SELECTIVE ATTENTION TO STATIC AND DYNAMIC FACES AND FACIAL CUES

by

Brenda Marie Stoesz

A Thesis submitted to the Faculty of Graduate Studies at the

University of Manitoba

in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology
University of Manitoba
Winnipeg, Manitoba, Canada

Abstract

Much of what is known about how we process faces comes from research using static stimuli. Thus, the primary goal of the present series of studies was to compare the processing of more naturalistic, dynamic face stimuli to the processing of static face stimuli. A second goal of the present series of studies was to provide insight into the development of attentional mechanisms that underlie perception of faces. Results from the eye-tracking study (Chapter 2) indicated that viewers attended to faces more than to other parts of the static or dynamic social scenes. Importantly, motion cues were associated with a reduction in the number, but an increase in the average duration of fixations on faces. Children showed the largest effects related to the introduction of motion cues, suggesting that they find dynamic faces difficult to process. Then using selective attention tasks (Chapters 3-5), interactions between the processing of facial expression and identity while participants viewed static and dynamic faces were examined. When processing static faces, viewers experienced significant interference from task-irrelevant cues (expression or identity) while processing the relevant cues (identity or expression). Age-related differences in interference effects were not evident (Chapter 3); however, biological sex and perceptual biases did contribute to the levels of interference seen with static faces (Chapters 4-5). During dynamic trials, however, viewers (regardless of age, sex, or perceptual bias) experienced negligible interference from task-irrelevant facial cues. Taken together, these findings stress the importance of using dynamic displays when characterizing typical face processing mechanisms, using the same methods across development, and of considering individual differences when examining various face processing abilities.

Acknowledgements

Thank you to my husband, Keith, and my children, Piper and Kiera, who supported me throughout the process of completing the research for this dissertation. Thanks to my extended family and friends for providing childcare and other types of support (e.g., shoulders to cry on, proofreading, etc.).

Thank you, Dr. Lorna Jakobson, for being my advisor and friend (for the past 10 years, even though it only feels like 8 years) and providing me with rich, challenging, and rewarding undergraduate and graduate experiences. I learned much from your expertise and skillful guidance. I look forward to collaborating with you in the future and continuing our friendship. Thanks also to my committee members, Drs. Jim Hare, Pauline Pearson, and Jonathan Marotta who provided me with valuable feedback on my candidacy exams, my research, and my progress.

I acknowledge the contributions of my former lab-mates, Drs. Wanda Snow, Nicole Taylor, Andrea Piotrowski, and current lab-mates, Kate Williamson and Damien Dowd, who provided much appreciated emotional support. A special thanks to lab-mate, Sarah Rigby, for her help with data collection and for allowing me to practice my advising and mentoring skills. Thanks to all the children and adults who participated in the studies and made it possible to complete this research. I also thank Drs. Janine Montgomery, Shahin Shooshtari, Toby Martin, and C.T. Yu who provided me with additional opportunities to hone my research and writing skills.

Finally, I acknowledge the financial support I received while pursing my doctoral degree. I was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) postgraduate fellowship, a Canadian Institute of Health Research (CIHR) studentship, and a University of Manitoba Graduate Fellowship (UMGF). The research was funded by a Discovery Grant to L. S. Jakobson from NSERC.

Dedication

Dedicated to my loving husband Keith,

and my beautiful daughters Piper and Kiera.

Abbreviations

ANOVA	analysis of variance
ASD	autism spectrum disorder
FFA	fusiform face area
FIE	face inversion effect
fMRI	functional magnetic resonance imaging
fps	frames per second
IT	inferotemporal
MT	middle temporal
RTs	response times
STP	superior temporal polysensory
STS	superior temporal sulcus
pSTS	posterior superior temporal sulcus

Table of Contents

List of Tables

List of Figures

Permission Statement

Chapter 1 contains content printed in *Cognitive Neuropsychology*. The material in Figure 1.1 was originally printed in “The neural substrates of face processing models – A review” by M. Tovée and E. M. Cohen-Tovée, 1993, *Cognitive Neuropsychology, 10,* p. 512, and is copyrighted (*Copyright © Taylor & Francis*; www.tandfonline). The figure is reprinted in this thesis with permission. See Appendix A for letter indicating permission to use this content.

Chapter 2 contains content printed in *Frontiers of Perception Science*. This material is copyrighted (*Copyright © 2014 Stoesz and Jakobson*) and is reprinted here with permission from co-author, Dr. Lorna Jakobson. The manuscript as printed in *Frontiers of Perception Science* is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution, or reproduction in other forums is permitted, provided the original author(s) are credited and that the original publication in *Frontiers of Perception Science* is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms. The version presented in this thesis is not the final published version.

Stoesz, B. M., & Jakobson, L. S. (2014). Developmental changes in attention to faces and bodies in static and dynamic scenes. *Frontiers in Psychology, 5*(193). doi: 10.3389/fpsyg. 2014.00193

Chapter 4 contains content printed in *Journal of Vision*. This material is copyrighted (Copyright *© 2013 Association for Research in Vision and Ophthalmology)* with permission from co-author, Dr. Lorna Jakobson. The use, distribution, or reproduction in other forums is permitted, provided the original author(s) are credited and that the original publication in *Journal of Vision* is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms. The version presented in this thesis is not the final published version.

Stoesz, B. M., & Jakobson, L. S. (2013). A sex difference in interference between identity and expression judgments with static but not dynamic faces. *Journal of Vision, 13*(5):26, 1–14, http://www.journalofvision.org/content/13/5/26, doi:10. 1167/13.5.26.

Thesis Type

This dissertation was written using the *sandwich method* consisting of a collection of papers, which have been published (Chapters 2 and 4) or were submitted for publication (Chapters 3 and 5), and largely follows the formatting guidelines set out in the Publication Manual of the American Psychological Association. This thesis type was approved by my examining committee. Although there is some redundancy in the literature reviews and interpretation of the results across chapters, I have made an effort to reduce redundancies in the method sections of Chapters 3-5, as similar methods were used in these three studies. In order to reduce the number of pages in the thesis and because of the substantial overlap in the citations used across chapters, I have compiled one list of references at the end of the thesis rather than providing separate reference lists at the end of each chapter. How the chapters connect to each other the nature and extent of my contribution to each chapter are specified in the sections entitled, *My Contribution to the Publication* and *My Contribution to the Manuscript*.

CHAPTER 1: GENERAL INTRODUCTION

Faces are among the most complex and important visual stimuli in our environment. The ability to automatically and effortlessly attend to and process faces is present in infancy; nevertheless, these abilities continue to develop throughout childhood. A strong visual preference for faces and face-like stimuli (Downing, Bray, Rogers, & Childs, 2004; Langton, Law, Burton, & Schweinberger, 2008; Nummenmaa, Hyona, & Calvo, 2006; Valenza, Simion, Cassia, & Umilta, 1996) is present within several hours of birth (Nelson, 2001), and continues to develop throughout childhood (Elam, Carlson, Dilalla, & Reinke, 2010), adolescence (Freeth, Chapman, Ropar, & Mitchell, 2010), and adulthood (Bayliss & Tipper, 2005; Hershler & Hochstein, 2005). Children's ability to recognize facial identity and to extract social cues from faces, including cues that allow them to process facial expression, improves with experience. The ability to attend to and integrate facial cues is essential for the development of joint attention processes (e.g., Ricciardelli, Bricolo, Aglioti, & Chelazzi, 2002) and for inferring the intentions of others (see George & Conty, 2008). As such, impairments in attending to faces, processing one or more aspects of faces, or integrating facial cues may contribute to problems navigating the social world. Research involving individuals with prosopagnosia (Barton, 2003), autism spectrum disorders (ASD) (Barton, Hefter, Cherkasova, & Manoach, 2007), schizophrenia (Marwick & Hall, 2008), and Williams syndrome (Riby, Doherty-Sneddon, & Bruce, 2008) supports this view.

Although static stimuli have been employed in much of the existing face processing research, many researchers are beginning to incorporate dynamic face stimuli into their investigations (see Roark, Barrett, Spence, & O'Toole, 2003). Despite this, dynamic faces have rarely been used in studies involving children, and even fewer studies have reported using the same methods with children and adults (see Crookes & McKone, 2009; McKone, Crookes, Jeffery, & Dilks, 2012). As such, the aspects of face processing that have been investigated in children cannot be directly compared to the aspects observed in adults (Crookes & McKone, 2009). This thesis begins with a summary of selected literature describing why faces are considered a special class of visual stimuli and what is known about the processing of static and dynamic face stimuli. Following this, a summary of the development of face processing is provided.

Are Faces Special?

Faces are special in that they play a crucial role in social interaction and communication, but this highly salient and biologically significant set of stimuli is special for a number of other reasons as well (Farah, 1996; Langton et al., 2008; McKone, Kanwisher, & Duchaine, 2007; Sugita, 2009; Vuilleumier, 2000). One reason why faces are special is that they form a class of visual stimuli with low interstimulus perceptual variance. In other words, faces are extremely similar and share a common visual pattern – two eyes, a nose, and a mouth arranged in a particular spatial configuration (i.e., they share identical *first-order relations*) (Diamond & Carey, 1986). Faces also capture our attention in a way that other stimuli do not. Thus, experiments using visual search paradigms reveal that the detection of faces is quicker than the detection of objects regardless of the size of the search array, or when the background contains distracters that differ in object category, color, shape, or size (Hershler & Hochstein, 2005; Kuehn & Jolicoeur, 1994; Lewis & Edmonds, 2003; Tong & Nakayama, 1999). Although the variations between faces are subtle, we have the ability to recognize a familiar face in a crowd instantly. We are also able to recognize a familiar face despite changes in facial expression or movements due to speech (Schweinberger & Soukup, 1998). Other evidence suggesting that faces are special comes from research examining differences in the way we process faces versus objects, sensitivity to faces in early development, and the neural bases of face and object processing.

Holistic processing of faces. According to the traditional view of face processing, rapid recognition of faces occurs because we process upright faces using a specialized computational mechanism -- one that is different from that involved in the processing of objects. Specifically, human faces, to a larger extent than other types of stimuli, are processed holistically, as *gestalts* or undifferentiated wholes, rather than as a collection of isolated parts encoded independently of each other (Bartlett & Searcy, 1993; Farah, Tanaka, & Drain, 1995; McKone et al., 2007; Mondloch, Le Grand, & Maurer, 2002; Tanaka & Farah, 1993). Sensitivity to subtle differences in facial features (Rotshtein, Geng, Driver, & Dolan, 2007) and to the multiple spatial relationships among facial features (i.e., sensitivity to *second-order relations*) (Joseph & Tanaka, 2003; Searcy & Bartlett, 1996) allows one to recognize that one face is distinct from another. Evidence that we process faces holistically comes from unique behavioural signatures. For example, we are significantly impaired and slower at matching or making identity judgments concerning inverted faces compared to upright faces, regardless of their familiarity (Farah, Levinson, & Klein, 1995; Marotta, McKeeff, & Behrmann, 2002; Ross & Turkewitz, 1981), despite the fact that upright and inverted faces are identical in every way except for orientation. This decrement in performance, referred to as the *face-inversion effect*, is much larger than the inversion effect one sees with nonface objects (Yin, 1969), and is believed to arise from disruption of configural processing.

Another demonstration of the holistic nature of face processing comes from experiments exploring the *part-whole effect* – the finding that we are better at recognizing face parts when we see the face parts within the context of an intact whole face than when we see the face parts in isolation. The part-whole effect disappears when the face is inverted, again suggesting that upright faces are processed in a holistic manner (Farah, Tanaka, et al., 1995; Tanaka & Sengco, 1997). Farah et al. (1995) asked participants to study sets of upright whole faces and face parts, and to identify them at test. A robust inversion effect occurred for the faces that participants initially encoded as whole faces, but an inversion effect did not occur for the faces encoded as a set of parts. In another study, participants made same-different responses while viewing a specific sequence of stimuli: first, an upright face, an inverted face, a house, or a word; next, a mask; and finally, a second stimulus of the same type as the first (Farah, Wilson, Drain, & Tanaka, 1998). Participants' ability to match faces presented at the beginning and end of a sequence was disrupted to a greater extent by a whole-face mask than by a mask consisting of face parts, or by masks comprised of inverted faces, houses, or words (whole or fragmented). These results suggest that we typically process objects (e.g., inverted faces, houses, and words) in a piecemeal fashion whereas we process upright faces holistically.

The *composite-face effect* provides additional evidence for holistic processing of faces (Mondloch, Pathman, Maurer, Le Grand, & de Schonen, 2007; Susilo, Crookes, McKone, & Turner, 2009; Young, Hellawell, & Hay, 1987). The composite-face effect refers to the finding that participants are slower and less accurate in recognizing the top half of a familiar face when it is aligned with the bottom half of a different face than when the two halves are misaligned. The composite-face effect is absent or significantly reduced when aligned and misaligned composites are inverted. Indeed, individuals can recognize the top half of an inverted aligned composite face more quickly than the top half of an upright aligned composite face. The composite-face effect is thought to arise because holistic processing fuses the two halves of the aligned upright faces, thereby creating a new identity that is unfamiliar to the viewer (Young et al., 1987).

The research summarized above highlights the importance of holistic processing for extracting *facial identity* cues. Complementary studies suggest that holistic processing is also important for extracting *facial expression*. Calder, Young, Keane, and Dean (2000) have demonstrated that participants can more easily extract the expression of the top half of a composite face if it is misaligned with the bottom half of another expression than when the two half faces are aligned. As in the case of identity judgments (Young et al., 1987), the composite-face effect for expression judgments was reduced when the faces were inverted. Further evidence suggesting that expression analysis involves holistic processing comes from a study showing that inversion impairs participants' recognition of sad, fearful, angry, and disgusted expressions (McKelvie, 1995). The fact that participants' performance with happy expressions was not affected by inversion in this study led McKelvie (1995) to conclude that inversion is most disruptive in situations where it is necessary to process more than one salient feature (e.g., the shapes of the eyes *and* the mouth) to identify the expression correctly. Indeed, the data from some studies show that certain facial expressions may be represented and identified from a particular half of the face (Calder et al., 2000), or from a particular facial feature (e.g., eye or mouth shape) (Ellison & Massaro, 1997; Lipp, Price, & Tellegen, 2009). For example, anger, fear, and sadness are more readily recognized from the top half of the face, whereas happiness and disgust are more easily recognized from the bottom half (Calder et al., 2000). This work suggests that the extraction of expression requires greater reliance on local, as opposed to global, processing (see also Song & Hakoda, 2012), but this may vary depending on one's emotional state. Specifically, individuals focus more on parts of objects (Derryberry & Reed, 1998) and faces (Curby, Johnson, & Tyson, 2012) when they are experiencing negative emotional states than when their mood is more positive.

A number of studies have explored the processing of visual speech. Articulation produces changes in the second-order relationships between facial features. If inversion disrupts the processing of these spatial relationships, then inversion should also affect speech perception. Indeed, the results of studies exploring an audiovisual illusion called the McGurk effect (McGurk & MacDonald, 1976) suggest that visual speech is processed holistically (Jordan & Bevan, 1997; Massaro & Cohen, 1996; Rosenblum, Yakel, & Green, 2000). Listeners experience the McGurk illusion when they are presented with incongruent auditory and visual speech; under these circumstances, what they perceive is a blend of the phoneme (or speech sound) and the viseme (the facial posture or movement that occurs with the voicing of a phoneme). Interestingly, facial orientation affects the strength of the McGurk effect (Jordan & Bevan, 1997; Massaro & Cohen, 1996; Rosenblum et al., 2000). For example, when participants viewed visemes with the largest visual articulations (e.g., /ma/), face inversion resulted in the loss of the McGurk effect that was observed with upright stimuli (Jordan & Bevan, 1997). This result suggests that some visemes are processed holistically.

Inversion also disrupts the processing of eye gaze (Farroni, Johnson, & Csibra, 2004; Senju & Hasegawa, 2006; Senju, Kikuchi, Hasegawa, Tojo, & Osanai, 2008). In a series of experiments, Senju and Hasegawa (2006) examined whether face and eye orientation would influence performance on a visual search task that required participants to find the test stimulus within an array that had the same gaze direction as that depicted in the target face. The detection of a match was significantly faster when the target exhibited a direct gaze as opposed to an averted gaze, but only for upright eyes. Inversion of the eye region resulted in the elimination of this effect, even when the rest of the face remained upright. Senju and Hasegawa (2006) proposed that eye inversion eliminates or reduces the saliency of direct gaze, rather than interfering with our ability to discriminate between direct and averted gaze, and that some form of configural processing of the eyes (and eye region) occurs during gaze judgments.

Sensitivity to faces in early development. Another indication that faces are special comes from the idea that attending to and processing faces have earlier developmental precursors than attending to and processing other objects (see Susilo et al., 2009). Typically developing infants are born with a bias to orient toward faces rather than other objects (Teresa Farroni et al., 2005; Morton & Johnson, 1991; Turati, Valenza, Leo, & Simion, 2005). Klein and Jennings (1979) described four types of reactions (i.e., visual focus, smile, vocalization, and movement of head, arms, legs, or trunk) that infants made to two social stimuli (mother and a stranger) and to an inanimate stimulus (a musical mobile). By 4 weeks of age, infants showed greater movement responses to the faces than to the mobile, and by 12 weeks of age, infants moved, smiled, and vocalized more in response to faces than to the mobile. Interestingly, infants also show an attentional bias for point-light displays of humans or animals versus point-light displays of random dot motion (Bardi, Regolin, & Simion, 2011; Simion, Regolin, & Bulf, 2008) or moving vehicles (Arterberry & Bornstein, 2002). These results suggest that infants may have a general bias to attend to cues signalling animacy.

Eye and head movement studies have shown that newborns will follow a schematic face further into the periphery than other types of stimuli (Easterbrook, Kisilevsky, Muir, & Laplante, 1999; Goren, Sarty, & Wu, 1975; Johnson, Dziurawiec, Ellis, & Morton, 1991). Goren et al. (1975) presented different two-dimensional, head-shaped forms to healthy newborns soon after birth. Infants as young as 3 minutes old turned their heads and eyes further for forms depicting a face than for those depicting a scrambled face or a blank pattern, suggesting that the arrangement of facial features is especially important in attracting attention. Interestingly, newborns also turn their heads more to view a scrambled face than to view a blank stimulus, indicating that facial features are important for preferential tracking (Johnson et al., 1991). In addition, newborns look longer at nonface patterns that resemble faces than at those that do not (Valenza, Simion, Cassia, & Umiltà, 1996), and prefer to look at faces that appear to be looking at them compared to faces that appear to be looking away (Teresa Farroni, Csibra, Simion, & Johnson, 2002). Additional research has shown that the bias to attend to faces continues throughout childhood (Kikuchi, Senju, Tojo, Osanai, & Hasegawa, 2009), adolescence (Freeth et al., 2010), and adulthood (Bayliss & Tipper, 2005; Hershler & Hochstein, 2005; Langton et al., 2008).

Perceptual expertise for faces. Infants are born never having seen a face, but by the time we reach adulthood, we are experts at processing the many types of information contained within faces. Some researchers have argued that faces themselves are not really special (e.g., Diamond & Carey, 1986; Gauthier, Behrmann, & Tarr, 1999), but that, because of the enormous exposure to faces occurring over the lifespan, face-processing *skills* are special. The *face-expertise theory* suggests that experience plays a major role in our ability to become proficient at fine-grained within-category discrimination between faces. Evidence for this theory comes from a number of lines of research. First, we can more easily recognize faces from races with which we are familiar than faces from races with which we have little experience (known as the own-race advantage [ORA]) (see Meissner & Brigham, 2001 for a review). Although the mechanisms for ORA are not well understood, some work suggests that weaker holistic processing (as described above) for the faces of other races compared to own-race faces contributes to the effects (Michel, Caldara, & Rossion, 2006). For example, the inversion (Stein, End, & Sterzer, 2014), part-whole (Degutis, Mercado, Wilmer, & Rosenblatt, 2013; Tanaka, Kiefer, & Bukach,

2004) and composite (Michel, Rossion, Han, Chung, & Caldara, 2006) effects are stronger for own-race than other-race faces. In addition, viewers are more sensitive to the distances between facial features for own-race compared to other-race faces, monkey faces, or houses (Robbins, Nishimura, Mondloch, Lewis, & Maurer, 2010). A growing body of research also shows that individuation training (i.e., practicing to recognize objects at a more specific category level) (McGugin, Tanaka, Lebrecht, Tarr, & Gauthier, 2011) and training in focusing attention on configural and holistic aspects of faces (Mercado, Cohan, & DeGutis, 2011) results in race-specific improvement in face discrimination.

A second line of evidence for the face-expertise theory comes from the work showing that the typical adult pattern of expertise for face processing does not emerge when visual input (due to bilateral congenital cataracts) is absent during infancy. Individuals who have experienced this type of visual deprivation can process facial features similarly to individuals with typical visual experience (LeGrand, Mondloch, Maurer, & Brent, 2001), but are impaired at discriminating faces that differ in the spacing between features (Robbins et al., 2010). These individuals also show a reduced composite face effect (i.e., they actually perform better than controls when faces are aligned), suggesting that they are unable to integrate information across the face into a gestalt (Le Grand, Mondloch, Maurer, & Brent, 2004). Together, these findings suggest that the development of expertise in face processing is dependent on exposure to faces during the first few months of life.

Neural substrates of face processing. Faces are processed in specialized cortical and subcortical networks (Haxby, Hoffman, & Gobbini, 2000; Johnson, 2005) that are separate from those supporting body and hand recognition (Downing, Jiang, Shuman, & Kanwisher, 2001; Kanwisher, McDermott, & Chun, 1997; Schwarzlose, Baker, & Kanwisher, 2005) and the recognition of objects (Tsao, Freiwald, Tootell, & Livingstone, 2006). According to Bruce and Young's (1986) cognitive model of face processing, attending to faces and the processing of facial identity, expression, and speech depend on parallel pathways that can function independently. Tovée and Cohen-Tovée (1993) updated Bruce and Young's model by adding a module dedicated to eye gaze detection and by mapping the cognitive framework onto the ventral and dorsal visual processing streams (Mishkin, Ungerleider, & Kathleen, 1983). Haxby et al. (2000) extended this work by suggesting that face processing is subserved by a core face-processing neural network that is supported and influenced by an extended system. More recently, Roark et al. (2003) proposed two speculative modifications to Haxby et al.'s model that may offer a clearer view as to how facial motion is processed. I describe briefly the key features of each of these models below.

Bruce and Young (1986). In their theoretical framework, Bruce and Young (1986) postulated that faces undergo structural encoding -- a type of processing that captures characteristics of facial features and the configuration of those features within a face. Following this, further processing occurs in four components serving various perceptual classification functions: (1) face recognition (familiar faces), (2) expression analysis, (3) facial speech analysis, and (4) directed visual processing. Each face recognition unit contains stored structural codes describing a face known to the viewer, which are linked to a separate store of identity-specific semantic codes (Person Identity Nodes or PINs) that describe the familiar person (e.g., the person's occupation and where we usually encounter the person) and that are accessed when a person, as opposed to a face, is recognized. PINs, in turn, are linked to name codes, which also have independent representations. In this model, face recognition is independent from the analyses of expression and speech. The configuration of various features leads to categorization of expression, whereas the visible movements of the mouth are categorized during the analysis of facial speech. Directed visual processing refers to the process by which we encode certain types of facial information selectively and strategically, as opposed to *passively* recognizing expressions, speech, and identities. Bruce and Young suggest, for example, that if we are going to meet a friend in a public place, we actively look for faces that resemble that of the friend. In addition, we use directed visual processing when we compare unfamiliar faces in laboratory experiments. In this model, each module shares connections with the rest of the cognitive system.

Tovée and Cohen-Tovée (1993). Tovée and Cohen-Tovée (1993) extended Bruce and Young's (1986) model by adding a module dedicated to the analysis of gaze information (i.e., processing information regarding another person's focus of attention; see Figure 1.1). The authors also summarized work that mapped various functional components in the human face processing system map onto specific brain regions. Studies carried out in nonhuman primates show that cells in the inferotemporal (IT) cortex[1] are involved in the processing of facial features and facial identity (Hasselmo, Rolls, & Baylis, 1989). In particular, certain cells respond preferentially when the viewer sees particular features (the mouth, hair, or eyes) or eyes looking in a particular direction, whereas other cells are sensitive to second-order spatial relations of facial features. Cells in the IT cortex show selectivity for identity, whereas those in the superior temporal sulcus (STS) region show selectivity for expression. The majority of face cells in the superior temporal polysensory (STP) region are responsive to the direction in which the head points, but cells in the banks and floor of the STS support the ability to perceive the direction of

[1] The inferotemporal (IT) cortex in nonhuman primates is homologous to the face-selective region in the lateral fusiform gyri in the human ventral temporal cortex. This face-selective region has been dubbed the "fusiform face area" (FFA) because it responds more strongly to faces than to any other object category (e.g., Kanwisher et al., 1997).

another person's gaze. The left STP appears to support lip reading ability, as evidenced by studies showing that individuals with right-sided lesions show impairments in facial identity and expression recognition but no impairment in lip reading ability, whereas individuals with left-sided lesions show the opposite pattern (Campbell, Landis, & Regard, 1986). Cells supporting directed visual attention (including but not limited to those involved in attention to faces) are located in the lower banks and floor of the STS. Finally, information associated with a particular face is stored in a number of regions including the inferior temporal lobe, amygdala, hippocampus, entorhinal cortex, and ventromedial prefrontal lobes. Although regions in the dorsal visual processing stream are implicated in several processes, Tovée and Cohen-Tovée place particular emphasis on the role that brain regions in the ventral visual stream play in the parallel processing of various types of facial information. The authors also suggest that face processing is no different from object processing, except in the amount of cortex that is devoted to the processing of faces.

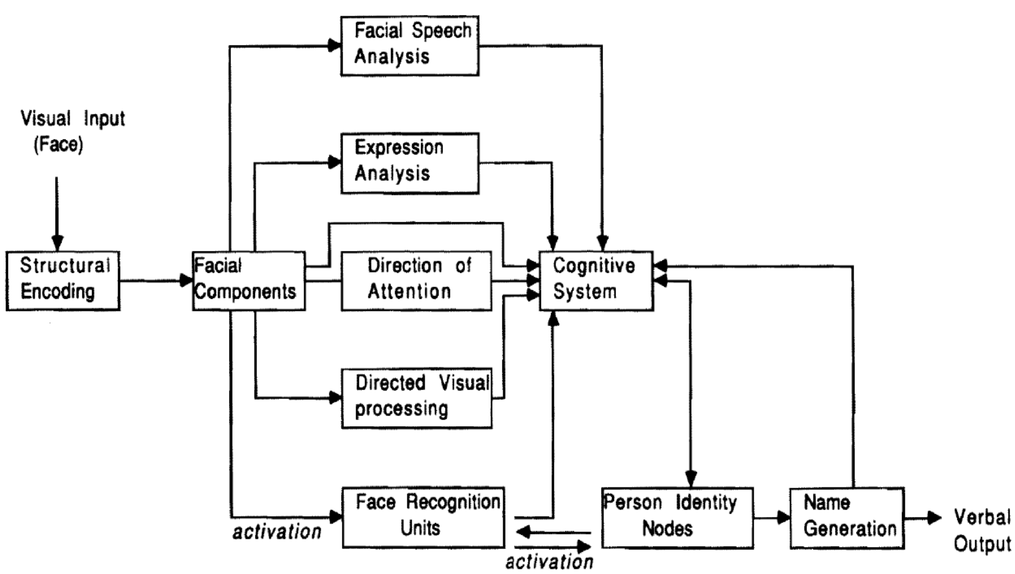

Figure 1.1. A modification of the face processing model proposed by Bruce and Young (1986), incorporating information obtained from neuropsychological studies. From "The neural substrates of face processing models – A review" by M. Tovée and E. M. Cohen-Tovée, 1993, *Cognitive Neuropsychology, 10*, p. 512. Copyright © Taylor & Francis; www.tandfonline.com. Reprinted with permission.

Haxby, Hoffman, and Gobbini (2000). Individuals can have selective impairments in face recognition, the analyses of expressions and facial speech, or the ability to judge gaze direction (Campbell, Landis, & Regard, 1986; Humphreys, Donnelly, & Riddoch, 1993). Further, results from early single cell recordings and neuroimaging studies suggest that different and largely independent neural structures subserve various aspects of face processing. This does not mean, however, that there is strict independence on a functional level, as specialized brain areas may affect each other within the face-processing and social-cognition systems (Beauchamp & Anderson, 2010; Kaufmann & Schweinberger, 2005). Indeed, after integrating the findings of neuropsychological and neuroimaging studies examining the various aspects of face processing, Haxby et al. (2000) proposed a *distributed neural system for face processing* that includes multiple, bilateral regions organized into a core and an extended system. The core system includes the inferior occipital gyrus (where the initial visual analysis of facial features occurs), with further processing taking place in the lateral fusiform gyrus [or the fusiform face area (FFA), which refers to all regions within the fusiform gyrus that show a functionally defined preference to faces] and the STS (Haxby et al., 1999; Hoffman & Haxby, 2000; Kanwisher et al., 1997; Yovel & Kanwisher, 2004).

Many researchers have demonstrated that the fusiform gyrus responds more strongly to faces than to a range of other types of objects, including cars, furniture, houses, and hands (Aguirre, Singh, & D’Esposito, 1999; Grill-Spector, Knouf, & Kanwisher, 2004; Kanwisher et al., 1997; McCarthy, Puce, Gore, & Allison, 1997), suggesting that the FFA is the primary site for face-specific processing (McKone et al., 2007). The FFA also responds more strongly to upright than inverted faces (Yovel & Kanwisher, 2005), and is activated during tasks requiring holistic processing (Schiltz & Rossion, 2006). Researchers propose that the FFA processes invariant aspects of a face needed for identity recognition (Gauthier & Logothetis, 2000; Hoffman et al., 2000; Hoffman & Haxby, 2000; Sergent, Ohta, & MacDonald, 1992). This proposal gains support from studies showing that individuals with fusiform damage are impaired in face recognition (Barton, Press, Keenan, & O'Connor, 2002; Sorger, Goebel, Schiltz, & Rossion, 2007; Williams, Berberovic, & Mattingley, 2007).

Unlike the FFA, the STS processes the *changeable* (or variant) features of faces that viewers can use to infer information about another person's state of mind (Allison, Puce, & McCarthy, 2000; Bonda, Petrides, Ostry, & Evans, 1996; Haxby et al., 2000; Hoffman & Haxby, 2000; Pelphrey, Morris, & McCarthy, 2005; Puce, Allison, Bentin, Gore, & McCarthy, 1998). Tasks requiring the processing of gaze direction (Hoffman & Haxby, 2000; Puce et al., 1998; Sato, Kochiyama, Uono, & Yoshikawa, 2008), facial expression (Furl et al., 2010; Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001; Olson, Gatenby, & Gore, 2002), lip reading (Calvert, 1997; Campbell et al., 2001; Puce et al., 1998), and facial motion in complex movie scenes (Bartels & Zeki, 2004; Hasson, Nir, Levy, Fuhrmann, & Malach, 2004) all elicit strong responses in the STS. The STS also responds more to observed eye and mouth movements than it does to various types of nonbiological motion (Puce et al., 1998). Narumoto et al. (2001) used a delayed match-to-sample procedure to determine the role of the STS in the processing of facial expressions. Face-responsive areas responded more to faces than to objects. However, only the right STS showed enhanced activation when participants matched faces based on expression relative to when they matched faces based on identity or contour, or when they matched images of objects. Thus, the STS appears to be particularly involved in social perception (Campbell et al., 2001; Hoffman & Haxby, 2000; Narumoto et al., 2001; Pelphrey, Morris, McCarthy, & Labar, 2007). Interestingly, Narumoto et al. (2001) used static images of facial expressions that convey implied motion, and perhaps the STS responded to this implied motion. Indeed, the STS and other motion-sensitive regions found near it (e.g., visual area MT) (Zeki et al., 1991) are involved in the processing of stimuli that imply motion (Kourtzi & Kanwisher, 2000; Krekelberg, Dannenberg, Hoffmann, Bremmer, & Ross, 2003; Senior et al., 2000).

Other brain regions also contribute to face perception by processing the significance of information picked up from the face. Auditory regions associated with processing speech sounds also process visual speech information (Calvert et al., 1997), facial expressions activate limbic regions involved in emotion processing (Kesler-West et al., 2001; Phillips et al., 1998), and parietal areas that are associated with spatial attention are involved in the perception of eye gaze direction (Hoffman et al., 2000). Although these secondary face processing regions are parts of neural systems involved in other cognitive functions, they facilitate the accurate recognition of changeable facial features when acting in concert with the core system (Haxby et al., 2000).

The *distributed neural system for face perception* (Haxby et al., 2000) shares features with the other face processing models, in that it identifies an overall hierarchical organization with distinct functional modules underlying the processing of identity, expression, facial speech, eye gaze, and directed visual attention (Bruce & Young, 1986; Tovée & Cohen-Tovée, 1993). Haxby et al.'s model, however, differs from the two earlier models in at least two ways. First, the *distributed neural system for face perception* places greater emphasis on the importance of the connections between the neural substrates associated with particular cognitive functions. Secondly, this model also emphasizes that, in addition to the ventral stream, the dorsal stream and the STS are necessary for the successful processing of social information, particularly information that is dynamic (changeable or variant) in nature.

Roark, Barrett, Spence, and O'Toole (2003). Roark et al ***(2003)***. proposed two modifications to the *distributed neural system for face perception* (Haxby et al., 2000). The first modification suggests that regions in the dorsal stream [specifically middle temporal (MT) cortex] and the STS together process facial motion *signatures* (i.e., idiosyncratic, facial movements that can be used for person identification – *supplemental information hypothesis*), along with other kinds of dynamic social communication cues. The second modification concerns the contribution of structure-from-motion analysis to face recognition. Roark et al. (2003) suggest that the benefits of motion in face recognition come from using motion to extract a more accurate representation of the invariant or static structure of a face (*representation enhancement hypothesis*). Thus, motion bootstraps the encoding of face structure. If this is the case, then the perceptual quality of participants' feature processing and identity extraction should improve when they view moving, rather than static, faces. In the next section, I summarize results from studies comparing viewers' responses during tasks employing static and dynamic faces.

Static versus Dynamic Face Processing

Natural facial movements convey changes in our expression of emotion or verbal information, and even subtle movements of our mouth and eyes provide a rich and powerful source of social information. Recognizing and understanding various facial movements are basic social skills required to interact with other people successfully. We have much more experience looking at naturally moving faces than we do looking at faces in photographs. Despite this, previous research has largely overlooked the role of motion in face processing and social cognition; however, behavioural studies are now emerging that have examined its effect on decoding facial information. Interestingly, despite the greater ecological validity of dynamic faces, some researchers have failed to find an advantage for processing dynamic over static faces

(Bruce, Henderson, Newman, & Burton, 2001; Christie & Bruce, 1998; Fiorentini & Viviani, 2011; Kamachi et al., 2001; O'Toole et al., 2011). Others report that we represent dynamic facial cues in a *qualitatively* different way than static facial cues (Ambadar, Schooler, & Cohn, 2005; Pilz, Thornton, & Bülthoff, 2006).

The magnitude of the performance differences observed when viewers process static versus dynamic faces may depend on the task and the specific stimuli used (Roark, O'Toole, Abdi, & Barrett, 2006). Differences in the processing of static and dynamic stimuli are typically observed when emotional (Fujimura & Suzuki, 2010; Pike, Kemp, Towell, & Phillips, 1997) or speaking (Everdell, Marsh, Yurick, Munhall, & Pare, 2007) faces are presented. When Fujimura and Suzuki (2010) asked participants to evaluate various expressions (excited, happy, calm, sleepy, sad, angry, fearful, and surprised), dynamic calm faces were rated as less activated (or arousing) than static faces, but there were no differences in static and dynamic activation ratings for the other seven emotions. Despite this, participants categorized dynamic excited, happy, and fearful faces (which were judged to have high- or mid-range activation expressions) more accurately than the corresponding static faces, a result suggesting that it is not the level of activation, but rather the emotional properties of the expression, that produced the dynamic advantage.

Others have argued that a key factor underlying the dynamic advantage is whether the movement is natural or unnatural. Horstmann and Ansorge (2009) reported that participants' search for dynamic angry target faces among distracters was more efficient than their search for static angry faces among distractors (Experiment 1). The researchers posited that the exaggerated vertical extension of the dynamic stimuli was greater for negative than for positive emotional expressions, making the angry faces more salient than the distractor faces. Further study using stimuli with more naturalistic facial movements revealed no dynamic advantage. Thus, exaggerated movements may confer an advantage because they make expression salient/attention grabbing.

Observable differences between the processing of static and dynamic face stimuli are also more likely when the processing demands of the task are higher (Kaufmann & Schweinberger, 2005; Thornton & Kourtzi, 2002). For example, Kaufmann and Schweinberger (2005) used a Garner's selective-attention paradigm (Garner, 1976) to examine the differences in processing visual speech while viewing dynamic and static faces. In the baseline condition, the relevant dimension varied (e.g., vowel articulation: /i/ or /u/), while the irrelevant dimension was held constant (e.g., identity). In the correlated condition, speaker A always articulated the /u/ phoneme, whereas speaker B always articulated the /i/ phoneme. In the orthogonal condition, both the relevant and the irrelevant dimensions were covaried (with multiple speakers articulating each phoneme). When dynamic faces served as stimuli (Experiment 1), adults took more time to make visual speech judgments in the orthogonal condition than in the control condition. This *Garner interference effect* suggests that the processing of speech involves the processing of identity and supports the idea that speech processing is dependent on identity. Interestingly, this effect was not evident for static-sequential or static presentation conditions (Experiment 3) – a finding that is inconsistent with other studies using static images (Schweinberger & Soukup, 1998). Overall, Kaufmann and Schweinberger (2005) suggest that their results do not support strictly separated modules for identity and facial speech (as described by Bruce & Young, 1986), but may instead support the idea that the processing of facial cues interacts within the extended system of the *distributed human neural system for face perception* (Haxby et al., 2000).

As with behavioural studies, most neuroimaging studies have investigated brain regions involved in face processing using static images, but there is a trend to incorporate moving face stimuli in this type of research. Results from recent research shows that the visual processing of faces from static and dynamic displays involve different neural networks and/or different levels of activation of the same brain regions (Arsalidou, Morris, & Taylor, 2011; Fox, Iaria, & Barton, 2009; Kessler et al., 2011; Sato, Kochiyama, Uono, & Yoshikawa, 2010; Schultz & Pilz, 2009). For example, Schultz and Pilz (2009) found that presentation of faces undergoing nonrigid facial motion (i.e., temporary deformations of the face that are seen in speech production and facial expressions) elicited stronger responses than presentation of static faces in lateral temporal areas corresponding to V5/MT cortex and the STS. Interestingly, static-face-sensitive regions in bilateral fusiform gyrus and left inferior occipital gyrus also responded more strongly to surprised and angry dynamic faces than to static faces depicting these same emotions. These results suggest that ventral temporal and inferior lateral occipital areas are involved in the integration of form and motion information during the processing of dynamic faces.

Kessler et al. (2011) found similar differences in brain activation between static and dynamic faces using a wider range of expressions. Adults passively viewed photographs and film clips (created using morphing techniques) of prototypical facial expressions of fear, disgust, sadness, and happiness. For all expressions, the STS, MT cortex, fusiform gyrus, thalamus, and other frontal and parietal areas showed increased activation bilaterally for dynamic versus static faces. These results confirm previous findings on neural correlates of the perception of dynamic facial expression processing (Furl et al., 2010; Narumoto et al., 2001; Olson et al., 2002; Phillips et al., 1998; Schultz & Pilz, 2009) and extend the literature demonstrating the involvement of the STS and MT cortex in the perception of other types of variant facial information (Calvert et al., 1997; Campbell et al., 2001; Puce et al., 1998; Sato et al., 2008). The fact that the fusiform gyrus was activated more for dynamic than static faces (Schultz & Pilz, 2009; Trautmann, Fehr, & Herrmann, 2009; but see Arsalidou et al., 2011; Lee et al., 2010) stands in contrast to classical models of face perception describing the importance of the FFA in processing invariant (and static) aspects of faces (i.e., identity) (Bruce & Young, 1986). Kessler et al.'s findings support the idea that the FFA receives structure-from-motion information from MT and STS (see Roark et al., 2003).

Not only do certain face processing regions show higher levels of activation for dynamic compared to static face stimuli, but the type of motion (natural vs. randomly sequenced or scrambled) is important as well (Furl et al., 2010). Furl et al. used magnetoencephalography to measure brain activity while participants viewed dynamic stimuli that resembled naturalistic transitions between fearful and neutral expressions. Evoked responses to stimuli showing natural movement were compared with those following presentation of scrambled stimuli that were unnatural and lacked a coherent trajectory. Visual areas including the STS showed increased activation for predictable movement relative to unpredictable movement. Specifically, the predictable movement in naturally unfolding facial expressions elicited early responses in primary visual cortex (165 ms), followed by increased activity in bilateral visual cortex, right posterior STS (pSTS) and posterior fusiform gyrus (237 ms). Furl et al.'s (2009) results are consistent with the fMRI and MEG results reported by Lee et al. (2010) showing higher levels of activation in STS for rigid facial motion (i.e., naturally turning heads) compared to static faces or scrambled movement.

Face Processing in Children

Studies involving children have provided fascinating insights into how our ability to process faces improves over the course of development (see Chung & Thomson, 1995). Research exploring the early development of facial identity coding has demonstrated that, in addition to showing a preference to attend to faces, 3-month-old infants have impressive face recognition capabilities (Pascalis, de Haan, Nelson, & de Schonen, 1998). Pascalis et al. investigated 3- and 6-month-olds' memory for faces using a visual paired-comparison task (Fagan, 1973). In this task, infants were presented with a photograph of a face for a brief familiarization period. After being habituated to the same face presented in various poses, their face recognition was assessed either 2 minutes or 24 hours later by presenting stimulus pairs comprised of one familiar face and one novel face. Both groups of infants exhibited novelty preferences after a delay (i.e., longer looking times at the novel compared to the familiar stimulus), suggesting that the infants had remembered the test faces.

Improvement in children's ability to recognize new faces is generally observed at 3-4 years of age and this ability reaches adult-levels by the age of 10-11 years (e.g., Carey & Diamond, 1977; Ge et al., 2008; Johnston & Ellis, 1995). During this time, developmental changes in face processing have been documented. For example, Carey and Diamond (1977) reported that older children use isolated features in their attempts to recognize people, whereas young children do not. Children ages 6, 8, and 10 years viewed photographs of different people wearing different paraphernalia (i.e., clothing, hats, hairstyle, or eyeglasses). After the presentation of each picture, the child viewed a pair of photographs of two different people and decided which of the two photographs depicted the person shown in the initial photograph. Younger children made matches based on paraphernalia, but older children were less likely to do so. These results are consistent with other research showing that children may be more distracted by isolated features or paraphernalia and, thus, have a tendency to use such cues to recognize people (Campbell & Tuck, 1995; Campbell, Walker, & Baron-Cohen, 1995; Freire & Lee, 2001) but that this tendency decreases with age (Ellis, Shepherd, & Davies, 1979).

Johnston and Ellis (1995) examined age-related differences in face memory in children ranging in age from 5-13 years and in one group of adults. In the learning phase, participants viewed a sequence of faces. During the test phase, participants viewed a larger set of faces and decided as quickly as possible which of these were new and which they had seen before. Overall, 5-year-olds were less accurate and slower at recognizing faces than older children and adults. In addition, unlike 5-year-olds, adults and children 9 years of age and older showed a distinctiveness effect; thus, atypical faces were easier to recognize than average-looking (typical) faces as evidenced by three out of four measures of performance [i.e., response latencies, sensitivity (*d'* scores), and number of false positives].

Carey and Diamond (1977) presented 6-, 8-, and 10-year-olds with sets of upright and inverted photographs of faces and houses. Recognition was then tested with pairs of photographs (the original stimulus and a new exemplar) shown in the same orientation as the original test stimulus. For 6- and 8-year-olds, the face inversion effect was comparable in magnitude to the inversion effect for houses. Ten-year-olds, however, showed a dramatically larger inversion effect for faces than houses -- the typical adult pattern (Yin, 1969). In addition, for stimuli presented in a given orientation, various age-related differences in performance occurred across the various classes of stimuli. First, recognition accuracy for upright faces improved significantly between 6 and 10 years of age, whereas performance on upright houses remained constant across the three ages. Secondly, performance on inverted faces remained constant (i.e., 6- and 8-year-olds did as well on inverted faces as adults), whereas performance with inverted houses improved (from chance level for 6-years-olds to above chance for 10-year-olds). These results suggest that by the age of 10 years a bias to process faces holistically replaces a part-based approach (Carey & Diamond, 1977).

Schwarzer (2000) instructed 7- and 10-year-old children and adults to classify schematic upright and inverted faces as either child or adult faces. The schematic faces varied in terms of four features: face shape, eyes, nose, and mouth. The prototypical child's face was round with large round eyes, a small round nose, and a small mouth, whereas the adult's face was narrow, and consisted of oval eyes, a narrow nose, and a wide mouth. Test stimuli consisted of the prototypical child and adult faces as well as ambiguous faces containing a mixture of child-like and adult-like features. All three groups categorized inverted, ambiguous faces based on individual features. With upright stimuli, however, 7-year-olds consistently used individual features (e.g., eyes), 10-year-olds used individual features on some occasions, but overall similarity on other occasions, and adults consistently used overall similarity to perform their categorizations. These findings suggest that 7-year-old children process both upright and inverted faces analytically, whereas a growing proportion of 10-year-olds, like adults, process upright faces using a holistic strategy (Schwarzer, 2000).

More recent studies have also reported age-related improvements in performance on face processing tasks involving sensitivity to second-order relations (see Maurer, Le Grand, & Mondloch, 2002, for a review). Mondloch, Geldart, Maurer, and Le Grand (2003) developed five face-matching tasks that tapped into this ability. Adults and children matched one of three static test faces to a static target face based on identity, or based on expression, gaze, or visual speech. Test and target faces could differ with respect to expression or head orientation. Compared to adults, 6-year-olds made more errors on every task, 8-year-olds made more errors when matching gaze direction and when matching facial identity whether expression or head orientation varied, and 10-year-olds made more errors only when matching facial identity when there was change in head orientation. Unfortunately, Mondloch et al. did not report inversion effects for the children, as they were shown only upright faces. However, adults did show an inversion effect for matching facial identity when there was a change in head orientation. As a whole, the results indicate that slow development of sensitivity to second-order relations contributes to children's especially poor identity recognition when they view a face in a new orientation (Mondloch et al., 2003).

The results summarized above suggest that younger children code faces in a *qualitatively* different way than adults, not utilizing a holistic processing style until later in development. Conversely, other research shows that 4-5-year-olds *do* display signs of using holistic processing, including the composite-face effect (Susilo et al., 2009) and the inversion effect (Gilchrist & McKone, 2003; Pellicano, Rhodes, & Peters, 2006; Picozzi, Cassia, Turati, & Vescovo, 2009), and sensitivity to exact spacing between facial features (Gilchrist & McKone, 2003; McKone & Boyer, 2006; Pellicano et al., 2006). For example, Picozzi et al. (2009) presented 3-4-year-olds with a target stimulus on a screen, followed first by a blank screen and then by a screen showing two images (the target and a distracter). Children decided which image in the final pair was the target. The children showed an inversion effect when the stimuli were faces but not shoes. Picozzi et al. (2009) suggested that preschoolers already possess the ability to extract the critical cues that lead to adults' efficient face recognition. These types of effects can also be quantitatively measured even in infancy (e.g., Gallay, Baudouin, Durand, Lemoine, & Lécuyer, 2006).

Given these findings, McKone and colleagues (Crookes & McKone, 2009; McKone et al., 2012) argue that face processing mechanisms *are* mature in early childhood. The authors suggests that the age-related improvements in performance on face processing tasks reported in many studies may reflect domain general improvements in perceptual skills (e.g., acuity), cognitive abilities (e.g., memory, sustained attention), and processing speed (Baenniger, 1994; Crookes & McKone, 2009; Taylor, Batty, & Itier, 2004), as these factors are known to improve substantially across childhood and adolescence (Betts, McKay, Maruff, & Anderson, 2006). In keeping with this suggestion, Spangler, Schwarzer, Korell, and Maier-Karius (2009) demonstrated that when participants (ages 5-6 years, 6-8 years, 9-11 years, and adults) sorted faces according to facial identity, accuracy did not differ across age groups, but response times (RTs) decreased significantly with increasing age. These results are consistent with the results from studies using evoked response potentials showing that a face-selective ‘N170’ over posterior temporal sites peaks at 170 ms after stimulus onset in adults, and peaks at progressively later times in earlier development (e.g., 185 ms in 10-11-year-olds, 270 ms in 4-5-year-olds) (Taylor et al., 2004).

Results from fMRI studies may also support the idea that children’s face processing abilities are mature. For example, a recent fMRI study involving participants ranging in age from 6 years to adult showed modest support for a developmental trend in the volume of the right FFA, and no developmental change in the intensity of activation in response to face stimuli (Haist, Adamo, Han Wazny, Lee, & Stiles, 2013). Specifically, when task demands were minimal (i.e., when the task did not include a memory component), activation in the middle portion of the right fusiform gyrus most commonly found in adults (as described above) was reliably detected in the youngest children, and was more likely to be included in the FFA with increasing age. This result is consistent with other studies reporting FFA activation in children as young as 6 years (Peelen, Glaser, Vuilleumier, & Eliez, 2009; Scherf, Luna, Avidan, &

Behrmann, 2011; Suzanne Scherf, Thomas, Doyle, & Behrmann, 2013). Outside of the FFA, nearly every part of the extended face processing system was hyperactivated in children compared to adults (Haist et al., 2013).

Despite the rich literature describing static face processing in children, their ability to process dynamic faces has not been described in much detail. In one of the few studies to directly compare static and dynamic face processing during development, Bahrick, Moss, and Fadil (1996) found that older infants (3-, 5-, and 8-month-olds) could discriminate their own face from that of a peer whether the faces were static or engaged in natural movement, however, 2-month-old infants succeeded only with moving stimuli. Bahrick, Gogate, and Ruiz (2002) then showed that 5-month-olds can discriminate and remember repetitive actions but not the faces of the people performing those actions, even though they were able to discriminate faces shown in static poses. These results may reflect earlier and greater attentional selectivity to actions than to faces (Bahrick et al., 2002).

Perhaps, younger infants are not performing as well as older infants on tasks involving dynamic faces because they do not attend as closely to faces. Frank, Vul, and Johnson (2009) found this may be the case. Infants 3-, 6-, and 9-months of age and adults viewed a series of short (4 s) clips from an animated cartoon (i.e., A Charlie Brown Christmas) and the proportion of time each group spent looking at the characters' faces in the clips was recorded with a Tobii ET-17 binocular corneal-reflection eye-tracker. Age-related differences in attention to faces were highly significant. Specifically, with increasing age, participants spent a larger proportion of time viewing the faces in the complex dynamic scenes. Age-related increases in attention to dynamic faces and general improvement in face processing task involving moving faces could be related to the finding that motion perception does not reach maturity until the early teenage years (e.g., Bogfjellmo, Bex, & Falkenberg, 2014).

Neuroimaging studies examining dynamic face processing in children or adolescents are even rarer than behavioural studies. In those that have explored the neural systems involved in processing of dynamic facial expressions, results suggested that adolescents (aged 11–17 years) exhibit larger responses (both activations and deactivations) in the left STS when they viewed fearful compared to happy expressions (Rahko et al., 2010). These results are consistent with the involvement of the STS in processing dynamic facial expression in adults (Furl et al., 2010; Narumoto et al., 2001; Olson et al., 2002; Phillips et al., 1998; Schultz & Pilz, 2009). Overall, however, our understanding of the functional systems underlying dynamic face processing in children is much less advanced than our understanding of the organization of these systems in adults (Haxby et al., 2000).

The Present Series of Studies

Given the limited research exploring the development of dynamic face processing, I designed two studies to examine age-related differences in performance on tasks requiring the processing of static and dynamic faces. In Chapter 2 (*Developmental Changes in Attention to Faces and Bodies in Static and Dynamic Scenes*), I describe the results of an eye-tracking experiment designed to investigate the development of attention to faces in children (6-7-year-olds), adolescents (12-13-year-olds), and adults (18-26-year-olds). Specifically, I examined whether children directed their attention toward faces in the same way as adolescents and adults when passively viewing complex static and dynamic visual scenes involving a single person or multiple people. In Chapter 3 (*Similarities in Selective Attention to Facial Expression and Identity across Development*), I describe a study that I designed to investigate how we extract information from a face during an active task requiring identity or expression judgments.

Specifically, I examined the ability of children (6-7-year-olds), adolescents (12-13-year-olds) and adults (18-21-year-olds) to selectively attend to particular cues in static and dynamic faces, using a Garner's selective attention task. Data obtained from this paradigm can be used to describe the interactions between the processing of facial expression and facial identity. When identity interferes with the processing of expression and vice versa, it has been suggested that there is cross-talk between the brain regions involved in processing these two facial cues. My aim was to determine if interference effects would be similar across ages, and with both static and dynamic faces. There was substantial overlap in the participants who took part in the studies described in Chapters 2 and 3.

Even in adult viewers, the magnitude of any observed interference effects may depend on individual differences in how faces are processed. In Chapter 4 (*A Sex Difference in Interference between Identity and Expression Judgments with Static but not Dynamic Faces*), I followed up on the study described in Chapter 3 by examining whether, in adults, the sex of the viewer influenced the magnitude, or the pattern, of interference effects seen during static and dynamic face processing. Data from the 27 adults who participated in the study described in Chapter 3 were combined with data from an additional 13 adults recruited specifically for this study, in order to obtain a large enough sample to allow me to examine sex differences. Finally, in Chapter 5 (*Perceptual Biases Influence Inversion Effects and Interference Between Expression and Identity with Static but not Dynamic Faces*), I describe the results of a study that I designed to explore how individual differences in participants' perceptual processing biases affect interference effects seen during static and dynamic face processing, and whether this changes after face inversion.

The order of the Chapters in this thesis does not follow the order that the manuscripts were prepared for publication. The order of preparation (and publication) was Chapters 4, 2, 5, and 3. They were reordered for this thesis in order to convey a coherent narrative. Also note that data from the familiarization (identity-matching) task described in Chapters 3 and 4 were analyzed in Chapter 4 but not in Chapter 3. In both chapters, this familiarization task was included primarily to allow participants to become acquainted with the expressions and identities that they would see in the Garner's selective attention task that followed. This procedure was followed because interference from task-irrelevant expression cues may be more evident during identity judgments when the viewer is familiar with the faces (Ganel & Goshen-Gottstein, 2004). However, when analyzing the data described in Chapter 5 (a study that did not include a familiarization task), I observed interference effects in the static Identity task. Thus, the familiarization task was unnecessary to produce the Garner interference effects (at least with static stimuli). For this reason, I decided to omit the results from the familiarization task from Chapter 3.

When designing the studies described in this thesis, my intention was to create a normative database that could be used in future studies involving various clinical groups, including individuals born very prematurely (< 32 weeks gestation). As such, when the data described in Chapters 2-4 were collected, the typically developing participants described in those chapters also completed several other tasks not described in this thesis, which I mention briefly here. Adult participants and parents/legal guardians completed a questionnaire designed to obtain relevant demographic (e.g., parental education) and developmental information. Participants also completed short visual and intellectual screening tests and two additional experiments. In one of these experiments, participants watched visible speech movements that were either congruent or incongruent with audible speech sounds that were presented simultaneously. Their task on each trial was to indicate what they had actually heard. This paradigm, modelled after work by McGurk and MacDonald (1976), allowed me to explore intersensory integration. In another experiment, participants watched static and dynamic faces and made judgments regarding perceived gaze-direction. All of the testing took place in a single experimental session that lasted approximately 1.5 hours (see Table 1). Data for the study described in Chapter 5 were collected separately, in a completely independent sample. The study protocols for the work described in Chapters 2-4 and Chapter 5 were approved by the Psychology/Sociology Research Ethics Board at the University of Manitoba (see Appendix B) and a sample consent form can be viewed in Appendix C.

Table 1.1
Experimental Measures for Participants in Chapters 2, 3, and 4

	Number of Trials	**Administration Time**
General Demographic Questionnaire		5 min
Intellectual Screening Measure		
Peabody Picture Vocabulary Test-Third Edition (PPVT-III)	variable	15-20 min
Visual Screening Measures		
Near-point acuity chart	variable	5 min
Worth 4-dot test	1	1 min
Stereotest	13	5 min
Experimental Measures (presented in pseudorandom order)		
Reaction time task	20	2 min
Familiarization (Identity-matching) Task	64	5 min
Garner’s selective attention task	320	15 min
McGurk task	81	15 min
Gaze-cueing task	84	10 min
Eye-Tracking Task	24	3 min
Total Session Duration (including consent, breaks, and debriefing)		~1.5 hrs

CHAPTER 2: DEVELOPMENTAL DIFFERENCES IN ATTENTION TO FACES AND BODIES IN STATIC AND DYNAMIC SCENES

Running head: ATTENTION TO FACES AND BODIES

Keywords: attention, cognitive load, development, dynamic faces, dynamic scenes, eye-tracking, fixations, motion, social attention

Except for some minor wording changes, Chapter 2 is reprinted here with permission as it appears in:

Stoesz, B. M., & Jakobson, L. S. (2014). Developmental changes in attention to faces and bodies in static and dynamic scenes. *Frontiers in Perception Science*, *5*, 1–9. doi:10.3389/fpsyg.2014.00193

My Contribution to the Publication

The topic for this paper was originally part of my thesis proposal. Given the limited research on the development of dynamic face processing, I designed this eye-tracking study to examine age-related differences in how children (6-7-year-olds), adolescents (12-13-year-olds), and adults (18-26-year-olds) passively view faces in scenes. Specifically, I examined whether children directed their attention to faces in the same way older participants did when viewing complex static and dynamic visual scenes involving a single person or multiple people. I conducted the literature review, designed the experiment (including creating the stimuli), and prepared the ethics submission. I recruited participants and collected the data. I determined the methods used for data analysis in consultation with Dr. Jakobson and conducted the analyses for this study. I then drafted and revised the manuscript based on feedback from my advisor. I selected the journal, *Frontiers in Perception Science*, to which Dr. Jakobson and I submitted the manuscript. Based on the reviewers' comments that we received, I revised the manuscript in consultation with Dr. Jakobson and resubmitted the manuscript along with the responses to the reviewers. When the manuscript was accepted I made any further required corrections.

Abstract

Typically developing individuals show a strong visual preference for faces and face-like stimuli; however, this may come at the expense of attending to bodies or to other aspects of a scene. The primary goal of the present study was to provide additional insight into the development of attentional mechanisms that underlie perception of real people in naturalistic scenes. I examined the looking behaviours of typical children, adolescents, and young adults as they viewed static and dynamic scenes depicting one or more people. Overall, participants showed a bias to attend to faces more than other parts of the scenes. Adding motion cues led to a reduction in the number, but an increase in the average duration of face fixations in single-character scenes. When multiple characters appeared in a scene, motion-related effects were attenuated and participants shifted their gaze from faces to bodies, or made off-screen glances. Children showed the largest effects related to the introduction of motion cues or additional characters, suggesting that they find dynamic faces difficult to process, and are especially prone to look away from faces when viewing complex social scenes – a strategy that could reduce the cognitive and the affective load imposed by having to divide one's attention between multiple faces. These findings provide new insights into the typical development of social attention during natural scene viewing, and lay the foundation for future work examining gaze behaviours in typical and atypical development.

Introduction

Typically developing individuals show a strong visual preference for faces and face-like stimuli (Downing et al., 2004; Langton et al., 2008; Nummenmaa et al., 2006; Valenza, Simion, Cassia, & Umilta, 1996). This preference is present within several hours of birth (Nelson, 2001), and throughout childhood (Elam et al., 2010), adolescence (Freeth et al., 2010), and adulthood (Bayliss & Tipper, 2005; Hershler & Hochstein, 2005). A tendency to attend to faces at the expense of attending to other objects is particularly evident when facial expressions are ambiguous, or when the stimuli are more realistic (Land & Hayhoe, 2001) and social (Foulsham, Walker, & Kingstone, 2011). This makes sense, as faces are a rich source of information that can help us to respond appropriately during social interactions (Domes, Steiner, Porges, & Heinrichs, 2013).

Studies exploring developmental changes in our attention to faces have shown that young infants look longer at static than at dynamic faces. Indeed, infants up to four months of age have been shown to fixate on the static faces of a toy monkey (Brazelton, Koslowski, & Main, 1974), a manikin (Carpenter, Tecce, Stechler, & Friedman, 1970), and a doll (Field, 1979; Legerstee, Pomerleau, Malcuit, & Feider, 1987) for longer periods than the dynamic faces of their own mothers. Looking away from the mother does not appear to reflect passive disinterest. Rather, when they look away, infants show expressions indicative of concentration, as if they were engaging in *time-outs* from the previous looking period (Field, 1979). Taking these time-outs may reduce infants' cognitive load by providing them with more time to process the rich information conveyed by moving faces (Doherty-Sneddon, Bruce, Bonner, Longbotham, & Doyle, 2002; Glenberg, Schroeder, & Robertson, 1998). This would be beneficial as infants are naïve perceivers of the world, for whom the processing of most stimuli is challenging and effortful (Bahrick et al., 2002).

Infants may reduce their cognitive load by shifting their attention from a moving face towards a moving body. Evidence in support of this idea comes from work showing that 5-month-olds can discriminate and remember repetitive actions (i.e., blowing bubbles, brushing hair, and brushing teeth) better than the faces of the people performing those actions (Bahrick et al., 2002). Like faces, bodies provide important social information but, because the movements typically occur at a grosser level, they may be less challenging for infants to process.

With increasing age, infants' periods of looking away from moving faces become shorter. For example, between 3 and 9 months of age, infants increase the amount of time they spend looking at the faces of talking cartoon characters depicted in complex dynamic scenes (Frank et al., 2009). This may reflect their growing understanding that faces are a significant source of social information (Frank et al., 2009; Frank, Vul, & Saxe, 2012), but it may also reflect the fact that they are becoming increasingly proficient at processing dynamic cues (e.g., Braddick, Atkinson, & Wattam-Bell, 2003; Wattam-Bell, 1996), and increasingly sensitive to intersensory redundancy (e.g., the match between speech sounds and moving mouths) (Bahrick & Lickliter, 2000). The fact that infants' attention to faces becomes especially marked when they are listening to a speaker (Smith & Mital, 2013; Tenenbaum, Shah, Sobel, Malle, & Morgan, 2013) supports the view that they use visual cues (lip movements) to facilitate speech perception (e.g., Bristow et al., 2009), although their ability to integrate visual and auditory speech cues is not as strong as that of adults (Desjardins & Werker, 2004).

Several studies have examined children's attention to faces as they listen and respond to questions posed by adults (Doherty-Sneddon et al., 2002; Doherty-Sneddon & Kent, 1996; Doherty-Sneddon & Phelps, 2005). These studies suggest that, by eight years of age, children (like adults; Glenberg et al., 1998) use gaze aversion to help them manage their cognitive load. Specifically, as the difficulty of the questions being posed increases, children show an increasing tendency to look away from the speaker both when the question is being posed, and when they are formulating and articulating their responses. This behaviour is evident whether children are engaged in face-to-face interactions or are viewing a speaker via video-link (Doherty-Sneddon et al., 2002; Doherty-Sneddon & Phelps, 2005), and suggests that processing the moving face of the speaker requires cognitive resources. Children’s tendency to engage in gaze aversion when being spoken to may explain why they show a significantly smaller McGurk effect (McGurk & MacDonald, 1976) than adolescents or young adults (Desjardins, Rogers, & Werker, 1997; Tremblay et al., 2007). The McGurk effect is an audiovisual illusion that occurs when an individual is presented with mismatched visual and auditory phonemes (e.g., ba and ga), but reports perceiving a third phoneme (e.g., da). Young children are less likely than older participants to experience the illusion, reporting instead the auditory phoneme that was presented (Tremblay et al., 2007) – a result that suggests young children are not attending closely to dynamic facial cues.

To the best of my knowledge, there have been no studies examining children’s gaze behaviours during *passive* viewing of naturalistic scenes (see Karatekin, 2007 for a review). This is unfortunate because adding task demands can lead to gaze behaviours that are quite different from those seen under passive viewing conditions (Smith & Mital, 2013), and age-related differences in task performance may obscure or alter age-related changes in deployment of attention (Scherf, Behrmann, Humphreys, & Luna, 2007). While studies examining passive viewing in children are lacking, some research involving *typical adolescents* suggests that they fixate significantly longer on faces than on bodies or objects while viewing movie clips of social interactions (Klin, Jones, Schultz, Volkmar, & Cohen, 2002), and while viewing static *and* dynamic scenes depicting single or multiple characters (Speer, Cook, McMahon, & Clark, 2007). These gaze behaviours differ from those made by adolescents with autism, who fixate longer on objects than on either faces or bodies (Klin et al., 2002), and who make shorter fixations on eye regions and longer fixations on bodies than typically-developing peers, particularly when viewing dynamic, multiple-character displays (Speer et al., 2007). Together, these results suggest that, whereas typical adolescents direct their attention toward moving faces during passive viewing of scenes, those with autism look away from faces -- perhaps in an effort to reduce their cognitive load.

Recently, a number of authors have examined the question of how adults control their attention to faces during various tasks. Although they do make more fixations on faces than on bodies, adults' person detection is improved when the whole person (i.e., face and body) is visible in a scene (Bindemann, Scheepers, Ferguson, & Burton, 2010). A similar effect has been reported for person identification, especially when the stimuli are moving – a result that supports the view that face *and* body movements are useful during the identification process (O'Toole et al., 2011; Pilz, Vuong, Bülthoff, & Thornton, 2011). Together, these findings suggest that, when the body is visible, introduction of dynamic cues may encourage adults to shift some of their attention from the face toward the body (O'Toole et al., 2011). Additional support for this idea comes from the finding that adults' analysis of facial expressions is affected by the presence of emotional body language (Hietanen & Leppänen, 2008), even when task demands encourage attention to be directed toward the faces (Meeren, van Heijnsbergen, & de Gelder, 2005).

The current study was designed to fill a gap in the literature by exploring how our attention to faces changes as a function of age. Specifically, I asked whether introducing dynamic cues or changing the number of people in a scene would have different effects on passive viewing behaviours, depending on the viewer's age. This question is of interest given that children's cognitive resources and processing efficiency are reduced compared to adults (e.g., Hale, 1990; Miller & Vernon, 1997); as such, I expected that my scene manipulations would place greater cognitive demands on younger viewers.

Face processing abilities, such as identity extraction, improve dramatically between 4 and 11 years of age (Carey & Diamond, 1977; Ellis & Flin, 1990; Ge et al., 2008; Johnston & Ellis, 1995; Mondloch et al., 2003). For this reason, I chose to compare the gaze behaviours of children whose ages were near the middle of this range (6-8 year-olds) to those of adolescents (12-14 year-olds) and young adults. I analyzed the average number and duration of fixations made in particular areas of interest (AOI: faces, bodies, background) as participants passively viewed naturalistic scenes. These variables were of interest as past research suggests that reductions in the number of fixations and increases in average fixation length reflect increasing processing demands (Henderson, 2003; Smith & Mital, 2013) and/or reduced processing efficiency (Açık, Sarwary, Schultze-Kraft, Onat, & König, 2010). I also measured the total time that viewers devoted to examining each AOI or glancing off-screen in each trial (dwell time). Dwell time algorithms combine time spent executing saccades and fixating within an AOI (Salvucci & Goldberg, 2000), and dwell time has been examined in other research to assess viewers' preferences for and attention to faces (e.g., Matsuda, Okanoya, & Myowa-Yamakoshi, 2013). I expected that children would find it more challenging than adults to process moving faces and multiple-character scenes, and thus be more likely to shift their attention away from faces in these conditions in an effort to reduce their cognitive load. Adolescents were expected to perform at near-adult levels. By breaking up the scenes into different AOIs, I was also able to determine if children were more likely than adults to redirect their attention from faces toward bodies, objects, or off-screen.

Method

Participants. Eighty-eight individuals participated in this study. I tested 32 children aged 6.0-8.0 years (*M* = 6.7, *SD* = .6; 13 boys, 19 girls), 26 adolescents aged 12.1-13.8 years (*M* = 12.8, *SD* = .6; 12 boys, 14 girls), and 30 young adults aged 18.1-26.8 years (*M* = 20.1, *SD* = 2.0; 17 men, 13 women). Children and adolescents were recruited via word-of-mouth and via local schools from Winnipeg and Altona, Canada. Young adults were recruited through the psychology participant pool at the University of Manitoba, Winnipeg, Canada. All participants were native English speakers and had normal or corrected-to-normal visual acuity.

Materials. The 24 stimuli in the eye-tracking experiment consisted of clips from several episodes of a television series (the *Andy Griffith Show*) that originally aired on CBS from 1960-1968. As outlined below, scenes were carefully chosen to meet certain criteria. First, the situations depicted were "realistic" in the sense that they were ones that individuals might experience in everyday life, and they took place in recognizable settings, such as a grocery store, a workplace, or on the street. In addition, scenes not only contained one or more people, but objects that one might naturally find in such situations (e.g., groceries, telephone, or park bench). I extracted twelve 4-s video clips. Six clips depicted a single character conversing with an off-camera character, and six clips depicted two or more characters engaged in a social interaction. All interactions were emotionally neutral. In all scenes, at least the upper half of characters' bodies were visible, to allow us to determine if viewers' attention was being drawn from a character's face toward his/her body, toward objects in the background, or off-screen. In all dynamic scenes, the primary motion cues came from nonrigid movements of the face and/or body of the character(s), the character(s) did not move into or out of the field of view, and the objects in the background were generally stationary. To create the static displays, I extracted one static image from each movie clip; as such, each static image depicted the same character(s) and objects present in the corresponding dynamic display. Thus, this experiment consisted of four conditions: (1) single-character-static, (2) multiple-characters-static, (3) single-character-dynamic, and (4) multiple-characters-dynamic, with each condition consisting of six trials. Stimulus size was standardized at 640 pixels (23.8 degrees of visual angle) wide and 480 pixels (18.0 degrees of visual angle) high. Photographs had a resolution of 72 pixels per inch and the video was shown at 29 frames per second (fps). No soundtrack accompanied the stimuli.

Procedure. The study protocol was approved by the Psychology/Sociology Research Ethics Board at the University of Manitoba. Adult participants and parents of each child/adolescent who participated in the study provided written informed consent. Children and adolescents also confirmed their assent. Participants were tested individually. Each participant was seated approximately 60 cm from the 17-inch computer screen of a Tobii 1750 binocular corneal-reflection eye-tracking system (0.5 degree precision, 50 Hz sample rate, 5 fps per second, 1280 x 1024 pixels resolution; Tobii Technology Inc., Fall Church, VA). Because this particular eye-tracking system compensates for large and rapid head movements, participants sat entirely unrestrained (i.e., did not wear helmets, chin-rests, or markers). Tobii Studio Enterprise experimental software controlled the stimulus presentation.

Before the experiment began, I carried out a short (approximately 15 s) 9-point calibration routine using the eye-tracker. Participants tracked a white dot moving on a black background. The dot moved slowly and randomly to nine locations on the screen. At each location, the dot appeared to grow and then shrink in size before moving to the next location.

Upon completion of the calibration trial, Tobii Studio Enterprise experimental software gave immediate feedback regarding the quality of the calibration. The calibration routine was repeated if the quality was poor initially. Participants then engaged in free-viewing task consisting of 24 trials. Each of the 24 trials consisted of a 2-s central white fixation point presented on a black background, followed by the presentation of the 4-s stimulus (see Figure 2.1). Trials were presented in a different random order for each participant. The experiment took approximately 2.4 min to complete. Participants were instructed to look at the fixation cross at the beginning of each trial, and then to passively view each of the 12 photographs and 12 movies that would be presented one at a time. See Appendix D for instructions given to participants.

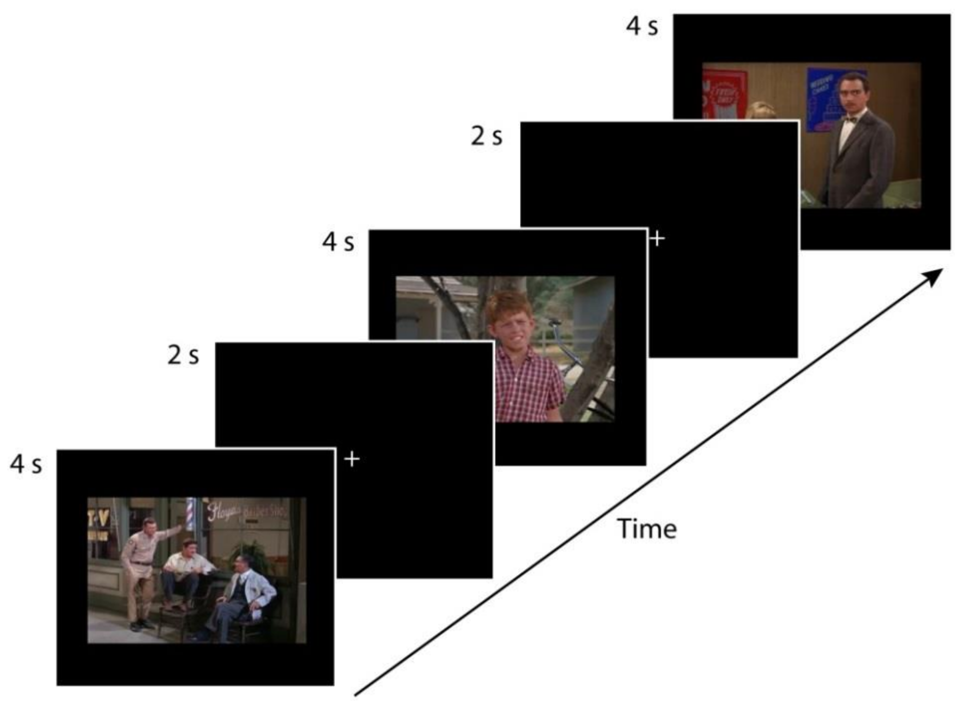

Figure 2.1. Each trial in the eye-tracking experiment consisted of a 2-s central white fixation point on a black background followed by the presentation of the 4-s stimulus. The figure depicts two trials from the static, multiple-people condition and one trial from the static, isolated-person condition. Trials were presented consecutively and in random order.

Analyses

Three areas of interest (AOI) were investigated: the face (or faces), the body (or bodies), and the background. Because the characters did not move rigidly across the screen in the dynamic scene, a complete frame-by-frame analysis was unnecessary. This also made it possible to make AOIs of identical sizes in static and dynamic displays. Due to differences in camera viewing angle, the area of individual face AOIs were smaller in multiple- compared to single-character scenes. I ensured, however, that the total (combined) area of all visible face AOIs did not differ by scene type [single character: M = 3.91% of the scene, SD = 2.06; multiple-character: M = 2.68 % of the scene, SD = 1.07; $t(10) = 1.31$, $p = .22$]. The total area devoted to body AOIs was also comparable in both types of scenes [single-character: M = 18.04% of the scene, SD = 6.88; multiple-character: M = 23.58% of the scene, SD =5.47; $t(10) = 1.54$, $p = .15$], as was the total area of the background [single-character: M = 78.04% of the scene, SD = 8.69; multiple-character: M = 73.75% of the scene, SD =6.12, $t(10) = 0.99$, $p = .35$]. Regardless of modality (static/dynamic) or scene type, the face AOI was smaller than the body AOI, which was smaller than the background AOI [$t(5) > 6.57$, $p < .001$, in all cases].

Using Tobii Studio Enterprise software, I extracted a series of measures of gaze behaviour within each AOI during each trial. The first was the number of fixations made within the AOI. Fixations were defined as any period where gaze stayed within a 30 pixel (0.9 degree of visual angle) diameter area for 200 ms or more. The second measure of gaze behaviour was mean fixation duration. The third measure was dwell time, which refers to the total time from the onset of the first fixation inside an AOI to onset of the first fixation outside the AOI. For the dwell time variable, I also calculated the amount of time that participants did not look at the screen by subtracting the total dwell time within the pre-defined AOIs from the total time each stimulus was on the screen (4 s). Finally, I computed the average for each variable across the six scenes within a condition.

The mean number of fixations and the mean fixation durations were entered into two separate 3 (Age Group: children, adolescents, young adults) x 2 (Scene Type: single-character, multiple-character) x 2 (Presentation Mode: static, dynamic) x 3 (AOI: faces, bodies, background) analysis of variance tests (ANOVAs), with repeated measures on the last three factors. Dwell time data were entered into a 3 (Age Group: children, adolescents, young adults) x 2 (Scene Type: single-character, multiple-character) x 2 (Presentation Mode: static, dynamic) x 4 (AOI: faces, bodies, background, off-screen) ANOVAs, with repeated measures on the last three factors. Variance assumptions for all comparisons were tested with Levene's test of equality of variances. Where violations of sphericity were observed, within-group effects were reported with Greenhouse-Geisser corrections. Follow-up multiple comparison tests on significant interactions were completed using Fisher's LSD tests. I analyzed the data using SPSS 22 (SPSS Inc., Chicago, IL, USA). Note that, before running the ANOVAs, I confirmed that age was not related to scores on any of the dependent variables in the sample of young adults. This step was deemed necessary because the age range in the adult group was larger than the age ranges in the other groups.

Results

Number of fixations. Overall, participants made more fixations within face AOIs, and fewer within body AOIs, than in the background [$F(1, 170) = 36.41$, $p < .001$, $\eta_p^2 = .30$]. Participants also made fewer fixations when viewing dynamic than static scenes [$F(1, 85) = 114.38$, $p < .001$, $\eta_p^2 = .57$], but this effect was: (a) larger in children than in adults [Presentation Mode x Age Group: $F(2, 85) = 3.13$, $p = .049$, $\eta_p^2 = .07$]; and (b) most pronounced for fixations occurring in the background [Presentation Mode x AOI: $F(2, 170) = 16.62$, $p < .001$, $\eta_p^2 = .16$]. In addition, I observed a significant three-way interaction between Presentation Mode, AOI, and Age Group [$F(4, 170) = 3.30$, $p = .02$, $\eta_p^2 = .07$; see Figure 2.2], and follow-up tests on this interaction revealed important age-related differences in the effect that adding dynamic cues had on the number of fixations made in face AOIs, specifically. On average, participants in all three groups made a similar number of fixations on faces during static trials but, as predicted, children and adolescents showed a significant drop in the number of fixations made in this AOI with the addition of dynamic cues [$t > 3.44$, $p < .003$, $d = .67$], whereas adults did not. This resulted in children making significantly fewer fixations on dynamic faces than adults [$t(60) = 2.80$, $p = .007$, $d = .71$].

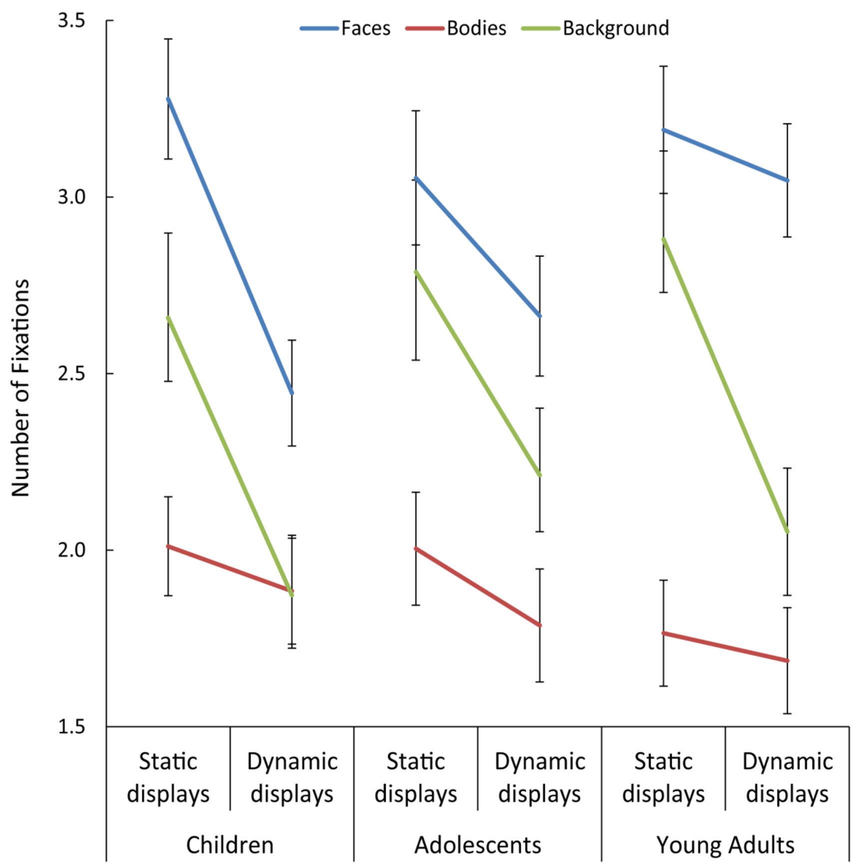

Figure 2.2. The number of fixations made by children, adolescents, and young adults within each area of interest (AOI: faces, bodies, and backgrounds) for the static and dynamic scenes.

In addition to the above, participants made more fixations when viewing multiple- compared to single-character scenes [$F(1, 85) = 317.431$, $p < .001$, $\eta_p^2 = .79$]. Although this Scene Type effect was smaller in face AOIs than in other regions [Scene Type x AOI: $F(2, 170) = 32.12$, $p < .001$, $\eta_p^2 = .27$], the impact of changing scene type on the number of fixations made on *faces* varied as a function of age [Scene Type x AOI x Age Group: $F(4, 170) = 2.44$, $p = .049$, $\eta_p^2 = .05$]. Specifically, as seen in Figure 2.3, adults increased the number of fixations they made on faces when additional characters were added to a scene [$t(29) = 2.86$, $p = .008$, $d = .52$], but children and adolescents did not.

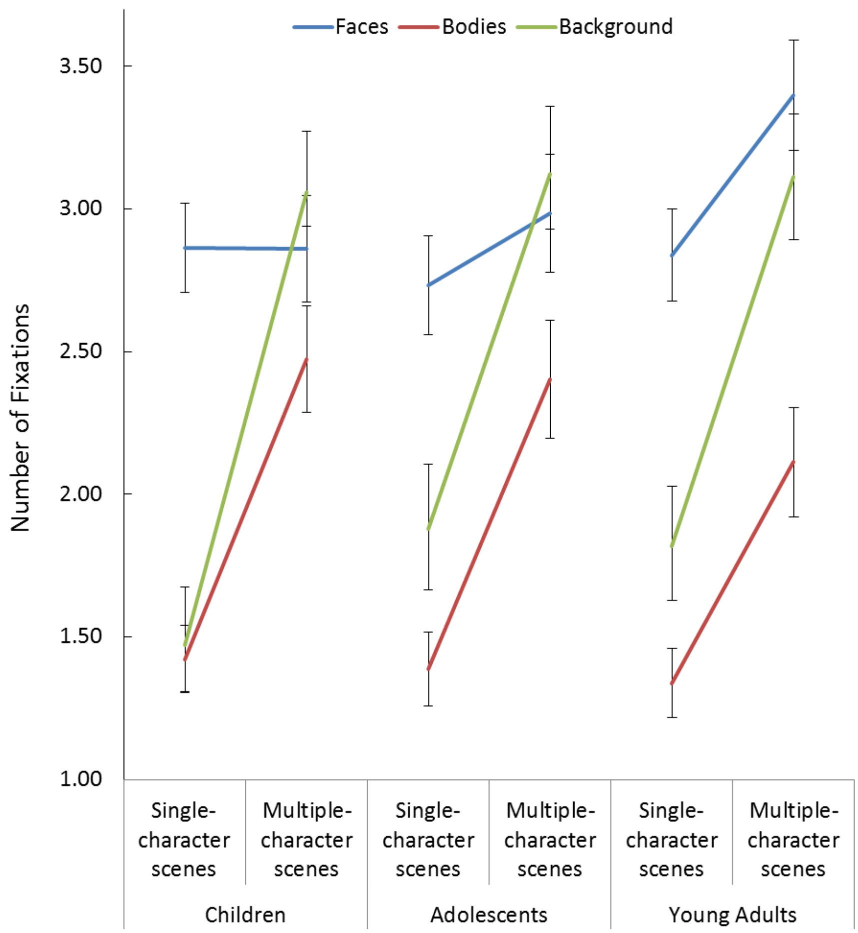

Figure 2.3. The number of fixations made by children, adolescents, and young adults within each area of interest (AOI: faces, bodies, backgrounds) for single- and multiple-character scenes.

Fixation duration. Overall, mean fixation duration was longer during viewing of dynamic compared to static scenes [$F(1, 85) = 40.15$, $p < .001$, $\eta_p^2 = .32$], and shorter during viewing of multiple- compared to single-character scenes [$F(2, 85) = 83.04$, $p < .001$, $\eta_p^2 = .49$]. Fixations made in face AOIs were also generally longer than those made in other regions [$F(2, 170) = 187.26$, $p < .001$, $\eta_p^2 = .69$]. I observed Scene Type x Presentation Mode [$F(1, 85) = 4.44$, $p = .04$, $\eta_p^2 = .05$], Presentation Mode x AOI [$F(2, 170) = 10.03$, $p < .001$, $\eta_p^2 = .11$], and Scene Type x AOI [$F(2, 170) = 56.31$, $p < .001$, $\eta_p^2 = .40$] interactions, but each of these interactions needed to be interpreted in light of a significant 3-way interaction involving Scene Type, Presentation Mode, and AOI [$F(2, 170) = 3.30$, $p = .02$, $\eta_p^2 = .07$] (see Figure 2.4). Follow-up tests performed on the interactions revealed two key findings. First, mean fixation duration increased with the introduction of dynamic cues across AOIs [$t(87) > 2.40$, $p < .02$, $d > .25$, in each case], however, this effect was largest for fixations made within face AOIs in single-character scenes. Second, the drop in mean fixation length seen with the introduction of additional characters was *only* evident in face AOIs [$t(87) = 8.23$, $p < .001$, $d = .88$].

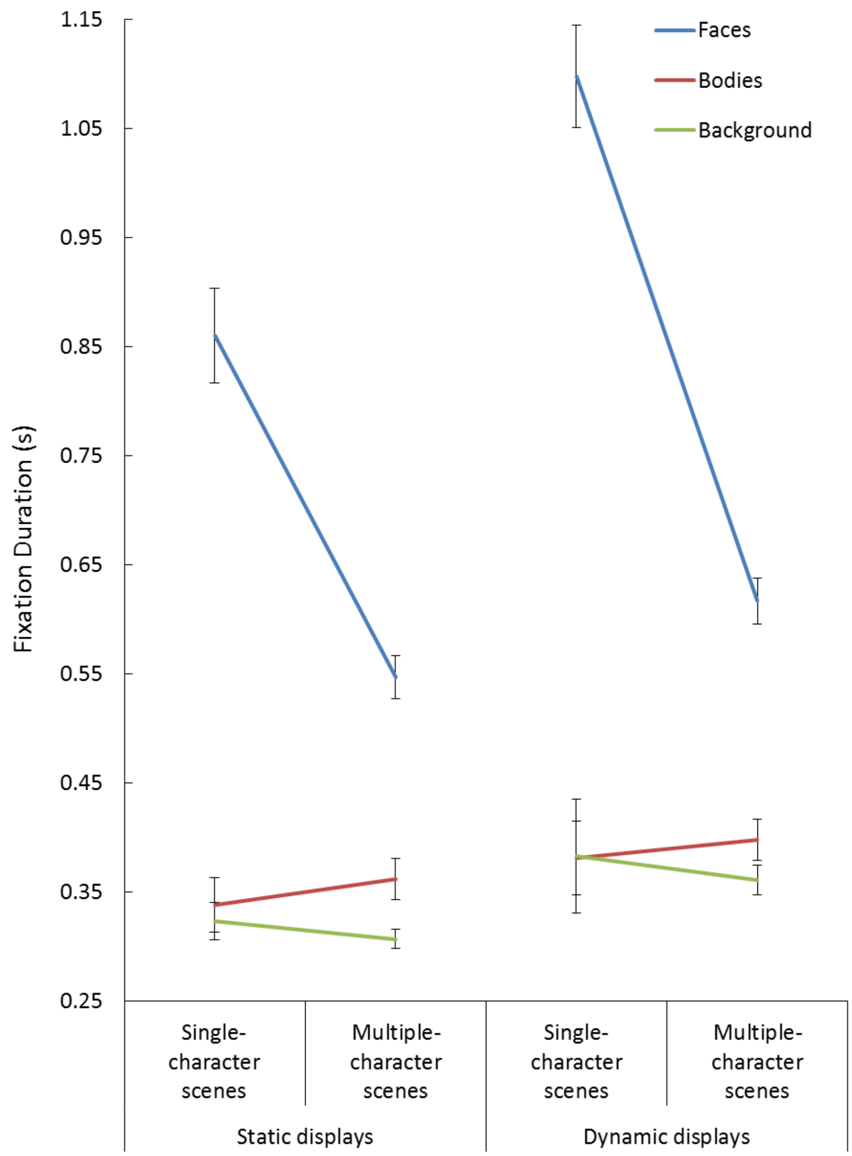

Figure 2.4. The mean fixation duration(s) for each area of interest (AOI: faces, bodies, backgrounds) while participants viewed static and dynamic displays in single- and multiple-character scenes.

Additionally, results indicated a significant Presentation Mode x Age Group interaction [F(1, 85) = 3.07, p = .05, η_p^2 = .07] (see Figure 2.5). As predicted, only children's mean fixation duration increased significantly with the addition of dynamic cues [t(60) = 2.09, p = .04, d = .53].

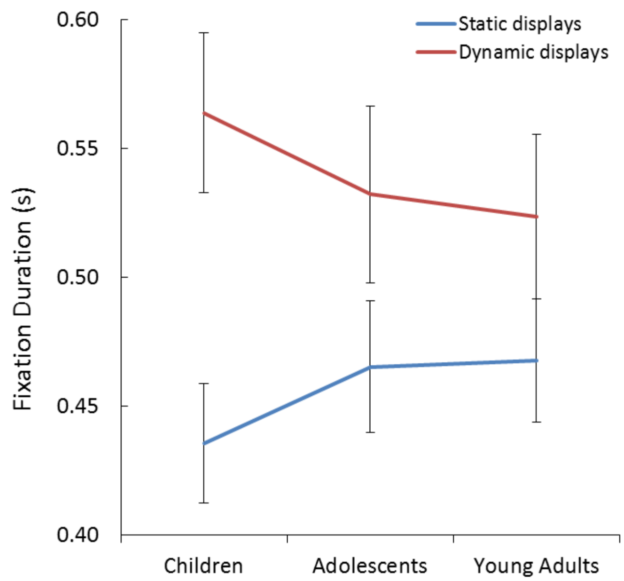

Figure 2.5. Mean fixation duration(s) while participants viewed static and dynamic displays.

Dwell time. In general, viewers spent more time looking at faces, and less time looking at backgrounds, than they did looking at bodies or off-screen [main effect of AOI: $F(3, 255) = 198.32$, $p < .001$, $\eta_p^2 = .70$]. This main effect varied depending on the number of characters in the scene [Scene Type x AOI: $F(3, 255) = 84.13$, $p < .001$, $\eta_p^2 = .50$]. Specifically, although the effect of AOI was present in both single- and multiple-character scenes [$t(87) > 3.62$, $p < .001$, $d > .70$, for all comparisons], adding more characters to a scene triggered participants to look less at faces [$t(87) = 11.91$, $p < .001$, $d = 1.27$] and more at bodies or off-screen [$t(87) > 6.75$, $p < .001$, $d = .72$, in both cases]. The Scene Type x AOI interaction was amplified when dynamic cues were added [Scene Type x AOI x Presentation Mode: $F(3, 255) = 9.06$, $p < .001$, $\eta_p^2 = .10$; see Figure 2.6]. This was primarily due to the finding that viewers were more drawn to examine moving than static faces in single- than in multiple-character scenes [$t(87) = 5.03$, $p < .001$, $d = .54$]. The Scene Type x AOI interaction also varied as a function of viewers' age [Scene Type x AOI x Age Group: $F(6, 255) = 2.25$, $p = .04$, $\eta_p^2 = .05$, see Figure 2.7]. Specifically, although adding more characters to the scene triggered all participants to shift their attention from faces to bodies or off-screen, these effects were more dramatic in children than in adults [$t(60) > 2.52$, $p < .02$, $d > .63$ for all comparisons].

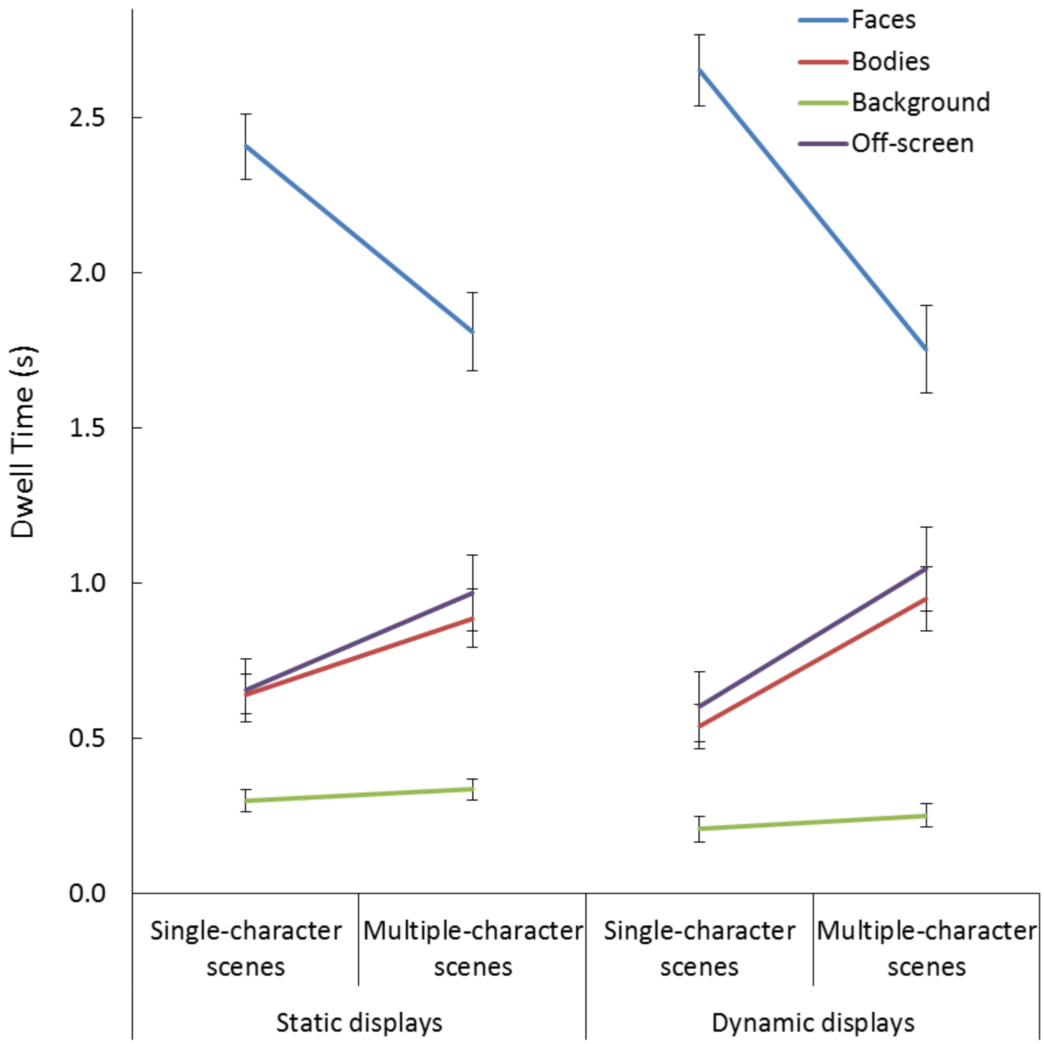

Figure 2.6. The mean dwell time (s) in each area of interest (AOI: faces, bodies, backgrounds) and off-screen while participants viewed static and dynamic, single- and multiple-character scenes. The Scene Type x AOI interaction seen with static scenes was amplified with the addition of dynamic cues. This was primarily due to the fact that viewers were more drawn to examine moving faces in single-than in multiple-character scenes.

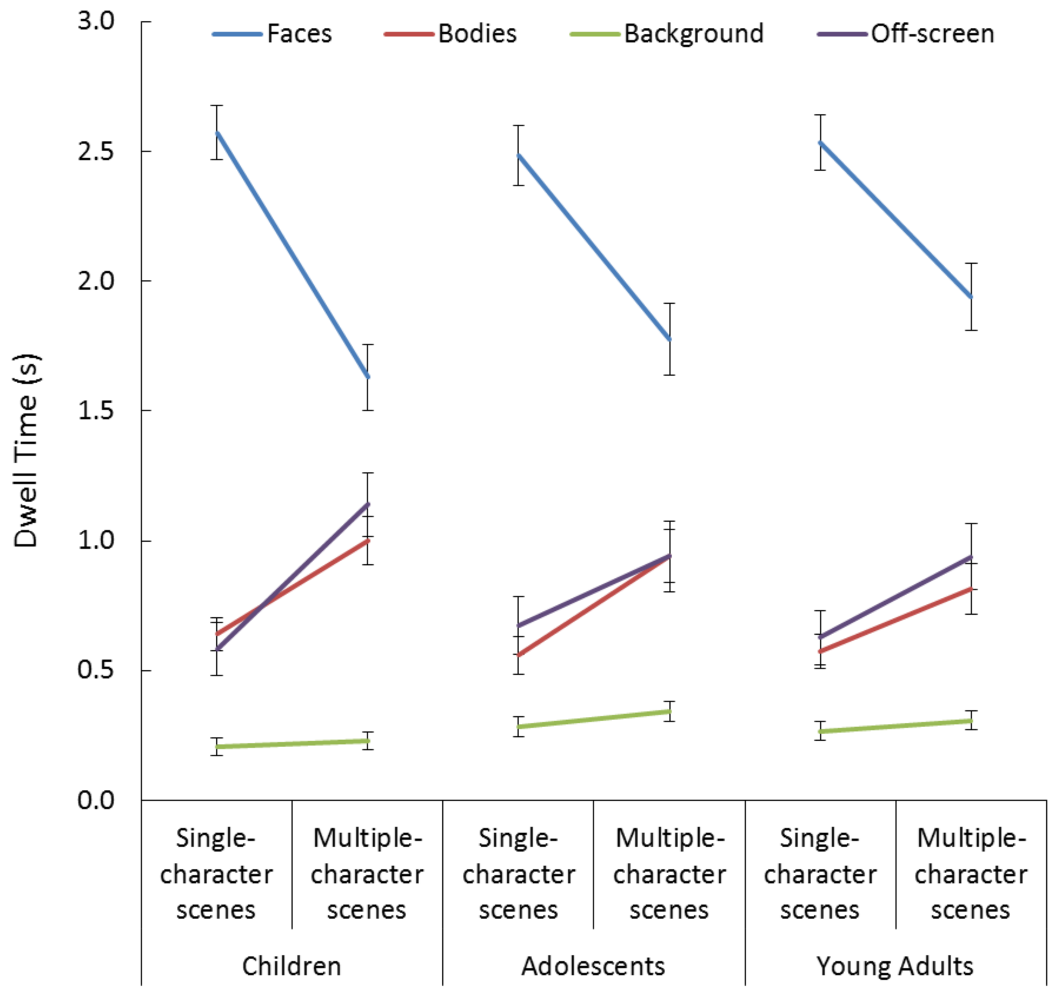

Figure 2.7. Age-related differences in mean dwell time (s) in each area of interest (AOI; faces, bodies, backgrounds) and off-screen, for single- and multiple-character scenes. Although adding more characters to the scene triggered all participants to shift their attention from faces to bodies or off screen, these effects were larger in children than in adults.

Discussion

The goal of the present study was to extend research on developmental differences in attention to faces by comparing the gaze behaviours of children, adolescents, and young adults as they viewed naturalistic scenes. I examined whether passive viewing behaviours in each age group would be affected by the introduction of motion and/or additional characters in scenes. I expected that each of these manipulations would make it more challenging for children, in particular, to attend to or process faces and that this would lead them to try to reduce their cognitive load by engaging in more *looking away* behaviour. In general, the results from the analyses of the eye-tracking data support these hypotheses. I discuss the findings below.

Despite the fact that the face AOIs were considerably smaller than any other regions, participants made more and longer fixations on faces than on other parts of the displays, which resulted in longer dwell times for faces. These results are consistent with recent eye-tracking studies (Bindemann et al., 2010; Birmingham & Kingstone, 2009; Rice, Phillips, Natu, An, & O'Toole, 2013) and other work showing that viewers of all ages are generally biased to attend to faces (e.g., Downing et al., 2004; Langton et al., 2008; Nummenmaa et al., 2006; Valenza, Simion, Cassia, & Umilta, 1996). Although viewers may have focused on faces because there was little or no movement occurring in the background to capture their attention, this would not explain why I found the face bias during viewing of static, as well as dynamic stimuli. A more likely explanation of the face bias is that faces automatically attract attention due to their high social significance (see Lavie, Ro, & Russell, 2003). As outlined below, however, factors such as the number of characters in a scene influence the way in which we divide our attention between faces and bodies. Additional insights into how we control our attention to faces when bodies are visible in a scene come from recent work on person detection (Bindemann et al., 2010) and person identification (O'Toole et al., 2011; Pilz et al., 2011).

Adding dynamic cues resulted in changes in participants' looking behaviours. In general, the addition of motion cues led to a reduction in the number of fixations and an increase in average fixation duration, but both of these effects were larger in children than in adults. As these effects are believed to reflect increasing processing demands (Henderson, 2003; Smith & Mital, 2013) and/or reduced processing efficiency (Açık et al., 2010), the present findings are consistent with the view that dynamic faces are more challenging for children than for adults to process. Children may also find dynamic faces more physiologically arousing, even when (as in the present study) the scenes are emotionally neutral. Interestingly, infants' arousal levels go down when their mothers slow down, simplify, or infantize their behaviours during interactions (Tronick, Als, Adamson, Wise, & Brazelton, 1978) – a result that supports the view that face processing is cognitively *and* affectively arousing for young viewers. In future work, it might be interesting to vary the *affective load* across scenes, and look for age-related differences in phasic changes in heart rate and respiration amplitude, and in passive gaze behaviours.

Another possible explanation for children making significantly fewer fixations on dynamic faces than adults is that this reflects age-related improvements in sensory and/or cognitive functions (Crookes & McKone, 2009). For instance, visual acuity (Skoczenski & Norcia, 2002), sustained attention (Betts et al., 2006), and the ability to narrow the focus of visual attention (Pastò & Burack, 1997) improve with age. Moreover, Betts et al. (2006) have reported rapid growth in sustained attention (for indices such as speed, errors, accuracy, and variability) occurring from 5 to 9 years of age, and a developmental plateau (with only minor improvements) occurring after the age of 10 years. Many studies examining the development of sustained attention have typically used tasks involving simple, static stimuli that represent restricted resemblance to the relevant features the real world (see Szalma, Schmidt, Teo, & Hancock, 2014). Thus, it is unclear whether improvements in sustained attention seen on these tasks would relate to gaze behaviours on dynamic faces. One way to test this idea would be to examine the association between age, gaze behaviours for dynamic faces, and sustained attention using tasks that incorporate dynamic stimuli, such as object tracking tasks (Fisher, Thiessen, Godwin, Kloos, & Dickerson, 2013) or tasks that are video-game like (Szalma et al., 2014).

Another explanation for the results is that children were less able than adults to use the direction of the actors' gaze to determine a plausible story line (i.e., to follow the plot of the silent scenes) (see Glenberg et al., 1998). It has been suggested that gaze cuing is reflexive or involuntary (Friesen, Ristic, & Kingstone, 2004; Hietanen, Nummenmaa, Nyman, Parkkola, & Hämäläinen, 2006), although this effect becomes larger with increasing age (Pruett et al., 2011) and may also be affected by the context (Dawel, McKone, Irons, O'Kearney, & Palermo, 2013; Noh & Isaacowitz, 2013). Thus, it may be that children in the present study relied less on gaze and more on the background to help them determine what was taking place in each scene.

As with the addition of motion cues, adding characters to a scene resulted in several changes in participants' gaze behaviour. First, adults (but not children or adolescents) made more fixations on faces when viewing multiple-character scenes. This was true despite the fact that, in order to match the total area of particular AOIs across scene types, individual faces were smaller in multiple- than in single-character scenes. Second, adding characters to a scene led viewers in all age groups to decrease the mean duration of face fixations. Together, these results may reflect a competitive push-pull interaction between two sources of social information (see Findlay & Walker, 1999). Specifically, when attending to multiple characters in a scene, a viewer's eyes may be pulled from one face to another, resulting in more frequent, but shorter fixations on faces. It is also possible that participants in the present study made shorter fixations on faces in multiple-character scenes simply because the individual faces were smaller, and, therefore, harder to resolve. In a related study, which involved static stimuli only, Birmingham, Bischof, and Kingstone (2008) found that their adult viewers made *longer* fixations on the eye region of characters' faces as the number of people in the scene increased. In this study, actors were photographed from a standard distance, which meant that the total AOI for eye or face regions in multiple-character scenes was much larger than the area of the corresponding AOI in single-character scenes. It is important to note, however, that exposure durations were also much longer in the Birmingham et al. study than in the present investigation (15 s vs. 4 s per trial). This may also have contributed to differences in the findings.

One strength of the present study is that I measured dwell times not just within particular AOIs, but also for off-screen glances. This proved to be important as these glances accounted for approximately 25% of total viewing times. Adding more characters to a scene resulted in viewers spending relatively less time attending to faces, and relatively more time attending to bodies or glancing off-screen. As one might expect if viewers found dynamic faces particularly difficult to process, these attentional shifts were especially evident when the characters were moving. In addition, shifts in attention from faces to bodies or off-screen were more pronounced in young children – supporting the view that children use gaze shifts like these to reduce their cognitive load. In future work, it would be interesting to study the effect that adding the soundtrack would have on viewers' gaze behaviours. This manipulation should increase the cognitive demands even further and, therefore, have a larger effect on children's than adults' gaze behaviours (see Doherty-Sneddon et al., 2002; Doherty-Sneddon & Kent, 1996; Doherty-Sneddon & Phelps, 2005).

Conducting studies with *information rich* displays that closely approximate naturalistic stimuli should be a priority for researchers interested in face processing, as much of the existing literature in this area has utilized static displays. Studies incorporating moving faces or whole bodies -- viewed in isolation or in the context of real-world scenes -- are providing new insights into how we process social information (O'Toole et al., 2011; Pilz et al., 2011). In other work, for example, I have used a Garner interference paradigm (Garner, 1976) to study how interference between the processing of facial identity and facial expression changes with the introduction of dynamic cues (see Chapters 2-5) (Stoesz & Jakobson, 2013). I replicated earlier findings of bidirectional interference between the processing of these cues with static faces (as in Ganel & Goshen-Gottstein, 2004), and then went on to show that interference dropped to negligible levels when moving faces were used as test stimuli – results that suggest that viewers may be better able to attend selectively to relevant facial cues when faces are moving than when they are static.

Like behavioural studies, most neuroimaging studies have investigated brain regions involved in face processing using static images, but this is beginning to change. Researchers have found that the visual processing of faces from static and dynamic displays involve different neural networks and/or different levels of activation of the same brain regions (Arsalidou et al., 2011; Fox, Moon, Iaria, & Barton, 2009; Kessler et al., 2011; Sato et al., 2010; Schultz & Pilz, 2009). Observations such as these lend weight to the suggestion that there is much to be gained from utilizing naturalistic, dynamic stimuli that are socially rich (see also Birmingham & Kingstone, 2009).

Exploring looking behaviours provides information on how components of our attentional system operate and what social interests we may have (Klin et al., 2002; Speer et al., 2007). Using eye-tracking technology to study eye movements and fixations has proven particularly useful for determining typical gaze behaviours in infants and adults, and contrasting these with gaze behaviours in various clinical groups. This study makes a unique contribution to the literature on social attention, and is one of the first to examine gaze behaviours in three different age groups of participants – children, adolescents, and adults – during passive scene perception. The results are significant in that they provide additional insights into age-related differences in the deployment of social attention in response to changing cognitive demands associated with the introduction of dynamic cues, or additional characters. This work also provides a foundation for future studies that involve children born prematurely at very low birth weight (< 1500 g). This group is known to be at risk for deficits in social perception and cognition. Williamson and Jakobson (2014) found, for example, that children born preterm show impairments in their ability to use nonverbal face and body cues to interpret the emotions of people engaged in naturalistic social situations. Incorporating eye-tracking in studies of this sort could help to determine if these deficits are associated with motion-processing problems and/or with gaze aversion or other atypical gaze behaviours. Knowing this may inform the development of interventions designed to improve social functioning in this at-risk population. Studies of this kind will also improve our understanding of the typical and atypical development of the social brain.

CHAPTER 3: SIMILARITIES IN SELECTIVE ATTENTION TO FACIAL EXPRESSION AND IDENTITY ACROSS DEVELOPMENT

Running head: ATTENTION TO FACIAL CUES ACROSS DEVELOPMENT

Keywords: development, dynamic faces, expression, identity, non-rigid motion, selective attention

My Contribution to the Manuscript

The topic for this paper was originally part of my thesis proposal. The study described in Chapter 2 examined how attention is directed to faces during passive viewing of scenes. In the present chapter, I explored how selectively viewers of different ages are able to attend to particular facial cues when actively making judgments about faces. To do this, I used Garner's selective attention task. Data obtained from this paradigm can be used to uncover interactions between the processing of facial cues. If the processing of one facial cue interferes with the processing of another, it is assumed that there is cross-talk between the brain regions involved in processing the two facial cues. The task was run using both static and dynamic faces, to allow me to see if the results were affected by the presence of dynamic cues. Many of the same children, adolescents, and adults who took part in the study described in Chapter 2 also completed this experiment, during a single testing session. I conducted the literature review, designed the experiment (including stimulus creation), and prepared the ethics submission. I recruited participants and collected the data. I determined the methods used for data analysis in consultation with Dr. Jakobson and conducted the analyses for this study. Dr. Jakobson and I wrote and submitted an abstract based on preliminary data analyses, and I created and presented these findings at the Vision Sciences Society (2013) conference. I then drafted and revised the manuscript based on feedback from my advisor. This manuscript is currently under review.

Abstract

The ability to extract facial expression and facial identity improves gradually, approaching adult levels in late childhood or early adolescence. The most common explanation given for children's poorer performance on these tasks is that it takes many years of visual experience to develop specialized *face-specific* perceptual processing mechanisms. In present study, I investigated the possibility that there are age-related differences in viewers' ability to attend selectively to facial expression and identity cues, using Garner's selective attention task (Garner, 1976). Children, adolescents, and adults made speeded judgments regarding a particular cue (identity or expression) while the task-irrelevant cue was held constant or varied. Participants experienced more interference from identity during expression processing than the reverse when static faces were viewed (as in Ganel & Goshen-Gottstein, 2004). Regardless of age, when viewing dynamic faces, participants seemed better able to ignore task-irrelevant information and direct their attention selectively to either expression or identity cues. These results suggest that face cues are processed in a similar manner across development – a finding that contradicts the traditional view that face-specific processing mechanisms undergo an extended period of development. Despite showing similar patterns of interference, children were less accurate and slower to respond than adults, which is consistent with the view that age-related differences in processing speed or memory may underlie group differences in performance on many face processing tasks (Baenniger, 1994; Crookes & McKone, 2009; McKone et al., 2012).

Introduction

Developmental studies suggest that the ability to extract facial expression and facial identity improves gradually, approaching adult levels in late childhood or early adolescence (Carey & Diamond, 1977; Chung & Thomson, 1995; de Sonneville et al., 2010; Herba, Landau, Russell, Ecker, & Phillips, 2006). These improvements are evident across various laboratory-based tasks, including those requiring face discrimination (e.g., Carey, Diamond, & Woods, 1980; Mondloch, Dobson, Parsons, & Maurer, 2004; Mondloch et al., 2002) and recognition memory (Carey et al., 1980; Carey & Diamond, 1977; Johnston & Ellis, 1995; Weigelt et al., 2014), and in those involving explicit and implicit processing (Herba et al., 2006; Mancini, Agnoli, Baldaro, Ricci Bitti, & Surcinelli, 2013). The most common explanation given for children's poorer performance on these tasks is that it takes many years of visual experience to develop specialized *face-specific* perceptual processing mechanisms (de Haan, Pascalis, & Johnson, 2002; de Heering, Rossion, & Maurer, 2012; Joseph et al., 2012). Researchers have suggested that, due to immaturity in these mechanisms, children use qualitatively different strategies to process faces than adults. In particular, it is thought that, compared to adults, children rely more on feature-based processing and information about facial contour, and less on holistic processing, when making judgments about faces (e.g., Carey & Diamond, 1977; Mondloch et al., 2002; Schwarzer, Huber, & Dümmler, 2005; Schwarzer, 2000). Children's identity judgments are also heavily influenced by non-face cues (e.g., paraphernalia or hairstyle) (Baenniger, 1994; Carey & Diamond, 1977; Diamond & Carey, 1977; Freire & Lee, 2001).

The idea that the development of configural processing lags behind the development of featural processing or the processing of external contour (Mondloch et al., 2002) has been challenged. Thus, some researchers have reported that, although children are less accurate and slower than adults, when investigators correct for these differences in performance, or equate task difficulty across age groups, both groups show similar face inversion (Gilchrist & McKone, 2003; Pellicano et al., 2006; Picozzi et al., 2009) and composite-face effects (Susilo et al., 2009), and comparable sensitivity to variations in the spacing between facial features (Gilchrist & McKone, 2003; McKone & Boyer, 2006; Pellicano et al., 2006). Given these findings, McKone and colleagues (Crookes & McKone, 2009; McKone et al., 2012) argue that face processing mechanisms *are* mature in early childhood, and that age-related improvements in performance on face processing tasks likely reflect domain general improvements in perceptual skills (e.g., acuity), cognitive abilities (e.g., memory, sustained attention), and processing speed (see also Baenniger, 1994).

In the present investigation, I looked for possible age-related differences in the processing of expression and identity using Garner's selective attention task (Garner & Feldoldy, 1970; Garner, 1976). This paradigm typically involves two experimental blocks: a *baseline block* (in which a relevant cue varies from trial to trial while an irrelevant cue remains constant) and an *orthogonal block* (in which the relevant and irrelevant cues vary randomly across trials). In both blocks, participants make speeded judgments about the relevant cue and their performance in the two blocks is compared. Comparable accuracy or RTs between blocks indicates that the ability to extract the relevant cue is not influenced by variations in the irrelevant cue. In contrast, significantly poorer performance in the orthogonal than in the baseline block suggests that the processing of the irrelevant cue interferes with the processing of the relevant cue (i.e., that the two cues cannot be processed independently).

To date, only a small number of studies have used Garner's selective attention task to explore age-related differences in face processing (Baudouin, Durand, & Gallay, 2008; Krebs et al., 2011; Spangler et al., 2009). In these studies, children and adults showed comparable levels of interference from identity during expression processing. In contrast, group differences in interference from expression during identity processing were reported by Baudouin et al., but not by Krebs et al. or Spangler et al. This apparent discrepancy in results may reflect differences in the way the stimuli were created. In addition, it is important to note that none of these investigators took group differences in baseline performance into consideration when assessing interference effects, so the impact of group differences in domain-general skills (e.g., processing speed and/or memory) were not controlled for. To rectify this, in the present study, I quantified interference in the orthogonal block in terms of the *percent change from baseline* accuracy or RT that occurred. I reasoned that, if children rely more on feature-based processing and less on holistic or configural processing than adults when judging static faces, they may not find trial-to-trial variations in the task-irrelevant dimension as distracting. As a result, children may experience less interference than adults, overall, once baseline differences in speed or accuracy are taken into account.

With moving faces, I expected to see a different pattern of results. Previous research has shown that adult viewers respond differently to static and non-rigidly moving faces (e.g., Everdell et al., 2007; Fujimura & Suzuki, 2010; Pike et al., 1997; Stoesz & Jakobson, 2014; Thornton & Kourtzi, 2002). In addition, I recently reported that adults show less interference between identity and expression processing in dynamic than in static testing conditions; indeed, interference dropped to negligible levels when faces were in motion (Chapter 4) (Stoesz & Jakobson, 2013). One interpretation of this result is that adults are better able to selectively attend to relevant cues when viewing moving faces, perhaps because they rely more heavily on feature-based processing in these circumstances. If children *and* adults use a feature-based approach to process moving faces, then interference levels may be low and comparable across age groups in dynamic testing conditions. However, children may also have more difficulty processing moving faces than adults. Indeed, motion perception skills follows a rather long developmental course, with global motion integration skills (for example) not reaching maturity until the early teenage years (e.g., Bogfjellmo et al., 2014). This may explain why adults are able to make use of dynamic facial information when processing expressions (Ambadar, Schooler, & Cohn, 2005; Arsalidou, Morris, & Taylor, 2011; Back, Jordan, & Thomas, 2008) and identity (Pilz, Thornton, & Bülthoff, 2006; Thornton & Kourtzi, 2002), but adolescents (Back, Ropar, & Mitchell, 2007) and young children (Gepner, Deruelle, & Grynfeltt, 2001) have difficulty doing so. If children do have particular difficulty processing moving faces, then the introduction of dynamic cues may lead to larger increases in RT and reductions in accuracy for children compared to adults. If these increases are particularly evident during orthogonal testing, children (unlike adults) may experience interference when making judgments about moving faces.

Method

Participants. Ninety-three individuals participated in this study. I tested 34 children aged 6.0-8.0 years (*M* = 6.7, *SD* = .6; 16 boys, 18 girls), 32 adolescents aged 12.0-13.8 years (*M* = 12.9, *SD* = .6; 16 boys, 16 girls), and 27 young adults aged 18.1-21.0 years (*M* = 19.1, *SD* = .8; 12 men, 15 women). Children and adolescents were recruited via word-of-mouth and through local schools in Winnipeg and Altona, Canada. Young adults were recruited through the psychology participant pool at the University of Manitoba, Winnipeg, Canada. All participants had normal or corrected-to-normal vision.

Materials. I obtained static face stimuli from researchers at the Max Planck Institute for

Biological Cybernetics, Germany (Pilz et al., 2006). During the production of these stimuli, actors wore a black cap and scarf that covered their hair and clothes, and sat against a black background. None of the actors wore glasses or jewellery. Each actor was filmed expressing several different emotions while being filmed at a frame rate of 25 fps. From the films made of four female actors, I selected sequences of 26 frontal-view, static images that captured the unfolding of two different expressions (surprised and angry). I used these images to create dynamic stimuli using QuickTime 7 Pro (Apple Inc., USA). My static stimuli consisted of the static image from each sequence that depicted the apex of the expression. Participants viewed the stimuli from a distance of 57 cm. At this viewing distance, the width and height of each stimulus face subtended 6.6° and 6.6°, respectively. The experiments were presented on the monitor of a PC computer and were executed with MATLAB (The MathWorks, Inc., MA).

Procedure. The study protocol was approved by the Psychology/Sociology Research Ethics Board at the University of Manitoba. Adult participants and parents of each child/adolescent who participated in the study provided written informed consent. Children and adolescents also confirmed their assent. Participants were tested individually. Each participant first completed a familiarization task that allowed them to become acquainted with the expressions and identities that they would see in the Garner's selective attention task that followed. This familiarization procedure was followed as it has been shown that interference from expression is more evident during identity judgments when faces are familiar to the viewer (Ganel & Goshen-Gottstein, 2004). Instructions given to participants are found in Appendices E and F.

Familiarization (identity-matching) task. On each trial, a central, white fixation cross appeared on a black background for 500-ms, followed by a 1,040 ms presentation of two faces presented side-by-side (see Figure 3.1). Participants completed a speeded, two-alternative forced-choice task, in which they judged whether the identities of the two faces were the same or different. Half of the participants pressed one key for a *same* judgment and another key for a *different* judgment, with the key assignments reversed for the remaining observers. The task consisted of 52 static trials and 52 dynamic trials, with an equal number of same and different trials in each presentation mode. Four different female faces, each showing two different expressions (surprise, anger), were shown equally often. To discourage use of picture-based strategies on *same* trials, the two faces always displayed different expressions. Static trials preceded dynamic trials and the order of trials within each presentation mode was randomized across participants.

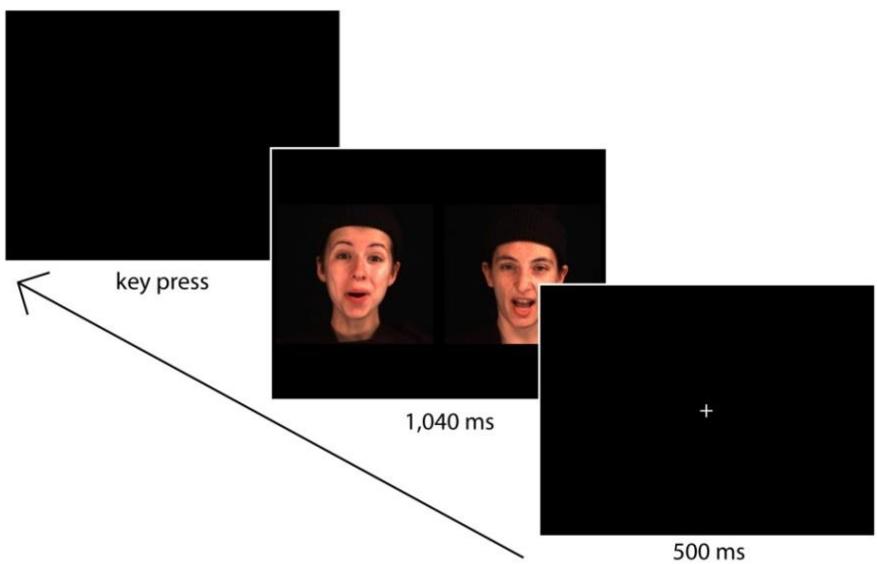

Figure 3.1. Presentation sequence for the Familiarization (identity-matching) task. Participants viewed two, simultaneously-presented static images (static condition) or dynamic sequences (dynamic condition) for 1,040 ms. Across trials, participants made same or different judgments via a key press, using their index fingers.

Garner's selective attention tasks. Two of the four faces used in the Familiarization task were used in this task. Participants completed two types of Garner tasks (Expression, Identity) in two presentation modes (static, dynamic), for a total of four conditions, the order of which was randomized across participants. Each condition consisted of a *baseline* block followed by an *orthogonal* block. In the baseline block (20 trials, randomly ordered), the relevant facial cue (e.g., Expression: surprised or angry) varied while the irrelevant facial cue (e.g., Identity: Jane or Anne) remained constant. The orthogonal block (40 trials, randomly ordered) consisted of all four combinations of the two dimensions (i.e., Jane surprised, Jane, angry, Anne surprised, Anne angry). Each trial began with a 500-ms central, white fixation cross on a black background, followed by a 1,040 ms central presentation of one stimulus face (see Figure 3.2). The participant made a speeded, two-alternative forced-choice response and the next trial began immediately after the response. Half of the participants pressed one key for *Jane* or *surprised* (depending on the task) and another key for *Anne* or *angry*, with key assignments reversed for the remaining participants. Participants completed five static practice trials and five dynamic practice trials, before beginning the experiment.

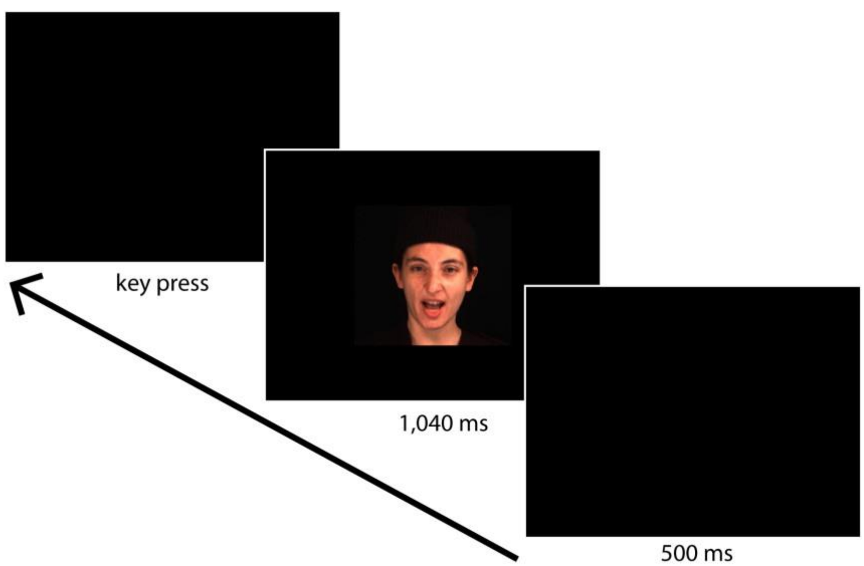

Figure 3.2. Presentation sequence for the Garner's selective attention tasks. Participants viewed a static image (static condition) or a dynamic sequence (dynamic condition) for 1,040 ms and responded with a key press using their index fingers. For the Identity judgment task, participants determined whether the face belonged to "Jane" or to "Anne." For the Expression judgment task, participants determined whether the expression was one of anger or surprise.

Results

For each participant, responses in the Garner's selective attention tasks outside the window of 200-5,000 ms after target onset were eliminated; this accounted for 0.27% of the responses in these tasks.

Baseline blocks.

Accuracy. Non-parametric tests were used to examine the accuracy in baseline blocks because the data from each of the age groups was generally not normally distributed. A series of Friedman tests for each age group revealed that accuracy was similar in the two tasks, and was not affected by the introduction of dynamic cues [$\chi^2(3) < 4.00$, $p > .26$, for all three tests]. However, a series of Krustal-Wallis tests revealed group differences within each condition [$\chi^2(2) > 32.00$, $p < .001$ in all cases]. Follow-up tests indicated that, in all four testing conditions, children performed more poorly than either adolescents or adults [Mann-Whitney $U > 337.0$, $p < .03$, in all cases]. In addition, adults performed more poorly than adolescents in three of four testing conditions [Mann-Whitney $U = 293.00$, $p = .005$; the single exception being that adolescents and adults performed similarly on the static identity task, $p = .08$]. See Table 3.1 for the median and ranges of accuracy scores for each group in each condition.

Table 3.1
Median Accuracy (% Correct) for Children, Adolescents, and Adults in the Baseline Block of Garner's Selective Attention Tasks

	Children		Adolescents		Adults	
	Mdn	Range (min-max)	*Mdn*	Range (min-max)	*Mdn*	Range (min-max)
Static displays						
Expression task	95	55-100	100	100-100	100	90-100
Identity task	95	35-100	100	90-100	100	50-100
Dynamic displays						
Expression task	90	60-100	100	90-100	100	85-100
Identity task	95	50-100	100	90-100	100	90-100

Response times. Median correct RTs (ms) from baseline trials were submitted to separate 3 (Age Group: children, adolescents, adults) x 2 (Task: expression, identity) x 2 (Presentation Mode: static, dynamic) ANOVAs, with repeated measures on the last two factors. Overall, I observed a decrease in median correct RTs with increasing age [$F(2, 90) = 34.10$, $p < .001$, $\eta_p^2 = .43$], with adolescents and adults responding more quickly than children ($p < .001$, for both contrasts). Participants also generally responded more quickly in the Identity than the Expression task [$F(1, 90) = 53.52$, $p < .001$, $\eta_p^2 = .37$], and when making judgments about static as opposed to dynamic faces [$F(1, 90) = 296.72$, $p < .001$, $\eta_p^2 = .77$]. Follow-up tests on the significant Task x Presentation Mode [$F(1, 90) = 17.25$, $p < .001$, $\eta_p^2 = .16$] and Age Group x Presentation Mode [$F(2, 90) = 8.38$, $p < .001$, $\eta_p^2 = .16$] interactions showed that the effect of presentation mode was larger in the Expression than the Identity task [$t(92) = 4.254$, $p < .001$, $d = .44$] (see Figure 3.4a), and smaller in children and adolescents than in adults [$t > 3.36$, $p \leq .001$, $d > 3.36$, for both contrasts] (see Figure 3.4b).

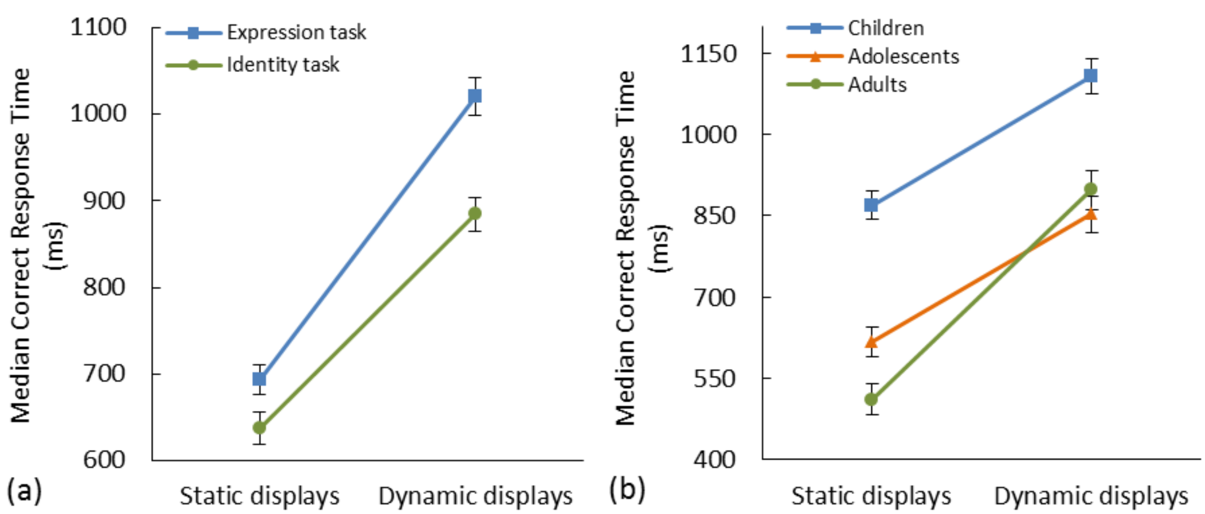

Figure 3.3. Median correct RT (ms) for the baseline blocks of the Garner's selective attention tasks. (a) significant Task x Presentation Mode interaction, RTs were longer for the Expression task than the Identity task in both presentation modes; (b) significant Age Group x Presentation Mode interaction. Standard error indicated.

Garner interference effect.

Accuracy data. Because adolescents' and adults' accuracy was at ceiling in the baseline and orthogonal blocks, it was not possible to measure interference effects in these groups. In analyzing the data from children, I corrected for baseline differences in accuracy across tasks and presentation modes by computing *corrected Garner interference scores,* using the formula: $\left(\frac{Baseline - Orthogonal}{Baseline}\right)(100)$. Corrected Garner interference scores reflect the percent change from baseline accuracy associated with the introduction of task-irrelevant changes in identity (Expression task) or expression (Identity task). Positive scores are indicative of interference.

Children's corrected Garner interference scores were submitted to a 2 (Task) x 2 (Presentation Mode) repeated measures ANOVA. Results indicated a main effect of Presentation Mode, with larger interference effects with static than dynamic displays [$F(1, 33) = 4.73$, $p = .04$, $\eta_p^2 = .13$]. Follow-up tests on the significant Task x Presentation Mode interaction [$F(1, 33) = 5.06$, $p = .03$, $\eta_p^2 = .13$] (see Figure 3.4) confirmed that children experienced more interference during static than dynamic testing in the Expression task [$t(33) = 3.05$, $p = .004$, $d = .52$], but low and comparable levels of interference with both types of displays when making identity judgments. One-sample *t*-tests verified that interference was significantly different from zero in the static Expression task only.

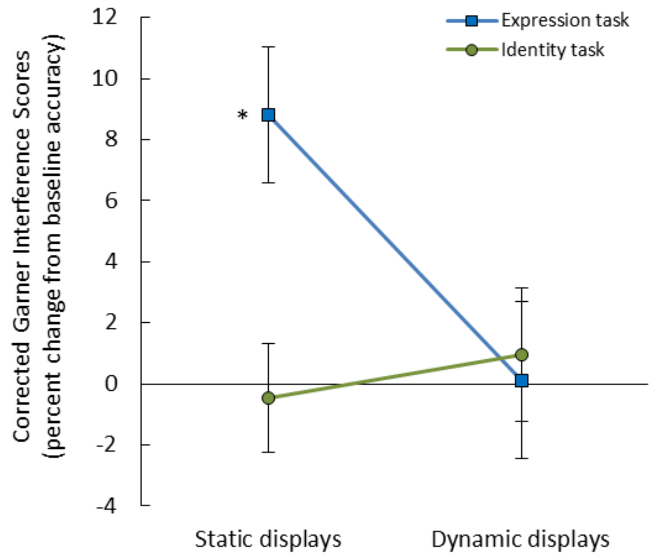

Figure 3.4. Children's mean corrected Garner interference scores (percent change from baseline accuracy) for the Expression and Identity tasks in the static and dynamic presentation modes. *Interference significantly greater than zero, $p < .001$. Standard error indicated.

Response time data. To correct for observed baseline differences in RT across groups, tasks, and modes of presentation, I computed *corrected Garner interference scores* using the formula: $\left(\frac{Orthogonal - Baseline}{Baseline}\right)(100)$. Positive scores indicate interference. I submitted these scores to a 3 (Age Group: children, adolescents, adults) x 2 (Task: expression, identity) x 2 (Presentation Mode: static, dynamic) ANOVA, with repeated measures on the last two factors.

In general, participants experienced more interference during the Expression than the Identity task [$F(1, 90) = 13.26$, $p < .001$, $\eta_p^2 = .13$] and more interference with static than dynamic displays [$F(1, 90) = 38.51$, $p < .001$, $\eta_p^2 = .30$]. Follow-up tests performed on the significant Task x Presentation Mode interaction [$F(1, 90) = 17.76$, $p < .001$, $\eta_p^2 = .17$] confirmed that the Task effect was evident in static testing only [$t(92) = 2.89$, $p = .005$, $d = .30$], and that the main effect of Presentation Mode was much larger for the Expression than the Identity task [$t(92) = 4.24$, $p < .001$, $d = .44$]. As is evident in Figure 3.5, this interaction took a similar form in all three age groups; indeed the three-way interaction did not approach significance [$F(1, 33) = 0.61$, $p = .54$]. The results of a series of one-sample *t*-tests confirmed that interference was significant when participants in all age groups viewed static faces [$t > 3.11$, $p < .005$, $d > .55$; the single exception being children's interference during the static identity task where a trend was observed, $p = .07$], but interference was negligible when they viewed dynamic faces.

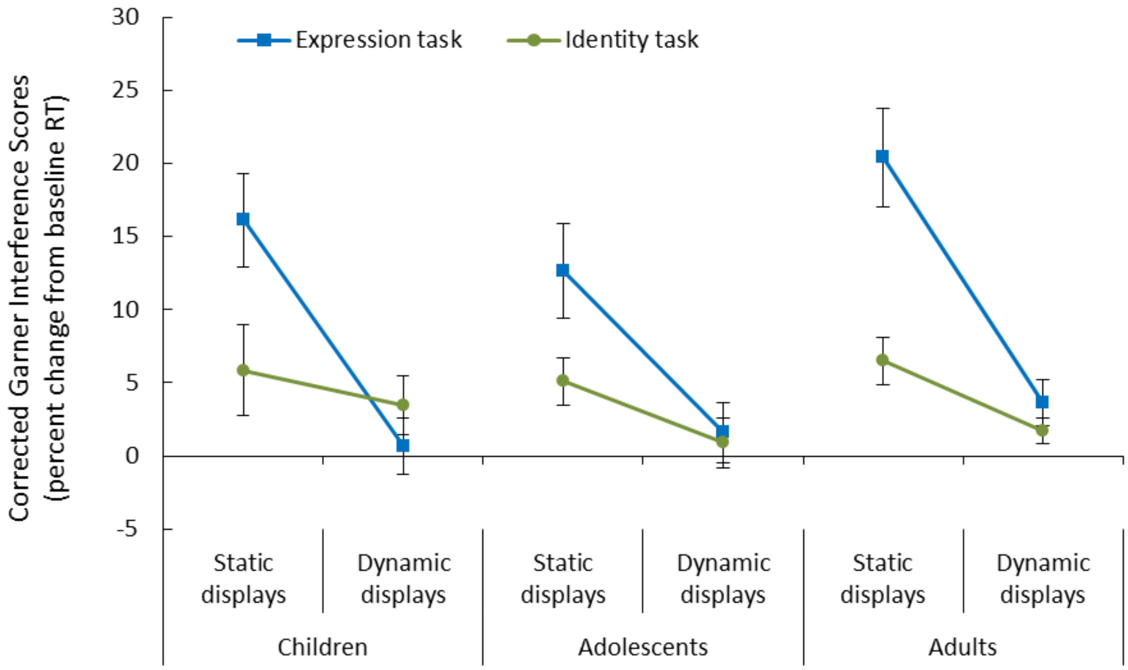

Figure 3.5. Mean corrected Garner interference scores (percent change from baseline RT) in the Expression and Identity Tasks for static and dynamic presentation modes in each age group. The pattern of asymmetrical interference was similar in all three age groups. Standard error indicated.

Discussion

The goal of the present study was to search for possible age-related differences in the ability to attend selectively to expression and identity cues in faces. When making judgments concerning static faces, children, adolescents, and adults showed more interference from identity during expression processing than the reverse. In contrast, when dynamic faces were viewed, participants in all age groups appeared to ignore task-irrelevant information and direct their attention selectively to either expression or identity. These results lend indirect support to the view that children, adolescents, and adults process face cues in a similar manner (Baenniger, 1994; Crookes & McKone, 2009; McKone et al., 2012), an interpretation that runs counter to the traditional view that the face-specific (e.g., holistic) processing mechanisms undergo a protracted period of development (Carey & Diamond, 1977; Chung & Thomson, 1995; de Sonneville et al., 2010; Herba et al., 2006). The fact that children were, nonetheless, less accurate and slower to respond than adults during baseline trials is consistent with McKone and colleagues' view that age-related differences domain-general abilities such as processing speed may underlie group differences in performance on many face processing tasks (Crookes & McKone, 2009; McKone et al., 2012; see also Baenniger, 1994).

Research indicates that motion processing follows a long developmental course (e.g., Bogfjellmo et al., 2014). Given this, I was surprised to find that adults showed a larger increase in baseline RTs with the addition of motion cues than children or adolescents. One possible explanation for this finding is that adults were processing and utilizing available motion cues to a greater extent than younger viewers. Support for this interpretation comes from previous research showing that, unlike adults (Back, Jordan, & Thomas, 2009), adolescents do not benefit from the availability of dynamic cues when they are required to use facial expression to judge mental states (Back et al., 2007). Similarly, young children's ability to correctly match the static image of a facial expression that they see on a television screen to a photograph is not better than when they match a movie of an unfolding facial expression on a television screen to a photograph, unless the motion is slowed (Gepner et al., 2001). In other work, Crookes and McKone (2009) and Taylor, Batty, and Itier (2004) have argued that the speed with which face perception mechanisms can resolve the identity of faces improves gradually with age.

In baseline testing, viewers showed a larger increase in RT with the introduction of dynamic cues during the Expression task than during the Identity task. This finding is consistent with other work showing that participants find it easier to extract identity than expression information, especially when faces are moving (see Chapters 3 and 4) (Stoesz & Jakobson, 2013). This makes sense considering that, with the static displays, diagnostic information for expression processing was immediately available to viewers at stimulus onset, but during dynamic testing the apex of the expression was not reached until 840 s after stimulus onset. Viewers may have had to wait to acquire enough information about the unfolding expression to be certain of its form; however, this additional information may not have improved the accuracy of their identity judgments substantially. Viewers may have also found it easier to extract identity than expression during dynamic baseline testing if they were relying on facial contour to make their identity judgments. Children may have been particularly likely to do this, given their tendency to use gaze aversion to control their cognitive load (see Chapter 2) (Stoesz & Jakobson, 2014), and given that they are more likely than adults to focus their attention on external paraphernalia (i.e., clothing, hats, hairstyle, or eyeglasses) and external contour (when visible) (Campbell et al., 1999; Carey & Diamond, 1977; Ellis et al., 1979; Freire & Lee, 2001; Mondloch et al., 2002). Indeed, some children in my study stated explicitly that they used jaw/chin shape, or the edge between the actor's black hat and her forehead, to make their identity judgments. If some viewers used contour information to make their judgments in the Identity task, this may also explain why interference was generally quite low in this task, even during static trials. It is important to note, however, that others have reported smaller composite face effects with moving as opposed to static face stimuli (Xiao, Quinn, Ge, & Lee, 2012, 2013). This finding supports the view that the drop in interference seen with the introduction of dynamic cues during the Identity task may reflect a shift toward use of a more feature-based approach, rather an increased tendency to utilize facial contour cues to make identity judgments. In future studies, the extent to which viewers of various ages utilize contour versus feature information could be explored by manipulating images in such a way that facial contour is held constant, or by incorporating eye-tracking technology.

The finding that participants in all age groups experienced interference during the static, but not the dynamic, trials of the Expression task is interesting and important. I speculate that the drop in interference that I observed with the introduction of dynamic cues in this task arises primarily because viewers relied more heavily on a feature-based processing strategy to extract expression from moving compared to static faces. The idea that action perception is reliant on featural processing gains support from the findings that, when discriminating between different actions, participants are more sensitive to changes in small-scale (local) actions than to configural changes in whole-body movements (Loucks & Baldwin, 2009). Using an approach that is more heavily feature-based (or even one that is more balanced) may afford greater resistance to interference, but may come at a cost -- specifically, an overall increase in processing time (Macrae & Lewis, 2002; Marzi & Viggiano, 2011).

In Garner’s selective attention paradigm, the absence of interference is typically interpreted as evidence that the relevant and irrelevant cues are being processed independently. I have suggested that independent processing of these cues might be achieved more easily in dynamic testing conditions if viewers shift toward more feature-based processing or focus more heavily on facial contour. It is also possible, however, that participants may have experienced negligible interference from task-irrelevant cues when viewing dynamic faces because they *integrate* expression and identity cues more efficiently under these circumstances. Efficient integration of expression and identity cues when naturalistic movements are present is more likely if the processing of these cues occurs within the same brain region. Indeed, some research suggests that invariant and changeable aspects of faces are integrated in a single region in the middle fusiform gyrus (Tsuchiya, Kawasaki, Oya, Howard, & Adolphs, 2008), whereas other work shows that the integration of these cues occurs in the STS (Hein & Knight, 2008; Puce et al., 2003; Rossion, Hanseeuw, & Dricot, 2012). There is speculation, however, that STS along with area MT extract identity cues from facial movement signatures and structure-from-motion information, which is then projected to FFA as static form information (Roark et al., 2003). This proposal is based on the findings that certain neurons in IT respond to the shape cues in motion-defined form stimuli (Sáry, Vogels, & Orban, 1993) and that IT-lesioned monkeys can discriminate shapes based on motion-defined form information, but cannot learn to discriminate shapes based on form-from-luminance cues (Britten, Newsome, & Saunders, 1992). Additionally, compared to static faces, dynamic faces produce greater activation in various regions within the core and extended face processing systems (e.g., Kessler et al., 2011; Kilts, Egan, Gideon, Ely, & Hoffman, 2003; Schultz & Pilz, 2009). Although future research is required to directly test this assertion, taken together these results suggest that the way in which various parts of the face processing network work together is altered when motion cues are present (Calder & Young, 2005).

Conclusions

Natural facial movements convey changes in our emotional state and verbal information, and even subtle movements of our mouth and eyes provide a rich and powerful source of social information. As a result, recognizing and understanding various facial movements are basic social skills required to interact with other people successfully. We have much more experience looking at naturally moving faces than we do looking at faces in photographs. Moreover, there is ample research showing that the cognitive mechanisms (e.g., Ambadar et al., 2005; Ceccarini & Caudek, 2013; Foley, Rippon, Thai, Longe, & Senior, 2012; Pilz et al., 2006) and neural networks (e.g., Arsalidou et al., 2011; Fox, Iaria, et al., 2009; Kessler et al., 2011; Kilts et al., 2003; Sato et al., 2010; Schultz & Pilz, 2009; Trautmann, Fehr, & Herrmann, 2009b) involved in processing moving faces are different from those involved in processing static images. The present findings reinforce the importance of using ecologically relevant, dynamic facial stimuli. I found different patterns of interference between the processing of facial identity and expression with static and dynamic stimuli, with interdependence seen in the former but not the latter case. Participants appeared to focus their attention on task-relevant cue(s) to a greater extent with dynamic than static faces, possibly by increasing their reliance on local processing or by integrating facial cues more efficiently. The absence of marked age-related differences in interference suggest that face-specific processing mechanisms *are* mature in early childhood, and that general cognitive development accounts for improvements in face processing with age (Baenniger, 1994; Crookes & McKone, 2009; McKone et al., 2012). However, additional research using a wider range of stimuli and approaches is needed to determine the precise factors that underlie performance in different age groups (McKone et al., 2012), and that explain differences in the ways static and moving faces are processed (Ceccarini & Caudek, 2013; Roark et al., 2006).

CHAPTER 4: A SEX DIFFERENCE IN INTERFERENCE BETWEEN IDENTITY AND EXPRESSION JUDGMENTS WITH STATIC BUT NOT DYNAMIC FACES

Running head: STATIC AND DYNAMIC FACE PROCESSING

Keywords: dynamic advantage, expression, identity, nonrigid motion, sex differences

Except for some minor wording changes and removal of some portions of the Methods section (to reduce redundancies across chapters), Chapter 4 is reprinted here (with permission) as it appears in:

Stoesz, B. M., & Jakobson, L. S. (2013). A sex difference in interference between identity and expression judgments with static but not dynamic faces. *Journal of Vision*, *13*, 1-14. doi: 10.1167/13.5.26.doi

My Contribution to the Publication

The study described in this paper was carried out as a follow-up to the work described in Chapter 3. In the present study, I examined sex differences in selective attention to expression and identity cues in faces. I designed the experiment, collected and analyzed the data, drafted the manuscript, and revised the manuscript based on feedback from my advisor. I selected the journal, *Journal of Vision*, to which Dr. Jakobson and I submitted the manuscript. Based on the reviewers' comments, I revised the manuscript in consultation with Dr. Jakobson and resubmitted the manuscript. When the manuscript was accepted I made any further required corrections.

Abstract

Facial motion cues facilitate identity and expression processing (Pilz et al., 2006). To explore this dynamic advantage, I used Garner's selective attention task (Garner, 1976) to investigate whether adding dynamic cues alters the interactions between the processing of identity and expression. I also examined whether facial motion affected women and men differently, given that women show an advantage for several aspects of static face processing (McClure, 2000). Participants made speeded identity or expression judgments while the irrelevant cue was held constant or varied. Significant interference occurred with both tasks when static stimuli were used (as in Ganel & Goshen-Gottstein, 2004), but interference was minimal with dynamic displays. This suggests that adult viewers are either better able to selectively attend to relevant cues, or better able to integrate multiple facial cues, when viewing moving as opposed static faces. These gains, however, come with a cost in processing time. Only women showed asymmetrical interference with static faces, with variations in identity affecting expression judgments more than the opposite. This finding may reflect sex differences in global-local processing biases (Godard & Fiori, 2012). Our findings stress the importance of using dynamic displays and of considering sex differences when characterizing typical face processing mechanisms.

Introduction

Faces are among the most complex and important visual stimuli in our environment. In the real world, faces move, and natural rigid motion (e.g., head turns and nods) and nonrigid changes in the shape of facial features over time (as in unfolding expressions) serve as important social signals (Barton, 2003). Dynamic displays of facial expression do not simply provide redundant static information; rather, the specific spatiotemporal information they convey leads to more accurate and faster recognition compared to that observed with displays that show form cues only (Knappmeyer, Thornton, & Bülthoff, 2003). This *dynamic advantage* is most evident when task demands are high (Horstmann & Ansorge, 2009), when form information is degraded or distorted as in point-light displays (Bassili, 1979) or morphed sequences (Kamachi et al., 2001), or when expressions are subtle (Ambadar et al., 2005; Bould & Morris, 2008) or synthetic (Wehrle, Kaiser, Schmidt, & Scherer, 2000). In each of these circumstances, the expressions are more difficult to identify and the addition of motion cues is beneficial. Behaviourally, the dynamic advantage may be lost when performance is at ceiling (Fiorentini & Viviani, 2011), but it may still be evident in the form of enhanced neural activation in the core and extended face processing network (Arsalidou et al., 2011; Kessler et al., 2011).

A dynamic advantage is also frequently observed during the processing of facial identity, although, as with expression processing, it is sometimes difficult to demonstrate with unaltered displays (Christie & Bruce, 1998; Knight & Johnston, 1997; Lander, Christie, & Bruce, 1999), or when the task is not sufficiently demanding (Lander et al., 1999). During the processing of familiar faces, a dynamic advantage can be observed under a variety of suboptimal viewing conditions (Knight & Johnston, 1997; Lander, Bruce, & Hill, 2001; Lander et al., 1999), with the greatest effects occurring in the presence of natural (as opposed to slowed or disrupted) movement (Lander & Bruce, 2000; Lander et al., 1999). One might also expect to find a dynamic advantage in a variety of other challenging situations, such as when viewers see faces briefly, match faces shown from different viewpoints, discriminate between unfamiliar individuals, or match faces displaying various types of movements. The results from investigations using non-degraded, unfamiliar faces (Pike et al., 1997; Pilz et al., 2006; Thornton & Kourtzi, 2002) support these predictions. Thornton and Kourtzi (2002) found that participants were quicker to match the identity of a static target to a dynamic prime (moving nonrigidly) than to a static prime when the expressions of the target and prime faces differed. In addition, Pilz et al. (2006) demonstrated a dynamic advantage with nonrigid motion regardless of changes in the viewpoint (i.e., front, left, or right facing) of the prime face, or whether the task involved sequential matching or visual search.

Researchers investigating the dynamic advantage in identity processing often utilize the nonrigid motion of expressive faces, and the benefit of motion may be particularly evident when viewers match the identities of faces expressing different emotions (Thornton & Kourtzi, 2002). One reason for this finding may be that viewers are better able to attend selectively to the identity or expression information in dynamic faces than in static faces, resulting in a greater resistance to interference between the processing of these different facial cues. Alternatively, viewers may be attending to multiple sources of information when processing identity and expression cues in dynamic, as opposed to static, displays – resulting in increased interdependence between the processing of these cues.

One way to study functional independence or, alternatively, interdependence between the processing of different facial cues is to use Garner's classification task (Garner, 1976). The vast majority of studies using this paradigm to study face processing have employed static displays only (e.g., Baudouin, Martin, Tiberghien, Verlut, & Franck, 2002; Ganel, Goshen-Gottstein, & Goodale, 2005; Ganel & Goshen-Gottstein, 2004; Schweinberger & Soukup, 1998; with the exception of Kaufmann & Schweinberger, 2005). Garner's task was originally designed to examine one's ability to process one dimension of a visual stimulus while ignoring another dimension of the same stimulus. The task typically involves presentation of stimuli in two experimental blocks: *baseline* (or *control*) and *orthogonal* (or *filtering*). In studies using faces, the baseline block comprises trials in which a relevant dimension (e.g., identity) varies while an irrelevant dimension (e.g., expression) is held constant. Participants make speeded judgments regarding the relevant dimension. Accuracy and response times (RTs) in the baseline block are then compared with performance in the orthogonal block, in which both relevant and irrelevant dimensions vary randomly. Equivalent performance in baseline and orthogonal blocks indicates that one's ability to extract the relevant facial dimension is not influenced by variations in the irrelevant dimension. In contrast, *Garner interference* occurs when significantly less accurate and/or slower responses occur in the orthogonal block compared to the baseline block; this pattern of results suggests that the relevant dimension cannot be processed independently of the irrelevant dimension[2].

Using unfamiliar static faces, researchers exploring the dependencies between identity and expression processing using Garner's selective attention task have generally found no interfering effects from expression when making identity judgments, but significant interfering effects from identity when making expression judgments. This asymmetrical pattern of results

[2] Sometimes a *correlated* block is also included, in which each level of the relevant dimension (e.g., each identity in an identity judgment task) is linked with only one level of the irrelevant dimension (e.g., a particular facial expression). Researchers have suggested that performance in correlated blocks does not provide information about whether two facial dimensions are dependent or independent (Schweinberger & Soukup, 1998). Instead, it appears to be strongly affected by differences in discriminability, and based on decisional strategies rather than on the perceptual relationship between the two face cues (see Schweinberger & Soukup, 1998). Because of this, researchers have either discarded the data from correlated blocks (Schweinberger & Soukup, 1998), or have not included these blocks in their experimental designs (Ganel & Goshen-Gottstein, 2002).

has been described in adults (Schweinberger, Burton, & Kelly, 1999; Schweinberger & Soukup, 1998) and typically-developing children (Krebs et al., 2011; Spangler et al., 2009), and suggests that systems supporting identity and expression processing may be interconnected, but there is only one direction of cross-talk between them. The pattern may change, however, when viewers process even somewhat familiar faces. Viewers may still experience asymmetrical Garner interference with familiar faces, but significant interference occurs in *both* directions, suggesting a *functional interdependence* between the processing of these two types of facial cues (Ganel & Goshen-Gottstein, 2004). These observations cast doubt on traditional face-processing models that suggest that the processing of identity and expression cues depend on parallel and functionally independent pathways (Bruce & Young, 1986), subserved by different and largely independent neural structures (Hoffman et al., 2000). Of course, the existence of specialized brain areas does not provide a strong argument for strict separation on a functional level, as specialized regions in the intact brain might influence one another within the face-processing network (Fox, Moon, et al., 2009) and within the broader social-cognition system (Beauchamp & Anderson, 2010). Functional interdependence also makes sense if one considers that a familiar person's idiosyncratic (i.e., characteristic) facial expressions (i.e., that individual's *facial motion signatures*) can aid in the determination of his or her identity, just as the unique structure of an individual's face can constrain the way that emotions are expressed (Ganel & Goshen-Gottstein, 2004).

A key goal of the present study was to determine if the addition of dynamic facial cues alters the strength or nature of the interactions between the processing of identity and expression information. If this was the case, it might explain why dynamic advantages have been observed in many studies of face processing. To investigate this question, I asked adult viewers to complete a familiarization (identity-matching) task and Garner's selective attention task, in that order. The former task was administered, in part, to familiarize viewers with the identities of the faces and facial expressions used in the Garner task. Given that the experiment involved the use of non-degraded stimuli, I expected a high level of performance on both tasks and, as such, I did not expect to see a dynamic advantage in terms of accuracy or reaction time. However, for the Garner task, I hypothesized that dynamic cues would alter the interactions between identity and expression processing, resulting in either increased resistance to interference (reflected in smaller Garner interference scores), or increased interdependence (reflected in larger Garner interference scores).

A second goal of this study was to examine the impact of participant sex on performance on static and dynamic face processing tasks. Studies comparing the processing of static and dynamic faces have not typically considered participant sex. Moreover, previous reports of asymmetrical Garner interference between the processing of identity and expression in static faces are based on studies in which sex ratios were unequal (with the majority of the participants being women, e.g., Ganel & Goshen-Gottstein, 2004; Kaufmann & Schweinberger, 2005), or in which small sample sizes precluded the exploration of sex differences (Schweinberger & Soukup, 1998). This is unfortunate as examining sex differences in Garner interference may provide valuable insights into why women outperform men when processing the identity of unfamiliar faces (Godard & Fiori, 2012; McBain, Norton, & Chen, 2009; Megreya, Bindemann, & Havard, 2011) and facial expressions (see McClure, 2000) [particularly when the faces being viewed are female (see Herlitz & Rehnman, 2009)]. Megreya et al. (2011) showed that this face processing advantage was not due to women showing a general superiority in episodic memory. There is evidence to suggest, however, that women's face processing advantage may be especially evident under more demanding task conditions, such as when displays are masked by visual noise (McBain et al., 2009) or two different facial cues are present. Importantly, as all of the work cited above involved static face stimuli, it is not clear whether sex differences in face processing will also be apparent with dynamic stimuli. In the present study, I looked for evidence of sex differences in the magnitude of any interference effects occurring during the Garner's selective attention task. Specifically, I wondered whether previous observations of asymmetrical interference between identity and expression processing with static faces might be more, or less, apparent with dynamic stimuli, and whether this might vary depending on participants' biological sex. Exploring these questions is important if one is to gain a deeper understanding of the factors that underlie individual differences in perceptual processing.

Method

Participants. The sample consisted of 20 women (aged 18-26 years, M = 20.1, SD = 2.3) and 20 men (aged 18-24 years, M = 19.9, SD = 2.0) from the psychology participant pool at the University of Manitoba, Winnipeg, Canada. All participants had normal or corrected-to-normal visual acuity.

Materials and procedure. The Human Research Ethics Board at the University of Manitoba approved the testing protocol. Participants provided written informed consent and received partial course credit. Participants were tested one at a time in a quiet room. Each participant completed the Familiarization (identity-matching) task first to allow us to: (a) examine static versus dynamic face matching, and (b) familiarize participants with the identities and expressions they would view in the Garner's task that followed. The familiarization process helped to ensure that the response options were clear when viewers completed their identity judgments in the Garner task. Note that becoming somewhat familiar with the faces may have increased the likelihood that viewers would be able to extract useful information about facial expression and structure during identity and expression processing, respectively (Ganel & Goshen-Gottstein, 2004). The detailed descriptions of the materials and procedures for the Familiarization task and Garner's selective attention task (including two figures) are provided in Chapter 3; as such, they have been removed from this chapter to minimize redundancies across chapters. Instructions given to participants are found in Appendices E and F.

Results

For each participant, trials in which responses occurred outside the window of 200-5,000 ms after target onset were eliminated. This represented 3.83% of trials in the Familiarization task, and 0.27% of the trials in the Garner tasks.

Familiarization (identity-matching) task. Median correct RTs were submitted to a 2 (Participant Sex: Women, Men) x 2 (Presentation Mode: Static, Dynamic) analysis of variance (ANOVA), with repeated measures on the last factor. Results revealed a main effect of Presentation Mode, $F(1, 38) = 620.83$, $p < .001$, $\eta_p^2 = .94$, indicating that participants were faster making judgments when viewing pairs of static faces ($M = 999$ ms; $SD = 212$ ms) compared to pairs of dynamic faces ($M = 1621$ ms; $SD = 247$ ms). See Figure 4.1. This did not appear to reflect a speed-accuracy trade-off as the results of a similar analysis conducted on accuracy scores revealed that accuracy was near ceiling (≥ 89%) and comparable in males and females across both static and dynamic testing conditions. However, a signal detection analysis did reveal that viewers' sensitivity to dynamic faces ($M_{d'} = 3.40$, $SD = .94$) was marginally better than their sensitivity to static faces ($M_{d'} = 3.04$, $SD = .89$) [$F(1,38) = 4.02$, $p = .052$, $\eta_p^2 = .096$].

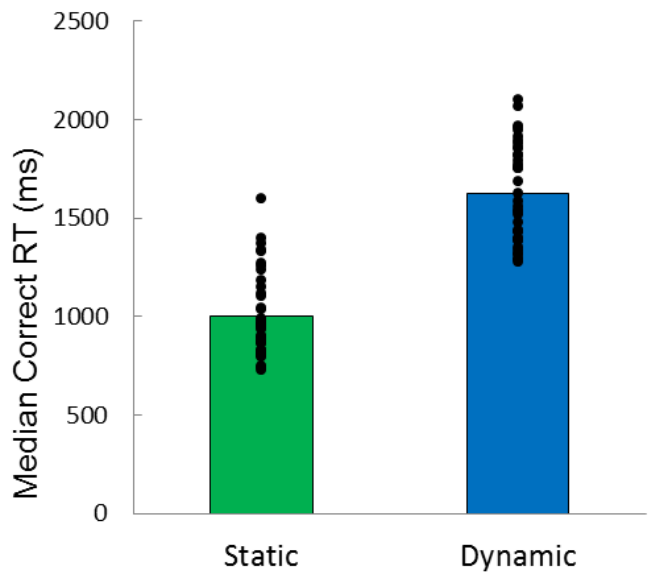

Figure 4.1. Median correct RT (ms) in the static and dynamic conditions of the familiarization (identity-matching) task. Dots represent data from individual participants, whereas bars represent group means.

Garner's selective attention tasks. Accuracy scores were submitted to a 2 (Participant Sex: Women, Men) x 2 (Task: Identity Judgments, Expression Judgments) x 2 (Presentation Mode: Static, Dynamic) x 2 (Condition: Baseline, Orthogonal) ANOVA, with repeated measures on the last three factors. Accuracy was slightly higher, overall, during baseline compared to orthogonal trials [97.4% vs. 96.1%; $F(1, 38) = 7.12$, $p < .05$, $\eta_p^2 = .158$], but this difference – while statistically significant -- was very small. Indeed, accuracy was essentially at ceiling in all conditions of the task for both men and women (≥ 94.3%), with no other significant main effects or interactions being observed. For this reason, I concluded that any main effects and interactions arising in the RT data were unlikely to reflect speed accuracy trade-offs. As such, results presented below focus on median RTs for correctly answered trials.

Baseline blocks. Median correct RTs were submitted to a 2 (Participant Sex: Women, Men) x 2 (Task: Identity Judgments, Expression Judgments) x 2 (Presentation Mode: Static, Dynamic) ANOVA, with repeated measures on the last two factors. Results revealed main effects of Task, $F(1, 38) = 17.00$, $p < .001$, $\eta_p^2 = .31$, and Presentation Mode, $F(1, 38) = 1051.21$, $p < .001$, $\eta_p^2 = .97$; comparisons of mean RTs confirmed that viewers were able to extract identity more quickly than expression, and made their judgments more quickly when viewing static as opposed to dynamic stimuli. These effects were mediated by a significant Task x Presentation Mode interaction, $F(1, 38) = 5.47$, $p < .03$, $\eta_p^2 = .13$. Follow-up tests on the interaction revealed that the Task effect (difference in RTs between Expression and Identity processing) was slightly larger with dynamic faces (M = 93 ms, SD = 145) than with static faces (M = 40 ms, SD = 97), and that the Mode effect (difference in RTs between Dynamic and Static stimuli) was slightly larger for the Expression task (M = 410 ms, SD = 124) than the Identity task (M = 358 ms, SD = 76) [$t(39) > 2.35$, $p < .03$ for both contrasts]. (See Figure 4.2).

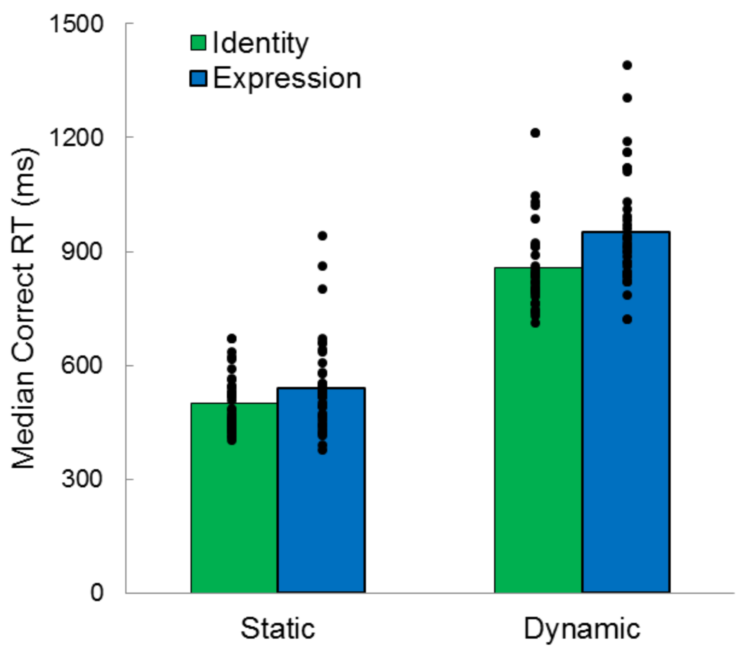

Figure 4.2. Median correct RT (ms) in the baseline conditions of Garner's selective attention tasks. Dots represent data from individual participants, whereas bars represent group means.

Garner interference effect. I computed Garner interference scores by subtracting the median RT for correct trials in the baseline block from that seen in the orthogonal block for each participant. Because there were significant differences in the baseline measures for the four test conditions, I divided each participant's Garner interference score in a given condition by his/her median correct RT in the corresponding baseline block. These *corrected Garner interference scores* reflect the percent change in RT from baseline.

I submitted corrected Garner interference scores to a 2 (Participant Sex: Men, Women) x 2 (Task: Identity Judgments, Expression Judgments) x 2 (Presentation Mode: Static, Dynamic) ANOVA, with repeated measures on the last two factors[3]. Results revealed main effects of Task [$F(1, 38) = 12.31$, $p = .001$, $\eta_p^2 = .25$] and Presentation Mode [$F(1, 38) = 36.00$, $p < .001$, $\eta_p^2 = .49$]. Overall, participants experienced less interference from expression when making identity judgments than vice versa, and less interference from the irrelevant dimension with dynamic than with static faces. These main effects were mediated by significant Task x Presentation Mode [$F(1, 38) = 6.30$, $p < .02$, $\eta_p^2 = .14$] and Task x Participant Sex interactions [$F(1, 38) = 4.71$, $p < .04$, $\eta_p^2 = .11$]. As these two-way interactions had to be interpreted in light of a significant Task x Presentation Mode x Participant Sex interaction [$F(1, 38) = 6.24$, $p < .02$, $\eta_p^2 = .14$], follow-up tests were limited to the three-way interaction, which is depicted in Figure 4.3.

[3] Note that submitting uncorrected Garner interference scores to the same analysis produced a very similar pattern of results. Given the differences in baseline performance across conditions, however, interpretation of the corrected scores is more straightforward and was, therefore, preferred.

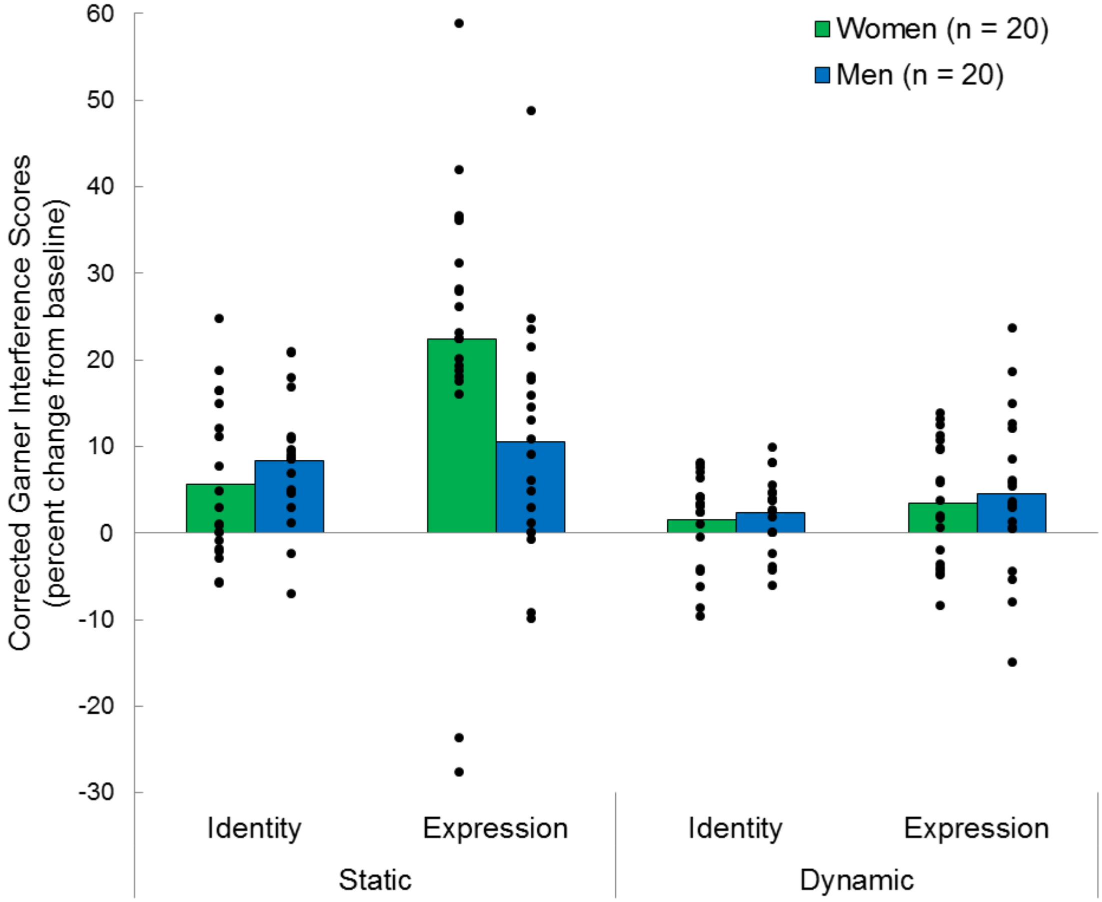

Figure 4.3. Women's and men's mean corrected Garner interference scores in Identity and Expression Tasks for static and dynamic presentation modes. Dots represent data from individual participants, whereas bars represent group means.

Inspection of Figure 4.3 suggested that the interactions arose because of sex differences in task performance during static trials. To test this, for each presentation mode, corrected Garner interference scores were entered into separate 2 (Participant Sex: Women, Men) x 2 (Task: Identity Judgments, Expression Judgments) ANOVAs, with repeated measures on the last factor. With static faces, I observed a main effect of Task, $F(1, 38) = 11.10$, $p = .002$, $\eta_p^2 = .23$, and a Task x Participant Sex interaction, $F(1, 38) = 6.53$, $p < .02$, $\eta_p^2 = .15$. Follow-up tests on the interaction revealed that only women showed an asymmetrical pattern of interference, experiencing four times as much interference from the irrelevant dimension when processing expression than when processing identity, $t(19) = 3.61$, $p = .002$. Men experienced similar levels of interference during the two tasks, and their performance was comparable to that of women completing the Identity task. In striking contrast to the results that I observed with static stimuli, the analysis conducted on data from dynamic trials did not reveal any significant main effects or interactions.

In additional follow-up tests, I compared static and dynamic trials for each task, for both men and women. Women experienced less interference in the dynamic compared to the static presentation mode when making expression judgments, $t(1, 19) = 4.83$, $p < .001$, whereas men experienced less interference in the dynamic than in the static presentation mode when making identity judgments, $t(1, 19) = 3.70$, $p < .003$.

In order to determine whether women or men experienced significant levels of Garner interference in any of the testing conditions, I conducted planned one-sample *t*-tests (comparing corrected Garner interference scores to zero). These tests confirmed that women and men showed significant interference effects for both tasks when viewing static stimuli [$t(19) \geq 2.78$, $p < .02$ in all cases]. Only men experienced significant interference when processing dynamic faces; this was the case whether men were making identity or expression judgments [$t(19) \geq 2.19$, $p < .05$ in both cases].

Supplementary analyses. As has been noted elsewhere (Garner, 1983; Schweinberger & Soukup, 1998), susceptibility to interference may depend on the *relative speed* with which relevant and irrelevant cues can be extracted. To explore the impact that this might have had on the present findings, I conducted supplementary analyses on the median RT data.

The first set of comparisons was designed to explore the possibility that differences in the relative speed of cue extraction could explain why women showed an asymmetrical pattern of interference in the static condition. As can be seen in Figure 4.3, there were large individual differences in the relative speed of cue extraction during static baseline testing. Inspection of the data revealed that 13 participants showed an *expression advantag*e (i.e., they judged expression faster than identity), 2 showed no advantage, and 24 showed an *identity advantage* (i.e., they judged identity faster than expression). If relative ease of cue extraction was a key factor determining interference levels then, in the Identity task, those showing an identity advantage might be able to make identity judgments before expression processing could cause substantial interference, but not the reverse (i.e., they might show an asymmetrical pattern of interference). In the Expression task, the opposite should be true. Thus, participants who show an expression advantage might be able to make expression judgments before identity processing could interfere with task performance, but not the reverse. Only 15 of the 24 participants showing an identity advantage (7 men, 8 women) showed less interference in the Identity task than in the Expression task. Only 2 of the 13 participants showing an expression advantage (both male) showed less interference in the Expression task than in the Identity task. Together, these results suggest that differences in relative ease of cue extraction might account (at least in part) for the behaviour of a portion of the sample during the static trials (17 of 40 individuals), however, they do not adequately account for the behaviour of the remaining participants.

Using a similar logic to that described above, I explored the possibility that the smaller Garner interference scores seen with dynamic faces might reflect differences in the relative ease of cue extraction between static and dynamic testing conditions. First, I considered the 16 participants who showed an identity advantage during static baseline trials that was *amplified* during dynamic baseline trials. During the Identity task, I might expect these participants to exhibit less interference from expression with moving faces than with photographs, compared to that seen in participants with a static identity advantage who did not show amplification (n = 8). What I found, however, was that both groups showed an equivalent drop in interference from expression when viewing moving as opposed to static faces (Mode main effect: $F(1,22) = 17.9$, $p < .001$, $\eta_p^2 = .45$).

Next, I considered the participants who showed an expression advantage during static baseline trials that was amplified during dynamic baseline trials. During the Expression task, I might expect these participants to exhibit less interference from identity with moving faces than with photographs. Only one individual (a women) showed amplification of her static expression advantage when tested with dynamic faces, and, in this case, the interference from identity was indeed lower in the dynamic than in the static condition. However, it is noteworthy that 11 of the remaining 12 participants who showed an expression advantage during static baseline trials actually showed an *identity* advantage with moving faces, and these individuals also experienced significantly less interference from identity when viewing dynamic compared to static faces (paired samples t-test: $t(10) = 4.4$, $p = .001$).

Discussion

The goals of this study were to determine if the addition of dynamic facial cues would affect the processing of identity and/or expression, and whether men and women would respond similarly to this manipulation. The results revealed three important findings: (1) there was no dynamic advantage in either the Familiarization (identity-matching) task or in the baseline blocks of the Garner task in terms of RT; (2) bi-directional Garner interference effects occurred between identity and expression processing in the static presentation mode but interference was minimal when dynamic faces were used as stimuli; and (3) in the static condition, only women showed an asymmetrical pattern of Garner interference, with changes in identity affecting expression judgments more than the reverse. Each of these results is discussed, in turn, below.

Identity-matching and baseline performance with static and dynamic faces. The fact that we found little evidence of a dynamic advantage with regard to RT in the matching task, or during the baseline block of Garner's paradigm, was not surprising given that participants were able to perform both tasks with a high level of accuracy. This may have been because the static images I presented were non-degraded and showed fully developed, intense emotional expressions (i.e., they captured the apex of each expression). Under these conditions it may be difficult to demonstrate a dynamic advantage (Fiorentini & Viviani, 2011; Lander et al., 1999). Key information needed to extract emotion and facial structure would have been immediately available to observers in these images, potentially making the processing of additional information from dynamic cues unnecessary. With dynamic displays, in contrast, the apex of the expression was reached some time after stimulus onset – a fact that may explain why RTs during dynamic trials were significantly *longer* than those seen during static trials. Interestingly, however, participants did not appear to wait until the apex of an expression was reached before initiating their responses on dynamic trials (see Fiorentini & Viviani, 2011 for a similar result). Indeed, responses on these trials were completed (on average) either before the apex of the expression or shortly thereafter (within 200 ms), suggesting that participants would have *initiated* their responses when the expressions were more subtle than those captured in the static stimuli. The movement of facial features may compensate for the "incompleteness" of an unfolding expression (Fiorentini & Viviani, 2011); this may explain why a dynamic advantage does occur when participants must discriminate between subtle expressions (Ambadar et al., 2005).

Interference effects with static and dynamic faces. Like previous research in adults

(Ganel & Goshen-Gottstein, 2004), I found significant interference effects for static stimuli when participants classified faces according to either identity or expression. With static stimuli, then, participants could not avoid computing expression when processing identity, or vice versa. These findings suggest that these aspects of face processing are interdependent, at least when the faces are somewhat familiar to viewers (see Ganel & Goshen-Gottstein, 2004).

In contrast to the results with static stimuli, I found little evidence of interference from irrelevant facial cues with dynamic faces – a result which, traditionally, would be interpreted as evidence for functional independence in the processing of those cues (Ganel & Goshen-Gottstein, 2004). This result was somewhat surprising as it seems reasonable to expect that facial motion signatures could provide useful, supplemental cues to facial identity, and that motion-enhanced recovery of the three-dimensional facial structure could improve one's ability to predict possible constraints on the way a given face moves (Calder & Young, 2005; Ganel & Goshen-Gottstein, 2004; O'Toole, Roark, & Abdi, 2002a). As dynamic faces should convey richer information about identity *and* expression than static images, viewers might find both types of cues more salient (i.e., more difficult to ignore) in dynamic faces, regardless of the task. Consistent with this, researchers have found greater interference from facial identity during the processing of facial speech with dynamic than with static displays (Kaufmann & Schweinberger, 2005). The fact that I observed a different pattern in the present study (with interference being significantly more evident with static than dynamic stimuli) supports the view that facial expressions are processed rather differently from facial speech (Fodor, 1983; Liberman & Mattingly, 1985). It remains to be seen if expression and speech processing interact with identity coding in different ways when dynamic stimuli are used.

The differences in interference effects I observed between static and dynamic faces suggest that photographs of faces are processed quite differently than dynamic faces. It could be that the use of photographs of social stimuli such as faces in research studies biases viewers to adopt strategies optimized for processing *images* rather than for dealing with complex, naturally moving stimuli. In this regard, there is a large body of evidence suggesting that static faces are processed using a global (holistic) approach (see Farah et al., 1998). One possible interpretation of the results is that the use of a global approach when processing static faces makes it difficult for viewers to ignore task-irrelevant information, leading to significant interference effects. If the introduction of dynamic cues caused participants to shift their attention to local facial features and ignore the global context, then this might explain the reduced interference scores I saw with dynamic displays. In support of this idea, Xiao, Quinn, Ge, and Lee (2012) showed that participants are better able to decompose faces into parts when processing dynamic as compared to static stimuli. In related work, Loucks and Baldwin (2009) found that the processing of local (small-scale) actions (i.e., “featural information in action,” p.87) is elevated relative to the processing of global movement patterns when viewers watch dynamic scenes depicting whole-body human actions. Loucks (2011) later reported that this was not the case with static displays. If increased reliance on a local processing strategy results in greater resistance to interference between identity and expression processing, it appears that this comes at a cost – specifically, an overall increase in processing time. This result is consistent with other research showing that the use of a part-based strategy can disrupt some aspects of performance (Macrae & Lewis, 2002; Marzi & Viggiano, 2011).

Although the present experiment does not directly address the question of whether there is a shift toward feature-based strategies (or, perhaps, toward the use of a more balanced or flexible approach) while processing dynamic facial cues, this would be a plausible mechanism through which a reduction in interference between the processing of identity and expression cues could be achieved. Specifically, switching to a feature-based processing approach may alter activation patterns within the broader face-processing network, minimizing the involvement of areas specialized for the global processing of facial cues. If such a shift did occur, it might underlie the increase in RTs seen in dynamic testing conditions.

While concluding that the reduction in interference with dynamic faces arises from enhanced attention to features would be consistent with the traditional interpretation of Garner interference, adopting this classic interpretation may be attractive simply because it fits with the popular view that the face processing system has a modular structure (see Calder & Young, 2005). Another way to interpret the negligible interference seen with dynamic displays is that viewers are better able to *integrate* multiple facial cues when viewing naturalistic, moving stimuli. Evidence supporting this idea comes from studies examining how invariant and changeable facial cues in dynamic displays interact. In a series of experiments, Knappmeyer et al. (2003) exposed participants to both form (i.e., identity) and nonrigid motion (i.e., expression) in morphed faces that represented a continuous transition between the identities of two learned faces. Characteristic, nonrigid facial motion associated with a particular form biased participants' identity decisions, suggesting that integration occurs when processing these two types of facial information in dynamic displays. Other research suggests that facial form and motion information are integrated in the STS (Puce et al., 2003), with nonrigid and rigid biological motion eliciting differential activation in an anterior-posterior gradient in the STS (Grèzes et al., 2001). Furthermore, dynamic faces produce more robust activation in multiple parts of the core and extended face processing systems than static faces (e.g., Kessler et al., 2011; Kilts et al., 2003; Schultz & Pilz, 2009). Together, these results suggest that the availability of motion cues may cause a shift in the way that different parts of the face processing and social cognition networks work together, which could explain why RTs were longer in the dynamic conditions. The notion that identity and expression cues may be integrated more successfully with dynamic than with static stimuli, resulting in a reduced interference, is consistent with Calder and Young's (2005) idea that separation between regions involved in processing invariant and changeable aspects of faces is relative rather than absolute.

I have suggested here that interference between the processing of different facial cues may be reduced or eliminated when dynamic stimuli are being viewed if participants focus more selectively on specific facial features, or if they integrate multiple cues more effectively. Both of these ideas hold merit and, indeed, it is possible that individual differences in processing style determine which strategy a given viewer will adopt. I tested this idea in Chapter 5. In future work, it will also be interesting, and important, to determine whether characteristics of the stimulus face not explored here, such as the *particular* expression that is displayed, impact performance on identity-matching and Garner tasks.

Sex differences in static face processing. The last major point that needs to be addressed relates to the fact that, in the static condition, only women showed an asymmetrical pattern of interference when viewing photographs of faces, with changes in identity affecting expression judgments more than the opposite. Men, in contrast, showed equivalent levels of interference from irrelevant cues in both tasks. Based on work with static stimuli, some researchers have argued that identity coding is obligatory but that expression coding is not (Palermo & Rhodes, 2007). If this applies more to women than men it could explain the sex difference in interference I observed in the Expression task. Women may be more attentive to multiple nonverbal cues when processing another individual's emotional state in an effort to improve their ability to infer the mental states of others, and/or to empathize with them. If so, this may explain (in part) why women and men show different patterns of neural activation when attempting to solve emotion tasks, including those requiring the recognition of facial expressions in photographs (Alaerts, Nackaerts, Meyns, Swinnen, & Wenderoth, 2011; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Derntl et al., 2010).

Some have argued that the extraction of identity involves global processing, whereas the extraction of expression requires greater attention to local features (Lipp et al., 2009; Song & Hakoda, 2012), at least in static displays. If, as suggested above, women are more likely than men to process identity *and* expression information while making their expression judgments, interference could arise from the simultaneous or sequential application of these two different processing strategies. Some preliminary evidence supporting this idea comes from the work of Proverbio and colleagues (Proverbio, Brignone, Matarazzo, Del Zotto, & Zani, 2006; Proverbio, Riva, Martin, & Zani, 2010) with face stimuli, and Kimchi, Amishav, and Sulitzeanu-Kenan (2009) with non-face stimuli.

It is important to remember that the sex difference I observed with static faces disappeared when dynamic cues were available. However, this does not necessarily mean that men and women are using similar approaches or the same neural networks to process dynamic facial information. Research involving other types of dynamic social stimuli speaks to this issue. For example, Pavlova, Guerreschi, Lutzenberger, Sokolov, and Krageloh-Mann (2010) showed that, although women and men exhibited ceiling-level performance in their ability to discriminate between random and "social" interactions between geometric shapes on the basis of their movement patterns, robust sex differences were apparent in the induced oscillatory response to these displays over left prefrontal cortex -- a region implicated in perceptual decision making (Heekeren, Marrett, & Ungerleider, 2008). Pavolva et al. speculated that women anticipate social interactions -- predicting others' actions ahead of their realization -- whereas men require accumulation of more sensory evidence before making decisions regarding the social meaning of actions. Amunts et al. (2007) have also shown sex differences in the cytoarchitecture of motion-sensitive complexes, and suggest that these brain structures work together in different ways in men and women to produce the same kind of behavioural performance in tasks involving the processing of motion. Given these findings, it might be interesting to explore differences in brain activation patterns during baseline and orthogonal blocks while women and men process identity and expression in static and dynamic facial images. In carrying out this work, it would be interesting to determine whether sex differences in activation vary depending on characteristics of the face stimuli, such as the sex of the individuals' depicted (recall that in the present study only female faces were used). Others have reported that a female advantage for static expression processing is particularly evident when the faces being viewed are female (see Herlitz & Rehnman, 2009). Investigations of this sort may shed light on factors underlying inter- and intra-individual differences in performance on face processing tasks.

Conclusions. I have shown that different patterns of interference between the processing of facial identity and expression are seen with static and dynamic faces (with interdependence seen in the former but not the latter case). It may be that, in order to make accurate judgments regarding complex, dynamic stimuli, viewers must focus their attention on task-relevant cue(s) to a greater extent than they do with static stimuli, resulting in less interference. This could be achieved by increasing their reliance on a local processing strategy. Alternatively, the lack of interference effects with dynamic faces may suggest that identity and expression information are being integrated more efficiently in moving faces. The fact that, with static displays, sex differences were observed in interference from identity on expression judgments suggests that men and women may use different strategies (or combinations of strategies) to extract various facial cues from photographs. These findings highlight the importance of using ecologically relevant, dynamic facial stimuli (e.g., Kilts et al., 2003), and of considering participant sex when characterizing mechanisms involved in face perception.

CHAPTER 5: PERCEPTUAL BIASES INFLUENCE INVERSION EFFECTS AND INTERFERENCE BETWEEN EXPRESSION AND IDENTITY WITH STATIC BUT NOT DYNAMIC FACES

Running head: PERCEPUTAL BIASES, INVERSION, AND INTERFERENCE

Keywords: dynamic faces, expression, identity, non-rigid motion, processing style, selective attention

My Contribution to the Manuscript

The study described in this paper was carried out as a follow-up to the work described in Chapters 3 and 4. In the present study, I examined the role of perceptual processing biases in selective attention to expression and identity cues in faces. I designed the experiment, created the stimuli, and prepared the ethics submission. I trained an undergraduate student, Sarah Rigby, to assist with the collection and extraction of the data. I determined the methods to be used for data analysis, analyzed the data, and drafted and revised the manuscript based on feedback from Dr. Jakobson. With input from Dr. Jakobson and Ms. Rigby, I wrote and submitted an abstract based on preliminary data analyses, and presented a poster describing the results at the Vision Sciences Society (2013) conference. This manuscript is currently under review.

Abstract

With static faces, the processing of identity interferes more with the processing of expression than vice versa; however, interference is negligible with dynamic faces (Stoesz & Jakobson, 2013). Reduced interference with dynamic displays may arise if viewers rely on a global (holistic) strategy to process static photographs of faces, but engage in a more local (feature-based) or balanced strategy when processing moving faces. To test this idea, I assessed participants' perceptual biases using hierarchical stimuli, and then asked them to complete a set of Garner's selective attention tasks in which they made expression and identity judgments of static and dynamic faces presented in upright and inverted orientations. Global processors showed a larger face inversion effect (FIE) when judging the expressions of static compared to moving faces, but only when identity cues were held constant. Unlike local processors, they also experienced a marginally significant drop in interference in the static Expression task after face inversion. Finally, global processors showed interference from expression cues when judging the identity of static faces, whereas local processors did not. As predicted, both groups showed negligible interference effects with dynamic displays under all testing conditions. Together, these findings highlight important differences in the way that static and dynamic faces are processed, and suggest that facial movement facilitates a shift in processing strategy from one that is more holistic to one that is more feature-based. The results also support the view that examining individual differences is essential for advancing face processing research.

Introduction

Our ability to attend to, extract, and integrate facial information is remarkable. According to classic and influential models of face processing, attending to faces and extracting facial identity, expression, speech, and gaze information depend on parallel pathways that function relatively independently (Bruce & Young, 1986; Tovée & Cohen-Tovée, 1993). More recently, Haxby, Hoffman, and Gobbini (2000) suggested that face processing is subserved by a core face-processing neural network that performs the visual analysis of a face, and an extended system that processes the significance of the information. This *distributed human neural system for face perception* includes discrete areas within the core system that are involved in feature extraction (inferior occipital gyri), the extraction of facial invariants useful for identity perception (lateral fusiform gyrus), and the processing of changeable aspects of faces, including expressions, lip movements, and gaze shifts (superior temporal sulcus, or STS). The face-processing network (Hoffman et al., 2000; Johnson, 2005) is considered to be separate from networks supporting body and hand recognition (e.g., Downing, Liu, & Kanwisher, 2001) and the recognition of objects (Tsao et al., 2006).

The idea that identity and expression are processed in separate pathways has not gone unchallenged. According to some, while we may be biased to process these cues in separate pathways, it may be that this bias can be strengthened or weakened under different circumstances; in other words, the separation may be relative, rather than absolute (see Calder & Young, 2005). One factor that may influence how independently identity and expression are processed is the viewer's individual processing style (i.e., whether the viewer is biased to attend to specific features or to process faces more globally). Another factor that may play an important role is whether the faces are stationary or moving. In the present study, we explored the role that processing style plays in viewers' ability to make speeded judgments concerning static and dynamic faces. In particular, we used two well-known paradigms -- face inversion and Garner's selective attention task (Garner, 1976) -- to explore the impact of these factors on the processing of expression and identity.

Inversion effects. Faces share a common structure, consisting of the same parts in the same spatial arrangement. Despite this, we are able to distinguish between even highly similar faces quickly and accurately – a skill that may be partly attributed to our ability to process faces holistically (globally) (Bartlett & Searcy, 1993; Farah, 1996). In holistic processing, the parts of a face are not seen and represented as discrete elements, but rather as a grouping of related components arranged in a particular way that is unique to each human face. Our sensitivity to the spatial relationships between face parts allows us to recognize that one face is distinct from another (Joseph & Tanaka, 2003).

Because faces are embedded with configural information (Lobmaier, Klaver, Loenneker, Martin, & Mast, 2008), face matching, recognition, and classification of expression (Derntl, Seidel, Kainz, & Carbon, 2009) or identity (Farah, Tanaka, et al., 1995) are highly sensitive to inversion. That is, turning faces upside-down disrupts the extraction of configural information, resulting in a *face inversion effect* (FIE) (Yin, 1969) – a proportionally larger decrement in performance attributable to shifting to a local or feature-based processing approach (Tanaka & Farah, 1993). While there is debate in the literature about how to interpret differences in performance during the completion of tasks involving face inversion (Robbins & McKone, 2007), one factor that influences the magnitude of the FIE is the viewer's perceptual processing style. In the general population, some individuals show a tendency to focus on specific features in a visual display rather than on its global aspects (*weak central coherence*; Happé & Frith, 2006), whereas others show a tendency to integrate features to extract the gestalt (*global precedence*; Navon, 1977). Using scores from a hierarchical letter identification task, Martin and Macrae (2010) grouped participants according to their processing style and examined their face-recognition accuracy. When viewing photographs of upright faces, participants with a strong global bias outperformed those with a weak bias. Although both groups showed similar sensitivity for inverted faces, the FIE was smaller for participants with a weak bias. These results complement other findings demonstrating that encouraging the use of a part-based approach can disrupt the accurate processing of upright faces, and attenuate the FIE (Macrae & Lewis, 2002; Perfect, Weston, Dennis, & Snell, 2008).

The FIE may also be attenuated when we extract identity from moving faces (Lander et al., 1999; but see Thornton, Mullins, & Banahan, 2011 when the biological sex of the face is processed). Indeed, Lander et al. (1999) showed that when nonrigid motion cues (expressive and speech movements) are available participants are more accurate at recognizing the identity of moving, compared to static, inverted faces, resulting in a smaller FIE. Smaller FIEs might be expected if viewers shift to a relatively more feature-based approach when faces are moving. Support for this idea comes from recent work showing that the composite face effect, which is thought to arise from disruption of configural processing, disappears when faces are moving (Xiao, Quinn, Ge, & Lee, 2012, 2013).

Interference effects. In addition to minimizing the size of the FIE, adopting a strategy more heavily weighted toward local processing might make it easier for viewers to ignore task-irrelevant facial cues that might otherwise prove to be distracting. One can measure the degree of interference arising from task-irrelevant cues using Garner's selective attention task (Garner, 1976). In this task, viewers typically complete separate baseline and orthogonal testing blocks.

In the baseline block, a relevant cue varies while an irrelevant cue remains constant, whereas in the orthogonal block, relevant and irrelevant cues vary randomly across trials. Participants make speeded judgments regarding the relevant cue in both blocks, and performance between blocks is then compared. When similar accuracy or response times (RTs) are seen across blocks, it indicates that the viewer's ability to extract the relevant cue is not influenced by variations in the irrelevant cue. An *interference effect* arises when significantly poorer performance occurs in the orthogonal block than in the baseline block – a pattern suggesting that the relevant and irrelevant cues cannot be processed independently.

Much of the research using Garner's selective attention task to study face processing has involved the use of static images (Baudouin et al., 2002; Ganel et al., 2005; Ganel & Goshen-Gottstein, 2004). In work investigating possible interactions between identity and expression processing, two different patterns of interference have been described. Specifically, some researchers report unidirectional interference, where identity interferes with expression processing but not the reverse (Schweinberger et al., 1999; Schweinberger & Soukup, 1998), but others report asymmetrical but bidirectional interference (*functional interdependence*), where interference from identity on expression processing is greater than interference from expression on identity processing, but both are significant (Ganel & Goshen-Gottstein, 2004). The latter group argued that expression might be expected to interfere with identity processing more when the degree of similarity between stimulus faces was high (as in their work) than when faces were quite distinct (Ganel & Goshen-Gottstein, 2004).

Recently, I replicated Ganel and Goshen-Gottstein's (2004) finding of asymmetrical and bidirectional interference in the processing of expression and identity cues in static faces (Chapters 3 and 4) (Stoesz & Jakobson, 2013). However, I went on to show that interference was negligible in both tasks when judgments are made regarding dynamic faces. I suggested that this drop in interference may have occurred because viewers were relying more on a feature-based approach to process moving faces, which may have made it easier for them to ignore distracting (task-irrelevant) cues.

The present study. The goal of the present study was to examine the impact of viewers' perceptual biases and stimulus inversion on interference between expression and identity cues during the processing of static and dynamic displays. I made several predictions. First, I expected to replicate the finding of asymmetrical and bidirectional interference between these cues with upright, static faces (Ganel & Goshen-Gottstein, 2004; Stoesz & Jakobson, 2013). Second, I predicted that when making judgments regarding static faces global processors, in particular, might find it easier to extract diagnostic, metric information about the configuration of face parts when task-irrelevant cues were held constant (in the baseline condition) than when they varied (in the orthogonal condition). If they were more likely to use a global than a local processing strategy during static baseline testing, one would also expect them to show (a) larger FIEs during the baseline than the orthogonal block, and (b) more interference when making judgments concerning upright than inverted faces. Unlike global processors, local processors were expected to show small and comparable FIEs in both blocks of testing, and relatively low levels of interference regardless of face orientation. Finally, I expected to find negligible interference effects (see Chapters 3 and 4) (Stoesz & Jakobson, 2013) and relatively small FIEs (Lander et al., 1999) when viewers processed dynamic faces. These predictions would support the idea that most viewers are biased to process moving faces using an approach weighted toward local processing.

Method

Participants. Sixty individuals from the psychology participant pool at the University of Manitoba, Canada participated in this study. For reasons outlined below, data from 10 of these individuals were excluded from the analysis, leaving a final sample of 50 participants (29 women, 21 men), 17-27 years of age (*M* = 18.7 years, *SD* = 1.5 years). All had normal or corrected-to-normal vision. Participants provided written informed consent and received partial course credit for their involvement in the study. The Psychology/Sociology Research Ethics Board at the University of Manitoba approved the testing protocol.

Materials and procedures. Participants completed a global-local task involving the speeded classification of hierarchical geometric shapes, and then a series of Garner's selective attention tasks involving judgments of the expression and identity, as described below. I tested participants individually, in a quiet room.

Global-local task. Stimuli and procedures for the global-local task were modeled after those used in other studies (Insch, Bull, Phillips, Allen, & Slessor, 2012; Miyake et al., 2000; Navon, 1977; Scherf, Behrmann, Kimchi, & Luna, 2009). I created images of hierarchical geometric figures (circle, diamond, square, triangle), in which the larger *global* shape was made up of smaller *local* elements. The global and local shapes could be the same (congruent trials) or different (incongruent trials). Thus, there were four types of stimuli: (1) global congruent, (2) global incongruent, (3) local congruent, and (4) local incongruent.

On each trial, participants viewed one hierarchical figure, presented at the centre of the monitor of a PC computer using E-Prime (Psychology Software Tools, Sharpsburg, PA). At a viewing distance of approximately 60 cm, the height and width of each global shape subtended visual angles of 15° x 10.5°, respectively, whereas each local element was approximately 1° in each dimension. The color of the figure served as a cue to identify either the global shape (blue) or the local elements (red). Each figure remained on the screen until the participant responded. Participants made a four-alternative forced-choice speeded response by pressing one of four keys (one for each shape). Trials were separated by a 500-ms delay. Participants completed 20 practice trials and then three blocks of 32 randomly-ordered experimental trials (8 congruent, 24 incongruent).

Garner's selective attention tasks. A detailed description of the materials (i.e., upright faces) is provided in Chapter 3. To generate stimuli for the inverted conditions, I rotated the stimuli 180° in the picture plane. I presented this task on the monitor of a PC computer using MATLAB (The MathWorks, Inc., Natick, MA) using the general procedures outlined in Chapter 3. Details specific to the present investigation are provided below.

Participants completed four testing phases. In each phase, participants completed static and dynamic versions of a given task (Expression or Identity) in which faces were presented in one of two orientations (Upright or Inverted). The order in which participants completed the static and dynamic trials within a phase was counterbalanced across participants, as was the order of the four test phases. Each of the eight testing conditions consisted of a baseline block followed by an orthogonal block. In the baseline block (20 trials, randomly ordered), the relevant task dimension (e.g., Expression: surprised or angry) varied while the irrelevant dimension (e.g., Identity: Jane or Anne) remained constant. The orthogonal block (40 trials, randomly ordered) consisted of all four combinations of the two dimensions (i.e., Jane surprised, Jane angry, Anne surprised, and Anne angry). Trials began with a central white fixation cross displayed on a black background for 500 ms, followed by the central presentation of one stimulus face for 1,040 ms. Participants made a two-alternative forced-choice speeded response by pressing one key for “Anne” or “angry” (depending on the task) and another for “Jane” or “surprised”. Key assignments were counterbalanced across participants. Participants completed five static and five dynamic practice trials before beginning each phase of testing. See Appendix F for detailed instructions given to participants.

Results

Classification of global and local processors. Participants were classified as global or local processors based on their performance on the global-local task. I discarded data from one participant because his accuracy in one of the four conditions was very low. For the remaining participants I used median RTs from correctly answered trials to compute (a) global interference during the local task (*Global-to-Local Interference = Local incongruent – Local congruent*) and (b) local interference during the global task (*Local-to-Global Interference = global incongruent – global congruent*). A positive score on either measure indicates that the viewer experienced interference from the irrelevant dimension (i.e., slower RTs in the incongruent than in the congruent condition); a negative score indicates that the viewer experienced a facilitation effect. Nine participants did not exhibit interference in either task; data from these individuals were excluded. The remaining 50 participants were classified as global processors if their global-to-local interference score was positive and greater than their local-to-global interference score (n = 19), or as local processors if their local-to-global interference score was positive and greater than their global-to-local interference score (n = 31).

Garner’s selective attention task. On a very small proportion of trials (0.13%) in the Garner speeded classification tasks, individuals made anticipatory responses (responding < 200 ms after target onset), failed to completely depress the response key, or failed to respond within 5,000 ms; these trials were eliminated.

Accuracy was near ceiling in static (*M* = 93.6%, *SD* = 5.8%) and dynamic (*M* = 95.5%, *SD* = 4.1%) conditions, thus, I concluded that any main effects and interactions arising in the RT data were unlikely to reflect speed/accuracy trade-offs. As such, the results described below focus on the RT data.

Upright baseline blocks. Before examining inversion effects and interference scores, I assessed participants' performance in the upright trials of the baseline blocks. I submitted median correct RTs to a 2 (Group: Global processors, Local processors) x 2 (Task: Expression, Identity) x 2 (Presentation Mode: Static, Dynamic) ANOVA, with repeated measures on the last two factors. Differences in perceptual processing style did not affect performance. Overall, participants responded more quickly with static than dynamic displays [$F(1, 48) = 1011.96$, $p < .001$, $\eta_p^2 = .95$]. Responses were also quicker during the Identity task than during the Expression task (i.e., most participants showed an *identity advantage*) [$F(1, 48) = 62.79$, $p < .001$, $\eta_p^2 = .57$], especially when making judgments of moving faces [Task x Presentation Mode: $F(1, 48) = 24.49$, $p < .001$, $\eta_p^2 = .34$] (see Figure 5.1).

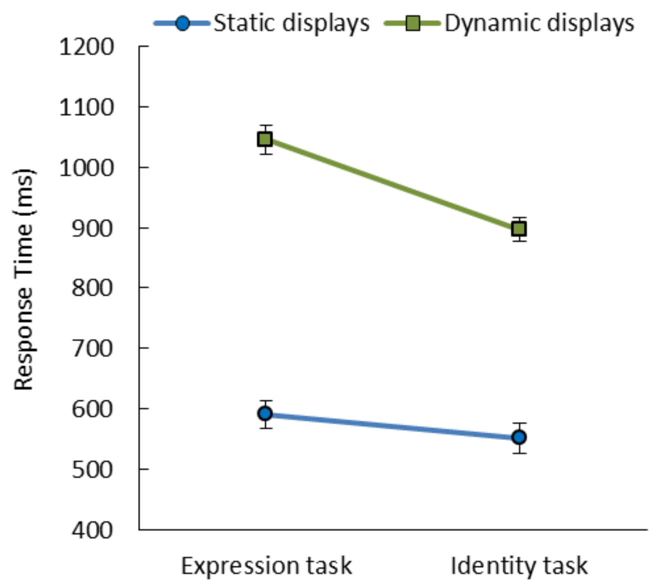

Figure 5.1. Median correct RT (ms) as a function of Task (Expression vs. Identity) and Presentation Mode (Static vs. Dynamic) for upright, baseline blocks of Garner's selective attention tasks. RTs were significantly faster for the Identity task than for the Expression task in both presentation modes ($p < .01$), but the difference was greater for dynamic faces ($p < .001$). Standard error indicated.

As an aside, with the static stimuli used here 64% of viewers showed an identity advantage. This suggests that, when task-irrelevant cues were held constant, identity cues were easier to extract than expression cues, even though the expression was immediately shown at its apex. The fact that the number of viewers displaying an identity advantage jumped from 64% to 92% with the introduction of dynamic cues makes sense if one assumes that, with these particular faces, it was not as crucial for viewers to wait for an expression to unfold to be certain of an identity judgment, as it would to be certain of an expression judgment. These results might be expected to change if the discriminability or familiarity of the faces being viewed was manipulated.

Face inversion effects. To correct for the RT differences between tasks and presentation modes during upright trials (see above), I computed *corrected FIE scores,* as follows: $\left(\frac{Upright-Inverted}{Upright}\right)(100)$. These scores, which reflect the percent change in median correct RT associated with inversion, were submitted to a 2 (Group: Global processors, Local processors) x 2 (Task: Expression, Identity) x 2 (Presentation Mode: Static, Dynamic) x 2 (Block: Baseline, Orthogonal) ANOVA, with repeated measures on the last three factors. I observed a significant main effect of Block, with larger FIEs occurring in the baseline block than in the orthogonal block [$F(1, 48) = 6.82$, $p = .01$, $\eta_p^2 = .12$]. There were also significant Task x Presentation Mode x Group [$F(1, 48) = 4.73$, $p = .04$, $\eta_p^2 = .09$] and Task x Presentation Mode x Block x Group [$F(1, 48) = 5.11$, $p = .03$, $\eta_p^2 = .10$] interactions. In the baseline block of the Expression task, the expected effects were observed. Thus, global processors showed a larger FIE when viewing static as opposed to moving faces [$t(18) = 2.77$, $p = .01$, $d = .64$], whereas local processors experienced relatively small and comparable FIEs regardless of the mode of presentation (see Figure 5.2). This pattern of results was much less evident in the orthogonal block of the Expression task. Contrary to predictions, group differences were not evident in the Identity task, and inversion effects were relatively small regardless of the presence or absence of dynamic cues during both blocks of testing.

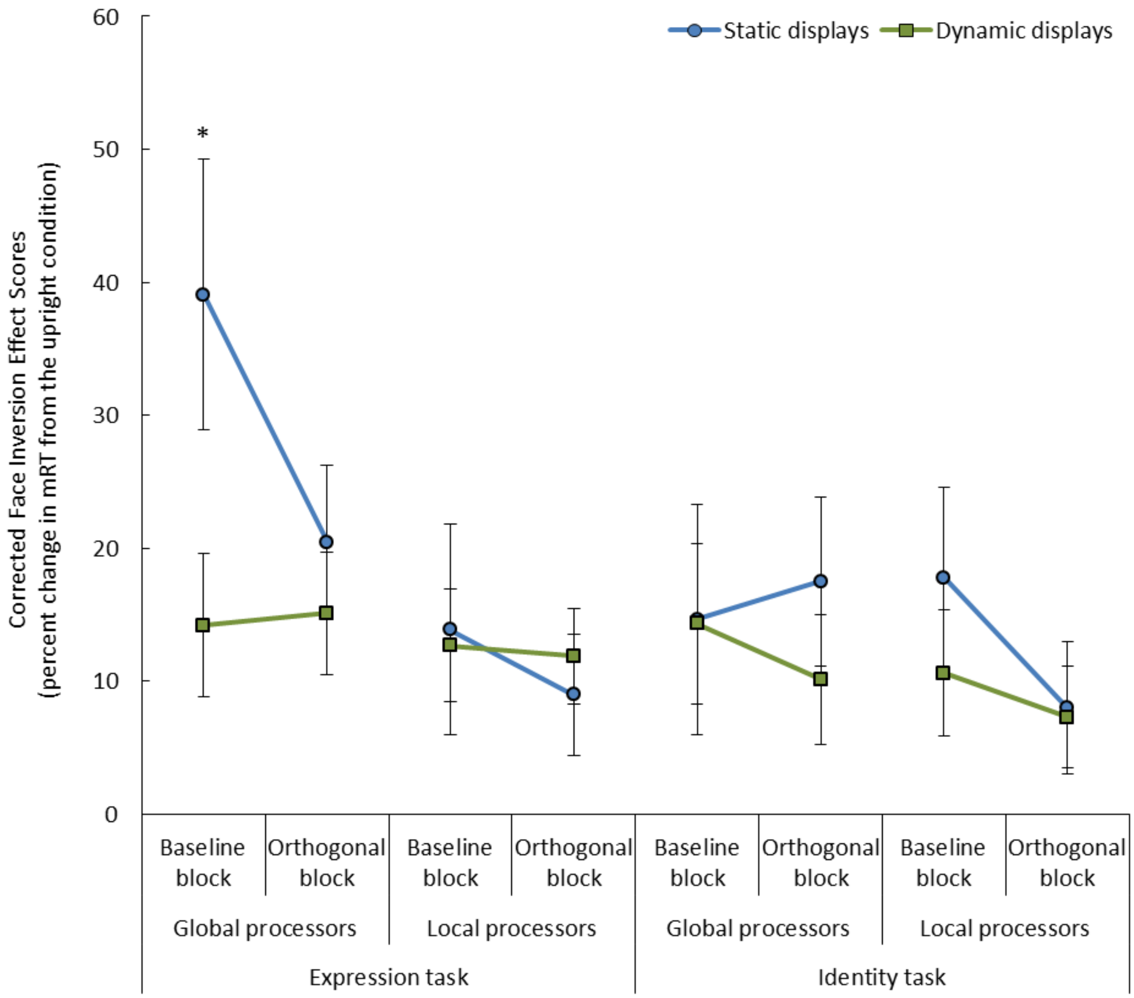

Figure 5.2. Global and local processors' mean corrected Face Inversion Effect (FIE) scores as a function of Block (Baseline vs. Orthogonal), Task (Expression vs. Identity), and Presentation Mode (Static vs. Dynamic). In the baseline block of the Expression task, the asterisk indicates that global processors experience significantly larger FIE with static than dynamic displays. Standard error indicated.

Re-running the analysis on uncorrected FIE scores produced a similar pattern of results. Given the differences in performance across tasks and presentation modes with upright displays, the interpretation of the corrected scores is more straightforward and was, therefore, preferred.

Garner interference effects. Given that participants' median correct RTs differed between tasks and presentation modes during baseline trials, I used correct median RTs to compute *corrected Garner interference scores,* as follows: $\left(\frac{Orthogonal - Baseline}{Baseline}\right)(100)$. These scores reflect the percent change in RT associated with the introduction of task-irrelevant changes in identity (Expression task) or expression (Identity task). I submitted corrected Garner interference scores to a 2 (Group: Global processors, Local processors) x 2 (Task: Expression, Identity) x 2 (Presentation Mode: Static, Dynamic) x 2 (Orientation: Upright, Inverted) ANOVA, with repeated measures on the last three factors. Participants experienced more interference during the Expression than the Identity task [$F(1, 48) = 16.82$, $p < .001$, $\eta_p^2 = .26$] and more interference with static than dynamic displays [$F(1, 48) = 57.85$, $p < .001$, $\eta_p^2 = .55$]. I also observed a significant Task x Presentation Mode interaction [$F(1, 48) = 7.24$, $p = .01$, $\eta_p^2 = .13$]. These main effects and the interaction were interpreted in light of the predicted, significant 4-way interaction [$F(1, 48) = 4.19$, $p = .046$, $\eta_p^2 = .08$] (see Figure 5.3); planned comparisons performed on this interaction allowed me to test the key predictions I had made regarding interference effects.

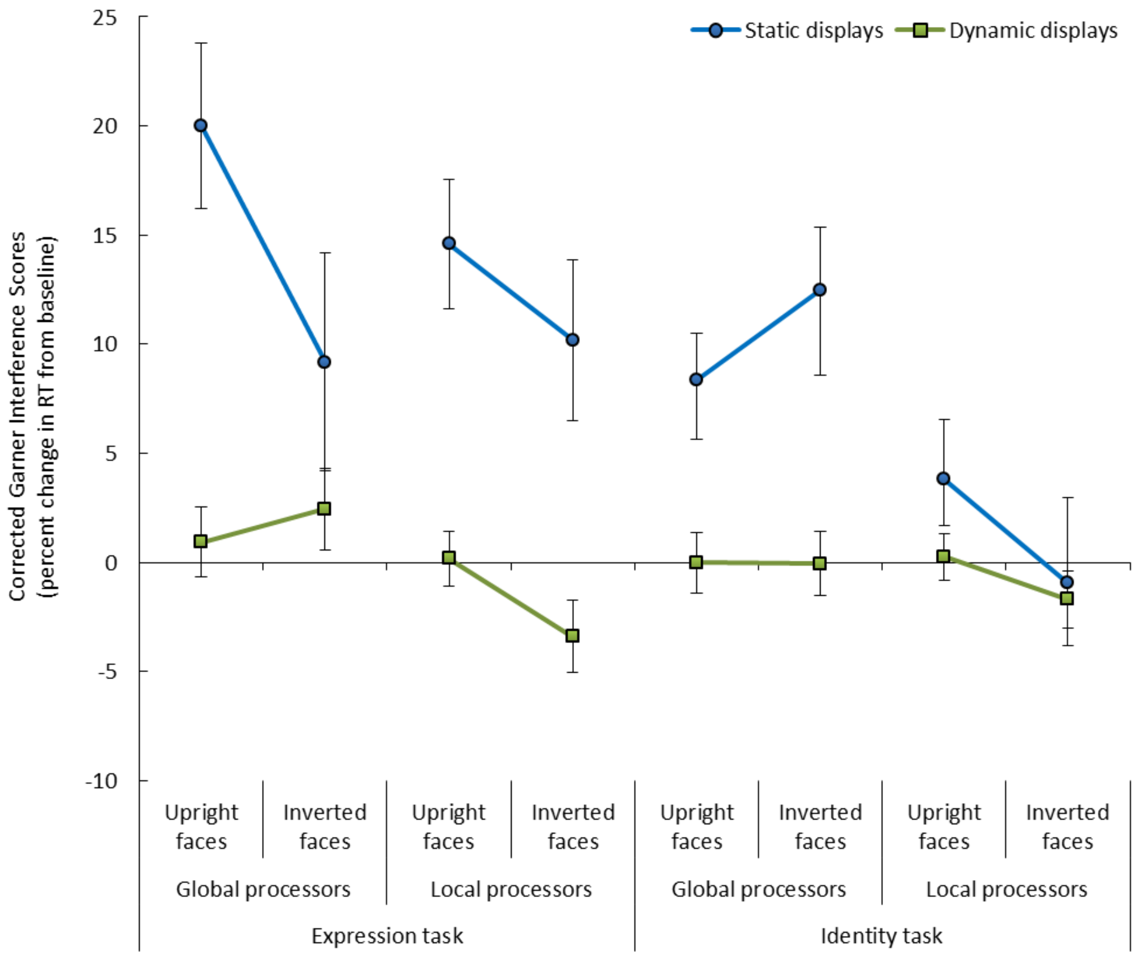

Figure 5.3. Global and local processors' mean corrected Garner interference scores as a function of Task (Expression vs. Identity), Presentation Mode (Static vs. Dynamic), and Orientation (Upright vs. Inverted). Standard error indicated.

My first prediction was that I would replicate earlier findings of asymmetrical and bidirectional interference when viewers made judgments regarding upright static faces. This prediction was supported in that, overall, participants showed nearly three times more interference from the irrelevant cue when judging expression than when judging identity (17.3 vs. 6.1% change from baseline) [Task: $F(1, 48) = 15.29$, $p < .001$, $\eta_p^2 = .24$]. Interestingly, however, local processors' interference was so low in the Identity task that it was not significantly different from zero (one-sample *t*-test, $t(30) = 4.92$, $p < .001$, $d = .88$]. Thus, while interference was asymmetrical in both groups, it was only bidirectional in those with a global processing bias.

My second prediction was that global processors would experience greater interference than local processors with static upright faces, but that this group difference would be eliminated after face inversion. Although the interaction was not significant [$F(1, 48) = 1.26$, $p = .27$], I did find that global processors showed a marginally significant drop in interference during the Expression task after face inversion [$t(18) = 1.86$, $p = .08$], whereas local processors did not [$t(30) = 0.42$, $p = .68$]. For the Identity task, global processors experienced higher levels of interference than local processors overall [$F(1,48) = 7.19$, $p = .01$, $\eta_p^2 = .13$], but, unexpectedly, this did not interact with face orientation [$F(1, 48) = 1.30$, $p = .27$]. Interestingly, one-sample *t*-tests confirmed that local processors did not experience significant amounts of interference in the static Identity task even when faces were presented in the upright orientation.

Third, I expected that participants would show negligible interference when processing dynamic faces. This hypothesis was supported. An ANOVA conducted on interference scores from dynamic trials revealed no significant main effects or interactions, and one-sample *t*-tests confirmed that interference was negligible in all conditions, for both groups.

Finally, I hypothesized that, with upright faces, the drop in interference with the introduction of dynamic cues would be larger for global than for local processors. This prediction was not supported in the Expression task, where both groups showed significant drops in interference [$t > 4.7$, $p < .001$, $d > .84$] that were of comparable magnitude. In the Identity task, however, the drop in interference *was* significant in the global processors [$t(18) = 2.60$, $p = .02$, $d = .60$], but not in local processors (who, as noted above, showed negligible interference even during upright trials).

As in the case of FIE scores, I re-ran the analyses described above on uncorrected Garner interference scores and observed a similar pattern of results.

Discussion

In the present study, my primary goal was to determine the impact of individual differences in perceptual processing biases on FIEs and on interference between expression and identity processing when viewing faces presented statically, or undergoing nonrigid motion. I obtained several key findings, which I discuss in turn, below.

Inversion effects. Past research suggests that a holistic approach may be the optimal strategy to use when extracting both facial identity (Love, Rouder, & Wisniewski, 1999; Richler, Mack, Gauthier, & Palmeri, 2009; Schwarzer et al., 2005) and facial expression (Derntl et al., 2009; Sato, Kochiyama, & Yoshikawa, 2011). In the present study, the strongest evidence for the use of a holistic approach was seen in global processors during baseline trials of the static Expression task. In this block, global processors may have exploited the fact that the metrics of a particular configural cue differed across expressions, making the configural information *diagnostic*. The fact that their FIEs became smaller when task-irrelevant changes in facial identity were introduced during orthogonal testing, and when motion cues were introduced, suggests that global processors shifted their processing style in these blocks to one that was more feature-based. This would make sense because in these conditions the metrics of the configural cues vary across and within trials, respectively, making them less reliable. Additional support for the idea that moving faces are processed in a way that is relatively more feature-based comes from the work of Xiao and colleagues (Xiao et al., 2012, 2013).

If global processing was not being used by many participants except (in the case of global processors) during the baseline trials of the static expression task, then it is perhaps not surprising that inversion effects were not *generally* larger during static than during dynamic trials. Indeed, my finding that inversion effects, though small, were still evident during dynamic testing makes sense, given that viewers should benefit from seeing local facial motion cues in their natural (canonical) orientation. In work examining the processing of body motion, Troje and Westhoff (2006) reported that viewers exhibit small, but reliable, inversion effects when making directional (facing) judgments of point-light walkers, even when the movements of individual dots comprising the walkers are spatially- and phase-scrambled. Because the scrambling eliminates all configural cues to body structure, the inversion effects seen in this situation likely arise from disruption of local biological motion processing.

Monitoring gaze behaviors while participants view upright and inverted faces may offer fresh insights into how global and local processors extract facial cues. When viewing upright static faces, global processors may exhibit gaze patterns that are consistent with holistic processing during baseline testing, with fixations clustering in the region just to the left of the nose (Hsiao & Cottrell, 2008; Schwarzer et al., 2005). In contrast, local processors may show higher concentrations of fixations on specific facial features (Schwarzer et al., 2005; Xu & Tanaka, 2013) during this block, although fixations are unlikely to be confined to single features (Guo, 2012), and the feature(s) that are most diagnostic likely vary with expression (Bimler, Skwarek, & Paramei, 2013; Calvo & Nummenmaa, 2008; Eisenbarth & Alpers, 2011). In addition to exploring expression-specific processing strategies, examining how inverting faces, changing the task demands, or adding dynamic cues alter gaze behaviours would be interesting avenues for future research.

Garner interference effects. As in previous work (Ganel & Goshen-Gottstein, 2004; Schweinberger et al., 1999; Schweinberger & Soukup, 1998; Stoesz & Jakobson, 2013), I found asymmetrical and (largely) bidirectional interference effects when viewers made judgments concerning upright, static faces – with interference from identity during expression processing being stronger than interference from expression during identity processing. This supports the idea that, during static face perception, there is *functional interdependence* between identity and expression processing (see Ganel & Goshen-Gottstein, 2004).

In the static Expression task, face inversion resulted in a marginally significant drop in interference in global processors, but not in local processors. While not a strong effect, this result was in the predicted direction and is consistent with the idea that expression and identity cues can be processed more independently when viewers shift to an approach that is weighted toward local processing. In the static Identity task, global processors experienced similar levels of interference regardless of face orientation, whereas local processors showed no significant interference. This result is consistent with the view that individuals with a strength in local processing can extract identity independent of task-irrelevant changes in expression.

More research is needed to explain why inversion did not reduce interference in the Identity task in participants with global biases, and to gain further insights into the factors that determine the weighing given to global/local strategies when viewers process static identity and expression cues. The limited effect that face inversion had on interference levels, particularly in global processors, may reflect a limitation in the way that I determined viewers' perceptual biases. In the present study, the stimuli used for determining group membership resulted in more viewers being classified as local than as global processors. This is an unusual finding, as in most reports the majority of typical viewers have been found to exhibit global precedence (Kimchi et al., 2009; Love et al., 1999; Navon, 1977; Scherf et al., 2009). It is important to note, however, that processing biases can be reduced or reversed by manipulating the position and saliency of the global and local elements of hierarchical stimuli (Ripoll, Fiere, & Pélissier, 2005). Some work has also revealed that the scores generated from various global-local tasks in the visual domain share relatively little variance (Dale & Arnell, 2013; Milne, Szczerbinski, & Sheffield, 2009). Thus, it would be of interest to use another global-local task to determine group membership and see whether a clearer pattern of results with regards to inversion and interference would emerge.

Importantly, I replicated my previous finding that interference was negligible during dynamic testing (see Chapters 3 and 4) (Stoesz & Jakobson, 2013). I then showed that, with upright faces, the drop in interference with the addition of motion cues was significant for both groups in the Expression task, and for global processors in the Identity task. As I predicted, viewers also took longer to make expression and identity judgments with dynamic compared to static displays. They may have needed more time to formulate hypotheses about the rich information contained in dynamic stimuli (Hegdé, 2008), but slower RTs are also expected when viewers shift to an approach more heavily weighted toward local processing (see also Love et al., 1999; Richler et al., 2009). It may be preferable to process moving faces at a local level given that, both within and across individual faces, moment-to-moment changes in the positioning of facial features can alter the precise metrics of a particular configural cue, making the metric information more difficult to compute. Given this, we might expect all viewers to approach dynamic identity *and* expression judgment tasks using a strategy that is weighted toward local processing. By directing one's attention to single features one would be less susceptible to distraction from task-irrelevant facial cues, resulting in low levels of interference.

Early work by Bassili (1979) supports the idea that viewers focus on changes that are occurring in a single region when attempting to make certain expression judgments about moving faces. Specifically, he found that movements of the upper face were most important for judging anger and fear, and that movements in the lower region of the face were particularly useful for judging happiness and disgust. However, even if tracking the movement of a particular feature is helpful, viewers may not fixate longer, or more frequently, on that feature when processing moving faces. Interestingly, I have found in other work that viewers exhibit more off-face fixations when passively viewing scenes containing dynamic, as opposed to static, faces (see Chapter 2) (Stoesz & Jakobson, 2014). This may reflect efforts to reduce both the cognitive load and the emotional (Sato et al., 2008, 2010) and physiological arousal (Skwerer et al., 2009; Soussignan et al., 2013) associated with viewing information-rich, dynamic displays. Viewers may find it especially difficult to maintain fixation on a dynamic face when viewing the unfolding of an angry or threatening expression (Terburg, Hooiveld, Aarts, Kenemans, & van Honk, 2011), opting instead to glance at a diagnostic feature (e.g., furrowed brow), look away, and then look back at the feature to re-examine it – steps that would increase RT. Consistent with this, supplementary analysis of the data from dynamic trials of the present study confirmed that processing angry expressions took significantly longer than processing surprised expressions (M_{angry} = 1057, *SD* = 138; $M_{surprised}$ = 997, *SD* = 159; $t(45) = 5.11$, $p < .001$). Future eye-tracking studies may shed light on the steps that viewers take when making judgments about different facial expressions.

Although the data presented here are largely consistent with the view that interference is reduced in dynamic testing conditions because viewers adopt a more feature-based style of processing, the drop in interference I observed could also reflect better integration of identity and expression cues under these conditions. Such integration would be useful in the natural world, particularly when the viewer is familiar with the individual (O’Toole et al., 2002), when the faces we are trying to distinguish between are very similar (e.g., siblings who share a strong family resemblance) or have characteristics that differ from those with which we are most familiar (e.g., "other races"), or when viewing conditions are suboptimal (Knight & Johnston, 1997; Lander et al., 1999). Under these circumstances, identity extraction might be enhanced by detection of an idiosyncratic movement signature (*supplemental information hypothesis*). Accessing this person-specific information might also help one to interpret subtle or ambiguous facial expressions, and to make to higher-order judgments such as whether a particular expression is genuine or posed, for example. O'Toole and colleagues (O'Toole et al., 2002; Roark et al., 2003) have suggested that the pSTS is involved in extracting these signatures, thereby providing an important, secondary route to identity perception. These authors have also suggested that, when we view a moving face, dorsal stream regions implicated in structure-from-motion processing (e.g., area MT) can provide information about its three dimensional structure (O'Toole et al., 2002). According to the *representation enhancement hypothesis,* the benefits of utilizing this kind of information for identity extraction should be evident even when the face is unfamiliar to the viewer (Christie & Bruce, 1998; Pike et al., 1997). Functional imaging studies are needed to determine if the drops in interference seen with the introduction of dynamic cues correspond with increased recruitment of areas such as STS and MT.

Conclusions. I replicated my earlier finding that different patterns of interference between the processing of expression and identity emerge when viewers process static and dynamic faces, with interference seen in the former but not the latter case (see Chapters 3 and 4) (Stoesz & Jakobson, 2013). Reduced interference with moving faces may occur because the processing of naturally moving faces is more heavily weighted towards a part-based approach. I acknowledge, however, that the switch from a holistic to a part-based strategy is unlikely to occur in an all-or-none fashion (see also Xiao et al., 2013). The shifting of processing strategy is likely to occur on a continuum, and is likely to be task- and stimulus-dependent. Importantly, the present findings emphasize that the mechanisms used for facial motion processing are different from those involved in static face perception (Longmore & Tree, 2013), particularly for global processors. Thus, further research exploring the role that individual difference factors play in face processing (Samson, Fiori-Duharcourt, Doré-Mazars, Lemoine, & Vergilino-Perez, 2014), and utilizing ecologically relevant, dynamic facial stimuli (Kilts et al., 2003), is essential for advancing face processing research.

CHAPTER 6: CONCLUSIONS

The ability to attend to faces and then process and integrate various facial cues, such as identity and expression, is essential for successfully navigating the social world. These abilities appear to be, in some respects, separate tasks for the perceptual system (Bruce & Young, 1986; Haxby et al., 2000; Tovee & Cohen-Tovee, 1993). Indeed, social, experimental, neuroimaging, and clinical studies have demonstrated that these aspects of face processing can be carried out more-or-less independently. Although current models of face processing have had considerable value in explaining how we process faces, much of what is known about face processing comes from research using static face stimuli. Static images of faces may imply motion (e.g., averted eyes may imply that a shift in gaze has occurred) (Langton, Watt, & Bruce, 2000), but do not allow the visual system to represent naturalistic movement trajectories (Furl et al., 2010), or form robust predictive models that allow for one to anticipate another person's feelings and behaviours. This is a major limitation in the field as the faces we see outside of the laboratory, in our everyday lives, are often moving – the people around us are expressing emotions, talking, and looking in one direction or another. Moment-to-moment changes in the shape and features of the face serve as important signals regarding another person's state of mind (Barton, 2003; Kamachi et al., 2001). Natural facial movements also help us to understand and anticipate another person's actions and intentions. Given the importance of facial motion, the reliance on static faces in studies designed to elucidate mechanisms behind face processing is surprising. Thus, understanding the cognitive mechanisms involved in processing naturally moving faces was the common theme threaded throughout this thesis.

Synthesis of Study Findings

Researchers have demonstrated that adults can use motion cues for recognizing facial identity (Knappmeyer et al., 2003; Pilz et al., 2006; Thornton & Kourtzi, 2002) and facial expressions (Ambadar et al., 2005; Arsalidou et al., 2011; Back et al., 2009). The extent to which children are able to use dynamic facial cues is largely unknown. Thus, one objective of this thesis was to examine age-related differences in how attention is passively directed to faces in static and dynamic scenes (Chapter 2). From here, I went on to explore how well children, adolescents, and adults are able to attend selectively to specific cues when making judgments regarding static and dynamic faces (Chapter 3). The task utilized in this latter study was one that both children and adolescents could complete with relatively high accuracy, so that their RTs could be compared directly to those of adults. Nonetheless, in this study I corrected for group differences in processing speed when calculating Garner interference scores. These features represent distinct strengths of the current work, as previous studies comparing the performance of children and adults on various aspects of face processing have not always taken these factors into account (Crookes & McKone, 2009; McKone et al., 2012).

In the work described in Chapter 2, I used eye-tracking technology to determine whether children directed their attention toward faces in the same way that adolescents and adults did when viewing complex static and dynamic visual scenes involving one or more people. I found that all participants attended to faces more than to any other parts of the scenes; however, when viewing single-character scenes, the addition of motion cues was associated with attention being directed away from the faces. When participants viewed scenes with multiple characters, gaze also shifted away from faces. Importantly, children were most affected by the introduction of movement or additional characters. Thus, children appeared to have difficulty processing dynamic faces, especially when viewing complex social scenes, and may have looked away from faces because they were attempting to reduce their cognitive and affective load. These results are consistent with the developmental literature showing an early competence for face processing that is refined with age (see Pascalis et al., 2011 for a review). This refinement in face processing abilities has largely been explained by a long period of maturation in face-specific mechanisms (Carey et al., 1980; Carey & Diamond, 1977; Ellis & Flin, 1990; Mondloch et al., 2002; Scherf et al., 2007). However, some researchers argue that age-related differences in face-processing are due to the development of general cognitive abilities, rather, than being domain specific (Baenniger, 1994; Crookes & McKone, 2009; McKone et al., 2012). I return to this point below.

Although I found that children looked away from faces more than adults when passively viewing scenes, I found little evidence to suggest that, when engaged in an active face processing task (Chapter 3), they process faces in a manner that is *qualitatively* different from the way adults process faces. Despite the fact that children were generally less accurate and slower than adults at processing faces during baseline trials of Garner's selective attention tasks, children, adolescents, and adults showed similar interference effects. Specifically, all participants were generally unable to attend selectively to expression or identity cues in *static* faces – a finding that is consistent with other research using the Garner paradigm (Baudouin et al., 2008; Ganel & Goshen-Gottstein, 2004). One suggestion in the literature is that increases in accuracy and speed with age, seen in the absence of age-related differences in holistic-face-processing effects, reflect age-related improvements in general sensory and/or cognitive functions (Crookes & McKone, 2009). For example, age-related improvements in visual acuity (Skoczenski & Norcia, 2002), sustained attention (Betts et al., 2006), and the ability to narrow the focus of visual attention (Pastò & Burack, 1997) have been reported. Data from ERP and fMRI studies offer additional support for this view. For example, Kuefner, de Heering, Jacques, Palmero-soler, and Rossion

(2010) found that the amplitude, latency, or topography of P1 and N170 did not differ for four types of stimuli (i.e., face, car, scrambled face, scrambled car) across development (ages 4-17 years and adults). Thus, it may be that age-related differences on face processing tasks reflect a general developmental trend that is not specific to faces. Additional research is required to test the opposing developmental theories of face processing. This work is important if we are to correctly interpret the data obtained from individuals in atypical populations, such as those born prematurely or with a diagnosis of ASD.

Findings of bidirectional interference between expression and identity processing with static faces (Chapters 3-5) contradict influential face processing models suggesting that these cues are processed in parallel (independent) systems (Bruce & Young, 1986; Tovée & Cohen-Tovée, 1993). Because interference from identity when processing expression is nearly always observed, but expression does not always interfere with the processing of identity, some researchers have proposed a parallel-dependent model of face processing (Calder, Burton, Miller, Young, & Akamatsu, 2001; Ellamil, Susskind, & Anderson, 2008; Haxby et al., 2000; Martens, Leuthold, & Schweinberger, 2010). This model allows for some cross-talk between the distinct neurological systems supporting the processing of expression and identity cues. The direction of the asymmetric dependencies is consistent with the assumption that invariant identity cues are more likely to provide a reference for computing information about expression (and other changeable facial information, such as speech and eye gaze) than vice versa (Martens et al., 2010). However, in the present thesis, this interpretation only held for interference effects with static faces. With dynamic faces, I consistently observed that interference effects were negligible in both Expression and Identity tasks, even in young children. I return to this point, below.

Having investigated the impact of viewer age on performance in the Garner task, in the remaining chapters of this thesis I went on to explore how other individual difference factors, such as sex and perceptual bias, would affect interference between the processing of expression and identity in adults (Chapters 4 and 5). I found that only women experienced asymmetrical interference with static faces, with variations in identity affecting expression judgments more than the opposite (Chapter 4). Men, in contrast, showed equivalent, and generally low, levels of interference from irrelevant cues in both tasks. As described in Chapter 4, it may be that women attend more closely to multiple nonverbal cues when processing other individuals' emotional states in an effort to empathize with them (Baron-Cohen et al., 2001). The sex differences that I observed on my behavioural tasks may also be related to the finding that women and men show different patterns of neural activation when they are completing emotion processing tasks (Derntl et al., 2010). Interestingly, I found no sex differences in interference effects with dynamic faces. Little work has been done to examine sex differences in neural activation during dynamic face processing. This is an important avenue for future research, especially when one considers that women are more accurate and quicker than men in discriminating between globally coherent biological motion patterns and scrambled motion, and in recognizing the 'emotional state' of male and female actors depicted in point-light displays (Alaerts et al., 2011).

In a recent eye-tracking study, Heisz, Pottruff, and Shore (2013) reported that women outperformed men on a face recognition-memory test. They speculated that this female advantage was directly related to the finding that women made more fixations when initially encoding faces. Interestingly, sex differences in face recognition and scanning patterns decreased when participants viewed the faces more often. This pattern suggests that some underlying perceptual process may drive sex differences in performance on face processing tasks. In Chapter 5, I explored whether individuals' perceptual processing style (as measured by the Navon task) would affect the magnitude of the FIE, or the amount of interference, that they experienced when making face judgments. Compared to their performance with dynamic faces, global processors displayed larger FIEs when judging the expressions seen in static faces, but only when identity cues were held constant (i.e., during baseline testing). After face inversion, global processors also showed a small drop in interference during the static Expression task – an effect that was not evident in local processors. During the static Identity task, global processors showed interference from expression cues, while local processors did not. These findings offer general support for the view that, during static face perception, individual differences in processing style influence how selectively (i.e., how independently) different facial cues can be processed.

Despite the interesting findings summarized above, I found no evidence of interference between facial identity and expression cues with dynamic faces in any of the work described in Chapters 3-5. This raises an important question. Does the absence of interference effects seen with moving faces provide evidence for independent processing of identity and expression cues in moving faces, or for more efficient integration of these cues? I have suggested in this thesis that greater independence in processing these cues might be achieved if viewers shift toward an approach that is more heavily weighted toward feature-based processing when analyzing moving faces, a conclusion supported by work in other laboratories (see Xiao et al., 2012, 2013). It is important to note that I do not believe that this shift occurs in an all-or-none fashion; rather, I speculate that the shift likely occurs on a continuum, until a balance is reached. Adopting a balanced approach to process naturally moving faces may allow for more efficient integration of multiple facial cues within or across particular facial features or regions, depending on the task and the nature of the stimuli. Recent imaging work provides some support for this hypothesis.

Tsuchiya, Kawasaki, Oya, Howard, and Adolphs (2008) recorded neuronal activity directly from the cortical surface in nine neurosurgical patients while they viewed static and dynamic facial expressions. They found that, compared to activation in the lateral temporal cortex, increased activation for both invariant and changeable aspects of faces occurred in a single region in the middle fusiform gyrus, suggesting a location for possible integration of these cues. In contrast, other research suggests that the pSTS plays a critical role in face perception because of its role in processing social perception in general. This region has been implicated in the processing of motion and various facial cues, including identity and expression (Hein & Knight, 2008; Puce et al., 2003; Rossion, Hanseeuw, & Dricot, 2012). Results from human studies mirror those seen in new and old world monkeys in that parts of the non-human primate STS (see Ungerleider & Haxby, 1994 for a review) receive input from the ventral object recognition system (the 'what' system) and from the dorsal spatial location–movement system (the 'where' system) (see Milner & Goodale, 1995), providing additional support for the idea that this region integrates form and movement cues (Oram & Perrett, 1996). Regardless of whether the FFA (Tsuchiya et al., 2008) or the pSTS (Allison et al., 2000; Lahnakoski et al., 2012) is primarily responsible for processing social information from *moving* faces, if the same brain region processes both cues cross-talk between brain regions would be unnecessary and interference would, as a result, be negligible. Future imaging work could test this further by comparing brain activations in both regions while participants complete Garner's selective attention tasks involving static and dynamic faces.

Limitations

The present series of studies have some potential limitations that could be addressed in future investigations. First, the scenes I used as stimuli for the study described in Chapter 2 were carefully selected because individual shots in these scenes were generally much longer than those seen in contemporary television program. It is possible, however, that outdated features of the scenes (e.g., rotary dial telephone, hair and clothing styles, etc.) affected how viewers' deployed their attention. Thus, it is possible that the use of a more recently recorded television program or the use of a video recording of a natural event would have resulted in a different pattern of results. Perhaps, use of a familiar television program would have resulted in more fixations on the faces if the actors' faces were familiar to the viewers, and this may also vary with age group. Alternatively, if the objects in a modern scene were more attention-grabbing (e.g., colourful toys, presence of particular food items, technology), there would have been fewer fixations on the actors' faces. Related to this point, the background of modern day television programs typically contain many more objects (and often particular objects for advertising purposes) than are seen in older television programs or in many natural settings. To follow up on the work described in Chapter 2, it would be interesting to examine gaze behaviours across recording type (i.e., older programs, modern programs, or natural setting).

A second limitation of the work described in all of the dynamic stimuli used in this thesis relates to the impoverished nature of video recordings and computer animations. While these kinds of stimuli are richer than static displays, there is still visual information missing from them that is present in the real world. Thus, although two dimensional recordings generally provide enough information (e.g., motion, perspective, relative size, occlusion, lighting, shading, edges) so that viewers can comfortably perceive depth, some information (such as texture and some depth cues) may not be directly related to the actual depth structure of a natural scene. Moreover, in the video clips used in the work described in Chapter 2, the focus of the camera limits where viewers can direct their attention within a scene. In studies of scene perception, recording participants' looking behaviours using eye tracking glasses with wireless control while they are either seated or moving within a natural setting may partly overcome these limitations. During perceptual studies such as those described in Chapters 3-5, the use of virtual reality might allow one to generate more compelling, three-dimensional displays.

Third, it is possible that the asymmetry in interference effects (i.e., more interference from identity when processing expression than the reverse) that I observed with static faces (Chapter 3-5) was due to a stimulus artifact. Although accuracy was at ceiling for both Expression and Identity tasks, I consistently observed an identity advantage for RTs during baseline testing, suggesting that viewers found it easier to extract identity cues than expression cues (i.e., that expression was less discriminable). This may have contributed to the asymmetrical pattern of interference I observed. Indeed, although more interference has been reported when viewers judge expression than identity when the physical discriminability of faces is not controlled for (Baudouin et al., 2008; Schweinberger et al., 1999), the direction of this effect can be reversed when discriminability is manipulated (Ganel & Goshen-Gottstein, 2004). For example, when Ganel and Goshen-Gottstein presented faces of people who looked very similar to one another, participants made expression judgments faster than identity judgments, and experienced greater interference from expression when making identity judgments than the reverse. These results provide support for the idea that a viewer can use the shape of an individual's face to aid them in making judgments about how particular emotions are expressed (*representation enhancement hypothesis*). Given the above, in future work it would be advisable to equate or directly manipulate task difficulty in investigations of this sort.

Fourth, the conclusions drawn in Chapters 3-5 of this thesis that relate to expression processing apply only to faces depicting *surprised* and *angry* expressions. Future research could examine the robustness of the effects I observed by carrying out studies in which various combinations of expressions are presented. The results of this work may show that the asymmetric pattern of interference generalizes to faces within the whole range of expressions. It is also possible, however, that patterns of interference seen with neutral faces and faces depicting expressions of happiness would different from those seen with faces depicting negative emotions. The basis for this prediction comes from various behavioural and imaging studies reporting similar patterns of results across negative valence expressions (e.g., fear, surprise, and anger), which are often different from those seen with neutral or positive valence expressions (e.g., happy) (Harris, Young, & Andrews, 2014; Kamachi et al., 2001; Kilts et al., 2003; Lassalle & Itier, 2013; Vrticka, Lordier, Bediou, & Sander, 2014). For example, facial expressions of fear and anxiety elicit greater activation in the amygdala than do joy and wonderment (Vrticka et al., 2014). In a similar vein, it would also be interesting to test whether interference effects would be seen if the faces of children or older adults were used as test stimuli. Previous research has shown that young adults (ages 20-29 years) rate the faces of older adults as being less distinctive than the faces of younger adults (Ebner, 2008). If this is true, one might expect that participants would experience larger interference effects from expression when judging the identity of older compared to younger adults, at least with static stimuli. To my knowledge, only the faces of young or middle-aged adults (male and female) have been used as stimuli in studies examining interference effects.

Finally, the present series of studies does not provide information about the time course of expression and identity processing, or about the specific stage of information processing where the interaction between these two facial cues takes place. Although some work has been done to address these two questions (Martens et al., 2010; Wild-Wall, Dimigen, & Sommer, 2008), these previous studies did not use dynamic face stimuli. Thus, future ERP studies could explore when in time the integration of cues in dynamic faces occurs.

Future Research

Although previous research has provided evidence that static faces are generally processed globally (Bartlett & Searcy, 1993; Farah, Tanaka, & Drain, 1995; McKone et al., 2007; Mondloch, Le Grand, & Maurer, 2002; Tanaka & Farah, 1993), the current research suggests that the likelihood of adopting this approach is affected by the specific task and stimuli. Less is known about the strategies required for processing faces in motion. If we shift more of our attention to local facial features (Xiao et al., 2012, 2013) and ignore the global context when viewing moving faces, this might explain why we are less influenced by task-irrelevant identity or expression cues. There is a need for more research exploring this idea. Thus, in the future, I plan to examine this question by using manipulations known to disrupt global (e.g., misalignment) and local processing (e.g., blurring). If viewers are more distracted by global information with static faces than with dynamic faces, they should show larger composite face effects (in terms of Garner interference) with static than dynamic displays. Blurring may result in increases in Garner interference between identity and expression with both presentation modes. Finally, it would be interesting to examine viewers' cognitive strategies for processing faces by recording eye movements. In simple recognition tasks involving static faces, viewers' eyes follow a systematic triangular sequence that locally samples the eyes and mouth (Althoff & Cohen, 1999), but when sampling global information, the center of the face (i.e. the nose) is fixated (Hsiao & Cottrell, 2008). The relationship between the type of gaze pattern and interference effects between various types of facial information, particularly identity and expression, when viewers are processing static and dynamic faces is currently unknown.

Future research exploring how individuals in various clinical populations known to have deficits in social function and communication attend selectively to specific facial cues is important. Below, I describe possible avenues for research involving individuals with ASD and individuals born prematurely.

Autism spectrum disorders. Research has shown that a contributing factor in social deficits in the ASD population is that individuals with ASD exhibit atypical face processing abilities (see Barton et al., 2007 for a review). Given this, Krebs et al. (2011) used a Garner's selective attention task to examine whether deficits in processing expression would affect identity judgments, or vice versa, in children with ASD. Unlike typical children, the children with ASD processed both types of static facial cues independently of each other – that is, they showed no interference effects. The authors suggested that different neuronal mechanisms are responsible for the interactive processing (or lack thereof) of invariant and changeable facial cues in typical children and children with ASD. An interesting question for future research would be to explore the interference effects in children and adults with ASD using dynamic stimuli. As I argued above, the lack of interference with dynamic displays may indicate that motion cues lead participants to shift their processing approach to one that is more feature-based or balanced. If this is the case, one might predict that individuals with ASD, who tend to be overly focused on small details (Happé & Frith, 2006; Mottron & Belleville, 1993; Shah & Frith, 1993), would show negligible levels of interference. On the other hand, significant interference with dynamic faces might be expected in this clinical group because individuals with ASD often show marked deficits in various tasks requiring the *integration* of social cues including those involving hybrid low- and high-spatial frequency facial stimuli (Corradi-Dell'acqua et al., 2014), visible and audible speech (Stevenson et al., 2013), and dynamic facial and vocal expressions (Charbonneau et al., 2013). Finally, because individuals with ASD have been shown to look at moving faces in scenes differently than typical individuals (Klin et al., 2002; Speer et al., 2007), it would be interesting to compare the looking behaviours in the two groups during selective attention tasks involving dynamic faces.

Preterm children. One area of face processing that has not received much attention in the literature, to date, is the abilities of individuals born prematurely at low birth weight to process faces. This is of interest because children born prematurely at extremely low birth weight (< 1000 g) with even mild brain damage show significant deficits on tasks requiring visual attention (Foreman, Fielder, Minshell, Hurrion, & Sergienko, 1997; Rose, Feldman, & Jankowski, 2001), visual navigation (Pavlova, Sokolov, & Krägeloh-Mann, 2007), and visuomotor skill (Foreman et al., 1997; Jakobson, Frisk, & Downie, 2006; Jakobson, Frisk, Knight, Downie, & Whyte, 2001). These children also have difficulty processing global motion and global cues signalling structure-from-motion in biological motion (point-light) displays (N. M. Taylor, Jakobson, Maurer, & Lewis, 2009; Williamson, Jakobson, Saunders, & Troje, n.d.). Performance on motor-free tasks of visual perception (e.g., static form perception), however, appears to be largely intact (Foreman et al., 1997; Jakobson et al., 2001; MacKay et al., 2005). Recently, Williamson and Jakobson (2014) reported that children born prematurely showed impairments in their ability to interpret the nonverbal face and body cues of people engaged in naturalistic social situations. Thus, one might predict that children and adults in this population would show greater deficits on tasks that involve processing the changeable aspects of faces (expression, visual speech, gaze) than on those that involve processing invariant facial information (age, sex, identity). Deficits on tasks involving the changeable aspects of faces may be particularly marked when dynamic face stimuli are used. Future research should test these possibilities.

Conclusion

Taken together, the results that emerged from the experiments described in this dissertation highlight the importance of incorporating naturalistic, dynamic facial displays when characterizing typical face processing mechanisms (Kilts et al., 2003). In particular, the results emphasize that the approach viewers use to process photographs of social stimuli may not be the same as that applied to moving social stimuli. The findings also highlight the importance of exploring the role that individual difference factors (e.g., age, sex, perceptual processing biases) play in face processing (Samson et al., 2014). The basic research findings from this dissertation may provide information for advancing research and development in face-recognition software. Understanding the typical face processing system in the dynamic realm also lays the groundwork for future research examining the atypical face processing abilities in clinical populations such as individuals with ASD (Barton et al., 2007), developmental prosopagnosia (Barton, 2003), schizophrenia (Marwick & Hall, 2008), and Williams syndrome (Riby et al., 2008). This is important work, as results from these types of studies could eventually lead to the development of better diagnostic and intervention tools.

APPENDIX A

Our Ref: LA/PCGN/P999

09 May 2014

Dear Brenda M Stoesz,

Thank you for your correspondence requesting permission to reproduce the following material from our Journal in your printed thesis at University of Manitoba Libraries entitled 'SELECTIVE ATTENTION TO STATIC AND DYNAMIC FACES AND FACIAL CUES DURING DEVELOPMENT' and to be posted in your university's repository.

Material Requested: 1 Figure – Fig 2 from "The neural substrates of face processing models: A review", by Martin J. Tovée, Esther M. Cohen-Tovée, published in Cognitive Neuropsychology, Vol10:6, pp505-528(1993)

We will be pleased to grant entirely free permission on the condition that you acknowledge the original source of publication and insert a reference to the Journal's web site: www.tandfonline.com

Please note that this licence does not allow you to post our content on any third party websites or repositories.

Thank you for your interest in our Journal.

Yours sincerely

Lee-Ann

Lee-Ann Anderson – Permissions Administrator
Taylor & Francis Group
3 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN, UK
Tel: +44 (0)20 7017 7932
Web: www.tandfonline.com
e-mail: lee-ann.anderson@tandf.co.uk

APPENDIX B

UNIVERSITY OF MANITOBA | OFFICE OF RESEARCH SERVICES
Office of the Vice-President (Research)

CTC Building
208 - 194 Dafoe Road
Winnipeg, MB R3T 2N2
Fax (204) 269-7173
www.umanitoba.ca/research

APPROVAL CERTIFICATE

November 9, 2011 NSERC

TO: **Brenda M. Stoesz** (Advisor L. Jakobson)
Principal Investigator

FROM: **Bruce Tefft, Chair**
Psychology/Sociology Research Ethics Board (PSREB)

Re: **Protocol #P2011:072**
"1. Static and Dynamic Face Processing in Typical Development and in Prematurity
2. Development of Form and Motion Processing"

Please be advised that your above-referenced protocol, as revised, has received human ethics approval by the **Psychology/Sociology Research Ethics Board**, which is organized and operates according to the Tri-Council Policy Statement (2). This approval has been issued based on your agreement with the change(s) to your original protocol required by the PSREB. It is the researcher's responsibility to comply with any copyright requirements. This approval is valid for one year only.

Any significant changes of the protocol and/or informed consent form should be reported to the Human Ethics Secretariat in advance of implementation of such changes.

Please note:

- **If you have funds pending human ethics approval, the auditor requires that you submit a copy of this Approval Certificate to the Office of Research Services, fax 261-0325 - please include the name of the funding agency and your UM Project number. This must be faxed before your account can be accessed.**

- **If you have received multi-year funding for this research, responsibility lies with you to apply for and obtain Renewal Approval at the expiry of the initial one-year approval; otherwise the account will be locked.**

The Research Quality Management Office may request to review research documentation from this project to demonstrate compliance with this approved protocol and the University of Manitoba Ethics of Research Involving Humans.

The Research Ethics Board requests a final report for your study (available at: http://umanitoba.ca/research/ors/ethics/ors_ethics_human_REB_forms_guidelines.html) **in order to be in compliance with Tri-Council Guidelines.**

Bringing Research to Life

APPENDIX C

Sample consent form

UNIVERSITY OF MANITOBA | Department of Psychology

190 Dysart Road
Winnipeg, Manitoba
Canada, R3T 2N2
Telephone (204) 474-9338
Fax (204) 474-7599

Research Project Title: Static and Dynamic Face Processing in Typical Development and in Prematurity

Adult Full-term Group (Introduction to Psychology Students)

Principal Investigator
Mrs. Brenda M. Stoesz, BEd, BSc, MA
PhD Candidate, Department of Psychology
Phone Number: 204-474-8354
Email Address: umstoes3@cc.umanitoba.ca

Research Supervisor
Dr. Lorna S. Jakobson, PhD
Professor of Psychology
Phone Number: 204-474-6980
Email Address: jakobson@cc.umanitoba.ca

Dear Participant,

This consent form, a copy of which will be left with you for your records and reference, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

Purpose of Research: The purpose of this research study is to examine our ability to recognize people, expressions, speech movements, and gaze direction in childhood and adulthood. The second part of my research involves looking at face processing in children and adults born preterm and at low birth weight. Past research has shown that preterm children sometimes have more trouble with certain aspects of their vision than children born full-term.

In this research study, information will be collected about the participant and their parents/legal guardians. You will have your vision checked with several eye tests, classify pictures, match and identify faces and facial expressions, look at and listen to people speaking, and examine scenes from a TV show presented on a computer screen while eye movements are recorded. Participants usually find most of these activities quite fun, and none of them take very long. You will also be asked to read and complete a *General Information Questionnaire* about parental education and your development. The entire testing session will be about **1.5 hours** and we will provide you with breaks when necessary throughout the testing, so that you do not become too tired.

Risks and Benefits: Taking part in this study will help you learn about the kind of research that psychologists do. The results will help us learn more about how vision develops, and specifically, how we look at faces and obtain information from faces. If we can figure out why many preterm children have difficulty with their vision, we may be able to think of better ways to help them develop their visual skills.

There is no known harm related to the completion of these tasks, however, if you feel uncomfortable and wish to discontinue the experiment at any point in time, you are free to do so.

Confidentiality: Confidentiality will be respected and no information about individuals will be released or published without consent unless required by law. The results of the tests described above will be used for research purposes only in the context of the study. The only people who will have access to this information are Dr. Jakobson and her students. All information will be stored in a secure place (locked filing cabinet and password protected electronic database) for up to 5 years post-publication of the results for paper forms and for an indefinite period in electronic form, in accordance with the guidelines of the University of Manitoba Research Ethics Review Boards. We will destroy identifying information (information that could be used to link a given participants to the electronic database) as soon as it is no longer needed. At no time will individual responses be reported.

If you would be interested in hearing about future opportunities to participate in research in our laboratory, however, we will ask for your permission to maintain your name and contact information in a secure place for up to 5 years. If you agree to this, you should know that you would be able to withdraw your consent at any point during this time, at which point we would destroy any records containing identifying information and remove your name and your child's name from our contact list.

Reimbursement: As a participant recruited from the Introduction to Psychology Participant Pool, you will receive **3 experimental credit**s for your **1.5 hour**s of participation in this study.

Participation: Your signature on this form indicates that you have understood, to your satisfaction, the information regarding participation in the research project, and agree to participate. In no way does this waive your legal rights nor release the researchers, sponsors, or involved institution from their legal and professional responsibilities. Participation in this study is voluntary. You are free to withdraw from the study at any time, and/or refrain from answering any questions you prefer to omit, without prejudice or negative consequences. Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation.

The University of Manitoba Research Ethics Board(s) and a representative(s) of the University of Manitoba Research Quality Management/Assurance office may also require access to your research records for safety and quality assurance purposes.

This research has been approved by the Psychology/Sociology Research Ethics Board at the University of Manitoba. If you have any concerns or complaints about this project you may contact any of the above-named persons or the Human Ethics Coordinator (HEC) at 474-7122. A copy of this consent form has been given to you to keep for your records and reference.

____
Participant's Name (PLEASE PRINT)

______________________________ ________________________
Participant's Signature Date

______________________________ ________________________
Researcher's or Delegate's Signature Date

APPENDIX D

Experiment in Chapter 2: Static and Dynamic Scenes with Single and Multiple Characters Eye-tracking Task

Verbal instructions to participants: "In this next experiment, a special camera will be recording where you look on the screen as you watch a two and a half minute series of pictures and short movies. Each picture or short movie will stay on the screen for 4 seconds and after each picture or short movie, you will see a plus sign (or crosshair) for 2 seconds. In this experiment, you will not press any buttons or make any decisions about what you see. Simply look at each of the pictures and short movies. Try not to look away from the screen until the experiment is over.

"Before I begin the experiment, I need to make sure that the camera can "see" your eyes. I will show you a black dot on a white background. The dot will move slowly to nine different spots on the screen. At each spot, the dot will look like its growing and shrinking before moving to the next spot on the screen. The computer will let me know if the camera "saw" your eyes. If the camera "saw" your eyes, we will start the experiment. If the camera did not "see" your eyes, then I will show the black dot moving on the white background to you gain."

"Do you have any questions before we begin?"

APPENDIX E

Static and dynamic face matching task: Chapters 3 and 4

Verbal instructions to participants: "In this experiment, you will see two pictures or two movies of the faces of actors saying a word (that you won't hear) on the screen for a very short time -- about 1 sec. You will look at the two faces on the screen and decide if the two faces are of the same person or are of different people. Try to be as accurate as you can be. You will press one key if the two faces are of the same person and press another key if the two faces are of different people. Before we start the experiment, I will let you practice so that you know which keys to press. If you have any questions, you may ask them while you are practicing."

APPENDIX F

Garner's selective attention tasks: Chapters 3, 4, and 5

Verbal instructions to participants: "This experiment has two parts."

Identity judgment task instructions: "In one part of this experiment, you will see one picture or one movie of an actor on the screen for a very short time -- about 1 sec. You have seen these two faces before in the previous experiment. You will look at the face on the screen and decide if the face you see is Anne or Jane. Try to be as accurate as you can be. You will press one key if the face is Anne and another key if the face is Jane. Before you start the experiment, I will let you practice a few times so that you know which keys to press. If you have any questions, you may ask them while you are practicing."

Expression judgment task instructions: "In another part of this experiment, you will see one picture or one movie of an actor on the screen for a very short time -- about 1 sec. This time, when you see the face, you will decide if the face is showing a surprised expression or an angry expression. Try to be as accurate as you can be. You will press one key if the face is surprised and another key if the face is angry. Before you start the experiment, I will let you practice a few times so that you know which keys to press. If you have any questions, you may ask them while you are practicing."

REFERENCES

Açık, A., Sarwary, A., Schultze-Kraft, R., Onat, S., & König, P. (2010). Developmental changes in natural viewing behavior: bottom-up and top-down differences between children, young adults and older adults. *Frontiers in Psychology*, *1*(November), 207. doi:10.3389/fpsyg.2010.00207

Aguirre, G. K., Singh, R., & D'Esposito, M. (1999). Stimulus inversion and the responses of face and object-sensitive cortical areas. *Neuroreport*, *10*(1), 189–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10094160

Alaerts, K., Nackaerts, E., Meyns, P., Swinnen, S. P., & Wenderoth, N. (2011). Action and emotion recognition from point light displays: An investigation of gender differences. *PloS One*, *6*(6). doi:10.1371/journal.pone.0020989

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences*, *4*(7), 267–278. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10859571

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memeory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *25*, 997–1010.

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science*, *16*(5), 403–10. doi:10.1111/j.0956-7976.2005.01548.x

Amunts, K., Armstrong, E., Malikovic, A., Hömke, L., Mohlberg, H., Schleicher, A., & Zilles, K. (2007). Gender-specific left-right asymmetries in human visual cortex. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *27*(6), 1356–64. doi:10.1523/JNEUROSCI.4753-06.2007

Arsalidou, M., Morris, D., & Taylor, M. J. (2011). Converging evidence for the advantage of dynamic facial expressions. *Brain Topography*, *24*(2), 149–63. doi:10.1007/s10548-011-0171-4

Arterberry, M. E., & Bornstein, M. H. (2002). Infant perceptual and conceptual categorization: the roles of static and dynamic stimulus attributes. *Cognition*, *86*(1), 1–24. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12208649

Back, E., Jordan, T. R., & Thomas, S. M. (2009). The recognition of mental states from dynamic and static facial expressions. *Visual Cognition*, *17*(8), 1271–1286. doi:10.1080/13506280802479998

Back, E., Ropar, D., & Mitchell, P. (2007). Do the eyes have it? Inferring mental states from animated faces in autism. *Child Development*, *78*(2), 397–411. doi:10.1111/j.1467-8624.2007.01005.x

Baenniger, M. (1994). The development of face recognition: Featural or configurational processing? *Journal of Experimental Child Psychology*, *57*, 377–396.

Bahrick, L. E., Gogate, L. J., & Ruiz, I. (2002). Attention and memory for faces and actions in infancy: the salience of actions over faces in dynamic events. *Child Development*, *73*(6), 1629–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12487483

Bahrick, L. E., & Lickliter, R. (2000). Intersensory redundancy guides attentional selectivity and perceptual learning in infancy. *Developmental Psychology*, *36*(2), 190–201. doi:10.1037/0012-1649.36.2.190

Bahrick, L. E., Moss, L., & Fadil, C. (1996). Development of visual self-recognition in infancy. *Ecological Psychology, 8*(3), 189–208.

Bardi, L., Regolin, L., & Simion, F. (2011). Biological motion preference in humans at birth: role of dynamic and configural properties. *Developmental Science*, *14*(2), 353–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22213905

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11280420

Bartels, A., & Zeki, S. (2004). Functional brain mapping during free viewing of natural scenes. *Human Brain Mapping*, *21*(2), 75–85. doi:10.1002/hbm.10153

Bartlett, J. C., & Searcy, J. (1993). Inversion and configuration of faces. *Cognitive Psychology, 25*, 281–316.

Barton, J. J. (2003). Disorders of face perception and recognition. *Neurologic Clinics*, *21*(2), 521–548. doi:10.1016/S0733-8619(02)00106-8

Barton, J. J. S., Hefter, R., Cherkasova, M. V., & Manoach, D. S. (2007). Investigations of face expertise in the social developmental disorders. *Neurology, 69*(9), 860–870.

Barton, J. J. S., Press, D. Z., Keenan, J. P., & O'Connor, M. (2002). Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*, *58*(1), 71–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11781408

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, *37*(11), 2049–58. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/521902

Baudouin, J.-Y., Durand, K., & Gallay, M. (2008). Selective attention to facial identity and emotion in children. *Visual Cognition*, *16*(7), 933–952. doi:10.1080/13506280701728990

Baudouin, J.-Y., Martin, F., Tiberghien, G., Verlut, I., & Franck, N. (2002). Selective attention to facial emotion and identity in schizophrenia. *Neuropsychologia*, *40*(5), 503–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11749980

Bayliss, A. P., & Tipper, S. P. (2005). Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology (London, England : 1953)*, *96*(Pt 1), 95–114. doi:10.1348/000712604X15626

Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin*, *136*, 39–64. doi:10.1037/a0017768

Betts, J., McKay, J., Maruff, P., & Anderson, V. (2006). The development of sustained attention in children: the effect of age and task load. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *12*(3), 205–21. doi:10.1080/09297040500488522

Bimler, D. L., Skwarek, S. J., & Paramei, G. V. (2013). Processing facial expressions of emotion: Upright vs. inverted images. *Frontiers in Psychology*, *4*(February), 54. doi:10.3389/fpsyg.2013.00054

Bindemann, M., Scheepers, C., Ferguson, H. J., & Burton, a M. (2010). Face, body, and center of gravity mediate person detection in natural scenes. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(6), 1477–85. doi:10.1037/a0019057

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008). Social attention and real-world scenes: the roles of action, competition and social content. *Quarterly Journal of Experimental Psychology (2006)*, *61*(7), 986–98. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18938281

Birmingham, E., & Kingstone, A. (2009). Human social attention. *Progress in Brain Research*, *176*(626), 309–20. doi:10.1016/S0079-6123(09)17618-5

Bogfjellmo, L., Bex, P. J., & Falkenberg, H. K. (2014). The development of global motion discrimination in school aged children. *Journal of Vision*, *14*(2), 1–12. doi:10.1167/14.2.19.doi

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience*, *16*(11), 3737–44. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8642416

Bould, E., & Morris, N. (2008). Role of motion signals in recognizing subtle facial expressions of emotion. *British Journal of Psychology*, *99*(Pt 2), 167–89. doi:10.1348/000712607X206702

Braddick, O., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: motion coherence and "dorsal-stream vulnerability." *Neuropsychologia*, *41*(13), 1769–1784. doi:10.1016/S0028-3932(03)00178-7

Brazelton, T. B., Koslowski, B., & Main, M. (1974). The origin of reciprocity: The early mother-infant interaction. In M. Lewis & L. Rosenblum (Eds.), *The effect of the infant on its caregiver*. Oxford, England: Wiley-Interscience.

Bristow, D., Dehaene-Lambertz, G., Mattout, J., Soares, C., Gliga, T., Baillet, S., & Mangin, J.-F. (2009). Hearing faces: How the infant brain matches the face it sees with the speech it hears. *Journal of Cognitive Neuroscience*, *21*(5), 905–21. doi:10.1162/jocn.2009.21076

Britten, K. H., Newsome, W. T., & Saunders, R. C. (1992). Effects of inferotemporal cortex lesions on form-from-motion discrimination in monkeys. *Experimental Brain Research*, *88*(2), 292–302. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1577103

Bruce, V., Henderson, Z., Newman, C., & Burton, A. M. (2001). Matching identities of familiar and unfamiliar faces caught on CCTV images. *Journal of Experimental Psychology. Applied*, *7*(3), 207–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11676099

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology1*, *77*(3), 305–327.

Calder, A. J., Burton, A. M., Miller, P., Young, A. W., & Akamatsu, S. (2001). A principal component analysis of facial expressions. *Vision Research*, *41*(9), 1179–208. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11292507

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

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology. Human Perception and Performance*, *26*(2), 527–51. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10811161

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., … David, A. S. (1997). Activation of Auditory Cortex During Silent Lipreading. *Science*, *276*(5312), 593–596. doi:10.1126/science.276.5312.593

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: Salient physical features guide effective visual search. *Journal of Experimental Psychology. General*, *137*(3), 471–94. doi:10.1037/a0012771

Campbell, R., Coleman, M., Walker, J., Benson, P. J., Michelotti, J., & Baron-cohen, S. (1999). When does the inner-face advantage in familiar face recognition arise and why ?, *6*(2), 197–216.

Campbell, R., Landis, T., & Regard, M. (1986). Face recognition and lipreading. A neurological dissociation. *Brain: A Journal of Neurology*, *109 (Pt 3*, 509–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3719288

Campbell, R., MacSweeney, M., Surguladze, S., Calvert, G., McGuire, P., Suckling, J., … David, a S. (2001). Cortical substrates for the perception of face actions: An fMRI study of the specificity of activation for seen speech and for meaningless lower-face acts (gurning). *Brain Research. Cognitive Brain Research*, *12*(2), 233–43. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11587893

Campbell, R., & Tuck, M. (1995). Recognition of parts of famous-face photographs by children: an experimental note. *Perception*, *24*(451-456).

Campbell, R., Walker, J., & Baron-Cohen, S. (1995). Identification, The development of differential use of inner and outer face features in familiar face. *Journal of Experimental Child Psychology*, *59*, 196–210.

Carey, S., & Diamond, R. (1977). From piecemeal to configurational representation of faces. *Science*, *195*(4275), 312–314. Retrieved from http://www.jstor.org/stable/1743270

Carey, S., Diamond, R., & Woods, B. (1980). The development of face recognition - a maturational component? *Developmental Psychology*, *16*, 257–269.

Carpenter, G. C., Tecce, J. J., Stechler, G., & Friedman, S. (1970). Differential visual behavior to human and humanoid faces in early infancy. *Merrill-Palmer Quarterly of Behavior and Development*, *16*, 91–108.

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

Charbonneau, G., Bertone, A., Lepore, F., Nassim, M., Lassonde, M., Mottron, L., & Collignon, O. (2013). Multilevel alterations in the processing of audio-visual emotion expressions in autism spectrum disorders. *Neuropsychologia*, *51*(5), 1002–10. doi:10.1016/j.neuropsychologia.2013.02.009

Christie, F., & Bruce, V. (1998). The role of dynamic information in the recognition of unfamiliar faces. *Memory & Cognition*, *26*(4), 780–90. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9701969

Chung, M.-S., & Thomson, D. (1995). Development of face recognition. *British Journal of Psychology*, *86*, 55–87.

Corradi-Dell'acqua, C., Schwartz, S., Meaux, E., Hubert, B., Vuilleumier, P., & Deruelle, C. (2014). Neural responses to emotional expression information in high- and low-spatial frequency in autism: evidence for a cortical dysfunction. *Frontiers in Human Neuroscience*, *8*(April), 189. doi:10.3389/fnhum.2014.00189

Crookes, K., & McKone, E. (2009). Early maturity of face recognition: No childhood development of holistic processing, novel face encoding, or face-space. *Cognition*, *111*(2), 219–47. doi:10.1016/j.cognition.2009.02.004

Curby, K. M., Johnson, K. J., & Tyson, A. (2012). Face to face with emotion: Holistic face processing is modulated by emotional state. *Cognition & Emotion*, *26*(1), 93–102.

Dale, G., & Arnell, K. M. (2013). Investigating the stability of and relationships among global/local processing measures. *Attention, Perception & Psychophysics*, *75*(3), 394–406. doi:10.3758/s13414-012-0416-7

Dawel, A., McKone, E., Irons, J., O'Kearney, R., & Palermo, R. (2013). Look out! Gaze-cueing is greater from fearful faces in a dangerous context for children and adults. *Journal of Vision*, *13*(9), 595–595. doi:10.1167/13.9.595

De Haan, M., Pascalis, O., & Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, *14*(2), 199–209. doi:10.1162/089892902317236849

De Heering, A., Rossion, B., & Maurer, D. (2012). Developmental changes in face recognition during childhood: Evidence from upright and inverted faces. *Cognitive Development*. doi:10.1016/j.cogdev.2011.07.001

De Sonneville, L. M. ., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2010). Facial identity and facial emotions: Speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology*, *24*(2), 200–213. doi:10.1076/jcen.24.2.200.989

Degutis, J., Mercado, R. J., Wilmer, J., & Rosenblatt, A. (2013). Individual differences in holistic processing predict the own-race advantage in recognition memory. *PloS One*, *8*(4), e58253. doi:10.1371/journal.pone.0058253

Derntl, B., Finkelmeyer, A., Eickhoff, S., Kellermann, T., Falkenberg, D. I., Schneider, F., & Habel, U. (2010). Multidimensional assessment of empathic abilities: neural correlates and gender differences. *Psychoneuroendocrinology*, *35*(1), 67–82. doi:10.1016/j.psyneuen.2009.10.006

Derntl, B., Seidel, E.-M., Kainz, E., & Carbon, C.-C. (2009). Recognition of emotional expressions is affected by inversion and presentation time. *Perception*, *38*(12), 1849–62. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20192133

Derryberry, D., & Reed, M. A. (1998). Anxiety and attentional focusing: trait, state and hemispheric influences. *Personality and Individual Differences*, *25*(4), 745–761. doi:10.1016/S0191-8869(98)00117-2

Desjardins, R. N., Rogers, J., & Werker, J. F. (1997). An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *Journal of Experimental Child Psychology*, *66*(1), 85–110. doi:10.1006/jecp.1997.2379

Desjardins, R. N., & Werker, J. F. (2004). Is the integration of heard and seen speech mandatory for infants? *Developmental Psychobiology*, *45*(4), 187–203. doi:10.1002/dev.20033

Diamond, R., & Carey, S. (1977). Development changes in the representation of faces. *Journal of Experimental Child Psychology*, *23*, 1–22.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107–117.

Doherty-Sneddon, G., Bruce, V., Bonner, L., Longbotham, S., & Doyle, C. (2002). Development of gaze aversion as disengagement from visual information. *Developmental Psychology*, *38*(3), 438–445. doi:10.1037//0012-1649.38.3.438

Doherty-Sneddon, G., & Kent, G. (1996). Visual signals and the communication abilities of children. *Journal of Child Psychology and Psychiatry*, *37*(8), 949–959. doi:10.1111/j.1469-7610.1996.tb01492.x

Doherty-Sneddon, G., & Phelps, F. G. (2005). Gaze aversion: A response to cognitive or social difficulty? *Memory & Cognition*, *33*(4), 727–33. doi:10.3758/BF03195338

Domes, G., Steiner, A., Porges, S. W., & Heinrichs, M. (2013). Oxytocin differentially modulates eye gaze to naturalistic social signals of happiness and anger. *Psychoneuroendocrinology*, *38*(7), 1198–202. doi:10.1016/j.psyneuen.2012.10.002

Downing, P. E., Bray, D., Rogers, J., & Childs, C. (2004). Bodies capture attention when nothing is expected. *Cognition: International Journal of Cognitive Science*, *93*, B27–B38. doi:10.1016/j.cognition.2003.10.010

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science (New York, N.Y.)*, *293*(5539), 2470–3. doi:10.1126/science.1063414

Downing, P., Liu, J., & Kanwisher, N. (2001). Testing cognitive models of visual attention with fMRI and MEG. *Neuropsychologia*, *39*(12), 1329–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11566315

Easterbrook, M. A., Kisilevsky, B. S., Muir, D. W., & Laplante, D. P. (1999). Newborns discriminate schematic faces from scrambled faces. *Canadian Journal of Experimental Psychology*, *53*(3), 231–241.

Ebner, N. C. (2008). Age of face matters: age-group differences in ratings of young and old faces. *Behavior Research Methods*, *40*(1), 130–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18411535

Eisenbarth, H., & Alpers, G. W. (2011). Happy mouth and sad eyes: Scanning emotional facial expressions. *Emotion (Washington, D.C.)*, *11*(4), 860–5. doi:10.1037/a0022758

Elam, K. K., Carlson, J. M., Dilalla, L. F., & Reinke, K. S. (2010). Emotional faces capture spatial attention in 5-year-old children. *Evolutionary Psychology*, *8*(4), 754–67. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22947832

Ellamil, M., Susskind, J. M., & Anderson, A. K. (2008). Examinations of identity invariance in facial expression adaptation. *Cognitive, Affective & Behavioral Neuroscience*, *8*(3), 273–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18814464

Ellis, H. D., & Flin, R. H. (1990). Encoding and storage effects in 7-year-olds’ and 10-year-olds' memory for faces. *British Journal of Developmental Psychology*, *8*, 77–92.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: Some implications for theories of face recognition. *Perception*, *8*(4), 431–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/503774

Ellison, J. W., & Massaro, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(1), 213–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9090153

Everdell, I. T., Marsh, H. O., Yurick, M. D., Munhall, K. G., & Pare, M. (2007). Gaze behaviour in audiovisual speech perception: Asymmetrical distribution of face-directed fixations. *Perception*, *36*, 1535–1545.

Fagan, J. F. (1973). Infants’ delayed recognition memory and forgetting. *Journal of Experimental Child Psychology*, *16*(3), 424–50. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4771431

Farah, M. J. (1996). Is face recognition “special”? Evidence from neuropsychology. *Behavioural Brain Research*, *76*(1-2), 181–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8734052

Farah, M. J., Levinson, K. L., & Klein, K. L. (1995). Face perception and within-category discrimination in prosopagnosia. *Neuropsychologia*, *33*(6), 661–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7675159

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology. Human Perception and Performance*, *21*(3), 628–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7790837

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

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602–5. doi:10.1073/pnas.152159999

Farroni, T., Johnson, M. H., & Csibra, G. (2004). Mechanisms of eye gaze perception during infancy. *Journal of Cognitive Neuroscience*, *16*, 1320–1326.

Farroni, T., Johnson, M. H., Menon, E., Zulian, L., Faraguna, D., & Csibra, G. (2005). Newborns' preference for face-relevant stimuli: Effects of contrast polarity. *Proceedings of the National Academy of Sciences*, *102*(47), 17245–17250.

Field, T. M. (1979). Visual and cardiac responses to animate and inanimate faces by young term and preterm infants. *Child Development*, *50*(1), 188–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/446203

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *The Behavioral and Brain Sciences*, *22*(4), 661–674. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11301526

Fiorentini, C., & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision*, *11*(3), 1–15. doi:10.1167/11.3.17

Fisher, A., Thiessen, E., Godwin, K., Kloos, H., & Dickerson, J. (2013). Assessomg selective sustained attention in 3- to 5-year-old children: Evidence from a new paradigm. *Journal of Experimental Child Psychology*, *114*(2), 275–294. doi:10.1016/j.jecp.2012.07.006.Assessing

Fodor, J. A. (1983). *The modularity of mind: an essay on faculty psychology*. Cambridge, MA: MIT Press.

Foley, E., Rippon, G., Thai, N. J., Longe, O., & Senior, C. (2012). Dynamic facial expressions evoke distinct activation in the face perception network: a connectivity analysis study. *Journal of Cognitive Neuroscience*, *24*(2), 507–20. doi:10.1162/jocn_a_00120

Foreman, N., Fielder, A., Minshell, C., Hurrion, E., & Sergienko, E. (1997). Visual search, perception, and visual-motor skill in "healthy" children born at 27-32 weeks' gestation. *Journal of Experimental Child Psychology*, *64*, 27–41.

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*(17), 1920–31. doi:10.1016/j.visres.2011.07.002

Fox, C. J., Iaria, G., & Barton, J. J. S. (2009). Defining the face processing network: Optimization of the functional localizer in fMRI. *Human Brain Mapping*, *30*(5), 1637–51. doi:10.1002/hbm.20630

Fox, C. J., Moon, S. Y., Iaria, G., & Barton, J. J. S. (2009). The correlates of subjective perception of identity and expression in the face network: An fMRI adaptation study. *NeuroImage*, *44*(2), 569–80. doi:10.1016/j.neuroimage.2008.09.011

Frank, M. C., Vul, E., & Johnson, S. P. (2009). Development of infants' attention to faces during the first year. *Cognition*, *110*(2), 160–70. doi:10.1016/j.cognition.2008.11.010

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy*, *17*(4), 355–375. doi:10.1111/j.1532-7078.2011.00086.x

Freeth, M., Chapman, P., Ropar, D., & Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct the attention of high functioning adolescents with ASD? Evidence from eye-tracking. *Journal of Autism and Developmental Disorders*, *40*(5), 534–47. doi:10.1007/s10803-009-0893-2

Freire, A., & Lee, K. (2001). Face recognition in 4- to 7-year-olds: Processing of configural, featural, and paraphernalia information. *Journal of Experimental Child Psychology*, *80*(4), 347–71. doi:10.1006/jecp.2001.2639

Friesen, C. K., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(2), 319–29. doi:10.1037/0096-1523.30.2.319

Fujimura, T., & Suzuki, N. (2010). Effects of dynamic information in recognising facial expressions on dimensional and categorical judgments. *Perception*, *39*, 543–552.

Furl, N., van Rijsbergen, N. J., Kiebel, S. J., Friston, K. J., Treves, A., & Dolan, R. J. (2010). Modulation of perception and brain activity by predictable trajectories of facial expressions. *Cerebral Cortex*, *20*(3), 694–703. doi:10.1093/cercor/bhp140

Gallay, M., Baudouin, J.-Y., Durand, K., Lemoine, C., & Lécuyer, R. (2006). Qualitative differences in the exploration of upright and upside-down faces in four-month-old infants: An eye-movement study. *Child Development*, *77*(4), 984–96. doi:10.1111/j.1467-8624.2006.00914.x

Ganel, T., & Goshen-Gottstein, Y. (2004). Effects of familiarity on the perceptual integrality of the identity and expression of faces: The parallel-route hypothesis revisited. *Journal of Experimental Psychology. Human Perception and Performance*, *30*(3), 583–97. doi:10.1037/0096-1523.30.3.583

Ganel, T., Goshen-Gottstein, Y., & Goodale, M. A. (2005). Interactions between the processing of gaze direction and facial expression. *Vision Research*, *45*(9), 1191–200. doi:10.1016/j.visres.2004.06.025

Garner, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, *8*, 98–123.

Garner, W. R. (1983). Asymmetric interactions of stimulus dimensions in perceptual information processing. In T. J. Tighe & B. E. Shepp (Eds.), *Perception, cognition and development: Interactional analyses* (pp. 1–38). Hillsdale, NJ: Erlbaum.

Garner, W. R., & Feldoldy, G. L. (1970). Integrality of stimulus dimentions in various types of information processing. *Cognitive Psychology*, *1*, 225–241. doi:10.1016/0010-0285(70)90016-2

Gauthier, I., Behrmann, M., & Tarr, M. J. (1999). Can face recognition really be dissociated from object recognition? *Journal of Cognitive Neuroscience*, *11*(4), 349–70. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10471845

Gauthier, I., & Logothetis, N. K. (2000). Is face recognition not so unique after all? *Cognitive Neuropsychology*, *17*(1), 125–42. doi:10.1080/026432900380535

Ge, L., Anzures, G., Wang, Z., Kelly, D. J., Pascalis, O., Quinn, P. C., … Lee, K. (2008). An inner face advantage in children's recognition of familiar peers. *Journal of Experimental Child Psychology*, *101*(2), 124–36. doi:10.1016/j.jecp.2008.05.006

George, N., & Conty, L. (2008). Facing the gaze of others. *Neurophysiologie Clinique = Clinical Neurophysiology*, *38*(3), 197–207. doi:10.1016/j.neucli.2008.03.001

Gepner, B., Deruelle, C., & Grynfeltt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders*, *31*(1), 37–45. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11439752

Gilchrist, A., & McKone, E. (2003). Early maturity of face processing in children: Local and relational distinctiveness effects in 7-year-olds. *Visual Cognition*, *10*(7), 769–793. doi:10.1080/13506280344000022

Glenberg, A. M., Schroeder, J. L., & Robertson, D. a. (1998). Averting the gaze disengages the environment and facilitates remembering. *Memory & Cognition*, *26*(4), 651–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9701957

Godard, O., & Fiori, N. (2012). Sex and hemispheric differences in facial invariants extraction. *Laterality: Asymmetries of Body , Brain and Cognition*, *17*(2), 202–216. doi:10.1080/1357650X.2011.556641

Goren, C. C., Sarty, M., & Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, *56*(4), 544–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1165958

Grèzes, J., Fonlupt, P., Bertenthal, B., Delon-Martin, C., Segebarth, C., & Decety, J. (2001). Does perception of biological motion rely on specific brain regions? *NeuroImage*, *13*(5), 775–85. doi:10.1006/nimg.2000.0740

Grill-Spector, K., Knouf, N., & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555–62. doi:10.1038/nn1224

Guo, K. (2012). Holistic gaze strategy to categorize facial expression of varying intensities. *PloS One*, *7*(8), e42585. doi:10.1371/journal.pone.0042585

Haist, F., Adamo, M., Han Wazny, J., Lee, K., & Stiles, J. (2013). The functional architecture for face-processing expertise: FMRI evidence of the developmental trajectory of the core and the extended face systems. *Neuropsychologia*, *51*(13), 2893–908. doi:10.1016/j.neuropsychologia.2013.08.005

Hale, S. (1990). A global developmental trend in cognitive processing speed. *Child Development*, *61*(3), 653–63. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2364741

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16450045

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

Hasselmo, M. E., Rolls, E. T., & Baylis, G. C. (1989). The role of expression and identity in the face-selective responses of neurons in the temporal visual cortex of the monkey. *Behavioural Brain Research*, *32*(3), 203–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2713076

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science (New York, N.Y.)*, *303*(5664), 1634–40. doi:10.1126/science.1089506

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10827445

Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, *22*(1), 189–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10027301

Heekeren, H. R., Marrett, S., & Ungerleider, L. G. (2008). The neural systems that mediate human perceptual decision making. *Nature Reviews. Neuroscience*, *9*(6), 467–79. doi:10.1038/nrn2374

Hegdé, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, *84*(4), 405–39. doi:10.1016/j.pneurobio.2007.09.001

Hein, G., & Knight, R. T. (2008). Superior temporal sulcus--It's my area: or is it? *Journal of Cognitive Neuroscience*, *20*(12), 2125–36. doi:10.1162/jocn.2008.20148

Heisz, J. J., Pottruff, M. M., & Shore, D. I. (2013). Females scan more than males: a potential mechanism for sex differences in recognition memory. *Psychological Science*, *24*(7), 1157–63. doi:10.1177/0956797612468281

Henderson, J. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498–504. doi:10.1016/j.tics.2003.09.006

Herba, C. M., Landau, S., Russell, T., Ecker, C., & Phillips, M. L. (2006). The development of emotion-processing in children: effects of age, emotion, and intensity. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *47*(11), 1098–106. doi:10.1111/j.1469-7610.2006.01652.x

Herlitz, A., & Rehnman, J. (2009). Sex differences in episodic memory. *Current Directions in Psycholocial Science*, *17*, 52–56. doi:10.1111/j.1467-8721

Hershler, O., & Hochstein, S. (2005). At first sight: A high-level pop out effect for faces. *Vision Research*, *45*(13), 1707–24. doi:10.1016/j.visres.2004.12.021

Hietanen, J. K., & Leppänen, J. M. (2008). Judgment of other people's facial expressions of emotions is influenced by their concurrent affective hand movements. *Scandinavian Journal of Psychology*, *49*(3), 221–30. doi:10.1111/j.1467-9450.2008.00644.x

Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R., & Hämäläinen, H. (2006). Automatic attention orienting by social and symbolic cues activates different neural networks: an fMRI study. *NeuroImage*, *33*(1), 406–13. doi:10.1016/j.neuroimage.2006.06.048

Hoffman, E., Haxby, J. V., & Gobbini, M. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*(6), 223–233. doi:10.1038/71152

Hoffman, E., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, *3*(1), 80–4. doi:10.1038/71152

Horstmann, G., & Ansorge, U. (2009). Visual search for facial expressions of emotions: A comparison of dynamic and static faces. *Emotion (Washington, D.C.)*, *9*(1), 29–38. doi:10.1037/a0014147

Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*(10), 998–1006. doi:10.1111/j.1467-9280.2008.02191.x

Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity, and it is computed separately for moving and static faces: neuropsychological evidence. *Neuropsychologia*, *31*(2), 173–81. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8455786

Insch, P. M., Bull, R., Phillips, L. H., Allen, R., & Slessor, G. (2012). Adult aging, processing style, and the perception of biological motion. *Experimental Aging Research*, *38*(2), 169–185. doi:10.1080/0361073X.2012.660030

Jakobson, L. S., Frisk, V., & Downie, A. L. (2006). Motion-defined form processing in extremely premature children. *Neuropsychologia*, *44*, 1777–1786.

Jakobson, L. S., Frisk, V., Knight, R. M., Downie, A. L., & Whyte, H. (2001). The relationship between periventricular brain injury and deficits in visual processing among extremely-low-birthweight (< 1000 g) children. *Journal of Pediatric Psychology*, *26*(8), 503–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11700335

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews. Neuroscience*, *6*(10), 766–74. doi:10.1038/nrn1766

Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1-2), 1–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1786670

Johnston, R. A., & Ellis, H. D. (1995). Age effects in the processing of typical and distinctive faces. *The Quarterly Journal of Experimental Psychology Section A*, *48*(2), 447–465. doi:10.1080/14640749508401399

Jordan, T. R., & Bevan, K. (1997). Seeing and hearing rotated faces: Influences of facial orientation on visual and audiovisual speech recognition. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(2), 388–403. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9104001

Joseph, J. E., Swearingen, J. E., Clark, J. D., Benca, C. E., Collins, H. R., Corbly, C. R., … Bhatt, R. S. (2012). The changing landscape of functional brain networks for face processing in typical development. *NeuroImage*, *63*(3), 1223–36. doi:10.1016/j.neuroimage.2012.08.021

Joseph, R. M., & Tanaka, J. W. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry*, *44*(4), 529–542. Retrieved from http://europepmc.org/abstract/MED/12751845/reload=0

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, *30*(7), 875–887. doi:10.1068/p3131

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–11. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9151747

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review*, *27*(3), 283–348. doi:10.1016/j.dr.2007.06.006

Kaufmann, J. M., & Schweinberger, S. R. (2005). Speaker variations influence speechreading speed for dynamic faces. *Perception*, *34*(5), 595–610. doi:10.1068/p5104

Kesler-West, M. L., Andersen, A. H., Smith, C. D., Avison, M. J., Davis, C. E., Kryscio, R. J., & Bonder, L. X. (2001). Neural substrates of facial emotion processing using fMRI. *Cognitive Brain Research*, *11*(2), 213–226. doi:10.1016/S0926-6410(00)00073-2

Kessler, H., Doyen-Waldecker, C., Hofer, C., Hoffmann, H., Traue, H., & Abler, B. (2011). Neural correlates of the perception of dynamic versus static facial expressions of emotion. *Psychosocial Medicine*, *8*, 1–8. doi:10.3205/psm000072

Kikuchi, Y., Senju, A., Tojo, Y., Osanai, H., & Hasegawa, T. (2009). Faces do not capture special attention in children with autism spectrum disorder: A change blindness study. *Child Development*, *80*(5), 1421–33. doi:10.1111/j.1467-8624.2009.01342.x

Kilts, C. D., Egan, G., Gideon, D. a., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *NeuroImage*, *18*(1), 156–168. doi:10.1006/nimg.2002.1323

Kimchi, R., Amishav, R., & Sulitzeanu-Kenan, A. (2009). Gender differences in global-local perception? Evidence from orientation and shape judgments. *Acta Psychologica*, *130*(1), 64–71. doi:10.1016/j.actpsy.2008.10.002

Klein, R. ., & Jennings, K. D. (1979). Responses to social and inanimate stimuli in early infancy. *Journal of Genetic Psychology*, *135*(1), 3–9.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, *59*(9), 809–16. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12215080

Knappmeyer, B., Thornton, I. M., & Bülthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, *43*(18), 1921–1936. doi:10.1016/S0042-6989(03)00236-0

Knight, B., & Johnston, A. (1997). The role of movement in face recognition. *Visual Cognition*, *4*(3), 265–273. doi:10.1080/713756764

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, *12*, 48–55.

Krebs, J. F., Biswas, A., Pascalis, O., Kamp-Becker, I., Remschmidt, H., & Schwarzer, G. (2011). Face processing in children with autism spectrum disorder: Independent or interactive processing of facial identity and facial expression? *Journal of Autism and Developmental Disorders*, *41*(6), 796–804. doi:10.1007/s10803-010-1098-4

Krekelberg, B., Dannenberg, S., Hoffmann, K. P., Bremmer, F., & Ross, J. (2003). Neural correlates of implied motion. *Nature*, *424*, 674–677.

Kuefner, D., De Heering, A., Jacques, C., Palmero-soler, E., & Rossion, B. (2010). Early visually evoked electrophysiological responses over the human brain (P1 , N170) show stable patterns of face-sensitivity from 4 years to adulthood. *Frontiers in Human Neuroscience*, *3*(January), 1–22. doi:10.3389/neuro.09.067

Kuehn, S. M., & Jolicoeur, P. (1994). Impact of quality of the image, orientation, and similarity of the stimuli on visual search for faces. *Perception*, *23*(1), 95–122.

Lahnakoski, J. M., Glerean, E., Salmi, J., Jääskeläinen, I. P., Sams, M., Hari, R., & Nummenmaa, L. (2012). Naturalistic FMRI mapping reveals superior temporal sulcus as the hub for the distributed brain network for social perception. *Frontiers in Human Neuroscience*, *6*, 233. doi:10.3389/fnhum.2012.00233

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, *41*(25-26), 3559–65. doi:10.1016/S0042-6989(01)00102-X

Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*, *12*(4), 259–272. doi:10.1207/S15326969ECO1204_01

Lander, K., Bruce, V., & Hill, H. (2001). Evaluating the effectiveness of pixelation and blurring on masking the identity of familiar faces. *Applied Cognitive Psychology*, *15*(1), 101–116. doi:10.1002/1099-0720(200101/02)15:1<101::AID-ACP697>3.0.CO;2-7

Lander, K., Christie, F., & Bruce, V. (1999). The role of movement in the recognition of famous faces. *Memory & Cognition*, *27*(6), 974–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10586574

Langton, S. R., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, *107*, 330–342. doi:10.1016/j.cognition.2007.07.012

Langton, S. R., Watt, R. J., & Bruce, I. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50–59. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10652522

Lassalle, A., & Itier, R. J. (2013). Fearful, surprised, happy, and angry facial expressions modulate gaze-oriented attention: behavioral and ERP evidence. *Social Neuroscience*, *8*(6), 583–600. doi:10.1080/17470919.2013.835750

Lavie, N., Ro, T., & Russell, C. (2003). The role of perceptual load in processing distractor faces. *Psychological Science*, *14*(5), 510–515. doi:10.1111/1467-9280.03453

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. *Psychological Science*, *15*(11), 762–8. doi:10.1111/j.0956-7976.2004.00753.x

Lee, L. C., Andrews, T. J., Johnson, S. J., Woods, W., Gouws, A., Green, G. G. R., & Young, A. W. (2010). Neural responses to rigidly moving faces displaying shifts in social attention investigated with fMRI and MEG. *Neuropsychologia*, *48*(2), 477–90. doi:10.1016/j.neuropsychologia.2009.10.005

Legerstee, M., Pomerleau, A., Malcuit, G., & Feider, H. (1987). The development of infants' responses to people and a doll: Implications for research in communication. *Infant Behavior and Development*, *10*(1), 81–95. doi:10.1016/0163-6383(87)90008-7

LeGrand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Neuroperception*, *410*(April), 2001. Retrieved from http://psych.mcmaster.ca/maurerlab/Publications/LeGrand_DeprivedFaces.pdf

Lewis, M. B., & Edmonds, A. J. (2003). Face detection: Mapping human performance. *Perception*, *32*(8), 903–920. doi:10.1068/p5007

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition: International Journal of Cognitive Science*, *21*(1), 1–36.

Lipp, O. V, Price, S. M., & Tellegen, C. L. (2009). No effect of inversion on attentional and affective processing of facial expressions. *Emotion (Washington, D.C.)*, *9*(2), 248–59. doi:10.1037/a0014715

Lobmaier, J. S., Klaver, P., Loenneker, T., Martin, E., & Mast, F. W. (2008). Featural and configural face processing strategies: evidence from a functional magnetic resonance imaging study. *Neuroreport*, *19*(3), 287–91. doi:10.1097/WNR.0b013e3282f556fe

Longmore, C. A., & Tree, J. J. (2013). Motion as a cue to face recognition: Evidence from congenital prosopagnosia. *Neuropsychologia*, *51*(5), 864–875. doi:10.1016/j.neuropsychologia.2013.01.022

Loucks, J. (2011). Configural information is processed differently in human action. *Perception*, *40*(9), 1047–1062. doi:10.1068/p7084

Loucks, J., & Baldwin, D. (2009). Sources of information for discriminating dynamic human actions. *Cognition*, *111*(1), 84–97. doi:10.1016/j.cognition.2008.12.010

Love, B. C., Rouder, J. N., & Wisniewski, E. J. (1999). A structural account of global and local processing. *Cognitive Psychology*, *38*, 291–316.

MacKay, T. L., Jakobson, L. S., Ellemberg, D., Lewis, T. L., Maurer, D., & Casiro, O. (2005). Deficits in the processing of local and global motion in very low birthweight children. *Neuropsychologia*, *43*, 1738–1748.

Macrae, C. N., & Lewis, H. L. (2002). Do I know you? Processing orientation and face recognition. *Psychological Science*, *13*(2), 194–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11934008

Mancini, G., Agnoli, S., Baldaro, B., Ricci Bitti, P. E., & Surcinelli, P. (2013). Facial expressions of emotions: Recognition accuracy and affective reactions during late childhood. *The Journal of Psychology*, *147*(6), 599–617. doi:10.1080/00223980.2012.727891

Marotta, J. J., McKeeff, T. J., & Behrmann, M. (2002). The effects of rotation and inversion on face processing in prosopagnosia. *Cognitive Neuropsychology*, *19*(1), 31–47. doi:10.1080/02643290143000079

Martens, U., Leuthold, H., & Schweinberger, S. R. (2010). Parallel processing in face perception. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(1), 103–21. doi:10.1037/a0017167

Martin, D., & Macrae, C. N. (2010). Processing style and person recognition: Exploring the face inversion effect. *Visual Cognition*, *18*(2), 161–170. doi:10.1080/13506280902868793

Marwick, K., & Hall, J. (2008). Social cognition in schizophrenia: A review of face processing. *British Medical Bulletin*, *88*(1), 43–58. doi:10.1093/bmb/ldn035

Marzi, T., & Viggiano, M. P. (2011). Temporal dynamics of face inversion at encoding and retrieval. *Clinical Neurophysiology*, *122*(7), 1360–70. doi:10.1016/j.clinph.2010.11.017

Massaro, D. W., & Cohen, M. M. (1996). Perceiving speech from inverted faces. *Perception & Psychophysics*, *58*(7), 1047–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8920841

Matsuda, Y.-T., Okanoya, K., & Myowa-Yamakoshi, M. (2013). Shyness in early infancy: approach-avoidance conflicts in temperament and hypersensitivity to eyes during initial gazes to faces. *PloS One*, *8*(6), e65476. doi:10.1371/journal.pone.0065476

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences*, *6*(6), 255–260. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12039607

McBain, R., Norton, D., & Chen, Y. (2009). Females excel at basic face perception. *Acta Psychologica*, *130*(2), 168–73. doi:10.1016/j.actpsy.2008.12.005

McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *9*(5), 605–10. doi:10.1162/jocn.1997.9.5.605

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, *126*(3), 424–453. doi:10.1037//0033-2909.126.3.424

McGugin, R. W., Tanaka, J. W., Lebrecht, S., Tarr, M. J., & Gauthier, I. (2011). Race-specific perceptual discrimination improvement following short individuation training with faces. *Cognitive Science*, *35*(2), 330–47. doi:10.1111/j.1551-6709.2010.01148.x

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *The British Journal of Social Psychology / the British Psychological Society*, *34 (Pt 3)*, 325–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7551775

McKone, E., & Boyer, B. L. (2006). Sensitivity of 4-year-olds to featural and second-order relational changes in face distinctiveness. *Journal of Experimental Child Psychology*, *94*(2), 134–62. doi:10.1016/j.jecp.2006.01.001

McKone, E., Crookes, K., Jeffery, L., & Dilks, D. D. (2012). A critical review of the development of face recognition: Experience is less important than previously believed. *Cognitive Neuropsychology*, *29*(1-2), 174–212. doi:10.1080/02643294.2012.660138

McKone, E., Kanwisher, N., & Duchaine, B. C. (2007). Can generic expertise explain special processing for faces? *Trends in Cognitive Sciences*, *11*(1), 8–15. doi:10.1016/j.tics.2006.11.002

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(45), 16518–23. doi:10.1073/pnas.0507650102

Megreya, A. M., Bindemann, M., & Havard, C. (2011). Sex differences in unfamiliar face identification: evidence from matching tasks. *Acta Psychologica*, *137*(1), 83–9. doi:10.1016/j.actpsy.2011.03.003

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*(1), 3–35. Retrieved from http://psycnet.apa.org/?&fa=main.doiLanding&doi=10.1037/1076-8971.7.1.3

Mercado, R. J., Cohan, S., & DeGutis, J. M. (2011). Training with Same-Race Faces Improves Holistic Processing of Other-Race Faces. *Journal of Vision*, *11*(11), 627–627. doi:10.1167/11.11.627

Michel, C., Rossion, B., Han, J., Chung, C.-S., & Caldara, R. (2006). Holistic processing is finely tuned for faces of one's own race. *Psychological Science*, *17*(7), 608–15. doi:10.1111/j.1467-9280.2006.01752.x

Michel, Caldara, R., & Rossion, B. (2006). Same-race faces are perceived more holistically than other-race faces. *Visual Cognition*, *14*, 55–73.

Miller, L. T., & Vernon, P. a. (1997). Developmental changes in speed of information processing in young children. *Developmental Psychology*, *33*(3), 549–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9149934

Milne, E., Szczerbinski, M., & Sheffield, S. (2009). Global and local perceptual style, field-independence, and central coherence: An attempt at concept validation. *Advaces in Cognitive Psychology*, *5*(0), 1–26. doi:10.2478/v10053-008-0062-8

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action.* Oxford: Oxford University Press.

Mishkin, M., Ungerleider, L. G., & Kathleen, A. (1983). Object vision and spatial vision: two cortical p hways. *Trends in Neuroscience*, *6*, 414–417.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100.

Mondloch, C. J., Dobson, K. S., Parsons, J., & Maurer, D. (2004). Why 8-year-olds cannot tell the difference between Steve Martin and Paul Newman: Factors contributing to the slow development of sensitivity to the spacing of facial features. *Journal of Experimental Child Psychology*, *89*(2), 159–81. doi:10.1016/j.jecp.2004.07.002

Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. *Journal of Experimental Child Psychology*, *86*(1), 67–84. doi:10.1016/S0022-0965(03)00102-4

Mondloch, C. J., Le Grand, R. Le, & Maurer, D. (2002). Configural face processing develops more slowly than featural face processing. *Perception*, *31*(5), 553–566. doi:10.1068/p3339

Mondloch, C. J., Pathman, T., Maurer, D., Le Grand, R., & de Schonen, S. (2007). The composite face effect in six-year-old children: Evidence of adult-like holistic face processing. *Visual Cognition*, *15*(5), 564–577. doi:10.1080/13506280600859383

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

Mottron, L., & Belleville, S. (1993). A study of perceptual analysis in a high-level autistic subject with exceptional graphic abilities. *Brain and Cognition*, *23*(2), 279–309. doi:10.1006/brcg.1993.1060

Narumoto, J., Okada, T., Sadato, N., Fukui, K., & Yonekura, Y. (2001). Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Brain Research. Cognitive Brain Research*, *12*(2), 225–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11587892

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(353-383).

Nelson, C. A. (2001). The development and neural bases of face recognition. *Infant and Child Development*, *10*(1-2), 3–18. doi:10.1002/icd.239

Noh, S. R., & Isaacowitz, D. . (2013). Emotional faces in context: Age differences in recognition accuracy and scanning patterns. *Emotion*, *13*(2), 238–249. doi:10.1037/a0030234

Nummenmaa, L., Hyona, J., & Calvo, M. G. (2006). Eye movement assessment of selective attentional capture by emotional pictures. *Emotion*, *6*, 257–268. doi:10.1037/1528-3542.6.2.257

O’Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., & Dunlop, J. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*(1), 74–83. doi:10.1016/j.visres.2010.09.035

O’Toole, A. J., Roark, D. A., & Abdi, H. (2002). Recognizing moving faces: A psychological and neural synthesis. *Trends in Cognitive Sciences*, *6*(6), 261–266. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12039608

Olson, I. R., Gatenby, J. C., & Gore, J. C. (2002). A comparison of bound and unbound audio-visual information processing in the human cerebral cortex. *Brain Research. Cognitive Brain Research*, *14*(1), 129–38. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12063136

Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, *76*(1), 109–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8836213

Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, *45*(1), 75–92. doi:10.1016/j.neuropsychologia.2006.04.025

Pascalis, O., de Haan, M., Nelson, C. a, & de Schonen, S. (1998). Long-term recognition memory for faces assessed by visual paired comparison in 3- and 6-month-old infants. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *24*(1), 249–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9438961

Pascalis, O., de Martin de Vivies, X., Anzures, G., Quinn, P. C., Slater, A. M., Tanaka, J. W., & Lee, K. (2011). Development of face processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(6), 666–675. doi:10.1002/wcs.146.Development

Pastò, L., & Burack, J. A. (1997). A developmental study of visual attention: Issues of filtering efficiency and focus. *Cognitive Development*, *12*(4), 523–535. doi:10.1016/S0885-2014(97)90021-6

Pavlova, M., Guerreschi, M., Lutzenberger, W., Sokolov, A. N., & Krägeloh-Mann, I. (2010). Cortical response to social interaction is affected by gender. *NeuroImage*, *50*(3), 1327–32. doi:10.1016/j.neuroimage.2009.12.096

Pavlova, M., Sokolov, A., & Krägeloh-Mann, I. (2007). Visual navigation in adolescents with early periventricular lesions: Knowing where, but not getting there. *Cerebral Cortex*, *17*(2), 363–9. doi:10.1093/cercor/bhj153

Peelen, M. V, Glaser, B., Vuilleumier, P., & Eliez, S. (2009). Differential development of selectivity for faces and bodies in the fusiform gyrus. *Developmental Science*, *12*(6), F16–25. doi:10.1111/j.1467-7687.2009.00916.x

Pellicano, E., Rhodes, G., & Peters, M. (2006). Are preschoolers sensitive to configural information in faces? *Developmental Science*, *9*(3), 270–7. doi:10.1111/j.1467-7687.2006.00489.x

Pelphrey, K. a, Morris, J. P., McCarthy, G., & Labar, K. S. (2007). Perception of dynamic changes in facial affect and identity in autism. *Social Cognitive and Affective Neuroscience*, *2*(2), 140–9. doi:10.1093/scan/nsm010

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain: A Journal of Neurology*, *128*(Pt 5), 1038–48. doi:10.1093/brain/awh404

Perfect, T. J., Weston, N. J., Dennis, I., & Snell, A. (2008). The effects of precedence on Navon-induced processing bias in face recognition. *Quarterly Journal of Experimental Psychology (2006)*, *61*(10), 1479–86. doi:10.1080/17470210802034678

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

Picozzi, M., Cassia, V. M., Turati, C., & Vescovo, E. (2009). The effect of inversion on 3- to 5-year-olds' recognition of face and nonface visual objects. *Journal of Experimental Child Psychology*, *102*(4), 487–502. doi:10.1016/j.jecp.2008.11.001

Pike, G. E., Kemp, R. I., Towell, N. A., & Phillips, K. C. (1997). Recognizing moving faces: The relative contribution of motion and perspective view information. *Visual Cognition*, *4*, 409–437.

Pilz, K. S., Thornton, I. M., & Bülthoff, H. H. (2006). A search advantage for faces learned in motion. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *171*(4), 436–47. doi:10.1007/s00221-005-0283-8

Pilz, K. S., Vuong, Q. C., Bülthoff, H. H., & Thornton, I. M. (2011). Walk this way: Approaching bodies can influence the processing of faces. *Cognition*, *118*(1), 17–31. doi:10.1016/j.cognition.2010.09.004

Proverbio, A. M., Brignone, V., Matarazzo, S., Del Zotto, M., & Zani, A. (2006). Gender differences in hemispheric asymmetry for face processing. *BMC Neuroscience*, *7*, 44. doi:10.1186/1471-2202-7-44

Proverbio, A. M., Riva, F., Martin, E., & Zani, A. (2010). Face coding is bilateral in the female brain. *PloS One*, *5*(6), e11242. doi:10.1371/journal.pone.0011242

Pruett, J. R., LaMacchia, A., Hoertel, S., Squire, E., McVey, K., Todd, R. D., … Petersen, S. E. (2011). Social and non-social cueing of visuospatial attention in autism and typical development. *Journal of Autism and Developmental Disorders*, *41*(6), 715–31. doi:10.1007/s10803-010-1090-z

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *18*(6), 2188–99. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9482803

Puce, A., Syngeniotis, A., Thompson, J. C., Abbott, D. F., Wheaton, K. J., & Castiello, U. (2003). The human temporal lobe integrates facial form and motion: evidence from fMRI and ERP studies. *NeuroImage*, *19*(3), 861–869. doi:10.1016/S1053-8119(03)00189-7

Rahko, J., Paakki, J.-J., Starck, T., Nikkinen, J., Remes, J., Hurtig, T., … Kiviniemi, V. (2010). Functional mapping of dynamic happy and fearful facial expression processing in adolescents. *Brain Imaging and Behavior*, *4*(2), 164–76. doi:10.1007/s11682-010-9096-x

Riby, D. M., Doherty-Sneddon, G., & Bruce, V. (2008). Exploring face perception in disorders of development: Evidence from Williams syndrome and autism. *Journal of Neuropsychology*, *2*(1), 47–64. doi:10.1348/174866407X255690

Ricciardelli, P., Bricolo, E., Aglioti, S. M., & Chelazzi, L. (2002). My eyes want to look where your eyes are looking : Exploring the tendency to imitate another individual ' s gaze, *13*(17). doi:10.1097/01.wnr.0000044227

Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013). Unaware person recognition from the body when face identification fails. *Psychological Science*, *24*(11), 2235–43. doi:10.1177/0956797613492986

Richler, J. J., Mack, M. L., Gauthier, I., & Palmeri, T. J. (2009). Holistic processing of faces happens at a glance. *Vision Research*, *49*(23), 2856–61. doi:10.1016/j.visres.2009.08.025

Ripoll, T., Fiere, É., & Pélissier, A. (2005). Relative Weight of Local and Global Properties Depends on both the Position of Local Elements and the Saliency of Global Form. *Experimental Psychology (formerly "Zeitschrift Für Experimentelle Psychologie")*, *52*(4), 272–280. doi:10.1027/1618-3169.52.4.272

Roark, D. A., Barrett, S. E., Spence, M. J., & O'Toole, A. J. (2003). Psychological and neural perspectives on the role of motion in face recognition. *Behavioural and Cognitive Neuroscience Reviews*, *2*(1), 15–46. doi:10.1177/1534582303251820

Roark, D. A., O'Toole, A. J., Abdi, H., & Barrett, S. E. (2006). Learning the moves: The effect of familiarity and facial motion on person recognition across large changes in viewing format. *Perception*, *35*(6), 761–773. doi:10.1068/p5503

Robbins, R. A., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, *103*(1), 34–79. doi:10.1016/j.cognition.2006.02.008

Robbins, R. A., Nishimura, M., Mondloch, C. J., Lewis, T. L., & Maurer, D. (2010). Deficits in sensitivity to spacing after early visual deprivation in humans: a comparison of human faces, monkey faces, and houses. *Developmental Psychobiology*, *52*(8), 775–81. doi:10.1002/dev.20473

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2001). Attention and recognition memory in the 1st year of life: A longitudinal study of preterm and full-term infants. *Developmental Psychology*, *37*, 135–151.

Rosenblum, L. D., Yakel, D. a., & Green, K. P. (2000). Face and mouth inversion effects on visual and audiovisual speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 806–819. doi:10.1037//0096-1523.26.2.806

Ross, P., & Turkewitz, G. (1981). Individual differences in cerebral asymmetries for facial recognition. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *17*(2), 199–214. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7285593

Rossion, B., Hanseeuw, B., & Dricot, L. (2012). Brain and cognition defining face perception areas in the human brain: A large-scale factorial fMRI face localizer analysis. *Brain and Cognition*, *79*(2), 138–157. doi:10.1016/j.bandc.2012.01.001

Rotshtein, P., Geng, J. J., Driver, J., & Dolan, R. J. (2007). Role of features and second-order spatial relations in face discrimination, face recognition, and individual face skills: behavioral and functional magnetic resonance imaging data. *Journal of Cognitive Neuroscience*, *19*(9), 1435–52. doi:10.1162/jocn.2007.19.9.1435

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Symposium on Eye Tracking Research & Applications - ETRA '00*, 71–78. doi:10.1145/355017.355028

Samson, H., Fiori-Duharcourt, N., Doré-Mazars, K., Lemoine, C., & Vergilino-Perez, D. (2014). Perceptual and gaze biases during face processing: related or not? *PloS One*, *9*(1), e85746. doi:10.1371/journal.pone.0085746

Sáry, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science (New York, N.Y.)*, *260*(5110), 995–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8493538

Sato, W., Kochiyama, T., Uono, S., & Yoshikawa, S. (2008). Time course of superior temporal sulcus activity in response to eye gaze: a combined fMRI and MEG study. *Social Cognitive and Affective Neuroscience*, *3*(3), 224–32. doi:10.1093/scan/nsn016

Sato, W., Kochiyama, T., Uono, S., & Yoshikawa, S. (2010). Amygdala integrates emotional expression and gaze direction in response to dynamic facial expressions. *NeuroImage*, *50*(4), 1658–65. doi:10.1016/j.neuroimage.2010.01.049

Sato, W., Kochiyama, T., & Yoshikawa, S. (2011). The inversion effect for neutral and emotional facial expressions on amygdala activity. *Brain Research*, *1378*, 84–90. doi:10.1016/j.brainres.2010.12.082

Scherf, K. S., Behrmann, M., Humphreys, K., & Luna, B. (2007). Visual category-selectivity for faces, places and objects emerges along different developmental trajectories. *Developmental Science*, *10*(4), F15–30. doi:10.1111/j.1467-7687.2007.00595.x

Scherf, K. S., Behrmann, M., Kimchi, R., & Luna, B. (2009). Emergence of global shape processing continues through adolescence. *Child Development*, *80*(1), 162–77. doi:10.1111/j.1467-8624.2008.01252.x

Scherf, K. S., Luna, B., Avidan, G., & Behrmann, M. (2011). "What" precedes "which": Developmental neural tuning in face- and place-related cortex. *Cerebral Cortex (New York, N.Y. : 1991)*, *21*(9), 1963–80. doi:10.1093/cercor/bhq269

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *NeuroImage*, *32*(3), 1385–94. doi:10.1016/j.neuroimage.2006.05.037

Schultz, J., & Pilz, K. S. (2009). Natural facial motion enhances cortical responses to faces. *Experimental Brain Research. Experimentelle Hirnforschung. Expérimentation Cérébrale*, *194*(3), 465–75. doi:10.1007/s00221-009-1721-9

Schwarzer, G. (2000). Development of face processing: The effect of face inversion. *Child Development*, *71*(2), 391–401. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10834472

Schwarzer, G., Huber, S., & Dümmler, T. (2005). Gaze behavior in analytical and holistic face processing. *Memory & Cognition*, *33*(2), 344–54. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16028588

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *25*(47), 11055–9. doi:10.1523/JNEUROSCI.2621-05.2005

Schweinberger, S. R., Burton, a M., & Kelly, S. W. (1999). Asymmetric dependencies in perceiving identity and emotion: Experiments with morphed faces. *Perception & Psychophysics*, *61*(6), 1102–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10497431

Schweinberger, S. R., & Soukup, G. R. (1998). Asymmetric relationships among perceptions of facial identity, emotion, and facial speech. *Journal of Experimental Psychology. Human*

Perception and Performance, *24*(6), 1748–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9861721

Searcy, J. H., & Bartlett, J. C. (1996). Inversion and processing of component and spatial-relational information in faces. *Journal of Experimental Psychology: Human*, *22*(4), 904–915.

Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., & Brammer, M. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. *Current Biology*, *10*, 16–22.

Senju, A., & Hasegawa, T. (2006). Do the upright eyes have it? *Psychonomic Bulletin & Review*, *13*(2), 223–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16892985

Senju, A., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2008). Is anyone looking at me? Direct gaze detection in children with and without autism. *Brain and Cognition*, *67*(2), 127–39. doi:10.1016/j.bandc.2007.12.001

Sergent, J., Ohta, S., & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain : A Journal of Neurology*, *115 Pt 1*, 15–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1559150

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *34*(8), 1351–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8294523

Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(2), 809–13. doi:10.1073/pnas.0707021105

Skoczenski, A. M., & Norcia, A. M. (2002). Late maturation of visual hyperacuity. *Psychological Science*, *13*(6), 537–41. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12430838

Skwerer, D. P., Borum, L., Verbalis, A., Schofield, C., Crawford, N., Ciciolla, L., & Tager-Flusberg, H. (2009). Autonomic responses to dynamic displays of facial expressions in adolescents and adults with Williams syndrome. *Social Cognitive and Affective Neuroscience*, *4*(1), 93–100. doi:10.1093/scan/nsn041

Smith, T. J., & Mital, P. K. (2013). Attentional synchrony and the influence of viewing task on gaze behavior in static and dynamic scenes. *Journal of Vision*, *13*(8), 1–24. doi:10.1167/13.8.16.doi

Song, Y., & Hakoda, Y. (2012). Selective attention to facial emotion and identity in children with autism: Evidence for global identity and local emotion. *Autism Research*, *5*, 282–285.

Sorger, B., Goebel, R., Schiltz, C., & Rossion, B. (2007). Understanding the functional neuroanatomy of acquired prosopagnosia. *NeuroImage*, *35*(2), 836–52. doi:10.1016/j.neuroimage.2006.09.051

Soussignan, R., Chadwick, M., Philip, L., Conty, L., Dezecache, G., & Grèzes, J. (2013). Self-relevance appraisal of gaze direction and dynamic facial expressions: Effects on facial electromyographic and autonomic reactions. *Emotion (Washington, D.C.)*, *13*(2), 330–7. doi:10.1037/a0029892

Spangler, S. M., Schwarzer, G., Korell, M., & Maier-Karius, J. (2009). The relationships between processing facial identity, emotional expression, facial speech, and gaze direction during development. *Journal of Experimental Child Psychology*, *105*(1-2), 1–19. doi:10.1016/j.jecp.2009.09.003

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism*, *11*(3), 265–77. doi:10.1177/1362361307076925

Stein, T., End, A., & Sterzer, P. (2014). Own-race and own-age biases facilitate visual awareness of faces under interocular suppression. *Frontiers in Human Neuroscience*, *8*. doi:10.3389/fnhum.2014.00582

Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., & Wallace, M. T. (2013). Brief Report: Arrested Development of Audiovisual Speech Perception in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-013-1992-7

Stoesz, B. M., & Jakobson, L. S. (2013). A sex difference in interference between identity and expression judgments with static but not dynamic faces. *Journal of Vision*, *13*, 1–14. doi:10.1167/13.5.26.doi

Stoesz, B. M., & Jakobson, L. S. (2014). Developmental changes in attention to faces and bodies in static and dynamic Scenes. *Frontiers in Perception Science*, *5*, 1–9. doi:10.3389/fpsyg.2014.00193

Sugita, Y. (2009). Innate face processing. *Current Opinion in Neurobiology*, *19*(1), 39–44. doi:10.1016/j.conb.2009.03.001

Susilo, T., Crookes, K., McKone, E., & Turner, H. (2009). The composite task reveals stronger holistic processing in children than adults for child faces. *PloS One*, *4*(7), e6460. doi:10.1371/journal.pone.0006460

Suzanne Scherf, K., Thomas, C., Doyle, J., & Behrmann, M. (2013). Emerging structure-function relations in the developing face processing system. *Cerebral Cortex*, 1–17. doi:10.1093/cercor/bht152

Szalma, J. L., Schmidt, T. N., Teo, G. W. L., & Hancock, P. a. (2014). Vigilance on the move: video game-based measurement of sustained attention. *Ergonomics*, (August), 1–22. doi:10.1080/00140139.2014.921329

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

Tanaka, J. W., Kiefer, M., & Bukach, C. M. (2004). A holistic account of the own-race effect in face recognition: evidence from a cross-cultural study. *Cognition*, *93*(1), B1–9. doi:10.1016/j.cognition.2003.09.011

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory & Cognition*, *25*(5), 583–92. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9337578

Taylor, M. J., Batty, M., & Itier, R. J. (2004). The faces of development: A review of early face processing over childhood. *Journal of Cognitive Neuroscience*, *16*(8), 1426–1442.

Taylor, N. M., Jakobson, L. S., Maurer, D., & Lewis, T. L. (2009). Differential vulnerability of global motion, global form, and biological motion processing in full-term and preterm children. *Neuropsychologia*, *47*(13), 2766–78. doi:10.1016/j.neuropsychologia.2009.06.001

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., & Morgan, J. L. (2013). Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. *Infancy*, *18*(4), 534–553. doi:10.1111/j.1532-7078.2012.00135.x

Terburg, D., Hooiveld, N., Aarts, H., Kenemans, J. L., & van Honk, J. (2011). Eye tracking unconscious face-to-face confrontations: Dominance motives prolong gaze to masked angry faces. *Psychological Science*, *22*(3), 314–9. doi:10.1177/0956797611398492

Thornton, I., & Kourtzi, Z. (2002). A matching advantage for dynamic human faces. *Perception*, *31*(1), 113–132. doi:10.1068/p3300

Thornton, I., Mullins, E., & Banahan, K. (2011). Motion can amplify the face-inversion effect. *Psihologija*, *44*(1), 5–22. doi:10.2298/PSI1101005T

Tong, F., & Nakayama, K. (1999). Robust representations for faces: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 10616–1035.

Tovée, M. J., & Cohen-Tovée, E. M. (1993). The neural substrates of face processing models - A review. *Cognitive Neuropsychology*, *10*(6), 505–528. Retrieved from www.isinet.com:WoK:UA

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009a). Emotions in motion: Dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, *1284*, 100–15. doi:10.1016/j.brainres.2009.05.075

Trautmann, S. A., Fehr, T., & Herrmann, M. (2009b). Emotions in motion: dynamic compared to static facial expressions of disgust and happiness reveal more widespread emotion-specific activations. *Brain Research*, *1284*, 100–15. doi:10.1016/j.brainres.2009.05.075

Tremblay, C., Champoux, F., Voss, P., Bacon, B. a, Lepore, F., & Théoret, H. (2007). Speech and non-speech audio-visual illusions: A developmental study. *PloS One*, *2*(1), e742. doi:10.1371/journal.pone.0000742

Troje, N. F., & Westhoff, C. (2006). The inversion effect in biological motion perception: evidence for a “life detector”? *Current Biology : CB*, *16*(8), 821–4. doi:10.1016/j.cub.2006.03.022

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant’s response to entrapment between contradictory messages in face-to-face interaction. *Journal of American Academy of Child Psychiatry*, *17*, 1–13.

Tsao, D. Y., Freiwald, W. a, Tootell, R. B. H., & Livingstone, M. S. (2006). A cortical region consisting entirely of face-selective cells. *Science (New York, N.Y.)*, *311*(5761), 670–4. doi:10.1126/science.1119983

Tsuchiya, N., Kawasaki, H., Oya, H., Howard, M. A., & Adolphs, R. (2008). Decoding face information in time, frequency and space from direct intracranial recordings of the human brain. *PloS One*, *3*(12), e3892. doi:10.1371/journal.pone.0003892

Turati, C., Valenza, E., Leo, I., & Simion, F. (2005). Three-month-olds' visual preference for faces and its underlying visual processing mechanisms. *Journal of Experimental Child Psychology*, *90*(3), 255–73. doi:10.1016/j.jecp.2004.11.001

Ungerleider, L. G., & Haxby, J. V. (1994). "What" and "where" in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157–65. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8038571

Valenza, E., Simion, F., Cassia, V. M., & Umilta, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 892–903. doi:10.1037/0096-1523.22.4.892

Valenza, E., Simion, F., Cassia, V. M., & Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology. Human Perception and Performance*, *22*(4), 892–903. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8756957

Vrticka, P., Lordier, L., Bediou, B., & Sander, D. (2014). Human amygdala response to dynamic facial expressions of positive and negative surprise. *Emotion*, *14*(1), 161–9. doi:10.1037/a0034619

Vuilleumier, P. (2000). Faces call for attention: Evidence from patients with visual extinction. *Neuropsychologia*, *38*, 693–700.

Wattam-Bell, J. (1996). Visual motion processing in one-month-old infants: habituation experiments. *Vision Research*, *36*(11), 1679–85. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8759468

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, *78*(1), 105–19. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10653509

Weigelt, S., Koldewyn, K., Dilks, D. D., Balas, B., McKone, E., & Kanwisher, N. (2014). Domain-specific development of face memory but not face perception. *Developmental Science*, *17*(1), 47–58. doi:10.1111/desc.12089

Wild-Wall, N., Dimigen, O., & Sommer, W. (2008). Interaction of facial expressions and familiarity: ERP evidence. *Biological Psychology*, *77*(2), 138–49. doi:10.1016/j.biopsycho.2007.10.001

Williams, M. A., Berberovic, N., & Mattingley, J. B. (2007). Abnormal FMRI adaptation to unfamiliar faces in a case of developmental prosopamnesia. *Current Biology : CB*, *17*(14), 1259–64. doi:10.1016/j.cub.2007.06.042

Williamson, K. E., & Jakobson, L. S. (2014). Social perception in children born at very low birthweight and its relationship with social/behavioral outcomes. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. doi:10.1111/jcpp.12210

Williamson, K. E., Jakobson, L. S., Saunders, D. R., & Troje, N. F. (n.d.). Local and global aspects of biological motion perception in children born at very low birth weight. *Child Neuropsychology*.

Xiao, N. G., Quinn, P., Ge, L., & Lee, K. (2012). Rigid facial motion influences featural, but not holistic, face processing. *Vision Research, 57*, 26–34. doi:10.1016/j.visres.2012.01.015

Xiao, N. G., Quinn, P., Ge, L., & Lee, K. (2013). Elastic Facial movement influences part-based but not holistic processing. *Journal of Experimental Psychology. Human Perception and Performance*. doi:10.1037/a0031631

Xu, B., & Tanaka, J. W. (2013). Does face inversion qualitatively change face processing: An eye movement study using a face change detection task. *Journal of Vision*, *13*(2), 1–16. doi:10.1167/13.2.22.doi

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

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747–59. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3454432

Yovel, G., & Kanwisher, N. (2004). Face perception: Domain specific, not process specific. *Neuron*, *44*(5), 889–98. doi:10.1016/j.neuron.2004.11.018

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology : CB*, *15*(24), 2256–62. doi:10.1016/j.cub.2005.10.072

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience*, *11*, 641–649.