compared one positive emotion (happiness) with several negative emotions, it is possible that the happiness superiority effect is an artefact of unbalanced valence of emotion included. For a more detailed discussion of this issue see Chapter 2 and 3. In line with this, the current ERP Chapter 5 only included one positive and one negative emotion and consequently could not find happiness superiority effects as reflected in neurological data. However, it is important to remember that the ERP Chapter 5 is a preliminary investigation only and further neurophysiological studies could extend the current findings.

Lastly, in the current thesis, consistent findings of similarities between modalities for emotion recognition dominated; however, for some emotions such as happiness, emotion recognition in faces was faster than in voices. In a natural setting, humans usually integrate sensory information from multiple modalities. Although we tend to perceive inputs across modalities and often regard them as occurring simultaneously, the time course of processing visual and auditory signals in natural multimodal stimuli depends on the modality. For example – and in concordance with the current findings - light travels faster than sound waves (Fain, 2003) and hence, this may aid emotion recognition within the visual modality. This could be especially helpful in situations where auditory information is limited such as in noisy environments. However, the subsequent chemical transduction of sound into electrical signals within the human brain is faster than the transduction of light (King, 2005), reducing the gap in processing time between visual and auditory stimuli. Indeed, in some realistic settings such as when talking to someone far away or in darkness, visual information is limited and auditory signals are the primary source of information. It has been stated that humans may integrate sensory information from different inputs by creating a temporal time-window in which the synchronization of sensory information can occur. During this time-window, the time-delay between visual and auditory information can be eliminated and information is synchronized (Sugita & Suzuki, 2003).

Hence, it is important to notice that during multimodal integration of emotional stimuli from several modalities, the processing of sensory information may occur with a time-lag which in turn has to be processed by the brain accordingly. Signals from faces and voices –for example during audio-visual speech – are often structured in a way so that the receiver perceives them simultaneously. For example, there seem to be a high correlation between perceived changes in mouth movements (visual input) and the auditory envelope (sound input) during the perception of audio-visual speech (Chandrasekaran, Trubanova, Stillittano, Caplier, & Ghazanfar, 2009). The current findings of high similarities in emotion recognition between faces and voices could possibly be a co-occurrence of successful integration of information during everyday life. Nonetheless, in the present thesis, similarities of emotion recognition from faces and voices were investigated within the laboratory setting where the information came from two separate sources: *e.g.* sound was delivered via headphones whilst vision was delivered via PC screens and signals were not presented simultaneously. Consequently, ecological validity may be somehow limited and future studies could address this issue.

6.4. Future Directions

To the best of the researcher's knowledge, no previous studies have investigated whether neural structures in children and adults are equally active during the processing of emotions in faces as well as in voices and whether those brain regions change over time. Since the present thesis has demonstrated that faces and voices show similarities in their behavioural rating profiles in adults, identifying the location of underlying brain activity in a within-subject design could help to answer the question about a modality-independent emotion core network. For example, is there a dissociation between regions that are active for emotion recognition in faces and regions that are active for emotion recognition in voices? Further, is there neurological evidence that the similarity of emotion recognition across modalities becomes stronger with age, as suggested by the data from the present thesis?

In terms of processing abnormalities, another future direction of the current thesis could be to look at emotion processing abilities across modalities in individuals with autism spectrum disorders. Autism has been found to lead to impairment in emotion recognition from faces and also from voices (Philip et al., 2010) and in multimodal human stimuli (Mongillo, Irwin, Whalen, Klaiman, Carter, & Schultz, 2008). It is possible that important emotion brain structures such as the STS (Gervais et al., 2004) or amygdala (Baron-Cohen, Ring, Bullmore, Wheelwright, Ashwin, & Williams, 2000) show reduced activation during emotion processing in autism. To the best of the researcher's knowledge, there is no previous study that has investigated emotion recognition from non-verbal affect bursts – rather than prosody – and face processing in a within-subject design in individuals with autism. Extending the type of stimuli used in emotion research could help to identify whether individuals with abnormal processing have general emotion processing deficits or whether this may depend on the type of stimulus presentation.

Further, assuming modality-independent emotion networks, it would be interesting to create interventions to increase emotion recognition abilities in individuals with emotion recognition deficits across a range of modalities. This has previously been proven effective for facial emotion recognition in children with high-functioning Autism (Baron-Cohen et al., 2009) and it has previously been found that giving specific instructions to pay attention to emotions expressed in faces as well as in voices enhanced neural brain activity in prefrontal areas in individuals with autism. However, there is a striking lack of auditory interventions which limits the conclusions about modality-independent mechanisms one can draw from existing autism-intervention studies.

In conclusion, behavioural results from the current thesis suggest a general core emotion network that may act independently of modality, as previously suggested (*e.g.* Peelen et al., 2010). This may be influenced by developmental or external factors such as culture. It seems that both faces and voices communicate emotion signals in parallel which allows us to recognise the other person's emotional state. Hence, in light on the introductory Shakespeare's quote from the Macbeth play (2010/1699), we may actually be able to find the mind's construction – not only in the face, but also in the voice.

Words: 67.081

References

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences, 3*(12), 469– 479. doi:10.1016/S1364-6613(99)01399-6

Adolphs, R., Tranel, D., & Damasio, H. (2001). Emotion Recognition From Faces and Prosody Following Temporal Lobectomy. *Neuropsychology*, 15(3), 396-404. DOI: 10.1037//0894-4105.15.3.396

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, *12*, 169–177. doi:10.1016/S0959-4388(02)00301-X

Aguado, L., García-Gutierrez, A., & Serrano-Pedraza, I. (2009). Research Articles. Symmetrical interaction of sex and expression in face classification tasks. *Attention, Perception & Psychophysics. 71*(1), 9–25. doi:10.3758/APP

Aguert, M., Laval, V., Lacroix, A., Gil, S., & Le Bigot, L. (2013). Inferring emotions from speech prosody: not so easy at age five. *PloS One*, *8*(12), e83657. doi:10.1371/journal.pone.0083657

Aguirre, G. K., Zarahn, E., & D'esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *NeuroImage*, *8*(4), 360–9. doi:10.1006/nimg.1998.0369

Alain, C., Arnott, S, Hevenor, S., Graham, S., & Grady. C. (2001). "What" and "where" in the human auditory system. *Proceedings of the National Academy of Sciences*, 98 (21), 12301–12306

Albanese, O., De Stasio, S., Di Chiacchio, C., Fiorilli, C., & Pons, F. (2010). Emotion comprehension: the impact of nonverbal intelligence. *The Journal of Genetic Psychology*, *171*(2), 101–15. doi:10.1080/00221320903548084

Anderson, A. K., & Phelps, E. A. (1998). Intact recognition of vocal expressions of fear following bilateral lesions of the human amygdala. *Neuroreport*, *9*(16), 3607–13. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9858368

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

Armony, J. L. & Sergerie, K. (2007). Own-sex effects in emotional memory for faces. *Neuroscience Letters. 426*, 1–5. doi:10.1016/j.neulet.2007.08.032

Ashley, V., Vuilleumier, P., & Swick, D. (2004). Time course and specificity of event-related potentials to emotional expressions. *Neuroreport*, *15*(1), 211–216. doi:10.1097/01.wnr.0000091411.19795.f5

Aubé, W. Angulo-Perkins, A., Peretz, I., Concha, L., & Armony, J. (2014). Fear across the senses: Brain responses to music, vocalizations and facial expressions. *Social Cognitive and Affective Neuroscience.* pii: nsu067 [Epub ahead of print]

Babchuk, W., Hames, R., & Thompson, R. (1985). Sex differences in the recognition of infant facial expressions of emotion: The primary caretaker hypothesis. *Ethology and Sociobiology*, *6*(2), 89–101. doi:10.1016/0162-3095(85)90002-0

Bach, D., Hurlemann, R. & Dolan, R. (2013). Unimpaired discrimination of fearful prosody after amygdala lesion. *Neuropsychologia*, *51*(11), 2070–4. doi:10.1016/j.neuropsychologia.2013.07.005

Baird, A., Gruber, S., Fein, D., Maas, L., Steingard, R., Renshaw, P., Cohen, B. & Yurgelun-Todd, D. (1999): Functional magnetic resonance imaging of facial affect recognition in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 38(2):195-9

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology,* 1(6), 589-595

Banissy, M. J., Sauter, D. A., Ward, J., Warren, J. E., Walsh, V., & Scott, S. K. (2010). Suppressing sensorimotor activity modulates the discrimination of auditory emotions but not speaker identity. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(41), 13552–7. doi:10.1523/JNEUROSCI.0786-10.2010

Banse, R. & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614–636. doi:10.1037//0022-3514.70.3.614

Baron-Cohen, S., Golan, O., & Ashwin, E. (2009). Can emotion recognition be taught to children with autism spectrum conditions? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 364(1535), 3567–3574. doi: 10.1098/rstb.2009.0191

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind” ? *Cognition*, *21*(1), 37–46. doi:10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1430), 361–74. doi:10.1098/rstb.2002.1206

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. R. (2000). The amygdala theory of autism. *Neuroscience & Biobehavioral Reviews*, *24*(3), 355–364. doi:10.1016/S0149-7634(00)00011-7

Barrett, H. C. & Bryant, G. (2008). Vocal Emotion Recognition Across Disparate Cultures. *Journal of Cognition and Culture*, *8*(1), 135–148. doi:10.1163/156770908X289242

Baseler, H. A, Harris, R. J., Young, A. W., & Andrews, T. J. (2014). Neural responses to expression and gaze in the posterior superior temporal sulcus interact with facial identity. *Cerebral Cortex, 24*(3), 737–44. doi:10.1093/cercor/bhs36

Batty, M. & Taylor, M. J. (2003). Early processing of the six basic facial emotional expressions, *Cognitive Brain Research*, *17*, 613–620

Baumeister, R. & Leary, M. (1995). The need to belong: Desire for interpersonal attachment as a fundamental human motivation. *Psychological Bulletin,* 117(3), 497-529

Bayer, M., & Schacht, A. (2014). Event-related brain responses to emotional words, pictures, and faces - a cross-domain comparison. *Frontiers in Psychology, 5*, 1–10. doi:10.3389/fpsyg.2014.01106

Becker, D., Anderson, U., Mortensen, C., Neufeld, S., & Neel, R. (2011). The face in the crowd effect unconfounded: happy faces, not angry faces, are more efficiently detected in single- and multiple-target visual search tasks. *Journal of Experimental Psychology. General*, *140*(4), 637–59. doi:10.1037/a0024060

Becker, D., Kenrick, D., Neuberg, S., Blackwell, K., & Smith, D. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*(2), 179–90. doi:10.1037/0022-3514.92.2.179

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129–35. doi:10.1016/j.tics.2004.01.008

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology (London, England : 1953)*, *102*(4), 711–25. doi:10.1111/j.2044-8295.2011.02041.x

Belin, P., Fillion-Bilodeau, S., & Gosselin, F. (2008). The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing. *Behavior Research Methods*, *40*(2), 531–539. doi:10.3758/BRM.40.2.531

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

Benton, A. L. (1980). The neuropsychology of facial recognition. *American Psychologist,* 35(2),176- 86.

Berndt, T. J., & Hoyle, S. G. (1985). Stability and change in childhood and adolescent friendships. *Developmental Psychology*, *21*(6), 1007–1015. doi:10.1037//0012-1649.21.6.1007

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger, *Brain.* 122, 883–893

Blais, C., Fiset, D., & Gosselin, F. (2013). Gender differences in the visual strategies underlying facial expression categorization. *Journal of Vision*, *13*(9), 598–598. doi:10.1167/13.9.598

Blank, H., Anwander, A., & von Kriegstein, K. (2011). Direct structural connections between voice- and face-recognition areas. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *31*(36), 12906–15. doi:10.1523/JNEUROSCI.2091-11.2011

Blau, V. C., Maurer, U., Tottenham, N., & McCandliss, B. D. (2007). The face-specific N170 component is modulated by emotional facial expression. *Behavioral and Brain Functions: BBF*, *3*, 7. doi:10.1186/1744-9081-3-7

Blonder, L. X., Bowers, D., & Heilman, K. M. (1991). The role of the right hemisphere in emotional communication. *Brain*, *114*(3), 1115-1127

Bodamer, J. (1947): Die Prosop-Agnosie. *Archiv für Psychiatrie und Nervenkrankheiten.* 179, 6-53

Bonora, A., Benuzzi, F., Monti, G., Mirandola, L., Pugnaghi, M., Nichelli, P., & Meletti, S. (2011). Recognition of emotions from faces and voices in medial temporal lobe epilepsy. *Epilepsy & Behavior: E&B*, *20*(4), 648–54. doi:10.1016/j.yebeh.2011.01.027

Bornstein, M. H. & Arterberry, M. E. (2003). Recognition, discrimination and categorization of smiling by 5-month-old infants. *Developmental Science, 6*(5), 585–599. doi:10.1111/1467-7687.00314

Borod, J. C., Pick, L. H., Hall, S., Sliwinski, M., Madigan, N., Obler, L. K., … Tabert, M. (2000). Relationships among Facial, Prosodic and Lexical Channels of Emotional Perceptual Processing. *Cognition and Emotion*, 14 (2), 193-211

Bosking W., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of Neuroscience*, 15;17(6):2112-27

Bostanov, V., & Kotchoubey, B. (2004). Recognition of affective prosody: Continuous wavelet measures of event-related brain potentials to emotional exclamations. *Psychophysiology, 41*(2), 259–268. doi:10.1111/j.1469-8986.2003.00142.x

Bracovic, A., de Heering, A., & Maurer, D. (2012). The Effect of Starting School on Preschoolers' Ability to Recognize Child and Adult Faces. *Journal of Vision*, *12*(9), 23–23. doi:10.1167/12.9.23

Brendel, E., DeLucia, P. R., Hecht, H., Stacy, R. L., & Larsen, J. T. (2012). Threatening pictures induce shortened time-to-contact estimates. *Attention,74,* 979–987

British Chamber of Commerce in Germany e.V., (2014). British German trade at record level in 2013. Retrieved from http://british-chamber.blogspot.co.uk/2014_04_01_archive.html

Brodbeck, F. C., Frese, M., Akerblom, S., Audia, G., Bakacsi, G., Bendova, H., … Wunderer, R. (2000). Cultural variation of leadership prototypes across 22 European countries. *Journal of Occupational and Organizational Psychology*, *73*(1), 1–29. doi:10.1348/096317900166859

Brown, J. R, & Dunn, J. (1996). Continuities in emotion understanding from three to six years. *Child Development*, *67*(3), 789–802. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8706526

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

Brüne M. (2005). Emotion recognition, 'theory of mind,' and social behavior in schizophrenia. *Psychiatry Research,* 133(2-3):135-47

Buck, R. (1980). Nonverbal behavior and the theory of emotion: the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *38*(5), 811–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7381683

Bundesverband der Natur – und Waldkindergaerten in Deutschland e.V. (2014). Retrieved from http://bvnw.de/

Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly*, *25*(2), 140–165. doi:10.1016/j.ecresq.2009.11.001

Cairns, R. B., Leung, M. C., Buchanan, L., & Cairns, B. D. (1995). Friendships and social networks in childhood and adolescence: fluidity, reliability, and interrelations. *Child Development*, *66*(5), 1330–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7555219

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience, 3*(11), 1077–8. doi:10.1038/80586

Calder, A. J. & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews. Neuroscience*, *6*(8), 641–51. doi:10.1038/nrn1724

Cantalupo, G., Meletti, S., Miduri, A., Mazzotta, S., Rios-Pohl, L., Benuzzi, F., … Cossu, G. (2013). Facial emotion recognition in childhood: the effects of febrile seizures in the developing brain. *Epilepsy & Behavior : E&B*, *29*(1), 211–6. doi:10.1016/j.yebeh.2013.07.007

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism: The International Journal of Research and Practice*, *9*(4), 428–49. doi:10.1177/1362361305056082

Castro, V. L., Halberstadt, A. G., Lozada, F. T., & Craig, A. B. (2014). Parents' Emotion-Related Beliefs, Behaviours, and Skills Predict Children's Recognition of Emotion. *Infant and Child Development* (ahead of print), doi:10.1002/icd.1868

Ceccarini, F. & Caudek, C. (2013). Anger superiority effect: The importance of dynamic emotional facial expressions. *Visual Cognition*, *21*(4), 498–540. doi:10.1080/13506285.2013.807901

Charest, I., Pernet, C. R., Rousselet, G. a, Quiñones, I., Latinus, M., Fillion-Bilodeau, S., … Belin, P. (2009). Electrophysiological evidence for an early processing of human voices. *BMC Neuroscience*, *10*, 127. doi:10.1186/1471-2202-10-127

Children and Family Court Advisory and Support Service CafCass (2009). Annual report and accounts 2009/2010. Children and Family Court Advisory and Support Service. An executive non-departmental public body. HC123, London:The Stationery Office. Available at: http://www.cafcass.gov.uk/media/6428/Cafcass%20Annual%20Report%20and%20Accounts%202009-10%20web%20pdf.pdf

Chronaki, G., Broyd, S., Garner, M., Hadwin, J. a., Thompson, M. J. J., & Sonuga-Barke, E. J. S. (2012). Isolating N400 as neural marker of vocal anger processing in 6-11-year old children. *Developmental Cognitive Neuroscience*, *2*(2), 268–276

Cohen, M., Prather, A., Town, P., & Hynd, G. (1990). Neurodevelopmental differences in emotional prosody in normal children and children with left and right temporal lobe epilepsy. *Brain and Language*, *38*(1), 122–34. doi:10.1016/0093-934X(90)90105-P

Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia*, *48*(1), 220–5. doi:10.1016/j.neuropsychologia.2009.09.007

Cutting, A. L. & Dunn, J. (1999). Theory of Mind, Emotion Understanding, Language, and Family Background: Individual Differences and Interrelations. *Child Development*, *70*(4), 853–865. doi:10.1111/1467-8624.00061

Dalrymple, K. A, Gomez, J., & Duchaine, B. (2013). The Dartmouth Database of Children's Faces: acquisition and validation of a new face stimulus set. *PloS One*, *8*(11), e79131. doi:10.1371/journal.pone.0079131

Damjanovic, L., Pinkham, A. E., Clarke, P., & Phillips, J. (2014). Enhanced threat detection in experienced riot police officers: cognitive evidence from the face-in-the-crowd effect. *Quarterly Journal of Experimental Psychology (2006)*, *67*(5), 1004–18. doi:10.1080/17470218.2013.839724

Darwin, C. (1965). *The expression of the emotions in man and animals.* Chicago: University of Chicago Press. (Original work published 1872)

De Bellis, M. D., Keshavan, M. S., Beers, S. R., Hall, J., Frustaci, K., Masalehdan, a, … Boring, a M. (2001). Sex differences in brain maturation during childhood and adolescence. *Cerebral Cortex, 11*(6), 552–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11375916

De Heering, A., Houthuys, S., & Rossion, B. (2007). Holistic face processing is mature at 4 years of age: evidence from the composite face effect. *Journal of Experimental Child Psychology*, *96*(1), 57–70. doi:10.1016/j.jecp.2006.07.001

Dellacherie, D., Hasboun, D., Baulac, M., Belin, P., & Samson, S. (2011). Impaired recognition of fear in voices and reduced anxiety after unilateral temporal lobe resection. *Neuropsychologia*, *49*(4), 618–29. doi:10.1016/j.neuropsychologia.2010.11.008

Demenescu, L., Mathiak, K.A., & Mathiak, K. (2014). Age- and gender-related variations of emotion recognition in pseudowords and faces. *Experimental Aging Research*, 40(2),187-207. doi: 10.1080/0361073X.2014.882210.

Denham, S., Mitchell-Copeland, J., Strandberg, K., Auerbach, S., & Blair, K. (1997). Parental Contributions to Preschoolers ' Emotional Competence : Direct and Indirect Effects. *Motivation and Emotion, 21*(1), 65–86.

Department of Health (2014). NHS 'SmokeFree'. Retrieved from http://www.nhs.uk/smokefree

Department for Transport (2014). 'Think!'. http://think.direct.gov.uk/index.html

Derek, M., Isaacowitz, D., & Stanley, J. (2011). Bringing an Ecological Perspective to the Study of Aging and Recognition of Emotional Facial Expressions: Past, Current, and Future Methods. *Journal of Nonverbal Behaviour.* 35, (4), 261-278

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial Identity and Facial Emotions: Speed , Accuracy , and Processing Strategies in Children and Adults. *Journal of Clinical and Experimental Neuropsychology*. 24(2), 200-213

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping, 15*(2), 95–111. doi:10.1002/hbm.10010

Du, S. & Martinez, A. (2013). Wait , are you sad or angry ? Large exposure time differences required for the categorization of facial expressions of emotion. *Journal of Vision*, *13*, 1–14. doi:10.1167/13.4.13.doi

Duchaine, B. & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience,* 17(2): 249-261

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young Children's Understanding of Other People's Feelings and Beliefs: Individual Differences and Their Antecedents. *Child Development.* 62(6), 1352-1366. DOI: 10.1111/j.1467-8624.1991.tb01610.x

Ebner, N. C. & Johnson, M. K. (2009). Young and older emotional faces: are there age group differences in expression identification and memory? *Emotion (Washington, D.C.)*, *9*(3), 329–39. doi:10.1037/a0015179

Egger, H. L., Pine, D. S., Nelson, E., Leibenluft, E., Ernst, M., Towbin, K. E., & Angold, A. (2011). The NIMH Child Emotional Faces Picture Set (NIMH-ChEFS): a new set of children's facial emotion stimuli. *International Journal of Methods in Psychiatric Research*, *20*(3), 145–56. doi:10.1002/mpr.343

Eimer, M. & Holmes, A. (2007): Event-related brain potential correlates of emotional face processing. *Neuropsychologia* 45 (2007) 15–31. doi:10.1016/j.neuropsychologia.2006.04.022

Eimer, M., Holmes, A., & Mcglone, F. P. (2003). The role of spatial attention in the processing of facial expression: An ERP study of rapid brain responses to six basic emotions. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(2), 97–110

Eimer, M. (2011). In : Calder A . J . et al . (eds). *The Oxford Handbook of Face Perception . Oxford* University Press , 2011 . The Face-Sensitive N170 Component of the Event-Related Brain Potential, 329–344

Engelien, A., Schulz, M., Ross, B., Arolt, V., & Pantev, C. (2000). A combined functional in vivo measure for primary and secondary auditory cortices. *Hearing research*, *148*(1), 53-160

Eisenberg, N., Cumberland, A., Spinrad, T. L., Eisenberg, N., Cumberland, A., & Spinrad, T. L. (2009). Parental Socialization of Emotion. *Psychological Inquiry : An International Journal for the Advancement of Psychological Theory.* 9(4), 241-273. doi:10.1207/s15327965pli0904

Ekman, P. & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology.* 17(2), 124-129

Ekman, P. & Friesen, W. V. (1975). *Pictures of facial affect*. Consulting Psychologists Press.

Ekman, P. & Friesen, W. V. (1975). *Unmasking The Face: A Guide To Recognizing Emotions From Facial Clues.* New Jersey: Prentice Hall, 1975. Reprint edition, 2003 Cambridge, Ma: Malor Books

Ekman, P. & Friesen, W. (1976): Measuring Facial Movement. *Environmental psychology and nonverbal behaviour.* 1(1), 56–75.

Elfenbein, H. A. & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, *128*(2), 203–235. doi:10.1037//0033-2909.128.2.203

Elfenbein, H. A., Beaupré, M., Lévesque, M., & Hess, U. (2007). Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions. *Emotion (Washington, D.C.), 7*(1), 131–46. doi:10.1037/1528-3542.7.1.131

Elfenbein, H. A. (2013). Nonverbal Dialects and Accents in Facial Expressions of Emotion. *Emotion Review*, *5*(1), 90–96. doi:10.1177/1754073912451332

Ellsworth, P. C. & Scherer, K. R. (2003). Appraisal processes in emotion. In R. Davidson, K. R.

Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 572_595). New York: Oxford University Press.

Ernst, M., Luckenbaugh, D., Moolchan, E., Temple, V., Jenness, J., Korelitz, K., … Kimes, A. (2010). Decision-making and facial emotion recognition as predictors of substance-use initiation among adolescents. *Addictive Behaviors, 35*(3), 286–9. doi:10.1016/j.addbeh.2009.10.014

Etcoff, N. &, Magee, J. (1992). Categorical perception of facial expressions. *Cognition.* 44(3):227-40

Ethofer, T., Anders, S., Wiethoff, S., Erb, M., Herbert, C., Saur, R., … Wildgruber, D. (2006). Effects of prosodic emotional intensity on activation of associative auditory cortex. *Neuroreport*, *17*(3), 249–53. doi:10.1097/01.wnr.0000199466.32036.5d

Ethofer, T., Bretscher, J., Gschwind, M., Kreifelts, B., Wildgruber, D., & Vuilleumier, P. (2012). Emotional voice areas: anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. *Cerebral Cortex (New York, N.Y. : 1991)*, *22*(1), 191–200. doi:10.1093/cercor/bhr113

Eurydice National Foundation for Educational Research (2013). Compulsory age of starting school in European countries. Available at http://www.nfer.ac.uk/nfer/index.cfm?9B1C0068-C29E-AD4D-0AEC-8B4F43F54A28

Fabbri-Destro, M. & Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *Physiology (Bethesda, Md.)*, *23*, 171–9. doi:10.1152/physiol.00004.2008

Fain, G. L., & Fain, M. J. (2003). *Sensory transduction* (pp. 101-111). Sunderland: Sinauer Associates

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychological Review*, *105*(3), 482–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9697428

Farrington, D. P. (1991). Longitudinal research strategies: advantages, problems, and prospects. *Journal of the American Academy of Child and Adolescent Psychiatry*, *30*(3), 369–74. doi:10.1097/00004583-199105000-00003

Fecteau, S., Belin, P., Joanette, Y., & Armony, J. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage* 36, 480–487. doi:10.1016/j.neuroimage.2007.02.043

Feldmar, G. (1979). Prosopo-affective agnosia as a symptom of cerebral organic disease. *Journal of the American Geriatrics Society*. 27(5):225-230

Felisberti, F., Chipi, E., & Trueman, J. (2014). Affective facial microexpressions and adult attachment. *Perception* 43, ECVP Abstract Supplement, page 157

Fitzgerald, D. a, Angstadt, M., Jelsone, L. M., Nathan, P. J., & Phan, K. L. (2006). Beyond threat: amygdala reactivity across multiple expressions of facial affect. *NeuroImage*, *30*(4), 1441–8

Flack, W. (2006). Peripheral feedback effects of facial expressions, bodily postures, and vocal expressions on emotional feelings. *Cognition & Emotion*, *20*(2), 177–195. doi:10.1080/02699930500359617

Frijda, N. H. (1986). *The emotions*. Cambridge University Press.

Frijda, N., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, 57(2), 212-228. http://dx.doi.org/10.1037/0022-3514.57.2.212

Gabard-Durnam, L. J., Flannery, J., Goff, B., Gee, D. G., Humphreys, K. L., Telzer, E., … Tottenham, N. (2014). The development of human amygdala functional connectivity at rest from 4 to 23 years: a cross-sectional study. *NeuroImage*, *95*, 193–207. doi:10.1016/j.neuroimage.2014.03.038

Gagnon, M., Gosselin, P., Hudon-ven der Buhs, I., Larocque, K., & Milliard, K. (2009). Children's Recognition and Discrimination of Fear and Disgust Facial Expressions. *Journal of Nonverbal Behavior*, *34*(1), 27–42. doi:10.1007/s10919-009-0076-z

Gallagher, H. ., Happé, F., Brunswick, N., Fletcher, P. ., Frith, U., & Frith, C. (2000). Reading the mind in cartoons and stories: an fMRI study of "theory of mind" in verbal and nonverbal tasks. *Neuropsychologia*, *38*(1), 11–21. doi:10.1016/S0028-3932(99)00053-6

Gallese, V. & Goldman, A. (1998). Mirror neurons and the mind-reading, *Trends in Cognitive science*, *2*(12), 493–501

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), 1645–54

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, a W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*(2), 191–7. doi:10.1038/72140

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014a). Cultural relativity in perceiving emotion from vocalizations. *Psychological Science*, *25*(4), 911–20, doi:10.1177/0956797613517239

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014b). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion*, *14*(2), 251

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Current Biology*, *18*(11), R457–60. doi:10.1016/j.cub.2008.03.030

Golan, O. & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology*, pp 591-617. doi:10.1017/S0954579406060305.

Goldman, A. I. & Sripada, C. S. (2005). Simulationist models of face-based emotion recognition. *Cognition*, *94*(3), 193–213. doi:10.1016/j.cognition.2004.01.005

Goldstein, E. B. (2002). *Sensation and Perception.* 6th edition. Belmont, CA: Wadsworth

Goleman, D. (1995). *Emotional intelligence.* Bantam Books, New York (1995).

Goodale, M. & Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*.15(1):20-5

Gosselin, N., Peretz, I., Johnsen, E., & Adolphs, R. (2007). Amygdala damage impairs emotion recognition from music. *Neuropsychologia*. 45,236–244

Gosselin, P. & Simard, J. (1999). Children's knowledge of facial expressions of emotions: Distinguishing fear and surprise. *The Journal of Genetic Psychology*. 160(2), 181-193

Grandjean, D., Sander, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). The voices of wrath: brain responses to angry prosody in meaningless speech. *Nature Neuroscience*, *8*(2), 145–6. doi:10.1038/nn1392

Grieve, A. (2010). "Aber ganz ehrlich": Differences in episodic structure, apologies and truth-orientation in German and Australian workplace telephone discourse. *Journal of Pragmatics*, *42*(1), 190–219. doi:10.1016/j.pragma.2009.05.009

Grossmann, T., Oberecker, R., Koch, S., & Friederici, A. (2010). The developmental origins of voice processing in the human brain. *Neuron.* 25; 65(6):852-8. doi: 10.1016/j.neuron.2010.03.001

Grossmann, T., Striano, T., & Friederici, A. D. (2007). Developmental changes in infants' processing of happy and angry facial expressions: a neurobehavioral study. *Brain and Cognition*, *64*(1), 30–41. doi:10.1016/j.bandc.2006.10.002

Hagan, C. C., Woods, W., Johnson, S., Calder, A. J., Green, G. G. R., & Young, A. W. (2009). MEG demonstrates a supra-additive response to facial and vocal emotion in the right superior temporal sulcus. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(47), 20010–5. doi:10.1073/pnas.0905792106

Halberstadt, A. G. & Lozada, F. T. (2011). Emotion Development in Infancy through the Lens of Culture. *Emotion Review*, *3*(2), 158–168. doi:10.1177/1754073910387946

Hall, J. & Matsumoto, D. (2004). Gender differences in judgments of multiple emotions from facial expressions. *Emotion (Washington, D.C.)*, *4*(2), 201–6. doi:10.1037/1528-3542.4.2.201

Hampson, E., Van Anders, S., & Mullin, L. (2006). A female advantage in the recognition of emotional facial expressions: test of an evolutionary hypothesis. *Evolution and Human Behavior*, *27*(6), 401–416. doi:10.1016/j.evolhumbehav.2006.05.002

Hansen, C. H. & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology*, 54(6), 917-924. http://dx.doi.org/10.1037/0022-3514.54.6.917

Harris, R. J., Young, A. W., & Andrews, T. J. (2014). Dynamic stimuli demonstrate a categorical representation of facial expression in the amygdala. *Neuropsychologia*, *56*, 47–52. doi:10.1016/j.neuropsychologia.2014.01.005

Harrison, V. & Hole, G. J. (2009). Evidence for a contact-based explanation of the own-age bias in face recognition. *Psychonomic Bulletin & Review*, *16*(2), 264–9. doi:10.3758/PBR.16.2.264

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review*, *10*, 47–66. doi:10.1207/s15327957pspr1001_3

Hatfield, E., Hsee, C. & Costello, J., SchalekampWeisman, M., & Denney, C. (1995). The Impact of Vocal Feedback on Emotional Experience and Expression, *Journal of Social Behaviour and Personality.* 10(2), 293-312.

Hauser, D. J., Preston, S. D., & Stansfield, R. B. (2014). Altruism in the wild: when affiliative motives to help positive people overtake empathic motives to help the distressed. *Journal of Experimental Psychology. General*, *143*(3), 1295–305. doi:10.1037/a0035464

Hawk, S., van Kleef, G., Fischer, A., & van der Schalk, J. (2009). "Worth a Thousand Words": Absolute and Relative Decoding of Nonlinguistic Affect Vocalizations. *Emotion (Washington, D.C.)*, 9(3), 293–305 DOI: 10.1037/a0015178

Haxby, J., Hoffman, E., & Gobbini, I. (2002). Human Neural Systems for Face Recognition and Social Communication. *Biological Psychiatry.* 51(1):59-67

Heerey, E. A, Keltner, D., & Capps, L. M. (2003). Making sense of self-conscious emotion: linking theory of mind and emotion in children with autism. *Emotion (Washington, D.C.)*, 3(4), 394–400. doi:10.1037/1528-3542.3.4.394

Henry, J., Phillips, L., Crawford, J., Ietswaart, M., & Summers, F. (2006). Theory of mind following traumatic brain injury: the role of emotion recognition and executive dysfunction. *Neuropsychologia*, *44*(10), 1623–8. doi:10.1016/j.neuropsychologia.2006.03.020

Herba, C. & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *45*(7), 1185–98. doi:10.1111/j.1469-7610.2004.00316.x

Hess, U., Adams, R. B., Grammer, K., & Kleck, R. E. (2009). Face gender and emotion expression: are angry women more like men? *Journal of Vision*, *9*(12), 19.1–8. doi:10.1167/9.12.19

Hess, U. & Thibault, P. (2009). Darwin and emotion expression. *The American Psychologist*, *64*(2), 120–8. doi:10.1037/a0013386

Hills, P. J. (2012). A developmental study of the own-age face recognition bias in children. *Developmental Psychology*, *48*(2), 499–508. doi:10.1037/a0026524

Hornak, J., Rolls, E.T., & Wade, D. (1996). Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. *Neuropsychologia*. 34 (4), 247–261

House, J. (2006). Communicative styles in English and German. *European Journal of English Studies*, 10(3), 249–267. doi:10.1080/13825570600967721

Iglesias, J. & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4- to 6-month-old infants. *Developmental Psychobiology*, 25: 411–425. doi: 10.1002/dev.420250603

Iredale, J., Rushby, J., McDonald, S., Dimoska-Di Marco, A., & Swift, J. (2013). Emotion in voice matters: Neural correlates of emotional prosody perception. *International Journal of Psychophysiology*, *89*(3), 483–490. doi:10.1016/j.ijpsycho.2013.06.025

Isaacowitz, D., Löckenhoff, C., Lane, R., Wright, R., Sechrest, L., Riedel, R., & Costa, P. (2007). Age differences in recognition of emotion in lexical stimuli and facial expressions. *Psychology and Aging*, 22(1), 147-159. http://dx.doi.org/10.1037/0882-7974.22.1.147

Isaacowitz, D., Wadlinger, H., Goren, D., & Wilson, H. (2006). Is there an age-related positivity effect in visual attention? A comparison of two methodologies. *Emotion (Washington, D.C.)*, 6(3), Aug 2006, 511-516. http://dx.doi.org/10.1037/1528-3542.6.3.511

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*, 288–299.

Izard, C. E. (2007). Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm. *Perspectives on Psychological Science*, *2*(3), 260–280. doi:10.1111/j.1745-6916.2007.00044.x

Jabbi, M., Swart, M., & Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *NeuroImage*, *34*(4), 1744–53. doi:10.1016/j.neuroimage.2006.10.032

Jack, R., Blais, C., Scheepers, C., Schyns, P., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current Biology : CB*, *19*(18), 1543–8. doi:10.1016/j.cub.2009.07.051

Jack, R., Garrod, O., & Schyns, P. (2014). Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time. *Current Biology* 24, 187–192, http://dx.doi.org/10.1016/j.cub.2013.11.064.

Johnstone, T., van Reekum, C. M., Oakes, T. R., & Davidson, R. J. (2006). The voice of emotion: an FMRI study of neural responses to angry and happy vocal expressions. *Social Cognitive and Affective Neuroscience*, *1*(3), 242–9. doi:10.1093/scan/nsl027

Jones, D. C., Abbey, B. B., & Cumberland, A. (1998). The Development of Display Rule Knowledge: Linkages with Family Expressiveness and Social Competence. *Child Development*, 69(4), 1209. doi:10.2307/1132370

Juslin, P. N., & Laukka, P. (2001). Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion (Washington, D.C.)*, *1*(4), 381–412. doi:10.1037//1528-3542.1.4.381

Juth, P., Lundqvist, D., Karlsson, A., & Ohman, A. (2005). Looking for foes and friends: perceptual and emotional factors when finding a face in the crowd. *Emotion (Washington, D.C.)*, *5*(4), 379–95. doi:10.1037/1528-3542.5.4.379

Kappenman, E. S., & Luck, S. J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, *47*(5), 888-904

Kesler-West, M., Andersen, A., Smith, C., Avison, M., Davis, C., Kryscio, R., Blonder, L. (2001). Neural substrates of facial emotion processing using fMRI. *Cognitive Brain Research,* 11(2):213-26

Ketelaars, M. P., van Weerdenburg, M., Verhoeven, L., Cuperus, J. M., & Jansonius, K. (2010). Dynamics of the Theory of Mind construct: A developmental perspective. *European Journal of Developmental Psychology*, *7*(1), 85–103. doi:10.1080/17405620903482081

Khoroshikh, V. V., Ivanova, V. Y., & Kulikov, G. A. (2012). The effect of unconscious color hue saturation on the emotional state of humans. *Human Physiology*, *38*(2), 129–136. doi:10.1134/S0362119712020090

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The Fusiform Face Area : A Module in Human Extrastriate Cortex Specialized for Face Perception, *Journal of Neuroscience*, *17*(11), 4302–4311

Kim, M. J., Gee, D. G., Loucks, R. a, Davis, F. C., & Whalen, P. J. (2011). Anxiety dissociates dorsal and ventral medial prefrontal cortex functional connectivity with the amygdala at rest. *Cerebral Cortex, 21*(7), 1667–73

Kiss, M. & Eimer, M. (2008). ERPs reveal subliminal processing of fearful faces. *Psychophysiology*, *45*(2), 318–326. doi:10.1111/j.1469-8986.2007.00634.x

Kohler, C. G., Turner, T., Stolar, N. M., Bilker, W. B., Brensinger, C. M., Gur, R. E., & Gur, R. C. (2004). Differences in facial expressions of four universal emotions. *Psychiatry Research*, *128*(3), 235–44. doi:10.1016/j.psychres.2004.07.003

Koeda, M., Belin, P., Hama, T., Masuda, T., Matsuura, M., & Okubo, Y. (2013). Cross-cultural differences in the processing of non-verbal affective vocalizations by Japanese and Canadian listeners. *Frontiers in Psychology*, *4*, 105. doi:10.3389/fpsyg.2013.00105

Kolb, B., Wilson, B., & Taylor, L. (1992). Developmental changes in the recognition and comprehension of facial expression: Implications for frontal lobe function. *Brain and Cognition*, *20*(1), 74–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1389123

Kornreich, C., Brevers, D., Canivet, D., Ermer, E., Naranjo, C., Constant, E., … Noël, X. (2013). Impaired processing of emotion in music, faces and voices supports a generalized emotional decoding deficit in alcoholism. *Addiction (Abingdon, England)*, *108*(1), 80–8. doi:10.1111/j.1360-0443.2012.03995.x

Kriegeskorte, N., Mur, M., & Bandettinii, P. (2008). Representational Similarity Analysis – Connecting the Branches of Systems Neuroscience. *Frontiers in Systems Neuroscience*, 2(4). doi: 10.3389/neuro.06.004.2008 PMCID: PMC2605405

Krieglmeyer, R., & Deutsch, R. (2013). Approach Does Not Equal Approach. Angry Facial Expressions Evoke Approach Only When It Serves Aggression. *Social Psychological and Personality Science* 4 (5), 607-61. doi: 10.1177/1948550612471060

Krolak-Salmon, P., Hénaff, M.-A., Isnard, J., Tallon-Baudry, C., Guénot, M., Vighetto, A., … Mauguière, F. (2003). An attention modulated response to disgust in human ventral anterior insula. *Annals of Neurology*, *53*(4), 446–53. doi:10.1002/ana.10502

Kryklywy, J. H., Macpherson, E. A, Greening, S. G., & Mitchell, D. G. V. (2013). Emotion modulates activity in the "what" but not "where" auditory processing pathway. *NeuroImage*, *82*, 295–305. doi:10.1016/j.neuroimage.2013.05.051

Kuefner, D., Macchi Cassia, V., Picozzi, M., & Bricolo, E. (2008). Do all kids look alike? Evidence for an other-age effect in adults. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(4), 811–7. doi:10.1037/0096-1523.34.4.811

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2014). Gender differences in emotion recognition: Impact of sensory modality and emotional category. *Cognition & Emotion*, *28*(3), 452–69. doi:10.1080/02699931.2013.837378

Lane, R., Chua, P., & Dolan, R. (1999). Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia,* 37(9). 989–997

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

Latinus, M., & Belin, P. (2011). Human voice perception. *Current Biology : CB*, *21*(4), R143–5. doi:10.1016/j.cub.2010.12.033

Laukka, P., Neiberg, D., & Elfenbein, H. A. (2014). Evidence for cultural dialects in vocal emotion expression: acoustic classification within and across five nations. *Emotion (Washington, D.C.)*, *14*(3), 445–9. doi:10.1037/a0036048

Leite, J., Carvalho, S., Galdo-Alvarez, S., Alves, J., Sampaio, A., & Gonçalves, O. F. (2012). Affective picture modulation: valence, arousal, attention allocation and motivational significance. *Int. Journal of Psychophysiology : Official Journal of the Int. Organization of Psychophysiology*, *83*(3), 375–81

Leitman, D. I., Wolf, D. H., Ragland, J. D., Laukka, P., Loughead, J., Valdez, J. N., … Gur, R. C. (2010). "It's Not What You Say, But How You Say it": A Reciprocal Temporo-frontal Network for Affective Prosody. *Frontiers in Human Neuroscience*, *4*,19. doi:10.3389/fnhum.2010.00019

Leppänen, J. M., Tenhunen, M., & Hietanen, J. K. (2003). Faster Choice-Reaction Times to Positive than to Negative Facial Expressions. *Journal of Psychophysiology*, *17*(3), 113–123. doi:10.1027//0269-8803.17.3.113

Liddell, B., Williams, L., Rathjen, J., Shevrin, H., & Gordon, E. (2004). A Temporal Dissociation of Subliminal versus Supraliminal Fear Perception: An Event-related Potential Study. *Journal of Cognitive Neuroscience* 16:3, pp. 479–486

Liddle, B. & Nettle, D. (2006). Higher-order theory of mind and social competence in school-age children. *Journal of Cultural and Evolutionary Psychology*. 4(3), 231-244

Lindner, J. L. & Rosén, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *36*(6), 769–77. doi:10.1007/s10803-006-0105-2

Lindquist, K. A, Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *The Behavioral and Brain Sciences*, *35*(3), 121–43. doi:10.1017/S0140525X11000446

Lithari, C., Frantzidis, C. A., Papadelis, C., Vivas, A. B., Klados, M. A., Kourtidou-Papadeli, C., … Bamidis, P. D. (2010). Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. *Brain Topography*, *23*(1), 27–40. doi:10.1007/s10548-009-0130-5

Lovén, J., Herlitz, A., & Rehnman, J. (2011). Women's own-gender bias in face recognition memory. *Experimental Psychology*, *58*(4), 333–40. doi:10.1027/1618-3169/a000100

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press

Ludemann, P. (1991). Generalized Discrimination of Positive Facial Expressions by Seven- and Ten-Month-Old Infants. *Child Development*, 62(1), 55–67

Malatesta, C. Z. & Haviland, J. M. (1982). Learning Display Rules : The Socialization of Emotion Expression in Infancy, *Child Development*, 53(4), 991–1003

Marinkovic, K., & Halgren, E. (1998). Human brain potentials related to the emotional expression , repetition , and gender of faces, *26*(4), 348–356

Markham, R., & Wang, L. (1996). Recognition of Emotion by Chinese and Australian Children. *Journal of Cross-Cultural Psychology*, *27*(5), 616–643. doi:10.1177/0022022196275008

Markowitsch, H., J. (1999). Differential Contribution of Right and Left Amygdala to Affective Information Processing. *Behavioural Neurology*. 11 (4), 233-244

Marsh, A., Adams, R., & Kleck, R. (2005). Why do fear and anger look the way they do? Form and social function in facial expressions. *Personality & Social Psychology Bulletin*, *31*(1), 73–86. doi:10.1177/0146167204271306

Mastropieri, D. & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology,* 35: 204–214. doi: 10.1002/(SICI)1098-2302(199911)35:3<204::AID-DEV5>3.0.CO;2-V

Matsumoto, D. & Fontaine, J. (2008). Mapping Expressive Differences Around the World: The Relationship Between Emotional Display Rules and Individualism Versus Collectivism. *Journal of Cross-Cultural Psychology*, *39*(1), 55–74. doi:10.1177/0022022107311854

Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology*, *96*(1), 1–10. doi:10.1037/a0014037

Mayer, J., (2000). Emotion, intelligence, emotional intelligence. In *The Handbook of Affect and Social Cognition*, ed. JP Forgas, pp. 410–31. Mahwah, NJ: Erlbaum

McCluskey, K. & Albas, D. (1981). Perception of the emotional content of speech by Canadian and Mexican Children, Adolescents and Adults. *International Journal of Psychology. 16*, 119–132

McCluskey, K., Albas, D., Niemi, R., Cuevas, C., & Ferrer, C. (1975). Cross-Cultural Differences in the Perception of the Emotional Content of Speech : A Study of the Development of Sensitivity in Canadian and Mexican Children. *Developmental Psychology,* 11(5), 551–555

McCluskey, K. & Albas, D. (1981). Perception of the emotional content of speech by Canadian and Mexican Children, Adolescents and Adults. *International Journal of Psychology. 16*, 119–132

Meltzoff, A. & Moore, M., (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology*, *25*(6), 954–962. doi:10.1037//0012-1649.25.6.954

Milesi, V., Cekic, S., Péron, J., Frühholz, S., Cristinzio, C., Seeck, M., & Grandjean, D. (2014). Multimodal emotion perception after anterior temporal lobectomy (ATL). *Frontiers in Human Neuroscience*, 5;8:275. doi: 10.3389/fnhum.2014.00275

Miller, N. E. & Dollard, J. (1941). *Social learning and imitation.* New Haven, CT, US: Yale University Press

Mills, K., Lalonde, F., Clasen, L., Giedd, J., & Blakemore, S.-J. (2014). Developmental changes in the structure of the social brain in late childhood and adolescence. *Social Cognitive and Affective Neuroscience*, *9*(1), 123–31. doi:10.1093/scan/nss113

Montirosso, R., Peverelli, M., Frigerio, E., Crespi, M., & Borgatti, R. (2010). The Development of Dynamic Facial Expression Recognition at Different Intensities in 4- to 18-Year-Olds. *Social Development*, *19*(1), 71–92. doi:10.1111/j.1467-9507.2008.00527.x

Morand, D. (1999). Family size and intelligence revisited: the role of emotional intelligence. *Psychological Reports* 84(2):643-9.

Moreno, C., Borod, J., Welkowitz, J., & Alpert, M. (1990). Lateralization for the expression and perception of facial emotion as a function of age. *Neuropsychologia*, 28(2), 199–209

Morris, J., Friston, K., Buchel, C., Frith, D., Young, A., Calder, A., & Dolan, R. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain,* 121, 47–57

Morris, J., Scott, S., & Dolan, R. J. (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia*, *37*(10), 1155–63

Morton, J. & Johnson, M. H. (1991). CONSPEC and CONLERN: a two-process theory of infant face recognition. *Psychological Review, 98*(2), 164–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2047512

Muñoz, L. C. (2009). Callous-unemotional traits are related to combined deficits in recognizing afraid faces and body poses. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(5), 554–62. doi:10.1097/CHI.0b013e31819c2419

Nakamura, K., Kawashima, R., Ito, K., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Nagumo, S., Kubota, K., Fukuda, H., & Kojima, S. (1999). Activation of the Right Inferior Frontal Cortex During Assessment of Facial Emotion. *Journal of Neurophysiology.* 82(3) 1610-1614

Neal, D. T. & Chartrand, T. L. (2011). Embodied Emotion Perception: Amplifying and Dampening Facial Feedback Modulates Emotion Perception Accuracy. *Social Psychological and Personality Science*, *2*(6), 673–678. doi:10.1177/1948550611406138

Neave, N., Laing, S., Fink, B., & Manning, J. T. (2003). Second to fourth digit ratio, testosterone and perceived male dominance. *Proceedings. Biological Sciences / The Royal Society*, *270*(1529), 2167–72. doi:10.1098/rspb.2003.2502

Nelson, N. L. & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology*, *110*(1), 52–61. doi:10.1016/j.jecp.2011.03.014

Nock, M., Wedig, M., Holmberg, E., & Hooley, J. (2008). The emotion reactivity scale: development, evaluation, and relation to self-injurious thoughts and behaviors. *Behavior Therapy*, *39*(2), 107–16. doi:10.1016/j.beth.2007.05.005

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2007). Face to face: blocking facial mimicry can selectively impair recognition of emotional expressions. *Social Neuroscience*, *2*(3-4), 167–78. doi:10.1080/17470910701391943

Office for National Statistics (2013). One in 8 of the usually resident UK population were born abroad. Retrieved from http://www.ons.gov.uk/ons/rel/migration1/population-by-country-of- birth-and-nationality/2012/sty-population-by-country-of-birth.html

Office for National Statistics (2014). UK Trade, November 2013. Retrieved from http://www.ons.gov.uk/ons/dcp171778_347845.pdf

Padmala, S., & Pessoa, L. (2008). Affective learning enhances visual detection and responses in primary visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(24), 6202–10. doi:10.1523/JNEUROSCI.1233-08.2008

Palermo, R., O'Connor, K. B., Davis, J. M., 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. doi:10.1371/journal.pone.0068126

Parker A., Mathis, E., & Kupersmidt, J. (2013). How Is This Child Feeling? Preschool-Aged Children's Ability to Recognize Emotion in Faces and Body Poses. *Early Education and Development*. 24 (2), 188-211

Paulmann, S., Jessen, S., & Kotz, S. A. (2009). Investigating the Multimodal Nature of Human Communication. *Journal of Psychophysiology*, *23*(2), 63–76. doi:10.1027/0269-8803.23.2.63

Paulus, A., Musial, E., & Renn, K. (2014). Gender of the expresser moderates the effect of emotional faces on the startle reflex. *Cognition & Emotion.* 28(8):1493-501. doi: 10.1080/02699931.2014.886557.

Peelen, M. V, Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(30), 10127–34. doi:10.1523/JNEUROSCI.2161-10.2010

Peirce, J. W. (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Method.* 162(1-2):8-13

Perner, J., Ruffman, T., Leekam, S. R., Url, S., Perner, J., Ruffman, T., & Leekam, S. R. (2013). Theory of Mind Is Contagious : You Catch It from Your Sibs Theory of Mind Is Contagious: You Catch It from Your Sibs, *Child Development*, *65*(4), 1228–1238.

Perrett, D. I., Burt, D. M., Penton-Voak, I. S., Lee, K. J., Rowland, D. A., & Edwards, R. (1999). Symmetry and Human Facial Attractiveness. *Evolution and Human Behavior*, *20*(5), 295–307. doi:10.1016/S1090-5138(99)00014-8

Phelps, E. A. & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, *48*(2), 175–87. doi:10.1016/j.neuron.2005.09.025

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., … Gray, J. a. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings. Biological Sciences / The Royal Society*, *265*(1408), 1809–17. doi:10.1098/rspb.1998.0506

Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. a., Johnson, R., … Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*(2), 127–152. doi:10.1111/1469-8986.3720127

Pitcher, D., Garrido, L., Walsh, V., & Duchaine, B. C. (2008). Transcranial magnetic stimulation disrupts the perception and embodiment of facial expressions. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(36), 8929–33. doi:10.1523/JNEUROSCI.1450-08.2008

Pittam, J. & Scherer, K. (1993). Vocal expression of communication of emotion. In '*Handbook of* emotions' (1993), Michael Lewis & Jeannette Haviland. The Guildford Press, New York, London

Pizzagalli, D. A, Lehmann, D., Hendrick, A. M., Regard, M., Pascual-Marqui, R. D., & Davidson, R. J. (2002). Affective Judgments of Faces Modulate Early Activity (∼160 ms) within the Fusiform Gyri. *NeuroImage*, *16*(3), 663–677

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*(2), 127–152. doi:10.1080/17405620344000022

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–34. doi:10.1017/S0954579405050340

Quam, C. & Swingley, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development*, *83*(1), 236–50. doi:10.1111/j.1467-8624.2011.01700.x

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences*, *97*(22), 11800-11806

Raven, J., Raven, J. C., & Court, J. H. (1998). Coloured progressive matrices. Hartcourt Assessments, Texas, USA.

Raven, J., Raven, J. C., & Court, J. H. (2000). Standard progressive matrices. NCS Pearson, Texas, USA.

Rellecke, J., Sommer, W., & Schacht, A. (2013). Emotion effects on the n170: a question of reference? *Brain Topography*, *26*(1), 62–71. doi:10.1007/s10548-012-0261-y

Rizzolatti, G. (2005). The mirror neuron system and its function in humans. *Anatomy and Embryology*, *210*(5-6), 419–21. doi:10.1007/s00429-005-0039-z

Rogier, O., Roux, S., Belin, P., Bonnet-Brilhault, F., & Bruneau, N. (2010). An electrophysiological correlate of voice processing in 4-to 5-year-old children. *International Journal of psychophysiology*, *75*(1), 44-47

Rozenkrants, B., & Polich, J. (2008). Affective ERP processing in a visual oddball task: arousal, valence, and gender. *Clinical Neurophysiology*, *119*(10), 2260-2265

Rossion, B. (2009). Distinguishing the cause and consequence of face inversion: the perceptual field hypothesis. *Acta Psychologica*, *132*(3), 300–12. doi:10.1016/j.actpsy.2009.08.002

Rothman, A. D. & Nowicki Jr., S. (2004). A Measure of the Ability to Identify Emotion in Children's Tone of Voice. *Journal of Nonverbal Behavior*, *28*(2), 67–92. doi:10.1023/B:JONB.0000023653.13943.31

Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, *110*(3), 426–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1758918

Salovey, P. & Mayer, J. (1990). Emotional Intelligence. *Imagination, Cognition and Personality*, 9(3), 185-211

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M. L., Scherer, K. R., & Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. *NeuroImage*, *28*(4), 848–58. doi:10.1016/j.neuroimage.2005.06.023

Sander, K., & Scheich, H. (2001). Auditory perception of laughing and crying activates human amygdala regardless of attentional state. *Brain Research. Cognitive Brain Research*, *12*(2), 181–98

Sato, W., Kochiyama, T., Yoshikawa, S., & Matsumura, M. (2001). Emotional expression boosts early visual processing of the face: ERP recording and its decomposition by independent component analysis. *Neuroreport*, *12*(4), 709–714

Sauter, D. A. & Eimer, M. (2009). Rapid Detection of Emotion from Human Vocalizations. *Journal of Cognitive Neuroscience,* 22(3):474-81. doi: 10.1162/jocn.2009.21215

Sauter, D. A., Eisner, F., Calder, A., & Scott, S. (2010). Perceptual cues in nonverbal vocal expressions of emotion, *The Quarterly Journal of Experimental Psychology*, 63:11, 2251-2272. *63*(11), 2251–72. doi:10.1080/17470211003721642

Sauter, D. A., Eisner, F., Ekman, P., & Scott, S. K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. *Proceedings of the National Academy of Sciences, 107(6)*, 2408-2412.

Sauter, D. A. & Scott, S. (2007). More than one kind of happiness: can we recognize vocal expressions of different positive states? *Motivation and Emotion.* 31, 192-199

Sauter, D. A., Panattoni, C., & Happé, F. (2013). Children's recognition of emotions from vocal cues. *The British Journal of Developmental Psychology*, *31*(Pt 1), 97–113. doi:10.1111/j.2044-835X.2012.02081.x

Sayette, M. A., Cohn, J. F., Wertz, J. M., Perrott, M. A., & Parrott, D. J. (2001). A Psychometric Evaluation Of The Facial Action Coding System For Assessing Spontaneous Expression. *Journal of Nonverbal Behaviour*. *25*(3), 167-185

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–318). Hillsdale, NJ: Erlbaum.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice : Official Journal of the Voice Foundation*, *9*(3), 235–48. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8541967

Scherer, K. R. (2007). Component Models of Emotion Can Inform the Quest for Emotional Competence. In G . Matthews, M. Zeidner, and R. D. Roberts (Eds.), *The Science of Emotional Intelligence: Knowns and Unknowns*. (pp. 101–126). New York: Oxford University Press.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model, *Cognition and Emotion*, 23:7, 1307-1351.

Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology*, *32*(1), 76–92. doi:10.1177/0022022101032001009

Scherer, K. R., Clark-Polner, E., & Mortillaro, M. (2011). In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. *International Journal of Psychology : Journal International de Psychologie*, *46*(6), 401–35. doi:10.1080/00207594.2011.626049

Schirmer, A. & Kotz, S. A. (2006). Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30. doi:10.1016/j.tics.2005.11.009

Schirmer, A., Zysset, S., Kotz, S., & von Cramon, Y. (2004). Gender differences in the activation of inferior frontal cortex during emotional speech perception. *NeuroImage,* 21(3), 1114-1123

Schneider, K. G., Hempel, R. J., & Lynch, T. R. (2013). That "poker face" just might lose you the game! The impact of expressive suppression and mimicry on sensitivity to facial expressions of emotion. *Emotion (Washington, D.C.)*, *13*(5), 852–66. doi:10.1037/a0032847

Schroeder, M. (2003). Experimental study of affect bursts. *Speech Communication.* 40, 99-116

Schupp, H. T., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2003). Attention and emotion: an ERP analysis of facilitated emotional stimulus processing. *Neuroreport*, *14*(8), 1107–1110. doi:10.1097/01.wnr.0000075416.59944.49

Schulte-Rüther, M., Markowitsch, H. J., Fink, G. R., & Piefke, M. (2007). Mirror neuron and theory of mind mechanisms involved in face-to-face interactions: a functional magnetic resonance imaging approach to empathy. *Journal of Cognitive Neuroscience*, *19*(8), 1354–72. doi:10.1162/jocn.2007.19.8.1354

Schyns, P. G., Petro, L. S., & Smith, M. L. (2009). Transmission of Facial Expressions of Emotion Co-Evolved with Their Efficient Decoding in the Brain: Behavioral and Brain Evidence. *PLoS ONE* 4(5): e5625. doi:10.1371/journal.pone.0005625

Scott, S., Young, A., Calder, A., Hellawell, D., Aggleton, J., & Johnson, M. (1997). Impaired auditory recognition of fear and anger following bilateral amygdala lesions. *Letters to Nature. Nature*, 385. 254-257

Seligman, M. (1971). Phobias and preparedness. *Behavior Therapy.* 2(3), 307–320

Shakespeare, W. (2010). *Macbeth*. Boston, MA: English Play Press. (Original work published 1699)

Shinkareva, S. V., Malave, V. L., Mason, R. A., Mitchell, T. M., & Just, M. A. (2011). Commonality of neural representations of words and pictures. *Neuroimage, 54*(3), 2418-2425

Simon-Thomas, E. R., Keltner, D. J., Sauter, D., Sinicropi-Yao, L., & Abramson, A. (2009). The voice conveys specific emotions: evidence from vocal burst displays. *Emotion (Washington, D.C.), 9*(6), 838–46. doi:10.1037/a0017810

Simpson, C., Pinkham, A. E., Kelsven, S., & Sasson, N. J. (2013). Emotion recognition abilities across stimulus modalities in schizophrenia and the role of visual attention. *Schizophrenia Research, 151*(1-3), 102–6. doi:10.1016/j.schres.2013.09.026

Soto, J. A., Levenson, R. W., & Ebling, R. (2005). Cultures of moderation and expression: emotional experience, behavior, and physiology in Chinese Americans and Mexican Americans. *Emotion (Washington, D.C.), 5*(2), 154–65. doi:10.1037/1528-3542.5.2.154

Smith, M. L. (2012). Rapid processing of emotional expressions without conscious awareness. *Cerebral Cortex*, *22*(8), 1748–1760. doi:10.1093/cercor/bhr250

Spiess, C., Büchel, F., & Wagner, G. (2003). Children ' s School Placement in Germany : Does Kindergarten Attendance Matter? *Early Childhood Research Quarterly*, 18 (2), 255-270

Sporer, S. (2001). Recognizing Faces of Other Ethnic Groups. *Psychology, Public Policy and Law*.

7(1), 36-97

Sprengelmeyer. R., Rausch, M., Eysel, U. T., & Przuntek, H. (1998). Neural structures associated with recognition of facial expressions of basic emotions. *Proceedings of the Royal Society of London B,* 265, 1927-1931

Sprengelmeyer, R., Schroeder, U., Young, A W., & Epplen, J. T. (2006). Disgust in pre-clinical Huntington's disease: a longitudinal study. *Neuropsychologia*, *44*(4), 518–33. doi:10.1016/j.neuropsychologia.2005.07.003

Sprengelmeyer, R., Young, A. W., Calder, A. J., Karnat, A., Lange, H., Homberg, V., … Rowland, D. (1996). Loss of disgust perception of faces and emotions in Huntington's disease, *Brain,* 119(5), 1647–1665

Statistisches Bundesamt (2014). Foreign Trade. Ranking of Germany's trading partners in foreign trade 2013. Wiesbaden. Available at https://www.destatis.de/EN/FactsFigures/NationalEconomyEnvironment/ForeignTrade/ForeignTrade.html

Stephani, C., Fernandez-Baca Vaca, G., Maciunas, R., Koubeissi, M., & Lüders, H. O. (2011). Functional neuroanatomy of the insular lobe. *Brain Structure & Function*, *216*(2), 137–49. doi:10.1007/s00429-010-0296-3

Sugita, Y., & Suzuki, Y. (2003). Audiovisual perception: Implicit estimation of sound-arrival time. *Nature*, *421*(6926), 911-911

Sun, Y., Gao, X., & Han, S. (2010). Sex differences in face gender recognition: An event-related potential study. *Brain Research*, *1327*, 69–76. doi:10.1016/j.brainres.2010.02.013

Sullivan S., Ruffman T., & Hutton S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62(1), 53-60

Sylva, K., Melhuish, E., Sammons, P., & Siraj-Blatchford, I. & Taggart, B. (2013). The Effective Provision Of Pre-School Education (Eppe) Project: Findings from Pre-school to end of Key Stage 1., 1–8. Institute of Education, University of London. *Retrieved from* http://eppe.ioe.ac.uk/eppe/eppepdfs/RBTec1223sept0412.pdf

Szymanowski, F., Kotz, S. A., Schröder, C., Rotte, M., & Dengler, R. (2007). Gender Differences in Processing Emotional Prosody. *Clinical Neurophysiology*, *118*(4), e102–e103. doi:10.1016/j.clinph.2006.11.239

Thomas, L., De Bellis, M., Graham, R., & LaBar, K. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science.* 10(5):547-58

Thomas, P. M., Jackson, M. C., & Raymond, J. E. (2014). A threatening face in the crowd: effects of emotional singletons on visual working memory. *Journal of Experimental Psychology. Human Perception and Performance*, *40*(1), 253–63. doi:10.1037/a0033970

Tonks, J., Williams, W. H., Frampton, I., Yates, P., & Slater, A. (2007). Reading emotions after child brain injury: a comparison between children with brain injury and non-injured controls. *Brain Injury*, *21*(7), 731–9. doi:10.1080/02699050701426899

Tranel, D, Damasio, A R,Damasio, H. 1988, "Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity", *Neurology*, 38(5), pp. 690

Uhls, Y. T., Michikyan, M., Morris, J., Garcia, D., Small, G. W., Zgourou, E., & Greenfield, P. M. (2014). Five days at outdoor education camp without screens improves preteen skills with nonverbal emotion cues. *Computers in Human Behavior*, *39*, 387–392. doi:10.1016/j.chb.2014.05.036

Van Beek, Y. & Dubas, J. S. (2008). Age and Gender Differences in Decoding Basic and Non-basic Facial Expressions in Late Childhood and Early Adolescence. *Journal of Nonverbal Behavior*, *32*(1), 37–52. doi:10.1007/s10919-007-0040-8

Vicari, S., Reilly, J., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expressions of emotions in school-age children: the intersection of perceptual and semantic categories. *Acta Paediatrica (Oslo, Norway : 1992)*, *89*(7), 836–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10943968

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. doi:10.1016/j.jecp.2008.04.005

Vink, M., Derks, J. M., Hoogendam, J. M., Hillegers, M., & Kahn, R. S. (2014). Functional differences in emotion processing during adolescence and early adulthood. *NeuroImage*, *91*, 70–6. doi:10.1016/j.neuroimage.2014.01.035

Vogel, M., Monesson, A., & Scott, L. (2012). Building biases in infancy: the influence of race on face and voice emotion matching. *Developmental Science*, *15*(3), 359–72. doi:10.1111/j.1467-7687.2012.01138.x

Volling, B. L. (2003). Sibling relationships. In Bornstein, M. H., Davidson, L., Keyes, C., Moore, K. (2003). *Well-being: Positive development across the life course. Crosscurrents in contemporary psychology.*, (pp. 205-220). Mahwah, NJ, US: Lawrence Erlbaum Associates

Publishers, xv, 601 pp.

Von Kriegstein, K. & Giraud, A.-L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *NeuroImage,* 22(2), 948-955

Vonsattel, J. P., & DiFiglia, M. (1998). Huntington disease. *Journal of Neuropathology and Experimental Neurology*, *57*(5), 369–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9596408

Van Lancker, D., Cummings, J., Kreiman, J., & Dobkin, B. (1988). Phonagnosia: A dissociation between familiar and unfamiliar voices. *Cortex,* 24(2), 195-209.

Wang, A. T., Lee, S. S., Sigman, M., & Dapretto, M. (2007). Reading Affect in the Face and Voice, *Archives of General Psychiatry,* 64(6):698-708.

Wege, B. V., Sánchez González, M., Friedlmeier, W., Mihalca, L., Goodrich, E., & Corapci, F. (2014). Emotion displays in media: a comparison between American, Romanian, and Turkish children's storybooks. *Frontiers in Psychology*, *5*(June), 600. doi:10.3389/fpsyg.2014.00600

Weimer, A. A., Sallquist, J., & Bolnick, R. R. (2012). Young Children's Emotion Comprehension and Theory of Mind Understanding. *Early Education & Development*, *23*(3), 280–301. doi:10.1080/10409289.2010.517694

Weymar, M., Löw, A., Ohman, A., & Hamm, A. (2011). The face is more than its parts--brain dynamics of enhanced spatial attention to schematic threat. *NeuroImage*, *58*(3), 946–54. doi:10.1016/j.neuroimage.2011.06.061

Whalen, P. J., Raila, H., Bennett, R., Mattek, a., Brown, a., Taylor, J., … Palmer, a. (2013). Neuroscience and Facial Expressions of Emotion: The Role of Amygdala-Prefrontal Interactions. *Emotion Review*, *5*(1), 78–83. doi:10.1177/1754073912457231

Wicker, B., Keysers, C., Plailly, J., Royet, J. P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. *Neuron*, *40*(3), 655–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14642287

Wierenga, L., Langen, M., Ambrosino, S., van Dijk, S., Oranje, B., & Durston, S. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage*, *96*, 67–72. doi:10.1016/j.neuroimage.2014.03.072

Wierzbicka, A. (1998). German `Cultural Scripts': Public Signs as a Key to Social Attitudes and Cultural Values. *Discourse & Society*, *9*(2), 241–282. doi:10.1177/0957926598009002006

Wiese, H., Schweinberger, S., & Hansen, K. (2008). The age of the beholder: ERP evidence of an own-age bias in face memory. *Neuropsychologia*, *46*(12), 2973–85. doi:10.1016/j.neuropsychologia.2008.06.007

Williams, B., Gray, K., & Tonge, B. (2012). Teaching emotion recognition skills to young children with autism: a randomised controlled trial of an emotion training programme. *Journal of Child Psychology and Psychiatry*, 53(12):1268-76. doi: 10.1111/j.1469-7610.2012.02593.x

Williams, L., Mathersul, D., Palmer, D., Gur, R., Gur, R., & Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*, *31*(3), 257–77. doi:10.1080/13803390802255635

Wouters, C. (2011). On Friendship, Equality and Introductions: Comparing English and German Regimes of Manners and Emotions. *Sociological Research Online*, 16 (1) 2. 10.5153/sro.2235

Yovel, G. & Belin, P. (2013). A unified coding strategy for processing faces and voices. *Trends in Cognitive Sciences*, 17 (6). pp. 263-271. ISSN 1364-6613

Zabel, T. A. & Chute, D. L. (2002). Educational neuroimaging: a proposed neuropsychological application of near-infrared spectroscopy (nIRS). *The Journal of Head Trauma Rehabilitation*, *17*(5), 477–88. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12802256

Appendix A – Adult Participants

Chapter 2 & 4: INFORMED CONSENT SHEET

Emotional Reactivity and typical emotional behaviour during emotional situations in adults across Germany and England

The Department of Psychology at Brunel University requires that all persons who participate in psychology studies give their written consent to do so. Please read the following and sign it if you agree with what it says.I freely and voluntarily consent to be a participant in the research project entitled “Emotional Reactions Across Different Countries” to be conducted with Lisa Kuhn and Prof. Taeko Wydell as supervisor or principal investigator. The broad goal of this research program is to explore how humans react to emotional situations across different countries.

Specifically, I have been told that I will be asked to complete the following short questionnaire and give some basic information about my age, country of birth and so on. The session should take no longer than 10 minutes to complete. I have been told that my responses will be kept strictly confidential. I also understand that if at any time during the session I feel unable or unwilling to continue, I am free to leave without negative consequences. That is, my participation in this study is completely voluntary, and I may withdraw from this study at any time. My withdrawal would not result in any penalty, academic or otherwise. My name will not be linked with the research materials, as the researchers are interested in cross cultural performances in general - not in any particular individual's behaviour in particular.

I have been given the opportunity to ask questions regarding the procedure, and my questions have been answered to my satisfaction. I have been informed that if I have any general questions about this project, or ethical issues relating to the project, I should feel free to contact Lisa Kuhn at lisa.kuhn@brunel.ac.uk. If I have any concerns or complaints regarding the way in which the research is or has been conducted I may contact Professor Taeko Wydell, Chair of the Psychology Research Ethics Committee, at taeko.wydell@brunel.ac.uk. I have read and understand the above and consent to participate in this study. * My signature is not a waiver of any legal rights. Furthermore, I understand that I will be able to keep a copy of the informed consent form for my records.

- Yes
- No

I have explained and defined in detail the research procedure in which the above-named has consented to participate. Furthermore, I will retain one copy of the informed consent form for my records.

_____________________ _____________________________ _______________

Principal Investigator Signature Please Print Date

Chapter 2 & 4: DEBRIEFING FORM:

Emotional Reactivity and typical emotional behaviour during emotional situations in adults across Germany and England

The aim of the study is to find out how human across different countries react to every day emotional situations. It is believed that humans across different countries show different emotional behaviour in emotional situations due to differences in display rules. More specifically, this study hopes to identify typical emotional reactions of participants depending on the country they are living in and to match those traits with the way we rate the emotional content of human voices and faces. Adults from Germany and England will participate in this study in order to compare the results cross-culturally and investigate differences or similarities in emotional reactions and detecting emotions within the human voice and face.

In short, the predicted aims of this study are as follows:

- Humans are able to detect basic emotions within non-human voice samples.
- The detection of basic emotions within voice should be universal across different countries; participants from Germany and England will equally be able to detect emotions within the human voice.
- Depending on the country of residency, humans could have different emotional reactions in certain situations which in turn could influence the way emotions are being perceived.

If you want to learn more about the idea of this research, the following

studies might be of interest to you:

Affleck, G. , Tennen, H., Urrows, S., *&* Higgins, P. *(*1994*).* Person and contextual features of daily stress reactivity: Individual differences in relations of undesirable daily events with mood disturbance and chronic pain intensity. *Journal of Personality and Social Psychology, 66, 329*-340.

Doherty, R. W. (1997). The emotional contagion scale: A measure of individual differences. *Journal of nonverbal Behavior*, *21*(2), 131-154.

Scherer, K. R., & Summerfield, A. B. (Eds.). (1986). *Experiencing emotion: A cross-cultural study*. Cambridge University Press.

Once again, thank you for taking part in this study.

ERS Online: INFORMED CONSENT SHEET

Emotional Reactivity and typical emotional behaviour during emotional situations in adults across Germany and England

The Department of Psychology at Brunel University requires that all persons who participate in psychology studies give their written consent to do so. Please read the following and sign it if you agree with what it says. I freely and voluntarily consent to be a participant in the research project entitled "Emotional Reactions Across Different Countries" to be conducted with Lisa Kuhn and Prof. Taeko Wydell as supervisor or principal investigator. The broad goal of this research program is to explore how humans react to emotional situations across different countries. Specifically, I have been told that I will be asked to complete the following short online questionnaire and give some basic information about my age, country of birth and so on. The session should take no longer than 10 minutes to complete.

I have been told that my responses will be kept strictly confidential. I also understand that if at any time during the session I feel unable or unwilling to continue, I am free to leave without negative consequences. That is, my participation in this study is completely voluntary, and I may withdraw from this study at any time. My withdrawal would not result in any penalty, academic or otherwise. My name will not be linked with the research materials, as the researchers are interested in cross cultural performances in general - not in any particular individual's behaviour in particular. I have been given the opportunity to ask questions regarding the procedure, and my questions have been answered to my satisfaction. I have been informed that if I have any general questions about this project, or ethical issues relating to the project, I should feel free to contact Lisa Kuhn at lisa.kuhn@brunel.ac.uk. If I have any concerns or complaints regarding the way in which the research is or has been conducted I may contact one of the Co-Chairs of the Psychology Research Ethics Committee, Dr. Bridget Dibb at bridget.dibb@brunel.ac.uk or Dr. Achim Schuetzwohl at achim.schuetzwohl@brunel.ac.uk

---

* Required

I have read and understand the above and consent to participate in this study. * My signature is not a waiver of any legal rights. Furthermore, I understand that I will be able to keep a copy of the informed consent form for my records.

- Yes
- No

I have explained and defined in detail the research procedure in which the above-named has consented to participate. Furthermore, I will retain one copy of the informed consent form for my records.

______________________ ______________________________ _______________

Principal Investigator Signature Please Print Date

Chapter 4: ERS

This questionnaire asks different questions about how you experience emotions on a regular basis (for example, each day). When you are asked about being "emotional," this may refer to being angry, sad, excited, or some other emotion. Please rate the following statements.

0 Not at all like me	1 A little like me	2 Somewhat like me	3 A lot like me	4 Completely like me

1	When something happens that upsets me, it's all I can think about it for a long time.	0	1	2	3	4
2	My feelings get hurt easily.	0	1	2	3	4
3	When I experience emotions, I feel them very strongly/intensely.	0	1	2	3	4
4	When I'm emotionally upset, my whole body gets physically upset as well.	0	1	2	3	4
5	I tend to get very emotional very easily.	0	1	2	3	4
6	I experience emotions very strongly.	0	1	2	3	4
7	I often feel extremely anxious.	0	1	2	3	4
8	When I feel emotional, it's hard for me to imagine feeling any other way.	0	1	2	3	4
9	Even the littlest things make me emotional.	0	1	2	3	4
10	If I have a disagreement with someone, it takes a long time for me to get over it.	0	1	2	3	4
11	When I am angry/upset, it takes me much longer than most people to calm down.	0	1	2	3	4
12	I get angry at people very easily.	0	1	2	3	4

13	I am often bothered by things that other people don't react to.	0	1	2	3	4
14	I am easily agitated.	0	1	2	3	4
15	My emotions go from neutral to extreme in an instant.	0	1	2	3	4
16	When something bad happens, my mood changes very quickly. People tell me I have a very short fuse.	0	1	2	3	4
17	People tell me that my emotions are often too intense for the situation.	0	1	2	3	4
18	I am a very sensitive person.	0	1	2	3	4
19	My moods are very strong and powerful.	0	1	2	3	4
20	I often get so upset it's hard for me to think straight.	0	1	2	3	4
21	Other people tell me I'm overreacting.	0	1	2	3	4

Other relevant questions/comments:

Chapter 2 & 4: Some facts about you

Now we would like to ask you for your time to tell us some facts about you. This will help us to get a better insight about some background information that may tell us more about your research results. Your information will be kept strictly confidential.

Gender:

____

Date of Birth:

____

Ethnicity:

____

Educational Status:

____

Country of Birth:

____

First Language:

____

Second Language (if applicable):

____

Indicate the regions where you have lived for significant parts of time in your life:

____

What culture do you feel a member of?

Translated words English-German

Happy – glücklich

Sad – traurig

Angry – wütend

Fearful – ängstlich

Surprised – überrascht

Disgusted – angeekelt

German versions of emotional reactivity scale (ERS) and demographic details:

ERS

Die folgende Umfrage bezieht sich auf Ihr Emotionsempfinden auf regelmässiger Basis (zum Beispiel täglich). Mit dem Wort “emotional,” könnte zum Beispiel 'traurig, wütend, aufgewühlt sein' oder ähnliche Emotionen gemeint sein.

	1 Trifft überhaupt nicht zu	**2 Trifft nur bedingt zu**	**3 Trifft etwas zu**	**4 Trifft zu**	**5 Trifft komplett zu**
1 Wenn etwas passiert, das mich aufregt, kann ich lange an nichts anderes denken.					
2. Ich bin schnell verletzt.					
3. Wenn ich Emotionen empfinde, spüre ich sie sehr stark und intensiv.					
4. Wenn ich aufgeregt bin, hat das physische Auswirkungen auf meinen ganzen Körper.					
5. Ich tendiere dazu, schnell emotional zu reagieren.					
6. Ich nehme Gefühle sehr stark wahr.					
7. Ich bin oft sehr ängstlich und beunruhigt.					
8. Wenn ich ein bestimmtes Gefühl empfinde, fällt es mir schwer mir vorzustellen, dass ich mich anders fühlen könnte.					
9. Sogar die kleinsten Dinge rufen Gefühle in mir hervor.					
10. Wenn ich mit jemandem eine Meinungsverschiedenheit habe, dauert es sehr lange, bis ich darüber hinweg bin.					
11. Wenn ich verärgert oder aufgeregt bin, benötige ich mehr Zeit als die meisten, um mich zu beruhigen.					
12. Ich regiere schnell verärgert auf andere Menschen.					
13. Ich störe mich oft an Dingen, welche andere nicht wahrnehmen.					

14. Ich bin schnell durcheinander.					
15. Meine Gefühle springen innerhalb von wenigen Augenblicken von neutral zu extrem.					
16. Wenn etwas Schlechtes passiert, ändert sich meine Stimmung sehr schnell. Man sagt mir, dass ich schnell die Beherrschung verliere.					
17. Menschen sagen mir oft, dass meine Gefühle nicht der Situation entsprechend sind.					
19. Meine Launen sind sehr ausgeprägt und stark.					
20. Ich bin oft so aufgewühlt, dass ich nichtmehr klar denken kann.					
21. Andere sagen mir, ich würde überreagieren.					

Demographische Daten & ERS für Erwachsene, Jan/Feb 2013

Folgende Daten werden uns helfen, ihre Ergebnisse besser zu interpretieren. Vielen Dank für Ihre Angaben, es wird natürlich alles streng vertraulich behandelt.

Geschlecht: ________________ Alter: ______________

Geburtsjahr: _______________ Geburtsland: ___________________

Akademische Ausbildung:
* bitte das höchste wählen (entweder abgeschlossen oder momentan durchführend).

○ Hauptschule/Mittlere Reife ○ Abitur ○ Bachelor ○ Master/Diplom/Doktor

○ Anderes: __________________

Land/Länder in dem Sie länger als 5 Jahre am Stück gelebt haben (ausser Ihrem Geburtsland):

____

Muttersprache: __________________ Zweite Muttersprache (falls zutreffend): _____________

Welcher Kultur fühlen Sie sich zugehörig?
* Normalerweise das Land in dem Sie geboren und aufgewachsen sind - außer in anderen Umständen.

____

Appendix B – Child Participants

Chapter 3 & 4: ***Agreement***

I am pleased to confirm my agreement that the student Lisa Kuhn can conduct the PhD research as part of the international research project *"Children's cross-cultural emotion detection in face and voice"* here at the school/club ________________________________ (name) in _____________________ (place).

The school/club agrees to provide Lisa Kuhn access to the institution/club named above and a selection of its pupils/members in order to conduct the research. Parental consent will be provided in advance.

Contact person at the school/club for the duration of this study will be ______________________.

Hiermit bestätige ich, dass die Studentin Lisa Kuhn einen Teil ihrer Doktorarbeit im Rahmen des internationalen Forschungsprojektes ‚*Emotionen erkennen in der Stimme und im Gesicht – eine zwischenkulturelle Entwicklungsstudie*` hier an der Schule/Verein _____________________________________(Name) in __________________ (Ort) durchführen kann.

Die Schule/der Verein gewährt Lisa Kuhn Zugang zu der oben genannten Institution/Verein und einer Selektion von Schülern/Mitgliedern um die Studie durchzuführen. Zustimmung der Eltern wird im Voraus bereitgestellt.

Ansprechpartner an der Schule/Verein für den Zeitraum der Studie wird sein: __________________.

________________________ ______________________ _______________________

Unterschrift/Signature Position Ort, Datum/Place, Time

Brunel University,

Uxbridge, UB8 3PH.

Tel: +44 (0)1895 274000
Fax: +44 (0)1895 232806

Lisa.Kuhn@Brunel.ac.uk

11th. October 2012

Dear Parents,

We (Brunel University) are conducting an international study funded by the Economics and Social Research council with primary school children in England as well as in Germany. We are interested in the way children across two different countries develop their emotion recognition abilities over age.

After several co-operations in previous years, St. Andrew CofE Primary school has very kindly provided us with access to the school once again and we are very happy to be able to conduct the study at the school this autumn during normal school times. For this study, children aged 5-10 will have the chance to take part in a fun computer based emotion recognition task of emotional faces and voices and a problem solving task.
If you are happy for your child to take part in this study and to help us collecting important research data, we would kindly ask you to sign and complete the two forms on the following pages (page 3 & 4). All data will be kept strictly confidential and won't be connected to any names. In case you have any questions please don't hesitate contacting me at Lisa.Kuhn@brunel.ac.uk.

Thank you very much for your help in advance!

Lisa Kuhn

Chapter 3 & 4: DEBRIEFING FORM: Investigation of human interaction

The aim of the study is to find out how children develop the ability to recognise emotions and whether the reaction times of doing so changes with age. Furthermore, this study hopes to identify the young children can reliably detect emotions within voice and face and this ability develops with age, in line with their Theory of Mind development, which means to understand that other people may have other feelings and perspectives about the world. More specifically, we also want to find out whether children across England as well as Germany judge emotions in a similar way or whether there are any differences in the timing and onset of the development.

In short, the predicted aims of this study are as follows:

- Children across England and Germany aged 5-10 can detect basic emotions within the voice and face
- We believe that this ability develops and improves with age, meaning that older children will be able to judge emotions faster and also assess perceived emotion intensity.
- There could be small differences between English and German children in their emotion understanding due to certain cultural differences or differences in educational systems.

If you want to learn more about the idea of this research, the following studies might be of interest to you:

Charles, S. & Campos, B. (2011): Age-Related Changes in Emotion Recognition: How, Why, and How Much of a Problem? *Journal of Nonverbal Behaviour*, 35:287–295.

Morton, J. and Trehub, S. (2001): Children's Understanding of Emotion in Speech, *Child Development*, Volume 72, Number 3, Pages 834–843.

Once again, thank you for taking part in this study!!

Chapter 3 & 4: INFORMED CONSENT SHEET

Parents of under aged participants: Children's emotion recognition

The Department of Psychology at Brunel University requires that all persons/their legal guardians who participate in psychology studies give their written consent to do so.

Please read the following and sign it if you agree with what it says. My child freely and voluntarily consents to be a participant in the research project entitled "**Children's emotion recognition"** to be conducted at ST. Andrews CofE school, with Lisa Kuhn and Prof. Taeko Wydell as supervisor or principal investigator.
The broad goal of this research program is to explore how children recognise basic emotions within the face and voice and how this ability develops with age.

Specifically, I have been told that my child will be asked to take part in a computer-based emotion recognition task of displayed faces and voices. Furthermore, my child will complete a short problem solving task and answer a few questions about people's belief after listening to a story. The session should take no longer than 40 minutes to complete. I have been told that my child's responses will be kept strictly confidential. I also understand that if at any time during the session he/she feels unable or unwilling to continue, he/she is free to leave without negative consequences. That is, my child's participation in this study is completely voluntary, and he/she may withdraw from this study at any time. His/her withdrawal would not result in any penalty, academic or otherwise.

My child's name will not be linked with the research materials, as the researchers are interested in social interactions in general - not in any particular individual's interaction/behaviour in particular. I/my child have been given the opportunity to ask questions regarding the procedure, and my questions have been answered to my satisfaction.

I have been informed that if I have any general questions about this project, or ethical issues relating to the project, I should feel free to contact Lisa Kuhn at lisa.kuhn@brunel.ac.uk. If I have any concerns or complaints regarding the way in which the research is or has been conducted I may contact Professor Taeko Wydell, Chair of the Psychology Research Ethics Committee, at Taeko.Wydell@brunel.ac.uk. I have read and understand the above and consent to my child's participation in this study. My signature is not a waiver of any legal rights. Furthermore, I understand that I can make a copy of the informed consent form for my records.

_______________________ _______________________ _______________

Participant's Signature Please Print Date

____

Child's Name

I have explained and defined in detail the research procedure in which the above-named has consented to participate. Furthermore, I will retain one copy of the informed consent form for my records.

____________________ ___________________________ ______________

Principal Investigator Signature Please Print Date

Chapter 3 & 4: To be filled out by Parents.

We would like to ask you for your time to tell us some facts about you and your child. This will help us to get a better insight about some background information that may tell us more about the research results.

Child's Gender: ______________________

Child's Age: ______________________

Number of Child's Siblings (if applicable): __________________________

Child's Ethnicity:

White British □ White Irish □ White other □

Asian □

Chinese □

Black British □ Black Caribbean □ Black African □ Black other □

Mixed □

Other □ please specify ________________________________

Child's Country of Birth: ________________________________

Child's First Language: ________________________________

Mother's Educational Status (if applicable):

GSCE level □

A Level □

Undergraduate Degree (BSc, BA, Diploma) □

Postgraduate Degree (MSc, MA, PhD) □

Other □ please specify __________________________________

Father's Educational Status (if applicable):

GSCE level □

A Level □

Undergraduate Degree (BSc, BA, Diploma) □

Postgraduate Degree (MSc, MA, PhD) □

Other □ please specify ___________________________________

Thank you very much for this information. All details will be kept strictly confidential and not be used in any connection with any names etc.

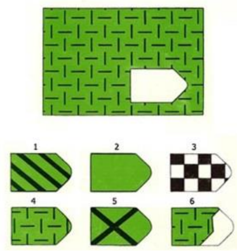
Chapter 3 & 4: INFORMED CONSENT SHEET – CHILD PARTICIPANTS:

Emotion recognition (Children, 5-10 years old)

<u>What I will do today:</u>

- Listen to voices and watch face on the computer and say how sad, happy etc they sound/look.

- Match some coloured patterns with missing ones, just like doing a puzzle.



- Take part in two very short play tasks.
- My reward will be stickers and other stationary.
- I can stop if I want to at any time.
- The session will take about 45 minutes
- Lisa Kuhn has explained to me details about the study.
- My parents and school teacher are happy for me to take part.

I would like to participate (circle):

YES **NO**

Chapter 3 & 4: Theory of Mind/Perspective Taking Task.

As used in: O'Brien et al. (2011): Longitudinal Associations between Children's understanding of Emotions and Theory of Mind. *Cognition and Emotion*, Volume 25, Number 6, pp. 1074-1086.

- Unexpected Contents (False Belief), developed by Astington & Gopnik, 1988
  Aim is to identify the child's own and second person's belief about the content of two items.

Procedure:

A cereal box containing pencils will be presented to the child. The child will be asked what he/she thinks is in the containers before the real content will be shown to them.

Task question 1: Before we opened the container, what did you think was in here?

Task question 2: What does this cuddle toy here think is in the container who hasn't seen the content?

Scoring: Do two trials, 1 point for each correct answer.

Risk assessment: Children's cross cultural emotion detection in face and voice

The research study is concerned with children's (aged 5-10) emotional understanding of visual as well as auditory emotional stimuli. Children in Germany as well as England will be asked to take part in a computerized task, including the emotional intensity and reaction time judgement of face (Pictures of Facial Affect, Ekman) and voice (Montreal Affective Voices) samples. Furthermore, they will be asked to complete a non-verbal IQ test (Coloured Progressive Matrices task) and two short Theory of Mind play tasks.

Safety during conducting research:

To the extent of my knowledge I am acting legally when planning and conducting my research. The experiment will be conducted at a local Primary school in London as well as Pforzheim (Ger) which is a neutral location where other people will be present. At no time will I be completely on my own with the participant. My family and my research supervisor will be informed when I am conducting the experiment. When children are taking part, teachers or parents/legal guardians will be asked to sit in the same room and there is a constant webmail/telephone contact to the research supervisor possible in case some advice is needed.

Materials & Methods:

A fully charged laptop will be used in order to conduct the computer based emotion recognition study. The non verbal IQ test is a paper based version and the Theory of Mind tasks will include cereal boxes filled with pencils and a bubble jar filled with straws, additionally to some picture cards. Material is safe as it has been previously tested (such as loudness, no electrical fault). The tests are non invasive and do not include any personal intrusive questions. Furthermore it is possible to withdraw any time. No special training for the participants will be required but there will be a short introduction on how complete the computer based task.

Possible risks:

All material used is non personal or intrusive. All tests are non offensive, written in basic terms so that it is easily understood by children. The tests used are widely accepted and validated. In case of young/less literate participants, questions and task instructions will be read out aloud or filled in by legal guardians. The reduced risk for a participant is a priority to the researcher and is kept at a minimum.

All in all, this study should not induce any form of distress in others and should take no longer than 40-50 minutes to participate.

I belief that the risk involved in this research is insignificant and unlikely to happen. All possible risks have been eliminated as much as possible. The possible risks involved should not exceed the distress of daily hassles of participants in any level.

Data protection & other requirements:

Participant's personal data about age, gender, country of birth, ethnicity, educational status will be stored in order to check for co-variation factors in research results.

However, data protection issues are considered. Therefore, personal data will only be used for research aims. They are only accessible by the researcher and will not be forwarded to a third party. No names will be mentioned at any time and the research data will not come in contact with individual personal data as we are only concerned with average data across all participants.

Since this study involves vulnerable participants such as children under the age of 16, a Criminal Record Bureau Check (enhanced disclosure) will be required for the researcher. Therefore, the researcher has been checked (13 March 2009) and cleared with no records.

This risk assessment will be shown to the University Research Ethics Committee, to the research supervisor, to parents of under aged participants or to participants themselves if they wish so.

Enhanced Disclosure

disclosure

Disclosure Number 001229265316

Date of Issue: 13 MARCH 2009

Applicant Personal Details

Surname: KUHN

Forename(s): LISA KATHARINA

Other Names: NONE DECLARED

Date of Birth: 10 MARCH 1987

Place of Birth: PFORZHEIM GERMANY

Gender: FEMALE

Employment Details

Position applied for:
RESEARCHER

Name of Employer:
BRUNEL UNIVERSITY

Countersignatory Details

Registered Person/Body:
BRUNEL UNIVERSITY

Countersignatory:
NARINDERPAL DHINGRA

Police Records of Convictions, Cautions, Reprimands and Final Warnings

NONE RECORDED

Information from the list held under Section 142 of the Education Act 2002

NONE RECORDED

Protection of Children Act List information

NONE RECORDED

Protection of Vulnerable Adults List information

NONE RECORDED

Other relevant information disclosed at the Chief Police Officer(s) discretion

NONE RECORDED

Enhanced Disclosure

This document is an Enhanced Criminal Record Certificate within the meaning of sections 113B and 116 of the Police Act 1997.

Continued on page 2

THIS DISCLOSURE IS NOT EVIDENCE OF IDENTITY

CRB Criminal Records Bureau, PO Box 165, Liverpool, L69 3JD Helpline: 0870 90 90 844

Chapter 3: ToM Task questions for 5-6 year olds

Story 1. BOBBY'S CHOCOLATE BARS

Bobby loves chocolate. Bobby's mum knows that chocolate is Bobby's favourite thing in all the world. He keeps lots of chocolate bars in the cupboard in his bedroom. Bobby's mum doesn't like him eating chocolate. It might spoil his tea! One day when he has gone to his friend's house, Bobby's mum moves the chocolate bars and she puts them into her pink shopping bag. When Bobby comes home, where does he think his chocolate is?

ToM Level 1:
a) Bobby thinks his chocolate is in his mum's shopping bag.
b) Bobby thinks his chocolate is in his cupboard.

Story 2. MUM'S BIRTHDAY

Anna wanted to get her mum a birthday present. Her brother Ben was out riding his bike so Anna decided to look around his room to see if she could find what present he had got for mum. Anna went in and found a big bunch of beautiful flowers with a little card that said: 'Happy Birthday Mum'. Anna thought to herself 'mum must want flowers for her birthday!' Anna didn't want Ben to know that she had been snooping round his room, so she said to Ben: "Ben, have you got mum a birthday present?" Ben thought for a minute, he didn't want Anna to copy him so he said: "Yes, Anna, I have got mum some perfume.

ToM Level 1:
a) Anna thinks Ben has bought mum some perfume.
b) Anna knows Ben has bought mum some flowers.

Story 4. THE SCHOOL FOOTBALL TEAM

Best friends Johnny and Bob both want to play on the school football team. Johnny thinks that he is not as good at football as Bob is. He thinks that the football manager is more likely to choose Bob for the football team. But the football manager thinks that both Johnny and Bob are good football players. He wants them both to play in the school football team. But the manager knows that Johnny doesn't think he will get on the team.

ToM Level 2:
a) Johnny doesn't know that the manager wants both him and Bob on the team.
b) Johnny thinks that the manager wants both him and Bob on the team.

Story 5. THE TEST

Class 4 were having a spelling test on Friday. Mrs Smith, the teacher, told all the children to work really hard. Kirsty wanted to do well. She learned all the words that she knew would be in the test. Kirsty told Mrs Smith that she had been learning the words all week. When the spelling test started, Mrs Smith turned to James first. James was a friend of Kirsty but sometimes he was rather lazy. "James," said Mrs Smith." How do you spell balloon?" James had not been learning his spellings. He did not know how to spell balloon. But James did remember that there was a poster in the classroom. He knew the word "balloon" was written on that poster. The poster was behind Mrs Smith. She could not see it. James cheated and spelt out that word "balloon" from the poster. Mrs Smith said: "Well done, James, that's correct."

ToM Level 2:
a) Mrs Smith thought that Kirsty wanted to do well on the test.
b) Mrs Smith didn't know that Kirsty wanted to do well on the test.

Chapter 3: ToM Task questions for 7-10 year olds

Story 4. THE SCHOOL FOOTBALL TEAM

Best friends Johnny and Bob both want to play on the school football team. Johnny thinks that he is not as good at football as Bob is. He thinks that the football manager is more likely to choose Bob for the football team. But the football manager thinks that both Johnny and Bob are good football players. He wants them both to play in the school football team. But the manager knows that Johnny doesn't think he will get on the team.

ToM Level 2:
a) Johnny doesn't know that the manager wants both him and Bob on the team.
b) Johnny thinks that the manager wants both him and Bob on the team.

ToM Level 3:
a) The manager thinks that Johnny knows he wants him to be on the football team.
b) The manager knows that Johnny doesn't know that he wants him to be on the team.

Story 5. THE TEST

Class 4 were having a spelling test on Friday. Mrs Smith, the teacher, told all the children to work really hard. Kirsty wanted to do well. She learned all the words that she knew would be in the test. Kirsty told Mrs Smith that she had been learning the words all week. When the spelling test started, Mrs Smith turned to James first. James was a friend of Kirsty but sometimes he was rather lazy. "James," said Mrs Smith." How do you spell balloon?" James had not been learning his spellings. He did not know how to spell balloon. But James did remember that there was a poster in the classroom. He knew the word "balloon" was written on that poster. The poster was behind Mrs Smith. She could not see it. James cheated and spelt out that word "balloon" from the poster. Mrs Smith said: "Well done, James, that's correct."

ToM Level 2:
a) Mrs Smith thought that Kirsty wanted to do well on the test.
b) Mrs Smith didn't know that Kirsty wanted to do well on the test.

ToM Level 3:
a) James thinks that Mrs Smith believes that he knows how to spell 'balloon'.
b) James thinks that Mrs Smith knows that he doesn't really know how to spell 'balloon'.

Taken from: Liddle, B. & Nettle, D. (2006): Higher-Order Theory Of Mind And Social Competence In School-Age Children. *Journal of Cultural and Evolutionary Psychology, 4(2006)3–4, 231–24.*

Chapter 3: INFORMED CONSENT SHEET

(Parents of under aged participants): Children's emotion recognition

The Department of Psychology at Brunel University requires that all persons/their legal guardians who participate in psychology studies give their written consent to do so. Please read the following and sign it if you agree with what it says. My child freely and voluntarily consents to be a participant in the research project entitled "**Children's emotion recognition"** to be conducted at ST. Andrews CofE school, with Lisa Kuhn and Prof. Taeko Wydell as supervisor or principal investigator.

The broad goal of this study is to explore how children recognise basic emotions within the face and how this ability develops with age. This study is a follow up to the study your child completed in 2012 in order to track development over a period of time. Specifically, I have been told that my child will be asked to take part in a computer-based emotion recognition task to judge faces displayed by either child or adult actors according to the displayed emotion. The session should take no longer than 20 minutes to complete.

I have been told that my child's responses will be kept strictly confidential. I also understand that if at any time during the session he/she feels unable or unwilling to continue, he/she is free to leave without negative consequences. That is, my child's participation in this study is completely voluntary, and he/she may withdraw from this study at any time. His/her withdrawal would not result in any penalty, academic or otherwise. My child's name will not be linked with the research materials, as the researchers are interested in social interactions in general - not in any particular individual's interaction/behaviour in particular.

I/my child have been given the opportunity to ask questions regarding the procedure, and my questions have been answered to my satisfaction. I have been informed that if I have any general questions about this project, or ethical issues relating to the project, I should feel free to contact Lisa Kuhn at lisa.kuhn@brunel.ac.uk. If I have any concerns or complaints regarding the way in which the research is or has been conducted I may contact Professor Taeko Wydell, Chair of the Psychology Research Ethics Committee, at Taeko.Wydell@brunel.ac.uk.

I have read and understand the above and consent to my child's participation in this study. My signature is not a waiver of any legal rights. Furthermore, I understand that I will be able to keep a copy of the informed consent form for my records.

________________________ ________________________ ______________

Participant's Signature Please Print Date

____

Child's Name

I have explained and defined in detail the research procedure in which the above-named has consented to participate. Furthermore, I will retain one copy of the informed consent form for my records.

____________________ ____________________________ ______________

Principal Investigator Signature Please Print Date

Dear Parents,

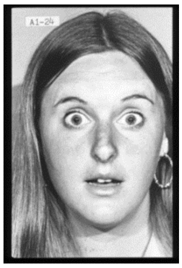
We (Brunel University) are conducting an international study funded by the Economics and Social Research council with primary school children in England as well as in Germany.

We are interested in the way children across two different countries develop their emotion recognition abilities over age. For this, we will return to St. Andrew CofE Primary school, which has very kindly provided us with access to the school once again. We are very pleased to conduct the study this March during normal school times.

Your child has kindly participated in the first part of this study in 2012 which gave us fantastic results. Now we are conducting a follow-up study with the same children in order to track their development over time. This will give us a better idea of how to interpret findings from the first study.

For this study, children aged 8-10 will have the chance to take part in a fun computer based emotion recognition task, rating emotional faces, such as:

or



If you are happy for your child to take part in this second part of the study and to help us collecting important follow up data, we would kindly ask you to sign the form on the back of this page and to fill in additional information below. All data will be kept strictly confidential. In case you have any questions, please contact me at Lisa.Kuhn@brunel.ac.uk.

Thank you very much for your help in advance!

<u>To be filled out by parents.</u>

<u>Child's</u> Gender: ________________________________

<u>Child's</u> Age: ________________________________

<u>Child's</u> Date of Birth: ______________________________

Number of <u>Child's</u> Siblings (if applicable): ______________________________

<u>Thank you very much for this information. All details will be kept strictly confidential and not be used in any connection with any names etc.</u>

Appendix C – EEG Chapter

	YES	NO
1. I have read the Research Participant Information Sheet.	☐	☐
2. I have had an opportunity to ask questions and discuss this study.	☐	☐
3. I understand that I am free to withdraw from the study:		
- at any time (Please note that you will unable to withdraw once your data has been included in any reports, publications etc)	☐	☐
- without having to give a reason for withdrawing	☐	☐
- without it affecting my future career	☐	☐
4. I understand that this study measures my brain activity which is a completely harmless procedure	☐	☐
5. I understand that I will not be referred to by name in any report/publications resulting from this study	☐	☐
6. I agree that I will be viewing faces and hear voices which I will be rating during the task	☐	☐
7. I agree to take part in this study	☐	☐

Research Participant Name:
Research Participant signature:
Date:
Principal Investigator name:
Principal Investigator signature:
Date:

Instruction EEG study

In this study, you will see either faces or hear short clips of voices.

The faces or voices will either be emotional or just neutral (so that means they don't express any emotions).

<u>There are two tasks for you:</u>

1) In the first task, you will see or hear briefly (under 1 second) a picture or sound. After it has disappeared, press '**1**' if it was an **emotional** face/voice and '**0'** if it was a **neutral** face/voice. So it is all about **emotions**!
   Be careful: Don't press any buttons if it is a neutral face/voice!

2) In the second task, you will see or hear exactly the same picture or sound. But this time, ignore the emotion they express and after it has disappeared, press '**1**' if it was a **female** face/voice and '**0**' if it was a **male** face/voice. So it is all about **gender**!

Just before the picture/sound appears, you will see a purple cross in the middle of the screen to tell you that the next stimulus is coming up. Please keep looking at it whenever you see it! The cross will stay on for the duration of the sound.

Most stimuli will be repeated several times. Can you find out how many times?

No panic, we will always tell you which task is coming next and remind you of which buttons to press.

Important: Please only click the button once the picture or sound has <u>disappeared</u>!

We will start with a short practice run before each task (one for emotion task, one for gender task) so get you used to pressing the buttons **after** the stimulus.

Any Questions?

Thank you & Enjoy!!

DEBRIEFING FORM:

Emotion Recognition in Human Voice and Face – An EEG Study

Results from our previous study indicated that participants rate happy emotions faster and with higher intensity than negative emotions such as anger or sadness. The aim of the present study is to investigate underlying brain mechanisms associated with rating either happy or angry emotions and to observe whether they correspond with our previous findings from behavioural studies.

Furthermore, we aim to find out whether the brain reacts faster to emotions displayed in either faces or voices and whether the same pattern of emotion recognition (*i.e.* react faster to angry or happy emotions) is replicated across modalities. If you have any questions or complaints about the execution of the study, you can contact Dr. Bridget Dibbs (bridget.dibbs@brunel.ac.uk) or Dr. Achim Schuetzwohl (achim.schuetzwohl@brunel.ac.uk) as co-chairs of the Brunel Social Science Ethics committee.

More specifically, from our behavioural data, we predict that:

1. Participants show enhanced brain activity to happy as compared to angry emotions

2a. This pattern is expected to be visible across both modalities (voices and faces)

2b. However, faces might show faster processing times than voices.

The following studies might be of interest to you:

- Becker et al., (2011): The Face in the Crowd Effect Unconfounded: Happy Faces, Not Angry Faces, Are More Efficiently Detected in Single- and Multiple-Target Visual Search Tasks. *Journal of Experimental Psychology.* 140(4), 637–659.
- Hansen, C. & Hansen, R. (1988): Finding the Face in the Crowd: An Anger Superiority Effect. *Journal of Peraonality and Social Psychology*, 54(6), 917-924.
- Kiss, M. & Eimer, M. (2008): ERPs reveal subliminal processing of fearful faces. *Psychophysiology*, 45, 318–326.

Once again, thank you for taking part in this study.

Information Sheet: Recording the electroencephalogram (EEG)

EEG is a recording of the electrical activity of the brain. It is a safe, non-invasive procedure. It involves wearing a cap that contains a number of electrodes (32 or 64) that are small ceramic discs with a silver coating. To make electrical contact with the scalp, these discs must be filled with saline gel (similar in consistency to hair gel). The gel is harmless to the hair and skin and is certified for this use. Once the electrodes have been filled, the cap is connected to the recording system to check whether the electrodes will record (impedance testing). There is usually some adjustment of the cap and re-filling of the electrodes to make sure that a good recording can be obtained.

In the experiment itself, a stimulus such as a picture or a sound is presented repeatedly, and/or you will carry out a simple task. The tiny electrical signals are passed into a powerful amplifier, and a computer records the EEG signal from each electrode. After the experiment, we average the activity to each stimulus to obtain a 3D map (two spatial dimensions plus time) of the brain's response to the stimulus, known as an event related potential (ERP).

The experiment will involve observing blocks of stimuli, and/or making simple repeated responses, while keeping still and concentrating. Each block would typically last 8 min, and there would be breaks in between the blocks in which you can move around, stretch, talk, drink some water, etc. The experiment part would last about 60 min and the preparation and removal of the electrodes would last about 30 min making about 90 min for the whole procedure.

With the gel method, it will be necessary to wash your hair soon after the experiment. There is a shower room nearby, with shampoo and a clean towel available, or if you prefer, you can remove most of the gel with a tissue and wash hair when you get home.

It is also advisable to wash hair on the morning or night before the experiment for your own comfort and to obtain good recordings. We cannot record from people wearing a hijab or other head covering, with hair extensions, weave, thick plaited hair, or hair styled using wax, hair spray or similar products.

Participants will be asked to read and sign a consent form. You will have the right to withdraw from the study at any time – either before or during the experiment. No explanation is necessary and there are no adverse consequences (academic or otherwise) if you do not wish to continue.

Examples of visual stimuli for the ERP study on emotion recognition

Female and male actors portrayed happiness, anger and neutral expressions. Expressions with open mouth/visible teeth (such as in happiness and anger expressions) were balanced.