Evaluating the Liking Effect Paradigm and the Impact of Realism in Face Processing Among Neurotypical and Autistic Populations

Alex Fraser

Somerville College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2019

Acknowledgements

First and foremost, I would like to thank Professor Gaia Scerif for her support throughout both my Masters and my DPhil. Without her help and guidance this project would not have been possible. Even though she is always running somewhere important she will find time for a catch up to put me on the right track. Gaia's willingness to be flexible and investigate experimental new methods of data collection meant that this project could adapt in ways I had not imagined when I first began. I look forward to working with her again in the future.

I would also like to thank all of the ABCD lab members, past and present. Advising me on practice presentations, teaching me how to use EEG, helping to collect data, and always providing fascinating conversations about their own research, they have all had some impact on my time at Oxford. Especially, I'd like to thank Brianna Doherty, Jacalyn Guy, and Megan von Spreckelsen who all directly contributed to the work in this thesis at different points. Kia Nobre also provided valuable support and advice throughout the project.

I would like to thank my parents who have guided me and supported me at all times. Both financially and emotionally, you've always been there when I was in trouble. It's not easy trying to help two children through doctorates simultaneously, especially when they both have to move back home at the same time! None of this would have been possible without your support.

I'm very grateful for the friends I have made at Oxford and the Somerville College MCR community. It has been a real pleasure being part of the MCR and it has been a constant source of comfortable working space, coffee, and relaxation. Again, there are too many people to thank to list here without offending somebody, but I'm going to give it a go. Thank you to Chloe Bracegirdle, Connor Scott, Gemma Seabright, Greg Simmonds, Kevin Judd, Oana Gurau, and Rowan

Nicholls. They have been with me from (almost) the beginning of my time at Oxford and the MCR would not have been as fun without them. I would also like to thank Gala and Kalil Copley, Jeff Martin, Fergus Cooper, Joanna Raisbeck, Ruben Janssen, Thomas Gaudelet, and James Hutton, who have all impacted my time at Oxford. In particular, I would like to thank Kate Finnie for your constant and unwavering support. I am not sure I would ever have been able to push through and finish this without you telling me that everything will be alright. I can't wait for everything that is yet to come.

Finally, I would like to thank Drs Andrew Bayliss and Robin Murphy for taking the time to examine this thesis and for their insightful suggestions. This project would not have been possible without the generous support from the Medical Research Council. I am also incredibly grateful for the transitional opportunities that have supported me as an early career researcher.

Abstract

The face plays an important role in human interactions, from identifying others to their to understanding their emotions and mental state. Particularly, eyes play a significant role in human communication. Individuals tend to follow the direction of another's eye gaze, resulting in faster more accurate processing. Furthermore, gaze cues may influence functional processing, such as improving memory and influencing affective judgments of objects. The observed preference for objects that have been cued by faces, referred to as the liking effect, only occurs when faces are used to cue objects. This notable difference between a gaze and a spatial cue opens the potential for this paradigm to investigate differences between faces of different levels of realism, a topic that could improve our understanding of face perception deficits in Autism Spectrum Disorders.

A series of experiments are presented across four chapters. In Chapter 2, the liking effect is investigated in long-term memory. Initially there were issues in replicating the cueing bias for visuospatial measures. To overcome this, the complexity of task was raised, emphasising the importance of the task in gaze cueing experiments. There was no explicit liking effect found, although there was a preference for items that were presented as part of learning trials. Experiments in Chapter 3 replicated the original experiment to investigate whether the effect occurs in both Schematic and Photographic faces. Initially, experiments failed to replicate the liking effect. Further experiments were able to replicate the liking effect with more detailed Schematic stimuli, but only if the target objects were cued by multiple Photographic faces. Chapter 4 was unable to replicate these findings across individuals with High and Low levels of autistic traits. The complexity of the design and limitations with sample size make it difficult to extrapolate the results beyond the current experiment. Finally, in Chapter 5 an eye-tracking

study is presented to investigate the effect in an ASD and a NT sample. It was found that the effect occurred in the NT group, but not in the ASD group. Insights from the eye tracking results suggest that explicit time observing each target may explain the effect and differences between the groups. In Chapter 6, findings are in the context of the wider literature. Conclusions are drawn about the efficacy of the liking effect as an index for studying the processing of facial realism in general and in ASD more specifically.

Chapter 1A: Literature Review

As a species, humans utilise faces as powerful signals to understand the intentions of others and to guide their behaviour (Burke & Sulikowski, 2013). Faces are our primary method of identifying other people, as well as understanding how they are feeling or what they may be thinking (Leopold & Rhodes, 2010). Previous research has found that observing the eye gaze of another person will influence how we orient our own attention (Friesen & Kingstone, 1998). However, there has been less focus on the influence that such social cues may have on subsequent cognition. This thesis will build on the current literature by utilising previously established paradigms that investigate the functional consequences of eye gaze on our affective judgments and our memory. In addition to developing our understanding of these functional consequences, the potential for paradigms to improve our current understanding of atypical face processing that has been observed in Autism Spectrum Disorder (ASD henceforth)will be investigated through both experimental evidence (Davies, Bishop, Manstead, & Tantam, 1994; Dawson, Webb, & McPartland, 2005; Jemel, Mottron, & Dawson, 2006), as well as anecdotal reports (Trevisan, Roberts, Lin, & Birmingham, 2017).

The first section of this chapter will review the broad literature around face perception, including classical models, the development of face processing abilities, and research into the pivotal role of eyes as a feature of the face. The second section will narrow the focus on the primary research questions of this thesis, specifically discussing how face perception is disrupted in ASD, as well as discussing potential theories of why this may occur. The final section will outline the potential for novel research about how faces are processed, and how they influence attention and social cognition, as well as proposing new behavioural methods to investigate face perception in both typical and atypical individuals.

Face Perception in Neurotypical Individuals

What is Face Perception?

Face perception refers to a range of abilities from simple visual processes such as detecting the presence of a face, to more complex processes (Bruce & Young, 1986; Rhodes, Calder, Johnson, & Haxby, 2012; Leopold & Rhodes, 2010; Young & Bruce, 2011). Face processing can broadly be separated into two categories - invariant and dynamic. Invariant processes typically include identity, gender and age; features that do not change in the short term. In contrast, dynamic processes are continually changing and adapting as required by their function - such as facial expressions, lip movements, and eye movements. Integrating the two categories of facial features allows us to efficiently build a perceptual representation of a face that will guide our subsequent behaviour. However, this raises the question: Are faces special?

The unique nature of facial processing can be outlined by several behavioural effects that demonstrate that faces are processed holistically as opposed to featurally. That is, faces are processed as a whole based on the configuration of features within the face; in contrast to featural processing, where objects are processed using specific features (J. Tanaka & Gordon, 2012). Three effects have been used to demonstrate holistic processing of faces in contrast to other non-face objects: the disproportionate inversion effect, the part-whole effect and the composite effect. The disproportionate inversion paradigm typically presents a target that can be either a face or a non-face object followed by two probe stimuli, that are either presented in the typical or inverted orientation. A participant will then select which probe matches the target. It has been found that individuals respond to faces that are inverted more slowly than faces that are upright (Susan & Rhea, 1977; Freire, Lee, & Symons, 2000; Yin, 1969). In contrast, there is little difference when participants are asked to identify objects, such as shoes (de Gelder, Frissen, Barton, & Hadjikhani, 2003), animals (Minnebusch, Suchan, & Daum, 2009), houses, and chairs (Boutet & Faubert, 2006).

The inversion effect is an indirect measure of holistic processing. As the inversion of the face stimuli results in individuals having to utilise a featural approach rather than the more efficient holistic method, the responses are slower. In the case of non-face objects, as the object is processed by individual features whether the stimuli is inverted or not, the inversion does not elicit a significant delay. Direct measures of holistic versus featural processing manipulate the features of the face itself. In the part-whole paradigm, as in the inversion effect paradigm, participants are shown a target, followed by two probe stimuli and are asked to identify which probe matches the target (J. Tanaka & Farah, 1995). The probes are either presented as a whole image or as a single feature or small section of the target. When the probes consist of only a single feature of the target face, accuracy drops significantly compared to when the probes consist of whole faces, whereas little difference is seen between the part or the whole probes when non-face objects are used as the targets. The composite face effect occurs when two distinct face identities are combined. That is, when the top half and the bottom half of a face are aligned, the image begins to take on a distinct identity. However, when the two halves are misaligned participants are better at distinguishing a difference between two identities; an effect that is not seen in non-face objects. (Young, Hellawell, & Hay, 1987).

Visual Preference for Face Shapes in Early Development

Faces appear to be highly attractive during the very early stages of human development. It has been observed that infants spend more time looking at faces when viewing scenes (Gliga, Elsabbagh, Andravizou, & Johnson, 2009), and they will even attend to more attractive faces for longer (Slater et al., 2000, 1998). There is even evidence that hours old neonates will follow simple face-like shapes but will not follow inverted versions of the same stimuli (Johnson, 2005; Johnson, Senju, & Tomalski, 2015; Morton & Johnson, 1991). See Figure 1.1 for an example of the stimuli used in these experiments. Although there are ongoing debates around how sophisticated neonates face detection is (Macchi Cassia, Valenza, Simion, & Leo, 2008; Quinn & Slater, 2003; Simion, Macchi Cassia, Turati, & Valenza, 2003), the fact that a simple face shape will draw the attention of neonates, who have limited vision and minimal experience, would suggest that there are mechanisms that accurately guide neonates to attend to face-like stimuli. In fact, this early attraction may subsequently guide how humans learn about faces (Johnson, Grossmann, & Kadosh, 2009; Johnson et al., 2015). These mechanisms may be grounded in fundamental perceptual preferences: there are even preliminary findings that foetuses will turn to follow dots of light, similar to the stimuli used by Morton and Morton and Johnson (1991), that is projected through the skin of the mother, but will not if the shape is inverted (Reid et al., 2017). However, further replication of these findings is required.

Figure 1.1: Example of simple stimuli used in neonatal research. It is found that infants will follow the image that is configured in a similar orientation to a face (left). However, when the stimuli are inverted (right), infants do not appear to follow the stimuli to the same extent. This behaviour may suggest an early preference to face like configurations necessary for the development of face perception abilities. Figure adapted from Johnson et al. (2015).

Although initial attraction to face shapes could suggest a level of holistic processing, this does not appear to be the case. Original findings suggested that infants utilised isolated features when observing faces, and it was only around the age of 10 that children would switch to more sophisticated holistic methods of face processing (Carey & Diamond, 1994). Named the encoding switch hypothesis, it has been challenged by findings of holistic style face processing in infants and children below the age of 10. For example, it has been found that children between the ages of 3-5 months develop a sensitivity to faces (Hayden, Bhatt, Reed, Corbly, & Joseph, 2007; Quinn & Slater, 2003). There is also further evidence that 4-year-olds (de Heering, Houthuys, & Rossion, 2007) and 6-year-olds (Mondloch, Pathman, Maurer, Le Grand, & de Schonen, 2007) demonstrate adult-like holistic face processing. It has been proposed that rather than encoding switching as originally proposed, individuals develop holistic processing through childhood and adolescence until full maturity (K. Lee, Anzures, Quinn, Pascalis, & Slater, 2012). It has been suggested that children do not reach the level of adult performance at holistic face processing tasks until approximately 15 years old (Mondloch, Le Grand, & Maurer, 2002).

Is Face Processing Domain Specific?

Holistic processing of faces is often used as an argument that face processing is a domain-specific behaviour, although there is still debate as to what extent this is the case (Gauthier & Nelson, 2001; McKone & Robbins, 2012). One argument against domain specificity is that studies of face processing vs object recognition often present a direct comparison between faces and a single non-face category. Therefore, it is not sufficient to claim faces are unique, as that assumes that a single object category represents all non-face objects (Gauthier & Nelson, 2001). Research utilising paradigms that compare multiple objects and faces in memory-based object recognition tasks identified two categories of objects that require unique processing abilities: Faces and cars (McGugin, Richler, Herzmann, Speegle, & Gauthier, 2012; Van Gulick, McGugin, & Gauthier, 2016). There is evidence to suggest that expertise predicts the use of holistic processing when recognising cars, relating to the level of neural activity in the Fusiform Face Area (FFA) in recognition tasks where the target stimuli are cars (Anderson, Gauthier, Skudlarski, & Gore, 2000; McGugin et al., 2012). However, contrary evidence found that specific car knowledge does not mediate object recognition of cars (Richler, Wilmer, & Gauthier, 2017), although the authors state the difficulty in specifying what classifies as expertise.

Proponents of the theory that face processing is a more domain-general process suggest that differences between holistic and featural processing in different object categories are driven by expertise. Evidence by Shakeshaft and Plomin (2015) has found that although intelligence may influence object recognition tasks, general visual ability appears to be independent of intelligence. These findings across multiple object categories, including both cars and faces, would suggest face processing is a more domain-general ability and holistic processing is evidence of expertise.

Similarities between the processing of face and car stimuli may suggest that face processing abilities are domain-general but are driven by a specific preference for face shapes. It has been found that recognition of faces and cars is highly heritable, more so than for other objects (Wilmer et al., 2010; Shakeshaft & Plomin, 2015). Although face processing may be heritable, it has also been found that face recognition increases alongside population density, which would suggest a level of improvement through exposure (Balas & Saville, 2016). Although it has not been studied directly it was proposed by Gauthier (2018) that a similar improvement in car recognition may occur in those who live around a large number of vehicles (e.g., those who live in cities). It has been found that the front of a car may show some similar patterns with faces, headlights comparable to eyes for example and the more similar a car's proportion is to a face, the higher the FFA activation is (Kühn, Brick, Müller, & Gallinat, 2014). If it is the case that a mechanism for rapid face detection using Low Spatial Frequency (LSF) face shapes drives efficient face processing, it is possible that, as a result of spatial similarities, car recognition has hijacked a specific face detection mechanism. Therefore, processing of cars may then develop through a domain-general style of object recognition that is typically only observed in faces. If this is the case, such a mechanism could also explain findings that face-like greebles elicit holistic style processing after a period of training (Chua, Richler, & Gauthier, 2015). Essentially, if there are specific mechanisms that guide individuals to attend to face-like shapes, subsequently building active recognition networks required for holistic processing; is it possible stimuli that match face-like patterns, such as cars and greebles, may also activate these mechanisms? This explains why faces consistently require more sophisticated processing styles, whereas holistic processing of non-face objects is more heterogeneous than faces. The next section will discuss models of face perception and provide a more detailed explanation of how an attraction to face shapes may develop and the knock-on effect on the domain-general mechanism for mature face perception.

Models and Measures of Face Processing

Behavioural Models of Face Perception in Adults

Models of face perception tend to divide elements of face processing into two categories, dynamic and invariant. The fundamental basis for this division came from Bruce and Young's (1986) model of face recognition, which split face perception into seven separate face modules (see Figure 1.2). The first module is defined as structural encoding, which accounted for the structure of the face and its expressive nature. This module then divides down two separate pathways. The first pathway accounts for invariant aspects of face processing and is made up of three modules; face recognition units, person identity nodes and name generation. The second pathway accounts for the dynamic aspects of the face and focuses on perceived expression and language with three distinct units: expression analysis, facial speech analysis and directed visual processing. The Bruce and Young (1986) model focuses purely on a cognitive interpretation but has been the backbone of research investigating the neural mechanisms of face perception, influencing the direction of face perception research 30 years later (Engell, Kim, & McCarthy, 2018). This model was updated in 2011 to discuss research that both supported and conflicted with the original findings (Young & Bruce, 2011). Specifically, it includes the addition of gaze perception, stating that current research provides evidence that gaze perception dissociates from other kinds of transient facial signal and would, therefore, warrant its specific module.

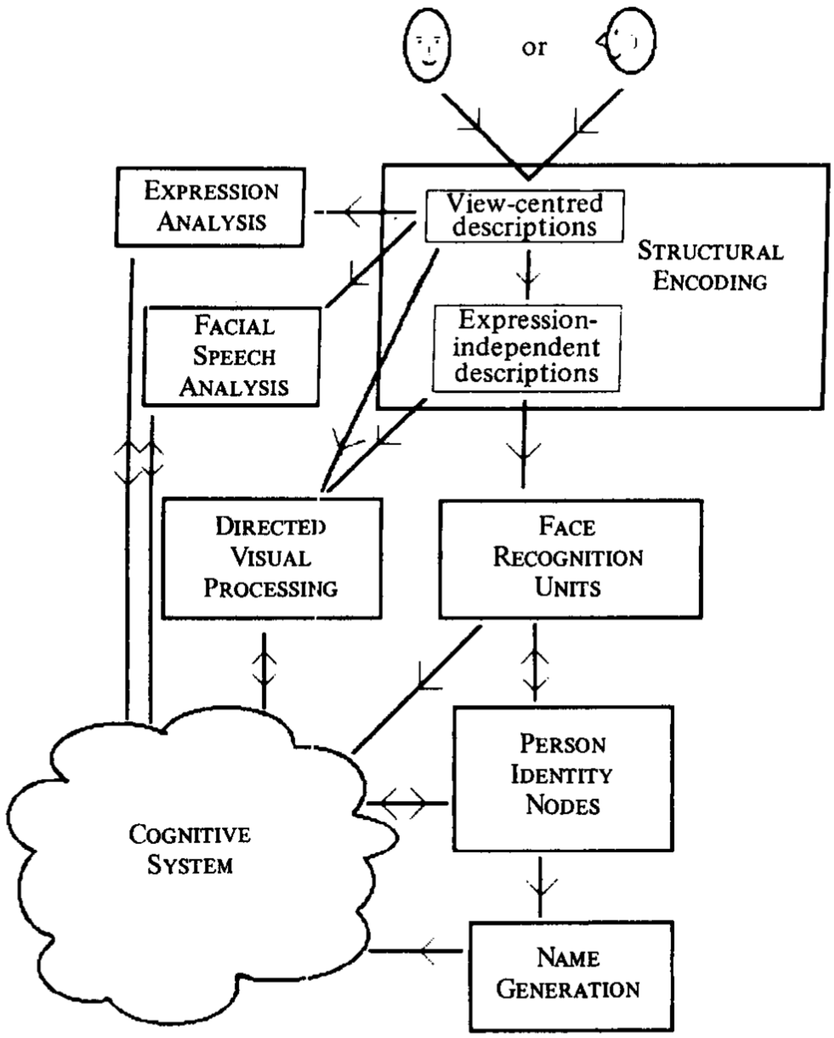

Figure 1.2: Original functional model of face recognition from Bruce and Young (1986). In this model, it is possible to see how the face modules interact with each other in two distinct pathways after structural encoding has occurred. The invariant pathway consisted of face recognition units, person identity nodes and name generation and the dynamic pathway, which consists of expression analysis, facial speech analysis and direct visual processing.

Measures of Neural Correlates of Face Processing

Investigations into the neural correlates of face perception have found that, when viewing faces compared to other, non-face stimuli, there is greater activity in the Occipital Face Area (OFA) and the Fusiform Face Area (Bernstein & Yovel, 2015; Haxby, Hoffman, & Gobbini, 2000). It has also been found that it is possible to selectively disrupt facial recognition without disrupting object or body recognition through Transcranial Magnetic Stimulation (TMS) over the right OFA

(Pitcher, Charles, Devlin, Walsh, & Duchaine, 2009). These two areas are part of what is often referred to as the Social Brain Network (SBN henceforth). The SBN consists of distributed networks of brain regions that are associated with a range of social processing. For example, the Superior Temporal Sulcus (STS) has been associated with the processing of biological motion, in faces (Campbell, Heywood, Cowey, Regard, & Landis, 1990), and the rest of the body (Grossman & Blake, 2001). There is even evidence that the STS appears to be active in the perception of human voices (Belin, Zatorre, Lafallie, Ahad, & Pike, 2000). The SBN also extends to subcortical regions that have been found to be associated with processing of emotions (Damasio et al., 2000; Ladouceur, Phillips, & Drevets, 2008; Wager, Davidson, Hughes, Lindquist, & Ochsner, 2008) - in particular the amygdala (Banks, Eddy, Angstadt, Nathan, & Luan Phan, 2007; Hare, Tottenham, Davidson, Glover, & Casey, 2005; Vuilleumier & Pourtois, 2007). The ventral striatum and the orbitofrontal cortex have been associated with processing facial attractiveness and other social information (Aharon et al., 2004; Winston, O'Doherty, Kilner, Perrett, & Dolan, 2007). Furthermore, the medial Prefrontal Cortex (mPFC) has been associated with the theory of mind judgments (Siegal & Varley, 2002; Stuss, Gallup, & Alexander, 2001). These are only a few examples of the extent of the research on the SBN, and as researchers move to more network-based approaches, our understanding of the SBN will expand (McCormick, van Hoorn, Cohen, & Telzer, 2018).

Neural Models of Face Perception

The most influential neural model of face perception was first described by Haxby et al. (2000). The model utilises specific neural regions to establish a distributed neural model of functional face perception (see Figure 1.3). In the Haxby model, early visual perception of facial features occurs in the OFA before splitting down the dorsal and the ventral pathway. The dorsal pathway projects to the STS and correlates with dynamic aspects of face perception such as eye direction, lip movement and facial expression. Information is then fed forward to extended systems, such as the Intraparietal Sulcus (IPS), the auditory cortex, and the amygdala and limbic system. The ventral pathway projects to the FFA and correlates with invariant identity processes before passing to the anterior temporal lobe to process biographical information about the individual. The Haxby model has been the dominant model of neural face processing, utilising the widely accepted two pathways associated with facial processing. However, it has been suggested that the two pathways do not align with the dynamic and invariable behaviours that were originally proposed (Bernstein & Yovel, 2015; Duchaine & Yovel, 2015). In a review of current research, they found that while identity was only processed in the FFA (ventral path), evidence of activity related to expression was found in both the FFA and STS (ventral and dorsal pathways). Therefore, they suggested that what had previously been subdivided into dynamic and invariable processes may represent signatures of form and motion processing. These suggestions align with findings by Calder and Young (2005) that expression and identity are processed using common mechanisms. Although there may still be some debate as to what extent each pathway contributes to each feature of face perception, there is a consensus that there are two pathways that project from the first perceptual features processed in the OFA.

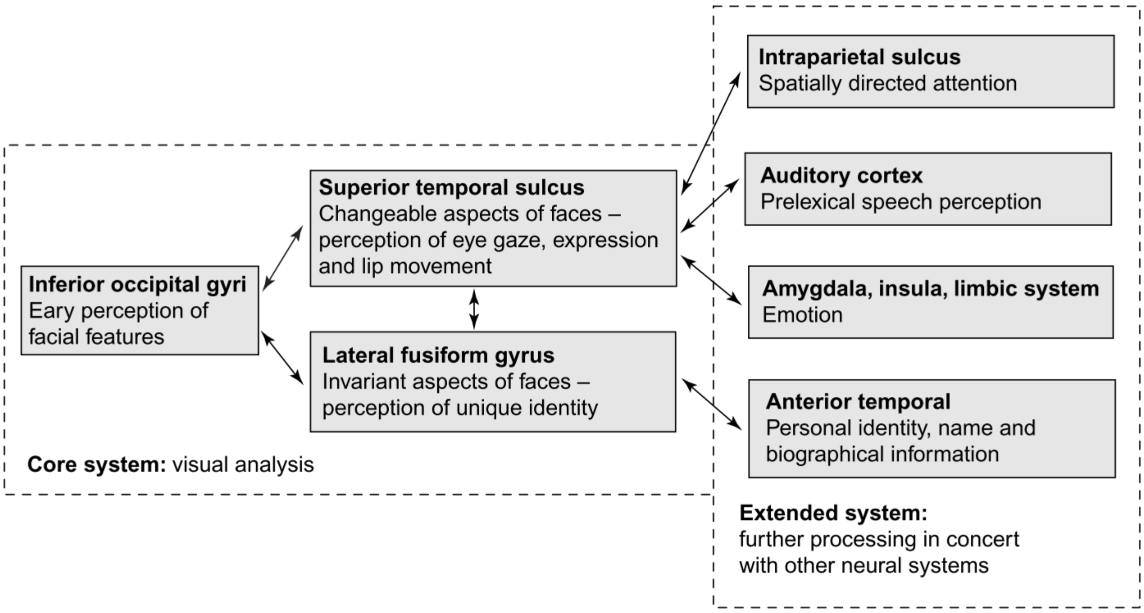

Figure 1.3: Model of a distributed human neural system for face perception from Haxby et al. (2000). In the Haxby model, it is possible to see, as in the Bruce and Young (1986) model, that the visual analysis is separated down two distinct pathways - Dynamic and Invariant. However, the Haxby model builds on the Bruce and Young model by providing the associated neural pathways that correlate with each stage of visual processing.

Measures of Temporal Dynamics in Face Processing

When viewing faces, we see patterns of neural activity that appear to be unique to faces emerging very early in processing. Electroencephalography (EEG) measures electrical activity in the brain through electrodes placed on the scalp. Although it is limited in its spatial resolution, it can collect data on the millisecond timescale (Luck, 2014), allowing researchers to investigate beyond observed behaviour by measuring EEG responses during or after the desired behaviour. One method used for measuring brain waves is Event-Related Potentials (ERP); a measured brain response that is the direct response to the onset of a stimulus (Luck, 2014). The N170, characterised by a negative peak at 170 ms after the onset of the stimuli, shows face-selective properties, demonstrated by a greater response to face stimuli over the occipitotemporal electrodes compared to other non-face stimuli (Bentin, Truett, Puce, Perez, & McCarthy, 1996; Rossion et al., 2003). Before the N170, the P1 (a positive peak 100 ms after stimulus onset) is also larger when viewing faces, although this may be influenced by participants' attention and task load (Hileman, Henderson, Mundy, Newell, & Jaime, 2011). These findings were corroborated using intracerebral EEG recordings, where electrodes are implanted directly onto the cortex at sites of interest. Barbeau et al. (2008) found similar patterns of neurological response that would correspond to the P1 – N170 in the fusiform gyrus, as well as the inferior frontal gyrus. They also found subsequent activity in the parahippocampal gyrus and hippocampus (240-480 ms) regions that are associated with distinguishing identity.

In addition to EEG, the use of eye-tracking can provide evidence of the temporal dynamics of facial processing. It has been found that 85% of people will attend to the eyes as their first fixation when shown a face for only 1 second, dropping to 75% for longer image presentations (Iskra & Gabrijelčič Tomc, 2017). They also found that in longer time frames of 4 seconds, 95% of participants would return their attention to the eyes after making secondary fixations. Furthermore, Chuk, Chan, and Hsiao (2014), found that first fixations would tend towards the centre of the face before shifting fixations towards an individual eye (predominantly the left eye, although there were individual differences) in subsequent fixations. Chuk and colleagues also deduced that individuals who focused towards the central eye regions and spent less time investigating other regions of the face were more likely to use holistic style processing than those who shifted their attention to multiple features of the face. Temporally sensitive techniques, such as eye-tracking, will be discussed in more detail in Chapter 5.

Development of Face Processing Abilities

Two Process Theory of Face Recognition

Attention to face shapes has been demonstrated in infants that are hours old using gaze following experiments. Morton and Johnson (1991) suggested a two-process theory of infant face recognition. The first process, CONSPEC, guides the innate preference for a face like configuration. In their experiments, they found that infants displayed preferential looking towards Low Spatial Frequency (LSF) shapes that replicated facial structures, but not if the shapes were inverted. Johnson's concept of CONSPEC established the ability for LSF face shapes to draw newborns' attention. One proposal was that the LSF of basic face structures makes them the most visually salient stimulus in neonates' limited visual field. This explanation was countered through experiments that found upright configuration was still the prominent influence on newborn's preference towards faces, even when the saliency was controlled (Morton & Johnson, 1991; Morton, Johnson, & Maurer, 1990). Subsequently, it was thought that there might be a more domain-general bias towards stimuli that were top-heavy as this may appeal to neonates limited visual capacity, but this does not appear to be the case (Chien, 2011). There have also been suggestions of evidence for more complex innate face processing behaviours (Quinn & Slater, 2003), for example, a preference for eyes within a face (Batki, Baron-Cohen, Wheelwright, Connellan, & Ahluwalia, 2000), and a preference for attractive faces (Slater et al., 2000, 1998). However, a high-level explanation is implausible, due to immaturity of the visual cortex at birth, but also because these behaviours can all be explained through LSF face configurations (Johnson, 2005). While CONSPEC describes early behaviours associated with following face shapes, the second process, CONLEARN, develops around two months (Morton & Johnson, 1991). CONLEARN describes the development of more advanced behaviours that come about through engagement with faces as a result of CONSPEC. Johnson and colleagues (1991) originally proposed that as cortical regions develop, CONSPEC processes are no longer needed and are replaced by CONLEARN. However, subsequent revisions have proposed that this may not totally be the case (Johnson, 2005), as will be discussed in more detail later in this chapter. The CONSPEC and CONLEARN approach does not inherently provide support for either an innate or an expertise-driven hypothesis. Although the rapid development of CONSPEC behaviours would suggest a level of innate process, the way that CONSPEC gives way to CONLEARN behaviours would be more indicative of an expertise-based model. In later modifications Johnson et al. (2015) suggest that CONLEARN cannot purely be a result of learnt experience due to the consistency of the specific networks between individuals, as well as developmental evidence of specific face-selective neural activity Johnson et al. (2015).

Subcortical Pathways and the Interactive Specialisation Account of Face Perception

Morton and Johnson (1991) proposed that infants move away from CONSPEC processes as they develop CONLEARN abilities. However, Johnson (2005) presented evidence of subcortical neural activity in response to LSF face shapes. It was found that activity in the superior colliculus, the pulvinar and the amygdala correlated with LSF face stimuli, which subsequently mediating later activity in the extrastriate area (Morris et al., 1998). These sub-cortical pathways have been found to activate in response to LSF rapidly face stimuli (Senju & Johnson, 2009a), suggesting that CONSPEC abilities do not disappear during development and remain as an early face detection process that also modulates later CONLEARN processes. Further evidence of the independence of the subcortical network comes from the activity and perceptual abilities in those with disrupted face perception, namely, activity in the presence of faces in individuals with blindsight and prosopagnosia (de Gelder et al., 2003; Morris, 2002).

Johnson proposed that the development of face processing can be explained in terms of the interactive specialisation hypothesis (Johnson et al., 2009, 2015; Johnson, 2000). Interactive specialisation proposes that there is an underlying architecture in the brain that develops prenatally. Then, through exposure and experience with external stimuli, individuals build and develop neural pathways for behaviour through a series of multidirectional, back-propagated interactions between genetics, brain and environment (Johnson, 2000). This approach has gained increasing attention due to advances in neuroimaging studies that conform with the interactive specialisation hypothesis (Amso & Casey, 2006; Johnson et al., 2009). The interactive specialisation hypothesis can explain why different individuals show similar networks and developmental trajectories, but also considerable heterogeneity as a result of the level of exposure received to the required stimuli (Johnson et al., 2009). Johnson used face perception as an example to demonstrate how subcortical activity resulting in attraction to LSF face shapes also results in more considerable experience with faces. The exposure to faces results in the development of a level of expertise, which would manifest as holistic processing (Johnson, 2000; Johnson et al., 2015) — possibly explaining the similarities between object recognition of cars and faces discussed earlier. If the LSF shape of a front-facing car matches that of a face, subcortical pathways may modulate further domain-general networks that allow for sufficient specialisation for processing cars through greater exposure. Therefore, a level of expertise for categorising cars may occur even among individuals who would not be considered car experts.

Pivotal Role of Eyes as a Feature of Face Perception

From birth, we find that individuals typically focus on eyes more than any other facial feature (Dalton et al., 2005; Hernandez et al., 2009). Eye contact is also perceived differently depending on context. Avoiding eye contact could be perceived as disrespectful, and prolonged eye contact could be perceived as aggressive, or romantic, depending on who the other person is and what facial expression they are displaying. These examples outline the intricate nature of perceived eye contact, but what is it about the eyes that are so special, and how are eyes processed differently from other facial features?

To understand how eyes are special, we need to understand what it is about eye contact that captures our attention. Multiple studies show clear attentional advantages that result from processing direct eye contact more efficiently, i.e., our ability to detect faces with direct eye gaze faster compared to faces with averted eye gaze (Senju, Kikuchi, Hasegawa, Tojo, & Osanai, 2008; Senju, Hasegawa, & Tojo, 2005). It is not just rapid detection that demonstrates how important eyes are for face perception; we use a combination of facial features to understand emotional expressions. However, people perform better at extrapolating emotional information from the eyes than they are from other features (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). For example, prolonged direct eye contact can be interpreted as aggressive (Adams & Kleck, 2005; Adams, Gordon, Baird, Ambady, & Kleck, 2003). Direct eye contact, coupled with an angry expression will result in spikes in amygdala activity, but when a fear expression is shown, spikes in amygdala activity are correlated with averted eye gaze (Adams & Kleck, 2005). Repeated direct eye contact combined with smiling can indicate attraction (Walsh & Hewitt, 2011). It has also been found that individuals making direct eye contact are rated as more attractive (Ewing, Rhodes, & Pellicano, 2010). Finally, direct eye gaze can subtly change an individual's behaviour, such as increasing levels of imitation between a participant and a model (Wang, Newport, & Hamilton, 2010). Referred to as the eye contact effect, these behaviours provide examples for the impact eye contact can have on social interactions. In the next section, different theories of the mechanisms behind the eye contact effect are discussed.

Theories of the Eye Contact Effect

It is clear that eyes are a particularly important stimulus, but why? What is it about eyes that draw our attention and what mechanisms give them their unique properties? The first theory is the affective arousal model. This model proposes that eye contact directly elicits the brains arousal system or elicits a strong emotional response, which in turn leads to enhanced perceptual and cognitive processing of faces (Hood, Macrae, Cole-Davies, & Dias, 2003; Kawashima et al., 1999). This model conforms with the findings that emotional responses trigger visceral autonomic responses (Pfaff, Ribeiro, Matthews, & Kow, 2008), which are consistent with physiological responses when viewing faces (Nichols & Champness, 1971). There are also suggestions that this model can account for differences between responses as a result of other face processes, such as expression and attractiveness (Vuilleumier & Pourtois, 2007; Winston et al., 2007). However, there are limitations in the generalised nature of this account. If the eye contact effect were the result of a general arousal mechanism, the activation would be more widespread than the areas associated with the SBN that are activated by faces in previous studies (Freiwald, Duchaine, & Yovel, 2016). Also, it has been found that the psychophysiological differences are incongruent with the emotional arousal measures. For example, measures of pupil dilation do not account for the raised activity in the medial prefrontal cortex (Kampe, Knut; Frith, Chris; Frith, 2003). Therefore, it appears unlikely that the affective arousal model can fully explain the eye contact effect.

An alternative suggestion is that eye contact may directly activate theory of mind processes that will prepare the individual for communication (Conty, N'Diaye, Tijus, & George, 2007; Pelphrey, Viola, & McCarthy, 2004), coined the communicative intention detector model. It has been found that many areas of the brain that are responsible for theory of mind calculations also overlap substantially with the areas of the social brain network which are activated through direct eye contact (Schultz et al., 2003; Conty et al., 2007; Pelphrey et al., 2004). This model offers a good account of the eye contact effect, but the two networks do not entirely overlap. That is, only parts of the theory of mind networks are activated during perceived eye contact. It is also notable that the varying levels of activation as a result of task demands and context contradict the suggestion that the eye contact effect is solely a result of the theory of mind activity (Senju & Johnson, 2009b).

An alternative model of the eye contact effect that provides a more comprehensive explanation has been proposed by Senju and Johnson (2009a). The fast-track modulator model proposes that the eye contact effect is mediated by the sub-cortical face detection pathway. The subcortical pathway is activated by LSF visual information that is best represented by faces. This route is fast compared to cortical activity, as shown by MEG data, from the pulvinar, amygdala and superior colliculus (Bailey, Braeutigam, Jousmäki, & Swithenby, 2005; Winston, Vuilleumier, & Dolan, 2003). Intra-cranial ERPs have also been used to show that there is subcortical activation that precedes cortical activation typically associated with the structural encoding of faces (Krolak-Salmon, Hénaff, Vighetto, Bertrand, & Mauguière, 2004). This activity is also correlated with greater activity in the fusiform gyrus (George, Driver, & Dolan, 2001; Kleinhans et al., 2008). There is also further activation in the medial prefrontal cortex and the orbitofrontal cortex that modulates the activation of the social brain networks areas that may account for activity proposed in the ToM model, instigating communication intent to allow for efficient interactions (Conty et al., 2007). The fast track activation model can account for the fast detection of eyes as well as early modulation of the SBN. These mechanisms also account for the findings in the theory of mind based communicative intention model. In doing so the fast track modulator model can account for many of the behaviours associated with eye gaze (Senju & Johnson, 2009b).

The subcortical fast track modulator model may also provide an adequate explanation of the characteristic infant behaviours associated with CONSPEC (Johnson, 2005), because subcortical structures are often associated with very early behavioural responses due to the relatively mature prenatal development (Sokolowski & Corbin, 2012). Therefore, it would make sense that rapid detection of LSF face shapes would appear very early in development. The fast track modulator model also fits with the proposed interactive specialisation theory (Johnson et al., 2009, 2015). That is, the fast track module uses LSF eye shapes to encourage interaction with faces (CONSPEC) from an early age. Early interactions with faces result in the development of connections within the specialised network that manifest in CONLEARN behaviours.

Perception of Averted Eye Gaze

As well as direct eye gaze, we utilise averted eye gaze to help ease interactions with others. One experimental paradigm that is often used to investigate how visuospatial cues will result in attentional shifts in observers is the Posner cueing paradigm (Posner, 1980). These paradigms as a whole utilise an attentional cue to shift the attention of an observer to a target and can be easily manipulated to test different styles of attentional cues, including perceived eye movements (Mari, Bennett, Scarpa, Brighetti, & Castiello, 1997; Morton & Johnson, 1991). Measurement of accuracy and reaction times demonstrate that the visuospatial attention of the participant has been shifted towards the cued location. That is, participants will show a positive bias of faster and more accurate responses towards congruently cued targets if the cue is shifting their attention.

When using cueing paradigms to measure attention, two types of cues are used; centrally presented endogenous cues and peripherally presented exogenous cues. These two types of cues shift attention in different ways. While exogenous cues cue attention in a reflexive bottom-up style, endogenous cues are more interpretive and top-down (Friesen & Kingstone, 1998; Friesen, Ristic, & Kingstone, 2004; Kingstone, Tipper, Ristic, & Ngan, 2004). As a result of this distinction, there are observable differences between how people suppress their responses to attentional cues. Typically, it is seen that the exogenous cues cannot be suppressed as these are a bottom-up, reflexive reaction to a peripheral stimulus. In contrast, endogenous cues, which are top-down reactions to a learnt stimulus, can typically be suppressed to some extent (Berger, Henik, & Rafal, 2005). In the context of faces, eye-gaze cues result in a typical cueing effect when an averted gaze from a face is used to cue a participant to an upcoming target location. Participants respond faster and more accurately when the gaze is directed congruently towards the target location (Bayliss & Tipper, 2005; Bayliss, Pellegrino, & Tipper, 2005; Friesen & Kingstone, 1998).

Adult humans are so efficient at detecting changes in the variation of eye gaze that a deviation of as little as 2.8°will be perceived as deviated gaze (Pick & Gibson, 1963). Human infants have been found to detect a deviated gaze at around 5°(Symons, Hains, & Muir, 1998). However, the age at which our ability to follow gaze indeed develops to be adult-like is not so easy to pin down. It has been suggested that children's ability to genuinely determine eye contact and gaze direction may not develop until the age of three (Doherty & Anderson, 2000). However, there are also examples of children younger than one year, and even as young as three months old, being able to follow the eye gaze of another person (Hood, Douglas Willen, & Driver, 1998; Scaife & Bruner, 1975). Of note, it has been suggested that infants' abilities are more straightforward than those of older individuals and rely on movement effects, more than processing the direction of the eye gaze. For example, Farroni, Massaccesi, Pividori, and Johnson (2004) found that neonates must observe the eye movements and a static image was not enough to establish gaze following. Further evidence found that in 9-month olds, those that had developed gaze following would also respond to static images, but those who did not display gaze following would not. Furthermore, 4-5-month-old infants would follow the direction in which a face moved, even if the eyes were directing the opposite way, whereas an adult would attend towards the direction of the eye gaze regardless of head direction (Bayliss & Tipper, 2005; Farroni, Johnson, Brockbank, & Simion, 2000). Therefore, more-adult-like abilities appear to emerge around the age of 12-18 months. Although infants' abilities are more simplistic than joint attention abilities in older children, they have been found to predict other behavioural abilities such as noun acquisition (Morales et al., 2000; Tomasello & Farrar, 1986), as well as the theory of mind abilities (Charman et al., 2000).

Functional Consequences of Eye Gaze Processing

Beyond the ability to cue visuospatial attention, eye gaze can also influence other aspects of human cognition. Combinations of eye gaze direction and facial expression will evoke different threat responses, demonstrated by amygdala activity. Spikes in amygdala activity are associated with a fearful face when the gaze is averted rather than direct. Conversely, when a face is displaying angry expression, spikes in amygdala activity are observed when making direct eye contact (Adams & Kleck, 2005). Humans will also use the eye gaze of others to make judgments about the individual. In experiments where faces would either congruently or incongruently gaze towards a target object, faces that consistently gazed away from the target were rated as less trustworthy, even if the participant is unaware that they were consistently averting their gaze (Bayliss & Tipper, 2006). Furthermore, it has been found that observed gaze cues can impact our interactions with our environment. For example, an object congruently cued by a face will be rated preferable compared to an object that was incongruently cued. However, if an arrow is used to cue attention, this effect is not seen (Bayliss, Paul, Cannon, & Tipper, 2006). These findings expand the effect of eye gaze away from the individual, demonstrating the power that a gaze shift may have on another person's cognition.

Neural Correlates of Eye Gaze Perception

A broad review of the networks responsible for the processing of eye gaze was recently conducted by (Carlin & Calder, 2013). They found three areas that showed specific activations to eye gaze: the anterior superior temporal sulcus (aSTS); the lateral parietal cortex and the medial prefrontal cortex (mPFC). They proposed that each area accounted for a different component of behaviour when processing averted eye gaze. The aSTS was involved in the perception of the gaze direction and appeared to distinguish averted gaze in line with various head directions Carlin and Calder (2013). The lateral parietal cortex appears to deal with attentional orienting, as activity in this areas has been observed in both gaze and non-gaze cueing experiments (Materna, Dicke, & Thier, 2008; Shepherd, Klein, Deaner, & Platt, 2009). It has also been proposed that the mPFC is active when engaging in joint attention (Schilbach et al., 2010). These findings would correspond with the proposed functions of direct eye contact. That is, in the Johnson et al. (2015) model of the fast track module, the mPFC is suggested to modulate communicative intent, which would also be a vital modulation when dealing with averted eye gaze as a form of communication. Both direct and averted eye gaze have incredible social importance and are crucial for smooth social interactions, but what would the impact be if there was interference with the face and gaze processing?

Summary

The current section has provided a broad overview of the current literature on face perception. Although there is still debate as to what extent face processing is either domain-specific or domain-general, the interactive specialisation account provides a good domain-general account, that can also explain the aspects of face perception that appear innate and more domain-specific in nature. Building on this debate, both historic and more recent models of face perception were discussed. Comparisons between the cognitive and neural models have identified two distinct pathways for face processing ability, with the most recent evidence suggesting that these pathways process form and motion independently. Furthermore, the two-process theory of face recognition is discussed to gain some insight into how face perception develops in infancy. The implications of recent additions of subcortical pathways may explain why infants are drawn to face-like shapes, as well as the potential for how interactive specialisation may account for the development of practical face processing abilities. Finally, the critical role that eyes play in face processing is discussed with a focus on how the observed eye gaze of another person can influence the attention and cognitive processing of observers. Now that we have established the typical methods of how face processing is investigated as well as current theories about the development of face processing and gaze processing, the subsequent section will discuss literature more relevant to the main research questions of the thesis. Namely, to what extent is face perception atypical in those with ASD and to what extent can any atypicality be explained by the visual properties and realism of the face.

Chapter 1B: Background to the Thesis

This section will discuss the current literature focused on the primary research aims of this thesis. The previous section has provided a broad overview of the literature of face perception and eye gaze perception. However, this has focused primarily on face perception in Neurotypical populations (NT henceforth). In this section, the evidence for atypical face perception in ASD will be considered and potential explanations discussed. Furthermore, this section will outline the evidence that autistic people do not suffer from similar atypicalities when processing faces with less realistic features. Finally, three research questions will be established that aim to investigate the potential for gaze cueing paradigms to disentangle the relationship between facial detail and face processing in ASD.

Disruption of Face and Gaze Perception in ASD

Autism spectrum disorder is a neurodevelopment disorder, based on the diagnostic criteria of the DSM V (American Psychiatric Association, 2013). ASD is diagnosed across two symptom pathways: Problems with social communication and social interactions; and restrictive and repetitive behaviours (American Psychiatric Association, 2013). Therefore, with their inherent social nature, faces can be an ambiguous stimulus for autistic individuals. One finding among autistic children is that they are less able to process faces in a holistic style at a comparable level with their peers (R. P. Hobson, Ouston, & Lee, 1988), although they may develop the ability later in adulthood (Nishimura, Rutherford, & Maurer, 2008). These deficits may impact how efficiently faces are processed in ASD (Behrmann, Thomas, & Humphreys, 2006), as well as emotion recognition (Baron-Cohen, Wheelwright, Hill, et al., 2001), which may have significant impacts on an individual's ability to interact with others. One commonly reported feature of face perception in autism is that individuals appear to have difficulty with making eye contact. While neurotypical individuals predominantly focus towards the eyes when viewing faces, autistic people spend less time looking at the eyes compared to other features and tend to spend more time looking at the mouth (Bayliss & Tipper, 2005; Jones &

Klin, 2013; Kliemann, Dziobek, Hatri, Baudewig, & Heekeren, 2012). It has been found that there are correlations between the amount of time an infant spends attending to the eyes and their broader developmental trajectory. For example, the amount of time an infant spends looking at eyes will predict the ability to join attention with others (Morales et al., 2000; Tomasello & Farrar, 1986), which in turn may impact on their development of the theory of mind in children with ASD (Charman et al., 2000).

There is still some debate about why there may be a notable deficit in how autistic people engage with faces, in particular, why there appears to be a reduced amount of time spent attending to the face and eyes of others; leading to two broad approaches that attempt to explain why there is a perceived deficit. In the first approach, it is suggested that the reason people attend faces and eyes is that they portray a large amount of social information and, as autistic individuals have deficits in social processing, faces and eyes are less attractive as stimuli in ASD (Moriuchi, Klin, & Jones, 2017). An alternative approach suggests that atypical processing of the face and eyes results in a specific pattern of avoidance, which is supported by anecdotal evidence from autistic people who report significant discomfort making direct eye contact (Trevisan et al., 2017). Furthermore, spikes in amygdala activity occurred in autistic individuals when they were presented with eyes level with a fixation cross, therefore focusing them directly on the eyes (Kliemann et al., 2012). Also, using the same paradigm, it was demonstrated using eye-tracking experiments that autistic participants would reflexively shift their gaze away from the eyes and towards the mouth (Dalton et al., 2005; Kliemann, Dziobek, Hatri, Steimke, & Heekeren, 2010). Together this would suggest a pattern of eye gaze avoidance. The proposal that faces, and in particular eyes, result in negative, avoidant behaviour in autistic people may also be consistent with the subcortical processing of eyes as proposed by (Senju & Johnson, 2009b). For example, if the efficient subcortical detection of eyes results in a negative processing response in ASD, as suggested by the spikes in amygdala activity demonstrated by Kliemann et al. (2012), it may develop into a learnt avoidance in some autistic individuals. Furthermore, based on the interactive specialisation account of face perception, avoidance of face and eye contact may have a knock-on effect in the development of active face processing networks.

Differences in Atypical Behaviour Across Photographic and Schematic Faces

Interestingly, there is evidence that autistic people do not show the same issues with the processing of less realistic faces. For example, the same style of holistic processing as neurotypical individuals has been observed in ASD when viewing cartoon faces, reverting to featural processes when looking at a photographic representation of a face (Rosset et al., 2008). There is also case study evidence using fMRI that found that an autistic child showed typical levels of activity in the fusiform gyrus and the amygdala when viewing animated faces of characters, with which they were familiar. However, again activity in these areas was reduced when viewing photographic faces (Grelotti et al., 2005), possibly explaining the relative success of interventions that use animated faces (Golan et al., 2010; Yan, Liu, Ye, & Liu, 2018), and why autistic people tend to choose animated videos over videos of faces (Dubey, Ropar, & Antonia, 2017).

Differential processing of faces with differing levels of realism has also been demonstrated in non-autism focused research, namely, the Uncanny Valley effect (Mori, MacDorman, & Kageki, 2012). First outlined by Mori, originally from the field of robotics, the uncanny valley effect outlines the uncomfortable sensation individuals have with robots that are perceived to be too realistic. Thus as an entity appears increasingly human, it becomes more endearing to an observer, but as it becomes increasingly realistic, there is a point where it appears too realistic, resulting in the uncanny sensation, which passes when the entity is a realistic representation of a human face. To date, there is minimal research into how autistic individuals interact with faces that fall within the uncanny valley. One study with children between 5 and 7 found that autistic children did not appear to show any uncanny valley effects, whereas children in a typically developing control group did (Feng et al., 2018). One explanation that has been used to explain the 'uncanny' feeling is the result of a disconnection between the perceptual information provided by a face (Kätsyri, Förger, Mäkäräinen, & Takala, 2015). That is, the uncanny valley effect in static faces is most extensive when there is a perceptual disconnect between the detail of some features and the stimuli as a whole (e.g., human eyes on an artificial face or atypical features; Kätsyri et al., 2015).

Although there is no evidence that the uncanny valley effect and autistic face perception are causally related, it may provide some insight into why autistic individuals do not show face processing deficits when viewing animated faces, which in turn may help to disentangle why face perception can be impaired in ASD. For example, in the perceptual mismatch hypothesis, conflicting visual information leads to a conflict between the bottom-up processes of interpreting the visual features and the top-down processes that indicate that it is not a human face. If it is the case that an alternative perceptual conflict occurs when viewing a human face in ASD, it may explain why gaze avoidance may occur. Thus it is crucial to understand how visual information is utilised during face processing in both NT and ASD populations to understand whether any perceptual conflict occurs.

Differences in Processing High and Low Spatial Frequency Face Information

Faces provide much information about another person, and as such, they also provide detailed visual information that needs to be processed. In spatial frequency theory, early visual information from an object is split into essential components based on their spatial frequency characteristics (De Valois & De Valois, 1988). These different components are then primarily separated down separate pathways, with LSF primarily conveyed by magnocellular cells and High Spatial Frequencies (HSF) primarily conveyed by parvocellular cells (Hubel & Livingstone, 1988; De Valois, Albrecht, & Thorell, 1982). This time course difference has been demonstrated in simple behaviour paradigms through longer response times to the perception of HSF than LSF when presented alone (Mihaylova, Stomonyakov, & Vassilev, 1999). Furthermore, EEG studies demonstrate early visual evoked potentials, such as the P1, have shorter latencies and greater amplitudes when viewing LSF compared to HSF stimuli (Proverbio, Zani, & Avella, 1996). As faces are such a complex stimulus, they present a wide range of information across the whole band of spatial frequency, with LSF information conveying holistic information about the face and HSF conveying fine details suitable for featural processing (Goffaux & Rossion, 2006). Recent evidence supports the coarse-to-fine hypothesis of processing, where processing is led by fast LSF information followed by progressive processing of mid to high spatial frequencies over time (Jeantet et al., 2019; Goffaux et al., 2011). Such evidence demonstrates that in processing tasks of LSF filtered faces degraded as the stimuli were presented for longer durations. However, performance increased for longer durations when HSF filtered images were used. The coarse style of processing is not seen in simple visual processing and has been proposed to only occur in more specific visual processes that involve categorisation of stimuli (Goffaux et al., 2011). Using faces as an example, this allows for effective detection and initial processing of the face, which subsequently informs slower HSF processing of finer detailed features required for accurate high-level face processing abilities such as identity or emotion recognition.

Although perception relies on a range of spatial frequency between the HSF and LSF bands, it has been found that in ASD, there is a preference for HSF processing (Kéïta, Guy, Berthiaume, Mottron, & Bertone, 2014; Deruelle, Rondan, Gepner, & Tardif, 2004). Preferential processing of HSF visual information has been observed in both non-face (Deruelle et al., 2004), and face stimuli (Kéïta et al., 2014; Yamasaki et al., 2017). A preference towards HSF over LSF processing may explain why autistic individuals maintain featural methods of face perception later in development compared to holistic processing seen in typically developing children (Deruelle et al., 2004). However, this does not explain the gaze avoidant behaviour, or why ASD children do not demonstrate such differences in processing cartoon faces.

These un-answered questions present an opportunity to build on our understanding of atypical face perception in ASD and the subsequent social deficits that may result. One possible explanation is that autistic individuals may experience difficulties when integrating the extreme high and low spatial frequencies that are associated with faces, or in particular the eyes. One possibility is that this complex integration of spatial frequencies may be akin to other over stimulatory issues that those with ASD experience. A commonly reported symptom of ASD is sensory overstimulation across multiple modalities (Haigh, 2018). Furthermore, face processing of human faces has been found to rapidly trigger activity across the wider SBN (Freiwald et al., 2016; Johnson, 2005). Overstimulation through both the integration of extreme spatial frequency information and the rapid activity of the SBN may account for a learnt avoidance of the faces of others. However, cartoon faces may not be as overstimulating for autistic individuals due to their reduced visual detail or their reduced social nature of them as a stimulus. If it is the case that avoidance of face stimuli is a learnt response and a result of overstimulation, it may be possible to reduce overstimulation through exposure, as has been attempted with other modalities (Gee, Thompson, & St John, 2014). Reducing avoidance behaviours may increase interactions with faces, allowing for the full development of the face processing networks required for efficient face processing.

A Potential Paradigm for Investigating Facial Realism

Throughout this chapter, evidence has been outlined that investigates how humans process faces, from early face recognition abilities to more advanced abilities such as how we use the direction of eye gaze to gain social information from others. These findings outline the vital role that eyes within the face play for both detection, as well as modulation of other face processing behaviours such as joint attention and emotion recognition. Furthermore, theories have been proposed about how attraction to eyes early in the developmental timeline will increase engagement with faces and the impact of the resulting development of mature face perception. In the context of an interactive specialisation account disruption of the implicit attraction to faces may impede social development, as may be the case in ASD. Early research has been introduced that suggests that autistic participants are less likely to present atypical face processing behaviours compared to NT individuals when viewing animated faces. Therefore, to build on this research, it is proposed that new paradigms should be explored that could be used to provide a comprehensive understanding of face processing abilities across a variety of visual realism. Face processing tasks could be developed that not only investigate the visual processing but expand into how eye gaze processing may alter participants' attention and subsequent changes in cognition. Utilising paradigms that can investigate the visual preferences as well as functional consequences as a result of social eye gaze, it will be possible to break down to what extent less realistic faces are processed in comparison to realistic representations of faces.

It has been found that faces of all levels of realism will elicit face processing behaviours and will even elicit second-order behavioural cues such as a gaze cueing bias for visuospatial attention (Friesen & Kingstone, 1998). Therefore, to understand more intricate differences between faces of different levels of realism, it is essential to move to paradigms that expand on the social consequences of eye gaze beyond visuospatial effects - for example, the liking effect. In a seminal study by Bayliss et al. (2006), it was found that individuals followed the eye gaze of a face to a target object. If the object was cued by eye gaze, they tended to rate the object higher than objects that had been incongruently cued. Although the difference between ratings was small, it was consistent across participants. However, when the same experiment was conducted with arrow cues instead of faces the effect was not found, demonstrating that it was not congruently cued attention alone that would result in the liking effect (See Figure 1.1 for a comparison between Face and Arrow results). Bayliss and colleagues suggested that the liking effect was an efficient form of objective evaluation. That is, if individuals attend to an object, they likely have a preference for that item and it would be an efficient form of processing to develop a preference for the gazed-towards object. To establish this, Bayliss, Frischen, Fenske, and Tipper (2007) manipulated the facial expressions of the faces used as cues and found that a positive facial expression boosted the effect, consistent with their proposed suggestion. However, there has been no research that investigates this effect in faces that are not photographic representations of human faces. As this effect has been demonstrated only to work in faces, it could be possible to separate the effects of the attentional cue into visuospatial measures (such as reaction time and accuracy) and functional consequences in the form of gaze cueing bias for object preferences.

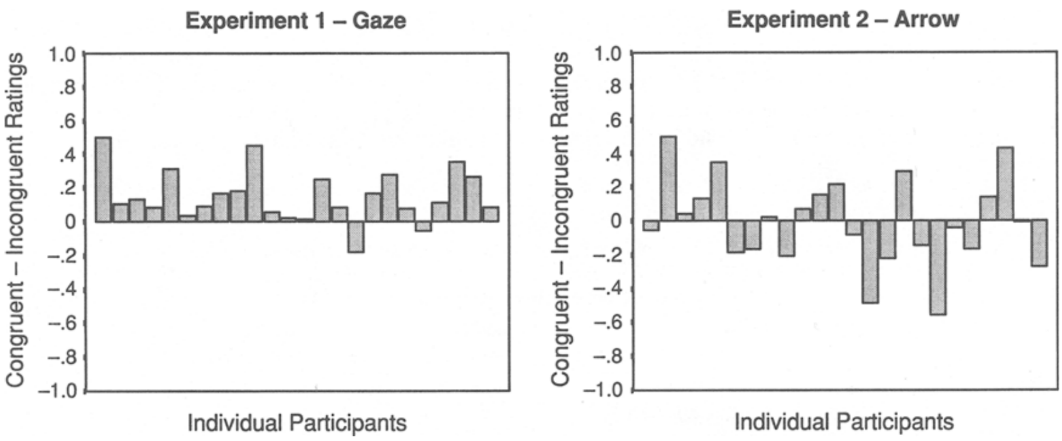

Figure 1.1: In this original figure from Bayliss et al. (2006) the differences between participants' rating of Congruent and Incongruent objects are displayed. These results demonstrate a consistent positive difference among participants in experiment 1 where a gaze cue is used. In experiment 2 when an arrow is used as a cue there is not the same level of consistency.

There are several possible outcomes for this effect using facial cues that will need to be considered. Firstly, is it possible that any face shape that cues attention may elicit the effect? As we have seen, it only takes fundamental visual representations of faces to trigger some face processing abilities. Therefore, it may be possible that a simple schematic face may be enough to elicit the effect. Alternatively, if the effect requires more extensive activation across the SBN, it may be possible that only detailed faces, presenting across a broad spatial frequency spectrum, will elicit adequate processing for the effect to occur; and as such the effect may not be seen in low realism schematic faces. If it is the case that the effect requires detailed facial features, it could open up this paradigm as a method to distinguish how we process facial realism. However, if the first possibility is correct and the effect occurs across both high and low detail faces, it may limit the value of using the paradigm to measure the effect in future.

Another consideration of the mechanisms involved in the liking effect is the requirement of memory for gaze cue associations that are necessary for the effect to occur. It has been suggested that preference as a result of gaze cueing bias is reliant on the recalled association between the eye gaze and the object (Bry, Treinen, Corneille, & Yzerbyt, 2011). It has been found that attention from eye gaze facilitates memory (Turk-Browne, Golomb, & Chun, 2013). One possibility that should also be considered is the fact that eye gaze may facilitate attention in a way that other spatial cues do not. If it is the impact of memory on attention that results in changes in affective judgments, it would be necessary to draw out that possibility. That is, to what extent is the liking effect related to attention driven memory? Moreover, if memory does play a role in the liking effect, to what extent does facial realism impact memory?

Specific Research Questions and Thesis Structure

The aim of this thesis is to investigate behavioural methods designed to understand face processing across a spectrum of facial realism by utilising the liking effect paradigm to separate the functional consequences of eye gaze from visuospatial consequences. Data will also be collected for this effect in the context of multiple memory modalities to understand the potential for memory to impact preferential judgments of items. The goal will be to establish whether the liking effect is a suitable and effective paradigm for distinguishing the effects of facial realism on face processing. Furthermore, data will be collected across the autistic spectrum, both through measures of the broader autism phenotype and directly from autistic participants, to understand how ASD may interact with this effect, as well as how different levels of realism impact face processing across the autistic spectrum. This broad aim has led to the development of three research questions:

1. To what extent can the liking effect be used to measure differences in how we process faces of different levels of realism?

2. Will autistic individuals utilise eye gaze in the same manner as neurotypical participants regarding the functional consequences of eye gaze, as measured by the liking effect?

3. Can this paradigm be used to distinguish differences in face processing in a spectrum of facial realism between ASD and NT participants?

A series of nine experiments were conducted to investigate the three research questions. Chapter 2 presents three experiments that attempt to study the functional consequences of eye gaze on affective judgments and memory. In Chapter 3, four experiments that attempt to replicate the original Bayliss et al. (2006) paradigm with extensions using schematic faces are presented. Chapter 4 investigates the relationship between facial realism and autistic traits directly in the context of the liking effect. In Chapter 5, eye-tracking data using the liking effect paradigm in NT and ASD samples are presented. Finally, in Chapter 6, the findings are discussed in the context of the wider literature.

Chapter 2:

Functional Consequences of Eye Gaze on Affective Judgments in Long Term Memory

Previous research into the liking effect has tended to investigate only the impact of gaze cues on affective judgments in short term memory paradigms (Bayliss et al., 2006, 2007; Strick, Holland, & van Knippenberg, 2008). Participants were presented an object alongside either a congruent or incongruent gaze cue. They were then asked to rate how much they liked the object. The effect of gaze cueing on affective judgements may have built up over the initial cueing trials, but in the rating stages, the participant is primed before the gaze cue. Therefore, this raises the question of how the memory of the object may impact the effect. One study by Bry et al. (2011) found that in a task with a delay between observing the gaze cue and a subsequent rating task, there was a gaze cueing benefit only if individuals recalled observing the item in the gaze cueing task, suggesting that memory may play a significant role in the gaze cueing bias for affective judgments. One possibility is that cueing of visuospatial memory towards an object may facilitate encoding that, in turn, results in the gaze cueing bias found in the liking effect. It has been demonstrated that attention during encoding will improve subsequent memory (e.g., Rock & Gutman, 1981). Therefore, to understand the mechanisms that drive the liking effect, it would be wise also to investigate the functional consequences of eye gaze on memory.

Turk-Browne et al. (2013) extended previous findings that attentional gaze cues will facilitate encoding in long term memory. Their primary goal was to investigate the neural correlates of facilitation and reorientation and to establish whether these could be used to predict memory recall. Participants were presented a face that either maintained a direct gaze; oriented their gaze away from a target location, or towards the target location. Participants were then rapidly presented with two images (a target and a distractor) and were asked to respond once they had detected the targets as fast as possible. It was found that participants responded faster when the eye gaze was congruent, facilitating attention towards the cue. It was also found that congruent gaze resulted in higher recall than both incongruent or direct gaze cues in subsequent recall tasks. The authors proposed that congruent cues resulted in attention to the target location, which facilitated memory, whereas incongruent and direct gaze cues result in reorientation that limits the participants' ability to encode effectively. Furthermore, activity in areas that have previously been found to correlate with facilitation was able to predict accurate recall of objects in later trials (Turk-Browne et al., 2013).

Turk-Browne et al. (2013) demonstrated that gaze facilitates memory, but they did not test whether the benefits of gaze cues extend beyond the realm of long-term memory recall to affective judgements associated with the gazed-at objects. Replication of the faciliatory effects of gaze cues on memory (Turk-Browne et al., 2013), in the context of a liking effect paradigm similar to that of Bayliss et al. (2006), could outline the mechanisms required for the liking effect to occur. Measures of memory would also provide a secondary effect to aid in the understanding of the consequences of eye gaze induced by faces of differing levels of realism.

In the current series of experiments, a paradigm was developed to investigate whether attention cued by human gaze has consequences for memory recall as well as the affective judgment of objects in long term memory. Following Turk-Browne et al. (2013), we prompted participants to respond to targets paired with a direct; congruent or incongruent gaze cue. In a second task, participants were then asked to state if they had seen the target before and to rate the certainty of memory recall. Two hypotheses were tested in the experiment:

> *Hypothesis 1*: Cueing effects on memory. Targets that are validly cued by the gaze of a face will be better remembered, as a result of more efficient top-down processing limiting the need for reorientation. In contrast, eye gaze that cues incorrect locations will hinder memory by requiring reorientation and therefore weaker target encoding.

> *Hypothesis 2*: Cueing effects on affective judgements. Targets that are cued by the gaze of a face will result in a more favourable appraisal because the facilitation effect engendered by valid cues will allow for efficient processing that, as well as improving memory, will impart information about social appraisal from the observer.

Methods

Three experiments were conducted as part of this chapter. An overview of the method that was utilised across all three experiments is discussed here. Specific participant details and procedures are discussed for each experiment independently.

Participants

Participants were either University of Oxford undergraduates, taking part for course credit, or were recruited using the Oxford Participant recruitment scheme and were compensated £7 for their time.

Ethical Considerations

All participants gave informed consent before the experiment commenced; they were made aware that they could withdraw at any time and that all data collected would remain anonymous. The experiment was given ethical approval by the University of Oxford Central University Research Ethics Committee (CUREC; R45285/RE001).

Apparatus

Tasks were presented using a computer running Matlab (64-bit version R2015b; The MathWorks, Natick, MA), images were displayed using Psychtoolbox (Kleiner, Brainard, & Pelli, 2007; Pelli, 1997; Brainard, 1997). Responses were collected via a keyboard press. Stimuli were presented on a 24-inch monitor with a resolution of 1920 x 1080 pixels.

Measures

To measure autistic symptomology, participants completed an Autistic Quotient questionnaire (Baron-Cohen, Wheelwright, Hill, et al., 2001). The AQ has been used to screen those who demonstrate Autistic traits and has demonstrated an 80% reliability for screening individuals with ASD at a critical cut-off score above 32 (Baron-Cohen, Wheelwright, Hill, et al., 2001; Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005). The Participants were also asked to complete a Liebowitz Social Anxiety Scale questionnaire to measure levels of social anxiety (SAS; Liebowitz, 1987). The SAS has demonstrated a highly significant correlation with several other self-report questionnaires for anxiety and phobia (Social

Interaction Anxiety Scale and the Social Phobia Scale; Heimberg et al., 1999). These questionnaires were used to compare individual differences between participants and their interaction with within-subject effects. These two measures of social anxiety and autistic traits are also used in future chapters of this thesis.

Stimuli

All images were presented on a mean grey background at full resolution (1920 x 1080 pixels).

Face Cues

Four face stimuli were used from the Oslo face database (Chelnokova et al., 2014). Faces were presented centrally with three variations of each face: left gaze, right gaze, and direct gaze. Each face image measured 276 x 312 pixels (approximately 76 x 86 mm) and was presented an equal number of times across trials, with all faces gazing congruently, incongruently and direct an equal number of times.

Target images

For Experiment 1, two stimuli sets were utilised as targets for the experiment and were counterbalanced across participants. Following from Turk-Browne et al. (2013), natural scenes were used as this had been found to elicit memory effects. It was also considered that participants were less likely to have pre-existing preferences to natural scenes that might impact participants' ratings of the targets in the experiment. However, it was possible that if the scenes were too abstract, preferential ratings might be less indicative. Therefore, a second stimulus set of objects was sourced to replicate that which been used by Bayliss et al. (2006). These target stimuli conditions were compared to decide if there were any differences and results are discussed in the Results section of Experiment 1.

For the scene images, two different stimuli sets were used. Ninety-six images were taken from the University of Texas Natural Image Database (Geisler & Perry, 2013); the remaining images were taken from UPenn Natural Image Database (Tkačik et al., 2011). These images were similar styles of natural images and scenes taken in either Austin, Texas or Botswana. Images were rejected if they contained pictures of animals, or if they were deemed too similar to the other images used. All images measured 400 x 276 pixels (approximately 110 x 76 mm).

Object images were collected from the Bank of Standardised Stimuli (BOSS; Brodeur, Guérard, & Bouras, 2014). In total, 180 images were used. Images were selected semi-randomly, removing images that were too similar to other images in the set. Items were also rejected if they contained any form of a face or face-like element (to avoid any confounding effects), and also if they contained any food or drink items to try to avoid any pre-preference for items during the rating task. As in Turk-Browne et al. (2013), blurred versions of the target images were created in Matlab to act as a distractor 'false' image. All images measured 300 x 300 pixels (approximately 83 x 83 mm).

All of the target images were blurred using a 2-D Gaussian filter in Matlab to create distractor images. The images were described to participants as either 'True' images for the target images or 'False' images for the blurred distractor images.

Counterbalancing

Each stimulus set was split in half and then further split into three to give a total of 6 groups. Groups one, two, and three were displayed in two halves: the first half in the Initial Learning phase and the second half as novel images in the testing phase. The gaze direction for each image was altered between the three groups so that all images were congruent, incongruent and direct for each image. The order was reversed for groups four, five and six so that each image was seen as novel or old an equal number of times for each participant. Images were presented in a randomised order and screen side. For Experiments 2 and 3, the gaze direction was counterbalanced for all images as well as the target letter in Experiment 3.

Procedure

Learning Phase Overview

In the initial phase, participants were asked to identify which image was the 'true' image (i.e., the non-blurry image) by using the corresponding directional arrow on a keyboard. After a fixation cross, participants were presented a face with a direct fixation for 500 ms before the gaze shifted either to the left or the right. After 500 ms an object appeared to either the left or right of the face stimuli. Simultaneously a distractor target appeared on the opposite side of the face. The participant was then asked to categorise the target item as to whether it would be typically found in the garage or the kitchen using the up and down keys. If the trial was Congruent, the gaze direction and the target object location was the same; if the gaze was towards the distractor, the trial was Incongruent. A 'neutral' Direct condition was also used, where the face would not make any gaze shifts. In the experiment, there were four face identities (two female, two male), and 12 target images (six garden and six kitchen items in the object group). Including presentation location and item type, the factors were balanced resulting in 192 trials. Trials were divided into three blocks of 64 trials, where the participant was asked to take a break between each block. Participants were put into two groups to counterbalance which objects were Congruent/Incongruent/Direct between participants. For each participant in the object group an equal number of garden and kitchen items were Congruent/ Incongruent/ Direct to avoid a general preference towards one category or the other. A timeline of the Learning procedure can be found in figure 2.1.

Test Phase Overview

In the test phase across all three experiments participants were presented with either a probe image and were first asked to respond if they had seen the item in the Learning phase using the directional keys (left for no, right for yes) to measure memory recall of the object. Subsequently, they were asked to rate how confident they were in their memory judgment on a scale of 1-6. Finally, to take a measure of participants affective judgments of the objects, participants were asked to rate 'how much they liked the image' on a scale of 1-6.

Data analysis

Cueing effects in the Learning phase

The two dependent variables for the Learning phase were Accuracy and Reaction Time (RT). For Accuracy, a proportion of correct and incorrect scores were calculated to provide Accuracy scores for analysis. Median RTs were calculated using trials where a correct response was given, with erroneous trials being discarded. Medians indexed reaction times better for these purposes, rather than means, as medians are less severely affected by extreme values, while they are still representative of the overall difficulty of the categorisation judgements in question. While using mean RTs is a common approach in studies that include neurotypical adults, it was less well justified for the thesis, as the interest was ultimately in investigating neurodiverse individuals, for whom excluding slow reaction times may mean excluding representative behaviour. Gaze direction of the face stimuli (Direct, Congruent or Incongruent) was the primary independent variable for all experiments. Analysis of both the Accuracy and Reaction time data using the Shapiro-Wilk test found that the data were not normally distributed.

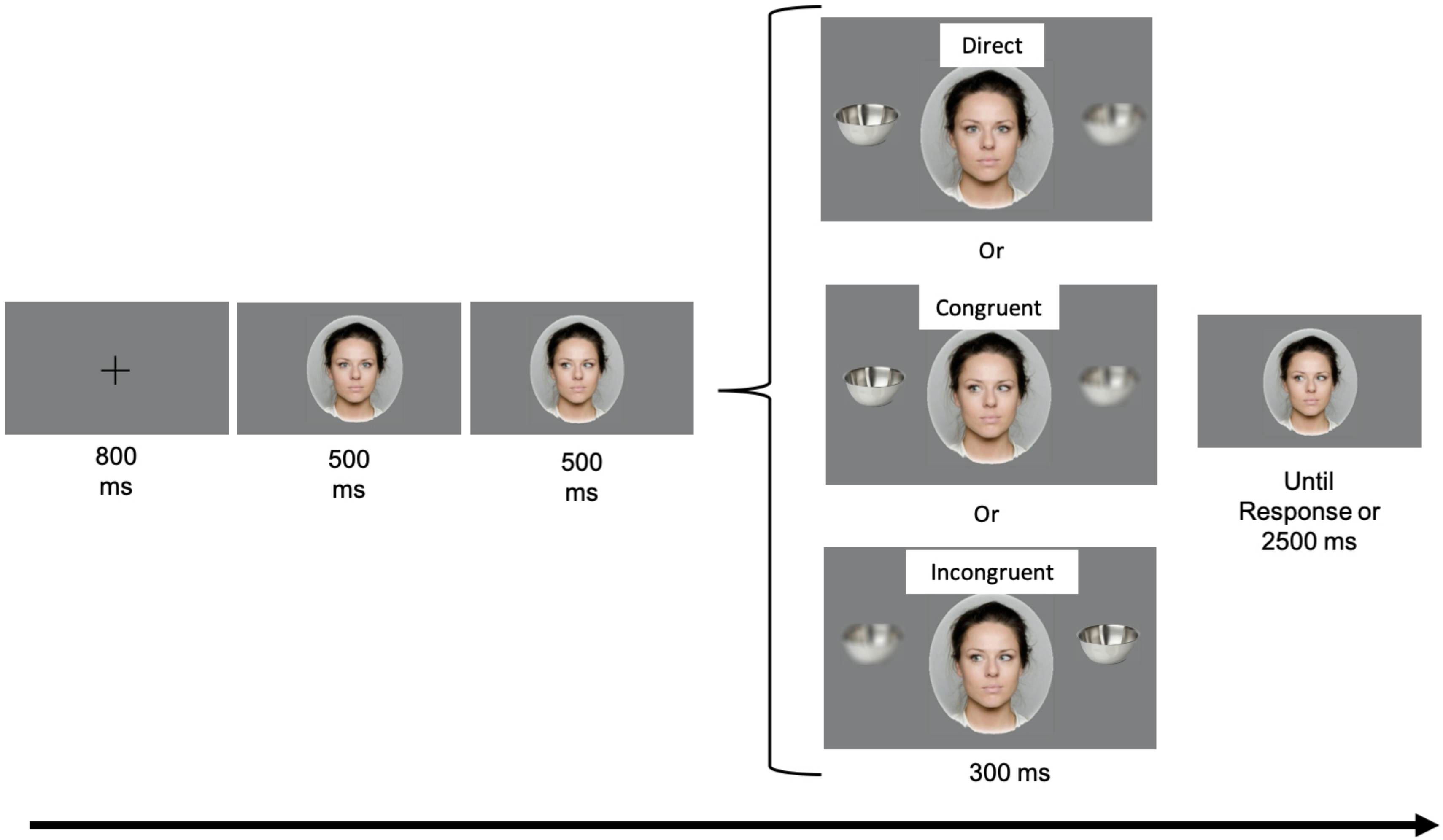

Figure 2.1: Timeline of the Learning phase for Experiment 1. Participants were presented with a fixation cross followed by the direct eye gaze for 500 ms. The image then shifts to a face with an averted gaze for 500 ms. Finally, both the target and distractor images appear simultaneously for 300 ms. The participants then respond to which side they believed the target image appeared, triggering the beginning of the next trial.

Data were, therefore analysed using non-parametric analysis methods. A Friedman test was used to analyse if there were any differences between the main factors of Congruency (Direct, Congruent and Incongruent) and Duration (Short and Long). If a significant difference was observed in either of the main factors then post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$ for both Accuracy and RT.

Memory recall

Memory Recall Accuracy was calculated using the proportion of correct and incorrect responses. Memory Recall responses were then divided into four groups based on the gaze condition in the Learning phase (Direct, Congruent, Incongruent) and Novel images. Again it was found that the data were not normally distributed, confirmed by a Shapiro-Wilk test, and as such non-parametric tests were applied. A one-sample Willcoxon signed-rank test was used to compare to hypothesised median of 0.5 (chance).

For experiments where memory for old objects was above chance, d' scores were calculated using proportion of hits and false alarms on the following trial types: Old objects with Direct gaze; Old objects with Congruent gaze; Old objects with Incongruent gaze. These scores were calculated to investigate participant sensitivity and were assessed using a Friedman test with gaze direction entered as the within-subject factor. If a significant difference was observed, then post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied.

Memory Confidence Judgments and Affective Ratings

Memory confidence was analysed using a Shapiro-Wilk test and was found to be non-normally distributed and was therefore examined using non-parametric statistics. Gaze direction (Direct, Congruent, Incongruent and Novel) was the primary independent variable and was compared using a Friedman test to establish and significant main effects. Post hoc analysis was conducted using a Willcoxon Signed-rank test with a Bonferroni correction. Due to a technical error, Memory confidence was only collected for a small subgroup of participants in Experiment 1 and will therefore not be presented for Experiment 1.

Affective rating data were found to meet the assumption of normality using the Shapiro-Wilk test and therefore were tested using repeated measure ANOVAs with Gaze Condition (Novel objects; Old objects with Direct gaze; Old objects with Congruent gaze and Old objects with Incongruent gaze) entered as a 4-level within-subject factor. In the event of significant main effects or interactions, a pairwise comparison was conducted with an applied Bonferroni correction.

Outlier Exclusion

It was decided that outlying data would not be removed from the experiment. As a primary goal of the experiment was to understand the impact of gaze cues on affective judgements - an inherently subjective judgment made by the participant - removal of extreme data as a result of one dependent variable may have unintended consequences on other dependent variables. Therefore, rather than trimming specific extreme data, the median was used to compare RT to compensate for potential extreme values without having to compromise other dependent variables.

Experiment 1

The aim of Experiment 1 was to investigate the impact that memory may have on the liking effect to clarify two questions. Firstly, would the impact of social gaze on affective judgments last in long term memory and secondly, if so, how much would explicit recall impact differences in affective judgment?

Methods

Participants

Thirty-two adult participants (mean age = 25.1, range = 19-50), of which 18 were female, took part in the experiment.

Procedure

In the initial phase, participants were asked to identify which image was the 'true' image (i.e., the non-blurry image) by using the corresponding directional arrow on a keyboard.

Results

Learning phase

A summary table of the descriptive statistics of both the Learning phase and the Rating phase can be found in Table (2.1.

Accuracy

To compare how accurate participants were at responding to the target stimuli a proportion of correct vs incorrect responses was calculated for the Direct (M = .978, SD= 0.071); Congruent (M = .981, SD = .033) and Incongruent (M = .975, SD= .047) Gaze conditions. These proportions were compared using a Friedmans test and it was found that there was no significant difference between the gaze conditions ($X^2(2) = 2.475$, $p = .290$).

Reaction Time

Reaction Times were calculated for trials where a correct response was recorded and were compared with a Friedmans test. It was found that there were significant differences between the Direct (M = 655.050 ms, SD= 515.352); Congruent (M = 591.687 ms, SD = 442.372) and Incongruent (M = 611.952, SD= 555.810; $X^2(2)$ = 21.063, $p < .001$). Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p <$ 0.017. It was found that there was a significant difference between the Congruent (Mdn(IQR) = 460.465 ms (405.182 ms to 629.38 ms)) and Direct (Mdn(IQR) = 483.431 ms (434.659 ms to 763.025 ms)) gaze conditions (Z = -3.628, $p < .001$), and the Incongruent (Mdn(IQR) = 444.255 ms (388.345 ms to 659.803 ms) and Direct condition (Z = -3.123, p = .002). However, there was no significant difference between Congruent and Incongruent gaze conditions (Z = -0.019, $p < .985$).

Table 2.1: *Mean and Standard Deviation for both Proportion Correct and Median RTs (ms) from the Learning phase of Experiment 1 N = 32).*

	Direct		Congruent		Incongruent	
	M	*SD*	*M*	*SD*	*M*	*SD*
Proportion Correct	0.978	0.072	0.981	0.034	0.976	0.047
Median RT (ms)	655.050	515.353	591.687	442.372	611.953	555.810

Test phase

A summary table of the descriptive statistics of the Test phase can be found in Table 2.2.

Memory Recall

To investigate whether Memory Recall was above chance a proportion of correct answers for Gaze conditions (Direct, M = .537, SD= .118; Congruent, M = .503, SD = .187; and Incongruent M = .525, SD= .174) and the Novel condition (M = .704, SD= .135) were compared against a chance variable fixed at 0.5. It was found that Memory Recall was not significantly above chance for any of the Gaze conditions - Direct (Z = -1.101, p = .271; Mdn(IQR) = .525 (.403 to .700)), Congruent (Z = -0.10, p = .992; Mdn(IQR) = .491 (.336 to .625)) or Incongruent items (Z = -.845, p = .398; Mdn(IQR) = .542 (.400 to .630)). However, it significantly above chance for Novel Items (Z = -4.862, $p < .001$; Mdn(IQR)(IQR = .692 (.602 to .810)).

Object Rating

Average participant Affective Rating scores were compared using a repeated measures ANOVA to gauge how preference varied across the different Gaze conditions. There was a significant main effect of condition on the Affective Rating scores (F (3, 93) = 4.385, p = .006, η_p^2 = .124), with the lowest scores for Novel items (M = 2.907, SD = 0.775) compared to Direct (M = 3.084, SD = 0.8), Congruent (M = 3.038, SD = 0.771) and Incongruent items (M = 3.143, SD = 0.824). Pairwise comparison, using a Bonferroni correction, found that there was a significant difference between Incongruent and Novel scenes (p = .035), There were no other significant differences between any of the other conditions.

Table 2.2: *Mean and Standard Deviation for Memory Recall and Object Rating scores from the Rating phase of Experiment 1* (N = 32).

	Direct		Congruent		Incongruent		Novel	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Memory Recall	0.538	0.188	0.503	0.187	0.526	0.174	0.705	0.135
Object Rating	3.136	0.870	3.211	0.768	3.122	0.771	3.009	0.777

Discussion

This chapter aimed to investigate the impact that gaze cueing has on memory recall and subsequent affective judgments. For the Affective Ratings, there was a main effect of cue type. However, this was driven by a lower score for Novel items rather than by differences within the gaze cues. These differences may support the proposal that it is memory that drives differences in affective judgments and that gaze cues facilitate memory. However, it was also found that, for all gaze conditions, differences in memory performance were not above chance. Therefore, if there is an effect of memory, it must be a trace amount that is not detected in an explicit recall task.

The more concerning finding was that there was no explicit RT benefit for congruently cued items. Although it appeared that both Congruent and Incongruent gaze conditions were significantly faster than Direct conditions, a RT bias for congruently cued items is consistently found across gaze cueing paradigms (for review see Frischen, 2007), raising the question: Why was no gaze cueing bias was found? One suggestion is that the timings of the attentional cues presented to participants may have interfered with the gaze cueing bias, a possibility that is explored in Experiment 2. Finally, there appeared to be no significant differences between natural scenes or object conditions. As such, it was decided that object stimuli would be used in future experiments as this would optimise the ability to replicate any gaze bias in affective judgments, as seen in Bayliss et al. (2006).

Experiment 2

One aspect of cueing paradigms that have been found to influence RTs is Inhibition of Return (IOR). IOR refers to an orienting mechanism by which, after the initial benefit gained from attending to a cued location, there is often a negative bias as individuals reorient away from the cued location (Posner & Cohen, 1984). In the 1984 study, Posner and Cohen demonstrated shorter RTs if a target appeared in a cued location. However, if there was a period between the initial cue and the appearance of the target stimuli, termed the Stimulus Onset Asynchrony (SOA), of more than 300 ms, there was no cueing benefit. Furthermore, there was a cost to speed of RT as the SOA was increased. It was proposed that this inhibitory effect aids in visual search as it prevents repeated examination of the same location, facilitating a thorough search of the whole environment.

Friesen and Kingstone (1998) investigated whether IOR would occur when faces were used to cue the attention of participants. They found that a cueing effect appeared very early (105 ms) in some response conditions and by 300 ms in all conditions. At longer SOA of 1005 ms, the cueing bias was no longer present. However, there was no RT cost, and therefore no IOR effect observed; confirmed in a comparable study by Driver et al. (1999) using more realistic faces. Frischen and Tipper (2004) identified several possible experimental differences between the peripheral style cues in experiments where IOR had been identified and the face cueing experiments where IOR had not. Firstly, they suggested that participants' attention in the original Posner IOR experiments was drawn back to the centre after the onset of the peripheral cue, a dimension not present in many gaze cueing experiments (Posner & Cohen, 1984; Frischen & Tipper, 2004). Furthermore, it was proposed that in non-face cueing experiments where IOR occurs, the cue is only presented very briefly (Collie, Maruff, Yucel, Danckert, & Currie, 2000; Frischen & Tipper, 2004). Frischen and Tipper (2004) ran a series of 10 experiments systematically investigating the influence that these factors may have on the occurrence of IOR in faces. They found that when the face cue was removed from view during the SOA, there was a small IOR effect, but only at the extended SOA of 2400 ms. In a follow-up experiment, they found that any salient change in the cue during the SOA, whether the face is visible or not, will elicit an adequate shift in attention to result in an IOR effect at a longer SOA of 2400 ms.

Although the timings of this experiment make it unlikely that there is an explicitly inhibitory effect, decay of facilitation may result in the non-significant cueing effects. It has been found that facilitation is optimal with an SOA between 300 and 600 ms (Friesen & Kingstone, 2003, 1998; Friesen, Moore, & Kingstone, 2005). Therefore, the facilitatory effect may be impeded at the longer combined stimulus presentation time of 300 ms and SOA of 500 ms; impacting the accuracy, RT, and memory advantages. To investigate how decay in the facilitatory effect may impact these measures, a Duration condition was added in Experiment 2. By altering the timing of the paradigm to create a Short (300 ms) and a Long (600 ms) condition, it was possible to investigate the null findings of Experiment 1, resulting in an additional hypothesis: Any effects as a result of a gaze cueing bias would only occur in shorter gaze cue presentations due to a decay in facilitation that may occur at the longer SOA.

Methods

Participants

Fourteen adult participants (mean age = 23.5, range = 18-30), of which eleven were female, took part in the experiment.

Procedure

The procedure for experiment 2 followed the same general procedure as Experiment 1, but the duration of the gaze cue was altered to be either 300 ms or 600 ms.

Results

For a summary table of the results of the Learning and Rating phase, please refer to Appendix B.

Learning Phase

A summary table of the descriptive statistics of both the Learning phase and the Rating phase can be found in Table 2.3.

Accuracy

Proportion of the Correct responses were calculated for the Direct (M = .897, SD= .241); Congruent (M = .904, SD = .263) and Incongruent (M = .849, SD= .333) Gaze conditions. The conditions were compared using a Friedman test and it was found that there was no significant difference between the gaze conditions ($X^2(2)$ = .558, p = .756). Proportion correct scores were also compared for the Duration conditions, Short (M = .881, SD= .266) and Long (M = .885, SD = .261). It was found that there were no significant differences between accuracy responses in the Long and the Short Duration conditions ($X^2(1)$ = 1.448, p = .836). Finally, interaction variables were compared using a Friedman test and it was found that there was no significant difference between any of the variables ($X^2(5)$ = 1.951, p = .856).

Reaction Time

Comparisons of RT between the three Gaze conditions found that there were significant differences between the three Congruency variables - Direct (M = 975.982, SD= 238.472), Congruent (M = 912.119 ms, SD = 190.905 ms) and Incongruent (M = 921.002 ms, SD= 202.834 ms; $X^2(2)$ = 8.714, p = .013). Post hoc analysis with a Wilcoxon signed-rank test was conducted with a Bonferroni correction applied, resulting in a required p = 0.017. Median (IQR) for the Direct, Congruent and Incongruent gaze conditions were 857.350 ms (851.988 ms to 1058.025 ms), 845.205 ms (775.545 ms to 1087.505ms) and 848.152 ms (805.158 ms to 1032.409 ms) respectively. It was found that there were significant differences between Direct and Incongruent conditions (Z = -2.417, p = .016). However, there was no significant differences between the Direct and Congruent (Z = -2.166, p = .30) or Congruent and Incongruent conditions Z = -0;408, p = .683.

Comparisons between the Duration conditions found that there was a significant difference between the two conditions ($X^2(1)$ = 14.000, p < .001), which appears to be driven by faster reaction times for the Short, M = 807.899 ms, SD = 368.361) compared to the Long Conditions M = 1008.544 ms, SD = 209.720).

Furthermore, there were significant differences between the interaction variables ($X^2(5)$ = 54.041, p < .001). A post hoc Wilcoxon signed-rank test was conducted. However, none of the analysed pairs managed to meet the required significance threshold of p = .008 once a Bonferroni correction had been applied.

Table 2.3: *Mean and Standard Deviation for both Proportion Correct and Median RTs (ms) from the Learning phase of Experiment 2 (N = 14). Scores are shown both as aggregate and by duration.*

		Direct		Congruent		Incongruent		Total	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Proportion Correct	Long	0.893	0.255	0.906	0.266	0.846	0.325	0.882	0.267
	Short	0.902	0.231	0.903	0.262	0.853	0.346	0.886	0.262
	Total	0.978	0.072	0.981	0.034	0.976	0.047		
Meadian RT (ms)	Long	1038.039	252.184	987.438	174.011	1000.158	211.124	1008.545	209.720
	Short	843.132	257.997	794.435	172.195	786.132	137.131	807.900	186.362
	Total	1047.390	386.033	1128.388	253.606	1349.046	245.850		

Test Phase

A summary of the descriptive statistics for the testing phase can be found in Table 2.4.

Memory Recall

Memory recall Accuracy was compared against chance (0.5) using a Wilcoxon signed-rank test and it was found that for Direct (Z = -2.383, p = .017; Mdn(IQR) = .400 (.225 to .475)) and Congruent (Z = -2.045, p = .041; Mdn(IQR) = .400 (.225 to .475)) accuracy was significantly below chance. Incongruently (Z = -1.635, p = .102; Mdn(IQR) = .433 (.225 to .571)) cued stimuli were not accurately recalled above or below chance. However, participants were able to distinguish Novel stimuli above chance (Z = -3.297, p = .001; Mdn(IQR) = .843 (.716 to .935)).

Memory Confidence

Analysis of the Confidence judgments found that there was a significant main effect of Gaze condition for memory confidence ($X^2(3) = 9.512$, p = .023; Direct, M = 2.709, SD = 0.901; Congruent, M = 2.735, SD = 0.943; Incongruent, M = 2.539, SD = 0.974; Novel, M = 2.278, SD = 0.762). Post hoc analysis was conducted using a Wilcoxon Signed-Rank test with an adjusted p value of .0125. It was found that there was a significant difference between the Confidence judgments of Congruently cued items and Novel items (Z = -2.691, p = .007; Congruent mdn (IQR) = 3.76 (2.488 to 3.279); Novel, mdn (IQR) = 2.301 (2.195 to 2.681). There was no significant difference between the Duration variables ($X^2(3) = 0.692$, p = .405; Short, M = 2.643, SD = 0.926; Long, M = 2.675, SD = 0.925). There was a significant difference between the interaction variables ($X^2(5) = 14.891$, p = .021), but post hoc analysis using a Wilcoxon Signed-rank test found that none of the pairings met the adjusted p = 0.125.

Affective Judgments

Average Affective rating scores were compared for the three Gaze Conditions (Direct, M = 3.179, SD = 1.063; Congruent, M = 3.011, SD = 1.253; Incongruent, M = 3.001, SD = 1.114). and the Novel condition (M = 2.581, SD = 1.926). It was found that there was a significant main effect of Gaze congruency on the liking effect. Post hoc analysis with a Bonferroni correction applied found no significant differences between any of the Old conditions. However, there were significant differences between the Novel and the Direct (p = .007) and Incongruent (p = .030) conditions.

Analysis of the differences between the Old variables when the Duration condition was applied found that there was no significant main effect of Congruency ($F(1.333, 17.327) = 1.304$, $p = .283$, $\eta_p^2 = .091$) or of Duration ($F(1, 13) = .435$, $p = .521$, $\eta_p^2 = .032$). There was also no significant interaction between the two factors ($F(1.431, 18.607) = .448$, $p = .581$, $\eta_p^2 = .033$).

Table 2.4: *Mean and Standard Deviation for Memory Confidence scores from the Rating phase of Experiment 2 (N = 14). Results are displayed by duration as well as an aggregate.*

		Direct		Congruent		Incongruent		Novel	
		M	SD	M	SD	M	SD	M	SD
Memory Recall		0.365	0.167	0.390	0.187	0.399	0.184	0.812	0.137
Confidence	Long	2.928	0.971	2.713	1.042	2.385	1.093	N/A	N/A
	Short	2.549	1.031	2.784	0.964	2.597	1.001	N/A	N/A
	Total	2.710	0.901	2.540	0.974	2.540	0.974	2.278	0.726
Object Rating	Long	3.244	1.138	3.005	1.333	2.873	1.062	N/A	N/A
	Short	3.165	1.278	3.069	1.281	3.056	1.214	N/A	N/A
	Total	3.179	1.063	3.011	1.253	3.001	1.115	2.581	0.927

Discussion

As in Experiment 1, it was found that while responses to Direct cues were significantly slower, there was no gaze cueing bias for congruently cued items. It should be noted that, while there were differences between Long and Short cue conditions in measures of spatial cueing, there were no differences between conditions when results were broken down between Gaze conditions. Therefore, although we cannot explicitly rule out any impact that Facilitation decay or inhibition may have on gaze cueing based on these findings, it would indicate that these factors do not explain why we have been unable to elicit a gaze cueing bias for visuospatial attention measures.

Another suggestion is that the task itself is not adequate for the detection of gaze cueing responses. A requirement for the gaze cueing paradigm is that a cue orients participants' attention to a location before the onset of a target stimulus. Then if the target is incongruent to the cued location, the participant must reorient to make a judgment and respond. If the participant can make an accurate judgment from both Congruent and Incongruent locations, then there is no need to reorient to make an accurate response, hence no differences would be expected in visuospatial attention measures. In Experiment 3, task demands were investigated as a possible explanation for the null findings in Experiments 1 and 2.

Experiment 3

In the original experiment by Turk-Browne et al. (2013), participants were only asked to make a single response when they detected the onset of the two images (target and distractor). In the current task, participants were asked to make a spatial judgment about the location of the target image. The spatial judgment task may create an issue for this task, as the correct response can be deduced from either cued location. That is, if the target appears in the cued location, the participant responds with that direction, or if it is the distractor that appears in the cued location participants can respond with the opposite direction. Therefore, there would be no need to reorient back to the target image if cued to the incongruent location.

Initially, the goal of responding to the target location was utilised to encourage more engagement with the target image to facilitate affective judgments. However, it may have had the unintended consequence of reducing engagement with targets. Therefore, in Experiment 3, the task difficulty was increased to ensure participants would have to attempt to reorient away from ncongruent cues to complete the task effectively. It was hypothesised that participants would respond less accurately and more slowly if incongruently cued to complete a complex task as participants would have to reorient before responding.

Methods

Participants

17 adult participants (mean age = 23.7, range = 19-30), of which fifteen were female, took part in the experiment.

Procedure

In Experiment 3, a letter was overlaid onto the 'true' image, and participants were asked to respond to the letter presented (up key for X down key for Y).

Results

Learning Phase

A summary of the results of the learning phase can be found in Table 2.5.

Accuracy

A Friedmans test was used to compare Accuracy among the three Gaze conditions. It was found that there was a significant difference between Gaze conditions ($X^2(2) = 6.310$, $p = .043$; Direct ($M = .952$, $SD = .426$); Congruent ($M = .943$, $SD = .483$); Incongruent ($M = .915$, $SD = .084$). Post hoc analyses was conducted using a Wilcoxon sign-ranked test with a Bonferroni correction (adjusted $p = .017$) and it was found that there was a significant difference between the Direct and Incongruent gaze conditions ($Z = -2.630$, $p = .009$; Direct Mdn(IQR) = .966 (.932 to .983); Incongruent Mdn(IQR) = .933 (.850 to .983). However there were no significant differences between the Direct and Congruent conditions ($Z = -1.404$, $p = .160$) or Congruent and Incongruent conditions ($Z = -1.527$, $p = .127$). There was no significant difference between Duration variables ($X^2(1) = 0.250$, $p = .617$). There were also no significant differences between any of the interaction variables ($X^2(5) = 5.032$, $p = .412$).

Reaction Time

For RT it was found that there was a significant difference between the Gaze conditions ($X^2(2) = 12.706$, $p = .002$; Direct $M = 1047.390$ ms, $SD = 386.033$; Congruent $M = 1128.387$ ms, $SD = 253.605$; Incongruent $M = 134.045$ ms, $SD = 245.850$). A Wilcoxon Signed-rank test with a Bonferroni correction (adjusted $p = .017$) was conducted post hoc. It was found that there were no significant differences between the Congruent and Direct gaze conditions ($Z = -1.586$, $p = .113$). However, there was a significant difference between Incongruent and Direct conditions ($Z = -2.864$, $p = .004$; Direct, Mdn(IQR) = 1225.760 ms (614.019 ms to 1369.540 ms); Incongruent, Mdn(IQR) = 1242.716 ms (1197.216 ms to 1503.126 ms), and the Congruent and Incongruent conditions ($Z = -3.385$, $p = .001$; Congruent, Mdn(IQR) = 1142.245 ms (913.928 ms to 1247.931 ms). There was no statistically significant difference between the Long and the Short conditions for the Duration ($X^2(1) = 2.882$, $p = .090$; Long ($M = 1083.619$ ms, $SD = 219.547$); Short ($M = 1225.467$ ms, $SD = 312.014$); Incongruent ($M = 134.045$ ms, $SD = 245.850$).

Analysis of interaction between the two factors found that there were significant differences ($X^5(5) = 38.983$, $p < .001$). Post hoc analysis with a Wilcoxon-signed rank test with a Bonferroni correction ($p = 0.008$) was conducted between factor pairs. In the Short condition, it was found that there were significant differences between the Congruent and Incongruent conditions (Z = -3.527, $p <$ .001; Congruent, Mdn(IQR) = 1003.207 ms (914.065 to 1106.790 ms); Incongruent, Mdn(IQR) = 1217.334 ms (112.997 ms to 1291.038 ms), as well as between the Direct and Incongruent conditions (Z = -3.527, $p < .001$; Direct, Mdn(IQR) = 1032.421 ms (616.419 ms to 1209.826 ms). For the Long condition there was a significant difference between the Incongruent and Congruent conditions (Z = -3.527, $p < .001$; Congruent, Mdn(IQR) = 1201.774 ms (868.381 ms to 1363.691 ms); Incongruent, Mdn(IQR) = 1321.876 ms (1216.893 ms to 1600.754 ms). None of the other pairings met the required level of significance (see figure 2.2).

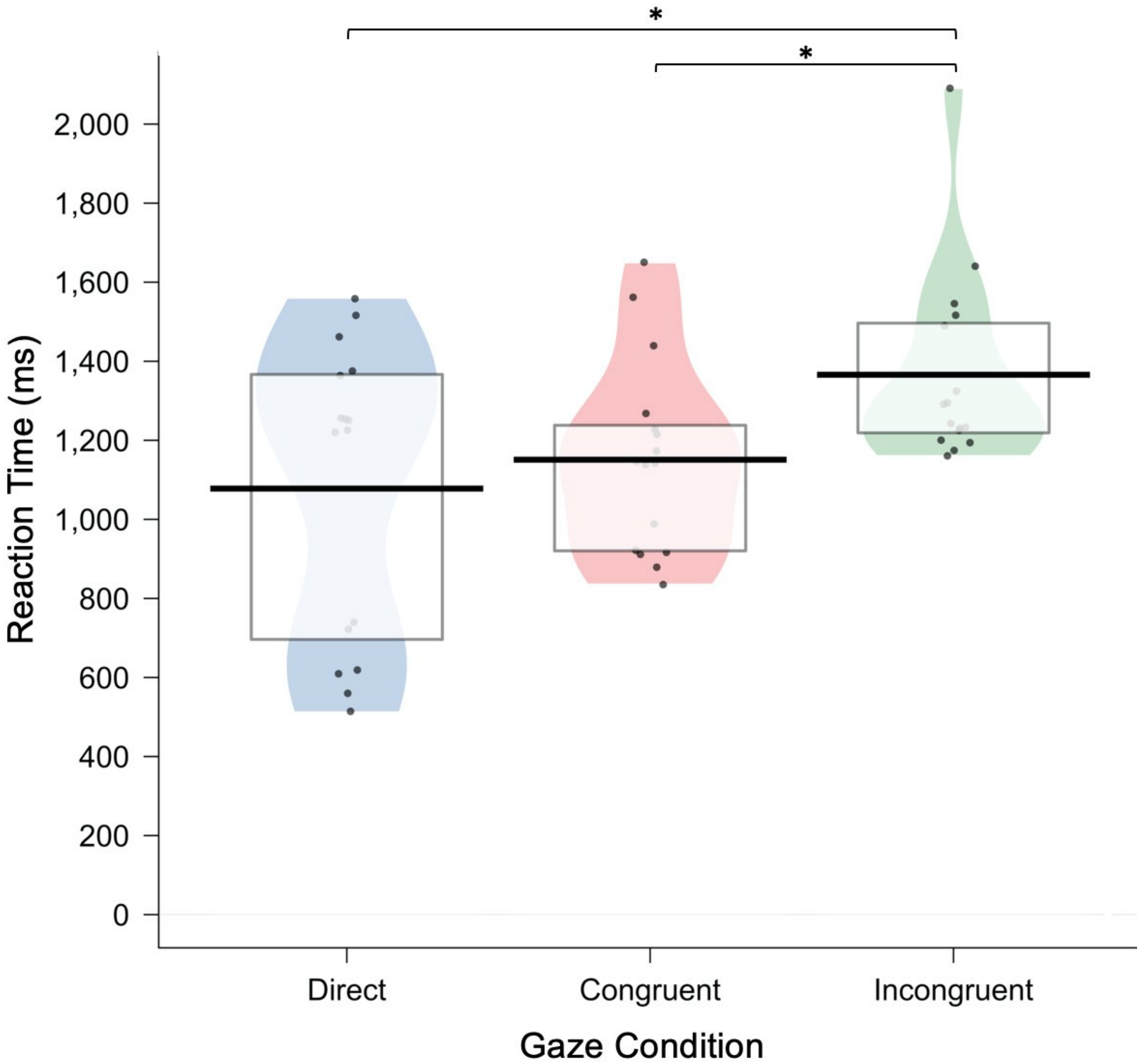

Figure 2.2: Mean RT by the Gaze condition. Error boxes represent IQR. It was found that there were significant differences between Direct vs Incongruent RTs and Congruent vs Incongruent RTs. $p < .05$ indicated by *.

Table 2.5: Mean and Standard Deviation for Memory Confidence scores from the Rating phase of Experiment 3 (N = 17). Results are displayed by duration as well as an aggregate.

		Direct		Congruent		Incongruent		Novel	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Memory Recall		0.561	0.194	0.630	0.216	0.538	0.211	0.779	0.126
Confidence	Long	2.848	0.863	2.853	0.899	2.871	0.846	N/A	N/A
	Short	2.915	0.881	2.934	0.883	2.958	0.842	N/A	N/A
	Total	2.893	0.841	2.925	0.801	2.925	0.801	2.174	0.708
Object Rating	Long	3.232	1.187	3.047	0.965	3.338	1.164	N/A	N/A
	Short	3.292	1.262	3.438	1.070	3.239	1.019	*N/A*	N/A
	Total	3.273	1.182	3.276	1.011	3.276	1.000	2.838	0.869

Test Phase

For a summary of the results of the Test Phase, see Table 2.6.

Memory Recall

Memory Recall Accuracy was significantly above chance for the Congruent (Z = -2.045, p = .041, Mdn(IQR) = .683 (.571 to .764) and for Novel stimuli only (Z = -3.575, p < .001, Mdn(IQR) = .766 (.692 to .882), but was not significantly different for either Direct (Z = -1.322, p = .186, Mdn(IQR) = .566 (.416 to .700) or Incongruent (Z = -.933, p = .351, Mdn(IQR) = .633 (.383 to .672) Gaze cues (see figure 2.3). To investigate these differences further, d' scores were calculated for Direct, Congruent, and Incongruent Gaze conditions. Wilcoxon Signed-rank test was used to compare pairings between Direct (mdn (IQR) = 1.068(0.612 to 1.334), Congruent (mdn (IQR) = 1.087(0.752 to 1.437), and Incongruent (mdn (IQR) = .802(0.490 to 1.387). It was found that there were no significant differences between Direct and Incognruent (Z = -1.194, p = .233), or Congruent and Direct (Z = -1.184, p = .236). There was a trend towards significance between Congruent and Incongruent, but it did not meet the threshold (Z = -1.915, p = .055).

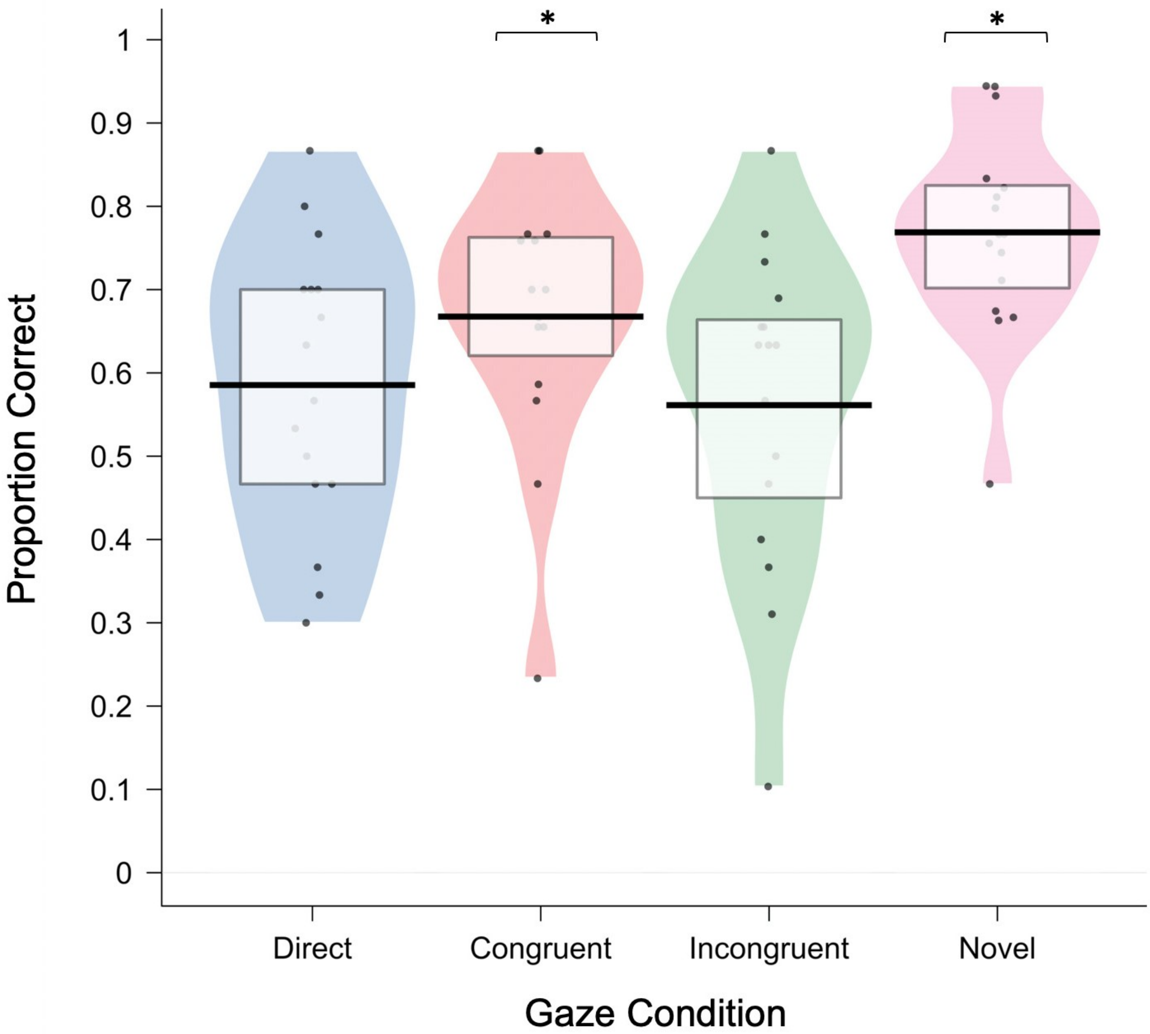

Figure 2.3: Memory recall by the Gaze condition. Error boxes represent IQR. It was found that participants Memory Recall accuracy above chance in both the Congruent and Novel conditions. $p < .05$ indicated by *.

Memory Confidence

Analysis of the Memory Confidence found that there was a significant difference between the Gaze conditions ($X^2(3) = 18.396$, $p < .001$; Direct (M = 2.893, SD= 0.841); Congruent (M = 2.90, SD = 0.811); Incongruent (M = 2.925, SD = 0.801; Novel, (M = 2.174, SD = 0.708). Post hoc analysis was conducted with a Wilcoxon Signed-Rank test with an adjusted $p = 0.012$. It was found that there was a significant difference between Direct (Z = -3.464, p = .001; *mdn* (IQR) = 2.933 (2.646 to 3.475)), Congruent (Z = -3.361, p = .001; *mdn* (IQR) = 3.000

(2.781 to 3.361), and Incongruent (Z = -3.361, p = .001; ; *mdn* (IQR) = 3.105 (2.873 to 3.289)), when compared to Novel items (*mdn* (IQR) = 2.214 (1.939 to 2.560). There was no significant difference between the Duration conditions ($X^2(1) = 2.250$, $p = .134$; Short (M = 2.935, SD= 0.814); Long (M = 2.857, SD = 0.803). There was a significant difference between the interaction variables ($X^2(5) = 26.269$, $p < .001$). However, post hoc analysis found that there were no significant differences between any of the paired conditions at the adjusted $p = 0.012$.

Affective Judgments

For Affective rating scores there was a significant main effect of Gaze condition [(F (3, 48) = 16.722 $p < .001$, $\eta_p^2 = .511$); Direct (M = 2.76, SD = 0.404); Congruent (M = 2.80, SD = 0.322); Incongruent (M = 2.75, SD = 0.368) and Novel objects (M = 2.32, SD = 0.420)]. There were statistically significant mean differences between Direct and Novel (p = .047), Congruent and Novel (p = .036), and Incongruent and Novel (p = .012). There were no significant differences between any of the Old gaze conditions (p = 1 for all comparisons). Interestingly, there was a significant main effect of Gaze Duration, with the Short (M = 3.322, SD = 1.117) condition resulting in a higher mean ratings then the Long condition (M = 3.172, SD = 1.105; F (1, 16) = 14.664, $p < .001$, $\eta_p^2 = .478$). There was no interaction between gaze cue Duration and the gaze cue Congruency ($F < 1$).

Table 2.6: *Mean and Standard Deviation for Memory Confidence scores from the Rating phase of Experiment 3 (N = 17). Results are displayed by duration as well as an aggregate.*

		Direct		Congruent		Incongruent		Novel	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Memory Recall		0.561	0.194	0.630	0.216	0.538	0.211	0.779	0.126
Confidence	Long	2.848	0.863	2.853	0.899	2.871	0.846	N/A	N/A
	Short	2.915	0.881	2.934	0.883	2.958	0.842	N/A	N/A
	Total	2.893	0.841	2.925	0.801	2.925	0.801	2.174	0.708
Object Rating	Long	3.232	1.187	3.047	0.965	3.338	1.164	N/A	N/A
	Short	3.292	1.262	3.438	1.070	3.239	1.019	*N/A*	N/A
	Total	3.273	1.182	3.276	1.011	3.276	1.000	2.838	0.869

Discussion

In Experiments 1 and 2, there was no gaze cueing bias for either visuospatial or functional measures of attention. However, by increasing the complexity of the task, it was possible to elicit a gaze cueing bias for both visuospatial and memory measures, but not for affective judgments. As expected, when the difficulty of the task is increased, participants must reorient away towards the object, ensuring engagement with the target stimuli; demonstrating similar patterns of memory recall accuracy to Turk-Browne et al. (2013). This methodology replicated the finding that congruently cued items were recalled more accurately than either congruently or incongruently cued items.

It is worth noting that even though congruently cued items were responded to significantly faster than incongruently cued items, the gaze was cueing attention as expected. RTs in the Direct condition were faster than the other two conditions, in direct contrast to Experiments 1 and 2, where the Direct condition was significantly slower. However, this is not reflected in the memory findings where only the congruently cued items showed a memory advantage.Although memory did appear to have some impact on affective judgments, the congruency of the gaze did not appear to translate to differences between affective judgments. However, implicit memory traces demonstrated by confidence scores could have had an impact on participants' preferences. Explanations for these findings will be considered in the main discussion of this chapter.

General Discussion

One of the most interesting findings in Experiments 1 and 2 was the lack of a gaze cueing effect for either visuospatial or functional measures of gaze cues. Accuracy results across all three experiments were approaching ceiling, and therefore a null result was perhaps unsurprising. However, the fact that gaze did not influence RT until the task was made more demanding may raise questions about how robust gaze cueing paradigms are. This is a surprising finding, in the light of previously reported robust benefits of gaze cues on both detection and discrimination (Friesen & Kingstone, 1998; Friesen et al., 2004; Frischen, 2007) and the fact that the primary task was modelled on Turk-Browne et al. (2013), who found a significant gaze cueing bias. The absent cueing effects on target processing also appear to explain the reduced memory performance in the recall tasks in Experiment 1 and 2. In both experiments, memory performance for previously cued items was below chance, whereas the ability to distinguish Novel items was significantly above chance. Although there was a significant difference in Affective Ratings of Incongruent vs Novel images in Experiment 1, the lack of any gaze cueing effects; the poor memory performance, and the inability to replicate it in Experiment 2 raises doubts as to whether the gaze cue was responsible for any differences in affective judgments. The same logic applies to gaze duration differences in Experiment 2. Even though there were significant effects of gaze duration in both Response Accuracy and RT, again there were no differences between gaze cueing conditions; resulting in null findings for both measures of functional attention in the subsequent rating tasks.

In Experiment 3, the task was made more challenging to increase participant engagement with the target images. The change in difficulty resulted in participants taking significantly longer to respond to incongruently cued items versus congruently cued items. These results conformed to the previous experimental findings and therefore suggests that task difficulty should not be overlooked in gaze cueing paradigms. Although response accuracy did not meet significance, there was a trend in the same direction as the reaction time findings.

In terms of memory recall accuracy, as in Experiment 1 and 2, it was found that participants' ability to respond to Novel items accurately was above chance. However, for Old items, only items that had been congruently cued were recalled significantly above chance. When d' scores were compared, it was found that d' for items in the Congruent gaze condition were higher than Incongruently cued items at a level that was approaching significance. Therefore, it is not possible to make inferences based on these results, but it may suggest that there is an opportunity to investigate these differences in future studies.

Gaze Cueing Paradigms for Measuring Visuospatial attention

Gaze cueing paradigms have been used since Posner (1980) to measure how attention is shifted due to external influences, whether the cues are exogenous, endogenous, faces, or arrows. These experiments typically find that attention is reflexively cued by faces and will even occur if participants are told to ignore the cue or that the cue is inaccurate (Friesen & Kingstone, 1998; Langton, Watt, & Bruce, 2000). However, the findings from Experiments 1 and 2 suggest that the effects require specific levels of task demands for the effect to occur. Although it is assumed that participants' attention is cued, there is no way of knowing this from the data collected in Experiments 1 and 2. In previous experiments, the simplest tasks were still able to elicit a cueing effect; for example, in the original Posner (1980) paradigm participants simply respond to the spatial location of the cue. Turk-Browne et al. (2013) used an even more straightforward task, asking participants to make a single response once the images were displayed. As demonstrated in Experiment 3, more complex discrimination tasks also resulted in the cueing effect. However, based on the first two experiments, the cueing effect may not be apparent if the task complexity falls between these simple and complex tasks. Put simply, a task which is complex enough to require secondary processing, but not complex enough to require reorienting, will not elicit a cueing effect. In Experiments 1 and 2 the task relied on distinguishing the side at which the 'true' (non-blurry) image appeared; participants would simply be able to distinguish the correct side regardless of cue direction. It could be proposed that the objective of distinguishing if the image is blurry or not was equally difficult no matter which image the participant was cued to, therefore eliminating the need to reorient and impacting the subsequent bias found in most gaze cueing paradigms. These surprising findings can have wide-reaching implications in terms of attention cueing paradigms. Adding other measures of visual attention, such as eye-tracking, will allow for more accurate interpretation of results. Were participants merely not the following gaze when the task was to detect the non-blurry image? Were they adopting a different strategy? We shall return to the use of complementary eye-tracking measures in Chapter 5.

The Effects of Congruent Gaze Cues on Memory Recall

The weak memory for old items across the board in Experiments 1 and 2, but not in Experiment 3, may again be the result of the Learning phase task because the task did not require full engagement with the object. Participants may have been unable to fully encode the objects, resulting in memory performance below chance. In Experiment 3, the participants were required to engage with the target image to discriminate the letter, resulting in better performance with congruently cued items, the only cued condition to perform above chance; providing further evidence of the importance of task definition in attentional cueing tasks.

Although the findings of Experiment 3 are based on a limited sample, and more data would need to be collected to replicate the finding to draw full conclusions, there did appear to be an effect of congruent eye gaze on memory; this would suggest that eye gaze influenced participants to focus on items for longer and more thoroughly, facilitating memory more than if the gaze was direct or incongruent. However, it is hard to conclude that this effect is unique to eye gaze. Subsequent experiments are required using non-social cues to investigate if cueing effects on memory result from simple attentional cueing or if there is something unique about directed eye gaze resulting in more reliable memory encoding.

Affective Ratings and Memory Confidence Judgments

Based on Bayliss et al. (2006) it was hypothesised that items gazed at by a face would be rated as more likeable than those which were not cued. Our results found that objects that were gazed towards were not significantly rated more highly than any other objects. Any preference for items based on eye gaze requires the participant to remember that the face looked at the item, as noted in Bry et al. (2011), but if recall of the object was already low, it is likely that any eye gaze associated with the object will not be subsequently remembered. Although they did not suggest an effect of gaze cues, the results did suggest that there were potential effects of prior exposure on the affective ratings and on the memory confidence judgements, both of which may have implications for memory recall. Ratings and confidence judgments were significantly higher for items that were seen, regardless of the direction of the gaze cue, even though explicit memory recall was not above chance for Direct or Incongruent gaze cues. This dissociation would suggest that at some level participants were encoding characteristics of the objects they saw during the Learning phase, even if this was not reflected in explicit recall results. Therefore, in future experiments, a design that accounts for both implicit and explicit memory recall would be appropriate.

Future Research Direction

The unique significance of faces is not limited to real human faces. There is evidence to show that gaze cueing effects can be prompted by animated faces, even those as simplistic as schematic drawings (Friesen & Kingstone, 1998). A proposed pipeline for testing the effects of face realism on atypical gaze processing by people with ASD is to identify a measurable effect of eye gaze on both visuospatial and functional measures of attention and then to use faces of varying levels of animacy to measure the point at which atypical behaviour appears. The goal of this study was to establish whether the liking effect had the potential to be able to disentangle the different styles of attention as well as to investigate the liking effect in a long-term memory paradigm. However, difficulties in eliciting even the visuospatial effects mean that this specific paradigm may not be ideal for testing the liking effect in future. Such paradigms that focus on replications of the liking effect are discussed in Chapter 3.

There were also some interesting findings regarding affective judgments and memory. Even though eye gaze did not appear to influence affective judgments, it is notable that items that were present in the Learning stage were rated preferable compared to novel items in Experiments 1 and 3. This preference is especially interesting when it is put into the context that memory recall was not significantly above chance for the majority of conditions across the two experiments; providing support to the findings of Bry et al. (2011) that memory is necessary to make connections between the gaze cue and the item to influence affective judgments. Our findings, as they are, were not adequate to make claims about the longevity of the effect, or how gaze cued attention may impact memory and the subsequent impact on affective judgments. To do so would require an experimental paradigm that was more directly comparable to Bayliss et al. (2006). The potential connections between memory and the liking effect will be discussed in more detail in the general discussion (Chapter 6).

Conclusion

In conclusion, the task used to measure the effects of attention cues is an important aspect of experimental design that must not be overlooked. These findings do show that provided a cueing effect on target processing can be demonstrated; eye gaze can influence memory. We speculate that this depends, crucially, on whether the task performed on target items involves participant engagement. Subsequent affective ratings did not fit our hypothesis that eye gaze should influence affective judgments, at least in the long-term memory realm. For future experiments to investigate the functional consequences of gaze on long term memory, both implicit and explicit recall should be assessed. Effects on affective judgements thus far have been demonstrated only in the short-term memory realm. Chapter 3 will build on these findings by aiming to replicate the findings of Bayliss et al. (2006) in a more directly comparable paradigm, and will also introduce the extension of using schematic faces to explore whether LSF gaze cues are adequate to elicit a gaze cueing bias for affective judgments.

Chapter 3: Replication and Extension of Bayliss, Paul, Cannon, & Tipper (2006)

Beyond cueing spatial attention, research by Bayliss et al. (2006) found that people rate objects that have been cued by the eye gaze of a photographic representation of a face as more likeable than objects that were not cued. The author's explanation as to why this effect occurred was that it was an efficient method of appraising objects in the environment. That is because people will typically attend towards objects they desire (Shimojo, Simion, Shimojo, & Scheier, 2003), basing one's affective feelings towards an object that has been gazed at would be an efficient method of appraising one's environment (Bayliss et al., 2006).

To further investigate this effect, Bayliss and colleagues also went on to investigate how facial expressions influenced this effect. They found that if the face was smiling, gazed-at objects were rated higher than if the face had a disgusted expression, but objects that were looked away from were rated equally for both conditions (Bayliss et al., 2007). These findings that positive facial expressions were associated with increases in affective judgments of objects were in line with their original explanations of the effect. If somebody is smiling when they attend to an object, it confirms the assumption that they like the object and therefore the effect would be higher. Furthermore, it has been found that trustworthiness of the face will also modulate this effect, with only faces that had been associated with a trustworthy description eliciting the effect (King, Rowe, & Leonards, 2011). It was also found by Strick et al. (2008) that the effect was only observed when more attractive individuals were presented with direct eye contact to the item. In the previous chapter it was established that as long as the task benefits from the attention being cued to a specific location, eye gaze cues will bias measures of visuospatial attention, as well as functional measures in the form of explicit memory tasks. Eye gaze has also been shown to demonstrate unique attentional properties compared to other spatial cues. It has previously been shown that the reflexive nature of eye gaze influences the attentional bias of individuals towards following eye gaze more than alternative attentional cues, such as arrows (for example, Friesen et al., 2004). One possibility is that the mechanisms for efficient face processing can facilitate attention in a way that non-social cues cannot. However, the reflexive nature of eye gaze has been reported in experiments utilising non-photorealistic faces (Friesen & Kingstone, 1998), which would be in accordance with the Johnson et al. (2015) theory of subcortical pathways for processing of eyes. That is, subcortical networks that efficiently process LSF face information may facilitate processes that enhance attentional cues from eye gaze when compared to non-social arrow cues.

Efficient processing of eye gaze utilising the subcortical pathways may explain a gaze cueing bias for visuospatial attention, but could this alone be sufficient to result in higher-order social cognition that would be required to influence affective judgments in this way? Or is the effect reliant on more social judgments that require the broader spatial frequency information? The findings that emotional expression, trustworthiness and attractiveness facilitate the effect would suggest that greater visual detail is necessary to enhance the effect. As of yet, no research available appears to specifically investigate the impact of the relative detail of faces regarding the liking effect. Investigating how these traits influence the functional consequences of gaze processing, from gaze cueing benefits, all the way to the liking effect may help us to understand how the cognitive and affective elements of gaze processing operate.

In addition to differences between the processing of faces across a spectrum of realism, individual differences, such as those along dimensions of Social Anxiety and Autism Spectrum Disorder (ASD), have been shown to have notable impacts on levels of eye contact in both Social Anxiety (Schneier, Rodebaugh, Blanco, Lewin, & Liebowitz, 2011), and ASD (Senju & Johnson, 2009b). A large body of evidence suggests that there is a complicated relationship between ASD and gaze following (Gillespie-Lynch, Elias, Escudero, Hutman, & Johnson, 2013; Pantelis & Kennedy, 2017). As previously discussed in Chapter 1, facial processing in individuals with ASD does not appear to be impeded when more simplistic, cartoon representations of faces are presented. The use of the liking effect, therefore, presents a potential dimension to investigate the processing of faces across a variety of realism in ASD. At one level, it will be possible to investigate to what extent eye-gaze following is impacted in simple face representations in both NT and autistic participants. In addition, it is possible to investigate how facial realism impacts more advanced social behaviours in the form of the liking effect, as well as how these impacts may differ between autistic individuals.

Experiment 4

In Experiment 4, the primary aim was to replicate the original finding that direct eye gaze from a photographic representation of a face would result in objects being rated higher than objects that were incongruently cued (Bayliss et al., 2006). Furthermore, these findings were extended by testing whether a simple Schematic representation of a face would influence people's affective judgments. The first between-subjects condition used a photograph of a face as in the previous experiments; the second condition used a simple schematic outline of a face. These types of schematic faces have been used in previous experiments and were found to result in similar cueing effects on reaction times and accuracy (Friesen & Kingstone, 1998), but no prior study has looked at how these schematic faces interact with the liking effect described by Bayliss et al. (2006). In addition to this experimental manipulation, measures Autistic Traits (Autistic Quotient (AQ); Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) and Social Anxiety measurements (Liebowitz Social Anxiety Scores (SAS); Heimberg et al., 1999) were taken. These two measures were selected as both ASD and social anxiety disorder have shown marked difficulties with processing eye gaze (for Social Anxiety and ASD respectively; Schneier et al., 2011; Senju & Johnson, 2009b). Gaze avoidance in ASD and social anxiety disorder do not necessarily share the same mechanisms, and therefore collecting both measures may allow us to differentiate differences between two potential dimensions of individual differences.

The replication was based on the original findings (Bayliss et al., 2006) using neutral expressions, rather than emotional expressions, to avoid any confounding effects in emotional processing that might interact with our measures of individual differences. Three hypotheses were developed:

Hypothesis 1: Participants' attention will be cued by the eye gaze of both Photographic and Schematic representations of faces, resulting in faster and more accurate responses to Congruently cued stimuli compared to Incongruently cued stimuli in both face conditions.

Hypothesis 2: Based on the findings of Bayliss et al. (2006) Congruently cued items will be liked more than Incongruently cued items when a Photographic face is used to cue attention. However, this difference will not be observed when the Schematic face is used to cue attention, as it will not be perceived as a social entity.

Hypothesis 3: – As autistic traits and socially anxious behaviours result in atypical patterns of attention when viewing faces, participants' reaction times, as well as ratings of objects, will be moderated by AQ and SAS scores.

Methods

Participants

Sixty participants (36 female, 22 male, 2 N/A) were recruited through the University recruitment system and were reimbursed either by £10 or course credit for their participation. The mean age of participants was 22.75 years (Range = 18-49). All participants gave informed consent before the experiment commenced. They were made aware that they could withdraw at any time and that all data collected would remain anonymous.

Ethical considerations

Informed consent was received by all participants, and they were made aware of their right to confidentiality and to withdraw at any time. There were no other specific ethical considerations raised before the experiment. The experiment was given ethical approval by the University of Oxford Central University Research Ethics Committee (CUREC; R43226/RE001).

Apparatus

Tasks were presented using a computer running Matlab (64-bit version R2015b; The MathWorks, Natick, MA), with images being displayed using Psychtoolbox (Kleiner et al., 2007; Pelli, 1997; Brainard, 1997). Responses were collected via keyboard input. Stimuli were presented on a 24-inch monitor with 1920 x 1080 pixel resolution.

Measures

To measure autistic symptomatology, participants completed an Autistic Quotient questionnaire (Baron-Cohen, Wheelwright, Skinner, et al., 2001). They were also asked to complete a Social Anxiety Scale questionnaire to measure the social anxiety of the participant (Heimberg et al., 1999). These were used to compare individual differences between participants and their interactions with within-subject effects.

Stimuli

All stimuli were presented on a mean grey background at full resolution. Before each trial, a fixation cross measuring 80 x 80 pixels (22mm x 22mm) was displayed.

As in the original experiment, target stimuli consisted of 36 household objects, 18 typically found in the kitchen, 18 typically found in the garage. All the objects originated from the Bank of Standardised Stimuli (Brodeur et al., 2014). These 36 objects were modified in Photoshop (Adobe; San Jose, CA) to appear in one of four colours (green, red, yellow or blue) giving a total of 144 possible targets. Images were presented in a box measuring 300 x 300 pixels (approximately 83mm x 83mm). The face stimuli consisted of 36 faces selected at random (18 male, 18 female) from the Oslo Face Database (Chelnokova et al., 2014). Faces were presented in a box measuring 332 x 337 pixels (approximately 90mm x 90mm; see Figure 3.1).

Figure 3.1: Two examples of faces from the Oslo Face Database (Chelnokova et al., 2014). Each model has been photographed with a neutral expression while maintaining a left, direct and right gaze.

Our stimuli were selected to be as similar as possible to the original experiment. The target objects belonged to the same categories of objects and were coloured in the same way. However, the faces were perceptually more complex than the images used in the original experiment as they were higher in quality and were also presented in colour. After completing the task, participants completed questionnaires in the same program, with questions presented at the top of the screen and the rating scales presented at the bottom. Participants' answers appeared in the centre of the screen before pressing enter, to confirm answers and moving to the next question.

Design

The study employed a 2x2 mixed design. The within-subjects factor was Gaze Congruency, and the two levels were defined by whether the gaze cue was Congruent or Incongruent with the upcoming target object location. The between-subjects factor was Cue Type, based on whether it was a Photographic or a Schematic representation of a face. Throughout the experiment, the response keys, the colour of the objects, and whether the gaze cue was Congruent or Incongruent were all counterbalanced between participants to avoid response bias.

Procedure

In the standard cueing trials of blocks 1 to 5, participants fixated on the fixation cross and were asked to refrain from eye movements, ignore the uninformative gaze cue and respond to the target as quickly as possible. At the start of each trial, the fixation cross appeared for 600 ms followed by a face with eyes gazing straight ahead for 1,500 ms; then, 500 ms before target onset, the eyes moved to the left or the right. The images were displayed for 2,500 ms or until a response was recorded. The screen then went blank for 1,500 ms before the next trial. At response, a tone gave feedback on performance (a bell for correct and a buzzer for incorrect/timeout). Participants used the H key and the space bar to respond to the category of the target object. The key which corresponded to Garage or Kitchen objects was counterbalanced between participants. Participants completed five blocks of 72 trials in which each target was displayed once per block. A timeline for each trial is shown in Figure 3.2A.

In the sixth block, the procedure was the same as in blocks one to five, but, after the participant had categorised the object, a blank screen was displayed for 500 ms, followed by the rating screen. Participants responded to the ratings using the keyboard but also responded aloud to an experimenter facing away from the screen. After the response, a blank screen was presented for 500 ms, followed by a fixation cross indicating the next trial. A timeline for the rating task is presented in Figure 3.2B.

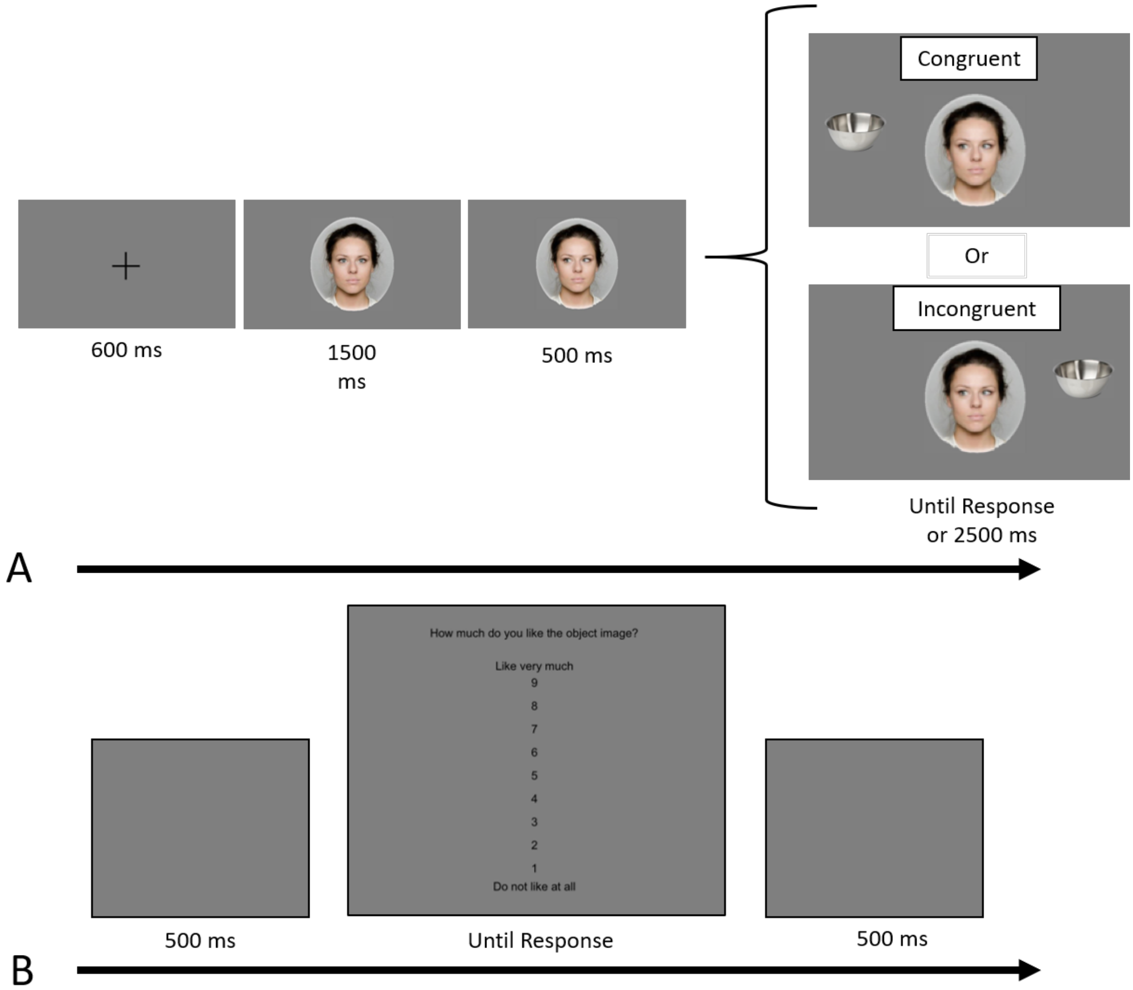

Figure 3.2: A timeline of the experimental procedure. Arrow A shows the timeline for the cueing trials in blocks 1-5. During the rating phase, participants would complete a trial as in the previous block (A), followed by a rating procedure where participants responded to the question "How much do you like the object image?".

Data analysis

A series of statistical models were applied to test each hypothesis.

Hypothesis 1: Effects of Gaze Congruency on Cueing Accuracy, Reaction Time in both Photographic and Schematic gaze cues.

To investigate how participants' attention was cued in both Schematic and Photographic faces, a series of two-way mixed ANOVAs were used to test effects on the dependent variables Response Accuracy and Median RTs. Response Accuracy scores were calculated as a proportion of correct and incorrect categorisation of the target stimuli. Median RTs were calculated based on accurate trials, removing RTs from erroneous trials. The between-subjects independent variable was whether the face cue was Photographic or Schematic (Face Type). The within-subjects independent variable was whether the gaze cue was Congruent or Incongruent (Gaze Congruency).

Hypothesis 2: Effects of Gaze Congruency on Object Ratings in both Photographic and Schematic gaze cues.

To investigate how participants' affective judgments of objects were influenced by gaze cues in both Schematic and Photographic faces, a two-way mixed ANOVA was used to test effects on the dependent variable of Object Ratings. The between-subjects independent variable was whether the face cue was Photographic or Schematic (Face Type). The within-subjects independent variable was whether the gaze cue was Congruent or Incongruent (Gaze Congruency).

Hypothesis 3: Autistic and social anxiety traits as moderators of gaze cueing effects

To investigate the effects of autistic traits and socially anxious behaviours in moderating response and rating behaviours, AQ and SAS were entered as covariates independently into a series of ANCOVAs to test effects on the dependent variables Response Accuracy, mean and Median RTs, and Object Ratings. As above, the between-subjects independent variable was the Face Type (Photo or Schematic). The within-subjects independent variable was the Gaze Congruency. The two covariates (AQ and SAS) were mean centred to overcome issues with the within-subject factors, as discussed in Thomas et al. (2009).

Outlier Exclusion

As in Chapter 2, it was decided that extreme values would not be removed from the data to minimise the impact of excluded data from one variable impacting the results of another variable. Furthermore, with the categorical nature of the task, extremes in data may be representative of processes. Therefore, RT data were compared using the median to limit the impact of extreme values without removing them from the comparisons.

Results

The mean and standard deviation for the sample n = 60 can be found in Appendix D.

Hypothesis 1: *Effects of Gaze Congruency on Cueing Accuracy, Reaction Time in both Photographic and Schematic gaze cues.*

Response Accuracy

Analysis of the proportion of correct responses found that there was no significant differences between the Congruent (M = .918, SD = .071) and Incongruent (M = .921, SD = .061) Gaze conditions ($F < 1$). Furthermore, there was no significant main effect of Face Type - Photo (M = .921, SD = .072) and Schematic (M = .918, SD = .060; $F < 1$). Analysis of the interaction between the two conditions found that there were no significant differences between the interaction variables ($F < 1$). A summary of the data can be found in Table 3.1.

Table 3.1: *Mean and Standard Deviation for Accuracy from Experiment 1 (N = 60).*

	Photo		Schematic		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Congruent	0.922	0.078	0.914	0.066	0.918	0.072
Incongruent	0.920	0.068	0.923	0.055	0.921	0.061
Total	0.921	0.072	0.918	0.060	N/A	N/A

Reaction Times

For Median RT it was found that there was a significant main effect of Gaze Congruency (F(1,58) = 12.314, $p < .001$, $\eta_p^2 = .175$), driven by faster Median RTs for Congruently cued objects (M = 670.06 ms, SD = 67.504) than Incongruent cued objects (M = 686.275 ms, SD = 71.507). There was no significant main effect between Photo (M = 662.686 ms, SD = 65.508) or Schematic (M = 694.362 ms, SD = 70.613) Face Types (F(1,58) = 3.52, p = .066, $\eta_p^2 = .057$). There was also no interaction between Gaze Congruency and Face Type ($F < 1$). A summary of the data can be found in Table 3.2.

Table 3.2: *Mean and Standard Deviation for RT (ms) from Experiment 1 (N = 60).*

	Photo		Schematic		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Congruent	656.257	62.664	684.756	70.171	670.506	67.504
Incongruent	668.581	68.743	703.970	70.919	686.275	71.507
Total	662.419	65.509	694.363	70.613	N/A	N/A

Hypothesis 2: *Effects of Gaze Congruency on Object Ratings in both Photographic and Schematic gaze cues.*

Analysis of the differences between the Congruent (M = 4.912 ms, SD = 0.828) and Incongruent (M = 4.809 ms, SD = 0.872) conditions found that there were no significant differences between the Gaze Congruency conditions (F(1,58) = 3.814, p = .056, η_p^2 = .004). Analysis of the main effect of Face Type found that there was no significant differences between the Photographic (M = 4.806 ms, SD = 0.863) or Schematic (M = 4.915 ms, SD = 0.865) conditions on Object Ratings ($F < 1$). Analyisis of the interactions found no significant differences between the two factors ($F < 1$). A summary of the data can be found in Table 3.3.

Table 3.3: *Mean and Standard Deviation for Object Rating from Experiment 1 (N = 60).*

	Photo		Schematic		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Congruent	4.869	0.866	4.957	0.801	4.913	0.828
Incongruent	4.745	0.872	4.873	0.882	4.809	0.872
Total	4.807	0.864	4.915	0.836	N/A	N/A

Hypothesis 3: *Autistic and social anxiety traits as moderators of gaze cueing effects*

A mixed model ANCOVA was used to investigate Response Accuracy, Median RTs and Object Ratings; while also controlling for AQ and SAS (independently). No notable differences were found between the ANCOVA and the ANOVA in Hypothesis 2 for any of the conditions. That is, once the variance accounted for by the covariate measures were removed from the model, there were no notable changes to the outcomes.

Discussion

It was possible to replicate the spatial gaze cueing effects found in Bayliss and colleagues' original work for both the Photographic and Schematic representations of faces. However, it was not possible to replicate the finding of influenced affective judgments in either condition, even when accounting for AQ and SAS scores. However, these unexpected null findings may result from an underpowered sample. Work done as part of the reproducibility project has demonstrated that many experimental designs may not be as robust as once thought (Open Science Collaboration, 2015). The sample size of 30 for the current experiment was based on the original study's sample of 24. Investigation of the original sample size and the effect sizes calculated from the information given ($dz = 0.196$ based on Mean and Standard deviation of the object rating findings in faces of Bayliss et al. (2006)), resulted in a post hoc power calculation of 0.15. The commonly accepted convention is that a study design should afford a minimum power of 0.8 and ideally should be 0.95 (Cohen, 1992). Based on these criteria, data from both the original experiment and the current replication were collected from underpowered samples, resulting in the increased likelihood of a type II error. The inability to replicate this effect does not mean that the effect does not exist, but instead that increased sample size is required to establish whether or not the effect holds.

Experiment 5

The unexpected null results seen in Experiment 4 raised the question of whether the sample size was adequate to detect the effect. The impact statistical power on research has been discussed in the context of the reproducibility crisis in psychology (Open Science Collaboration, 2015), raising questions on the validity of these types of effects. To ensure that this was not the cause of the null results found in Experiment 4, an experiment was designed that would be adequately powered to test for the effect.

An identical method to Experiment 4 was used to conduct this experiment online to allow for fast and easy participation from a wide variety of respondents. The hypotheses used for Experiment 5 were identical to those of Experiment 4. As the study was designed to test the robustness of the effects, the study was preregistered with the Open Science Framework (OSF; https://osf.io/25kds/) – this will be discussed in more detail throughout this report.

Methods

Power analysis

The sample size of Experiment 4 was based on the sample size on the original experiment by Bayliss et al. (2006), who used 24 participants per experiment. Subsequent power analysis found this to be under-powered. Therefore, the second set of experiments were designed to be conducted online to allow for adequate data collection. As the primary goal of the study was to replicate the findings related to affective judgements, the power analysis was based on the initial analysis of object ratings reported by Bayliss et al. (2006).

The effect size was calculated based on the mean, standard deviation and r2 value for the affective judgments scores in the face condition from the original experiment (Cohen's dz = 0.196). It was estimated that 276 participants were required to achieve the target power statistic of 0.9 based on calculations in g*Power (Faul, Erdfelder, Lang, & Buchner, 2007). A target of 300 participants was set for each condition. If after recruitment the number of participants who were below a minimum threshold for accuracy was less than 0.8 (N = 207), a further 50 participants would be collected until the number of participants had surpassed the minimum.

Participants

A total of 675 participants took part in the experiment. Participants were recruited and reimbursed through Prolific Academic (Prolific Academic: Oxford UK; Website: www.Prolific.ac.uk). Participants were expected to correctly categorise the target object in over 75% of trials to account for data quality issues associated with remote data collection and were given a £2 bonus for achieving the 75% minimum to encourage engagement. This cut off was chosen as it was felt this was the minimum acceptable level that would ensure participants had engaged with the target objects. This minimum criterion was met by 99% of participants in Experiment 4. After participants who failed to meet the minimum requirement were removed, 479 participants in total (277 Female, 200 Male, 2 Non-Binary; mean age, 35.67) provided data for our final analysis. Of the 481 participants: 235 were allocated to the Photographic face condition and 246 to the Schematic face condition. Therefore, based on the power calculations done before the experiment, the Photo face condition had a power statistic of 0.85, and the Schematic Face condition had a power score of 0.86, surpassing the minimum set before data collection began.

Ethical considerations

Informed consent was received by all participants, and they were made aware of their right to confidentiality and to withdraw at any time. Specific ethical issues regarding the use of online data collection and storage were considered to be resolved by the inherently anonymous nature of online data collection, as well as the legal data protection guarantees made by the software providers. The experiment was given ethical approval by the University of Oxford Central University Research Ethics Committee (CUREC; R43226/RE004).

Apparatus

Data was collected using Gorilla Experiment Builder (www.groilla.sc; Anwyl-Irvine, Massonnié, Flitton, Kirkham, & Evershed, 2019) between the 13th and the 21st of September 2017. Gorilla presents stimuli online via the user's browser. Both Gorilla and Prolific checked to ensure that people only took part in the experiment on desktop or laptop computers and were not able to take part on a mobile phone or tablet. Participants were asked to put their browser in full screen. However, there was no control over screen size and position. Participants responded using the H or Space key on their keyboard.

Measures

As in Experiment 4, participants were asked to complete an AQ and SAS questionnaire to collect individual differences in autism symptomology and social anxiety traits. These were presented as two separate questionnaires. In the AQ questionnaire participants completed a single page with all questions; for the SAS questionnaire participants completed the two parts for each question on a single page.

Stimuli

The same stimuli were utilised as in Experiment 4. However, due to the nature of the remote experiment, it was not possible to report the absolute size of the presented stimuli. All images were scaled so that they were presented in the same proportions as the previous experiment.

Procedures

Participants were recruited through Prolific Academic. Once they had confirmed that they would like to take part, participants were given a link to Gorilla. Here they were shown a full information sheet and were asked to consent to partake in the experiment. Once they had consented, participants were shown the instructions, and the experiment progressed as described in Experiment 4.

In addition to this, in Experiment 5, participants were asked two further questions:

> "Please list any factors that you think affected your ratings of the objects" – participants were then allowed to list any factors they wished.
>
> "How much do you think the face looking at the object affected your rating of the objects" – participants were displayed a slider bar beginning at 50 to say how they felt the eye gaze affected their decision between 0 and 100.

Data analysis

Data were analysed using the same methods as Experiment 4 to analyse Hypotheses 1 to 3. An Accuracy score was calculated based on the proportion of correct and incorrect responses to the target stimuli, and Median RTs were calculated from trials where a correct response was given. Further to this, an exploratory analysis was conducted to investigate the effects of Perceived Gaze Influence (participants self-report score between 1-100). As with the AQ and SAS, Perceived Gaze was entered as a covariate into a 2x2 ANCOVA to model the effects of individual differences in Perceived Gaze Influence.

Outlier Exclusion

Due to the nature of the online experiment participants were only selected for analysis if they achieved a minimum of 75%. As previously discussed, this was to ensure a minimum level of engagement with the study itself. Otherwise, data were treated as in previous experiments with no removal of extreme values.

Preregistration statement

The design, procedure and data analysis plan for Experiment 5 was registered on Open Science Framework (OSF) before data collection on 09/11/2017 (the registration document can be found at (https://osf.io/25kds/). Both Experiment 4 and 5 had the same design, procedure, and ultimately were analysed using the same analysis plan. However, it should be noted that only Experiment 5 was preregistered.

Results

The mean and standard deviation for the sample of 479 can be found in Appendix E.

Hypothesis 1: *Effects of Gaze Congruency on Cueing Accuracy, Reaction Time in both Photographic and Schematic gaze cues.*

Accuracy

Analysis of the proportion of correct answers for Congruent (M = .947, SD = .047) and Incongruent (M = .946, SD = .048) conditions found that there was no significant main effect (F (1,477) = 2.383, $p < .123$, $\eta_p^2 = .005$). There was also no significant difference between the Photographic (M = .946, SD = .051) or Schematic (M = .945, SD = .046) conditions ($F < 1$). Finally, there was no significant interaction between the two factors ($F < 1$). It was found that there was no significant main effect of Gaze Congruency, or Face Type($F < 1$). There was no significant interaction between the two factors. A summary of the data can be found in Table 3.4.

Table 3.4: Mean and Standard Deviation for Accuracy from Experiment 1 (N = 479).

	Photo		Schematic		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Congruent	0.947	0.050	0.946	0.046	0.948	0.047
Incongruent	0.945	0.050	0.945	0.048	0.946	0.049
Total	0.946	0.051	0.946	0.047	N/A	N/A

Reaction Times

For Median RT it was found that there was a significant main effect of Gaze Congruency (F (1,477) = 273.086, $p < .001$, $\eta_p^2 = .364$) driven by faster RT for Congruently cued objects (M = 828.740 ms, SD = 141.914) than Incongruently cued objects (M = 856.776 ms, SD = 139.980). However, there were no significant main effect of Face Type (Photogrpahic, M = 848.521 ms, SD = 140.239; Schematic, M = 844.581, SD = 147,789; $F < 1$). There was also no interaction effect between Gaze Congruency and Face Type ($F < 1$). A summary of the data can be found in Table 3.2.

Table 3.5: *Mean and Standard Deviation for Median RT (ms) from Experiment 1 (N = 479).*

	Photo		Schematic		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Congruent	833.370	141.403	831.202	148.985	831.962	142.980
Incongruent	863.672	139.077	857.961	146.595	859.917	140.486
Total	848.521	140.240	845.939	141.733	N/A	N/A

Hypothesis 2: *Effects of Gaze Congruency on Object Ratings in both Photographic and Schematic gaze cues.*

It was found that there was a trend towards significance difference between the Congruent (M = 5.218, SD = 1.143) and Incongruent (M = 5.263, SD = 1.105) conditions, but it did not reach statistical significance (F (1,477) = 3.594, p = .059, $\eta_p^2 = =$ .007). There was also no significant difference between the Congruent (M = 5.199, SD = 1.077) and Incongruent Face Types (M = 5.269, SD = 1.127; F (1,477) = 1.057, p = .813, $\eta_p^2 = <$ .001). There was also no significant interaction between the Face Type and Gaze Congruency (F < 1). A summary of the data can be found in Table 3.6.

Table 3.6: *Mean and Standard Deviation for Object Rating from Experiment 1 (N = 479).*

	Photo		Schematic		Total	
	M	*SD*	*M*	*SD*	*M*	*SD*
Congruent	5.184	1.085	5.245	1.153	5.218	1.144
Incongruent	5.216	1.070	5.295	1.102	5.264	1.105
Total	5.200	1.078	5.270	1.127	N/A	N/A

Hypothesis 3: *Autistic and social anxiety traits as moderators of gaze cueing effects.*

A mixed model ANCOVA was used to analyse Response Accuracy, RTs and Object Ratings while also controlling for AQ and SAS (independently), as well as the Perceived Gaze Influence of the eye gaze on participants' ratings. It was found that there were no notable differences between the ANCOVA and the ANOVA in Hypothesis 2 for either the AQ, the SAS, or the Perceived Gaze Influence. The exception was when AQ was entered as a covariate; there were no significant main effects of Congruency for Accuracy. For all other conditions, there were no differences once the variance accounted for by the covariates were removed.

Discussion

As in Experiment 4, it was possible to replicate the finding that Congruent gaze cues resulted in faster RTs for both face conditions, but for accuracy, only the Photo face condition resulted in greater accuracy for Congruent conditions. As in Experiment 4, no differences in affective judgements of objects based on the eye gaze of either face condition were found. It might be concluded that the gaze cue of another individual does not influence the affective judgments of an object. However, there are methodological subtleties which may explain why it was not possible to replicate this finding. Two proposed options were: 1) The interaction between face identity and objects. Capozzi, Bayliss, Elena, and Becchio (2015) found that there were interactions between the identity of the facial cues associated with the target items. They found that for the effect to occur items had to be associated with multiple faces across trials. If items were only cued by a single identity face, there were no differences between affective judgments of congruently and incongruently cued items. 2) Neutral face expression of cue stimuli. It has been suggested that faces that are presented as "neutral" are often perceived in more complex ways, typically in a negative manner (Collins, 2016; Cooney, Atlas, Joormann, Eugène, & Gotlib, 2006; E. Lee, Kang, Park, Kim, & An, 2008; Said, Sebe, & Todorov, 2009). If it is the case that the neutral expressions of these faces are being perceived as negative, and previous research has shown negative emotions may impact the effect (Bayliss et al., 2007), this may also be an explanation for the null effect reported here. These confounding factors are explored further in Experiment 6.

Furthermore, the methodological change from a lab-based setting to a home-based, online setting should not be ignored. Even with the controls put in place to ensure that participants adequately attended to the stimuli, there is always the possibility that the modality change could influence participants in ways not experienced in lab-based settings. However, the fact that the results in Experiment 5 matched those of Experiment 4, and the main finding of gaze cueing was present in Experiment 5, both argue that the online modality does seem to provide adequate results. In conjunction with lab-based testing, online experiments are an ideal way to collect results with large, diverse samples efficiently in future.

Experiment 6

The inability to replicate the effect would suggest that the effect might not be as robust as previously suggested. However, further review of the literature suggests that there may be methodological discrepancies accounting for the null results found in Experiments 1 and 2. Capozzi et al. (2015) modified the basic experimental paradigm by Bayliss et al. (2006) to investigate how the number of faces gazing towards the object that is being appraised affected the influence on the observer. They found that if a single face gazed at an object, the liking effect did not occur, and multiple faces gazing towards an object are required throughout the experiment to elicit the liking effect. Furthermore, it has been proposed that the 'neutral' expression most people display when asked to the model may appear contrary to an observer (Collins, 2016; Cooney et al., 2006; E. Lee et al., 2008; Said et al., 2009). It is possible the faces in Experiments 1 and 2 were perceived as negative and therefore, did not affect the appraisal of the object. This is in line with the findings of Bayliss and colleagues that faces displaying happy expressions were found to boost the effect of gaze cues on the liking effect compared to negative, disgusted expressions (Bayliss et al., 2007).

A third experiment was conducted to investigate if either of these factors might have influenced the findings, replacing the between-subjects factor, previously investigated facial realism (Photo vs Schematic), with manipulation of emotional expression (Neutral vs Happy). Another within-subjects factor was also tested to investigate whether a variety of faces is required to elicit the effect (Multiple faces vs Single face conditions). These additional factors aimed to give a good indication of whether either of these methodological differences may have influenced the liking effect, either independently or in combination. We hypothesised that:

Hypothesis 1: An object must be gazed at by multiple faces rather than an individual face to elicit a positive difference of affective judgments of objects.
Hypothesis 2: As found in Bayliss et al. (2007), the positive facial expression will result in a greater liking effect due to the association between the perceived positive attitude to the object.

Methods

Power analysis

Power analysis was conducted using g*Power (Faul et al., 2007), based on the effect size of the relationship between eye gaze and judgments scores in the original experiment as in Experiment 5. Initial a priori power analyses based on the t-tests used in prior experiments estimated a 173 participants required per between-subject condition for a power level of 0.8. Power analysis for the Mixed ANOVA gave an estimated power level of 0.9 for 346 participants, and this was therefore set as the target sample size.

Participants

As in Experiment 5, participants were included if they achieved overall accuracy over 75%. However, in this experiment recruitment continued until 343 participants were collected. As in Experiment 5, participants were recruited through Prolific (Happy: Mean age = 33.5(18-72) 102 female, 67 male, Non-binary 1, N/A 1; Neutral: Mean age = 36.1(18-69) 103 female, 69 male). Participants were paid a base rate of £3.25 and a bonus of £1.25 if they achieved the minimum score of 75%.

Ethical considerations

Informed consent was received by all participants, and they were made aware of their right to confidentiality and to withdraw at anytime. The experiment was given ethical approval by the University of Oxford Central University Research Ethics Committee (CUREC; R43226/RE004).

Apparatus

Data were collected using Gorilla Experiment Builder between 11/2/2018 - 14/2/2018. Stimuli were presented on the participants' PC or laptop, and it was not permitted on phones or tablets. Participants responded with the H or Space key on their keyboard.

Measures

As in the previous Experiments, the AQ and SAS questionnaires were presented through Gorilla as in Experiments 1 and 2.

Stimuli

Experiment 6 was shortened compared to Experiment 5. A reduced number of 48 object images (24 displayed in two colours) were used in the current experiment, compared to 72 images in Experiments 1 and 2. The reduction of the experiment was required due to expense issues and had the added benefit of retaining engagement by reducing the length of each block.

The Face images were also novel to this experiment. Eight models were photographed displaying two expressions: Neutral, where models were asked to present a neutral expression and Happy, where models were asked to smile without displaying any teeth. We used these bespoke images as it had been challenging to find a database of images which had both emotional expressions in combination with left-right gaze cues. It should be noted that unlike the Oslo Face Database (Chelnokova et al., 2014), these images were not controlled for factors that may influence their visual salience, but it was deemed that would be acceptable in these circumstances as long as the gaze direction was sufficient to ensure a gaze cueing bias towards the object and the expression was identifiable (see Figure 3.3 for an example of the face stimuli used for Experiment 6)

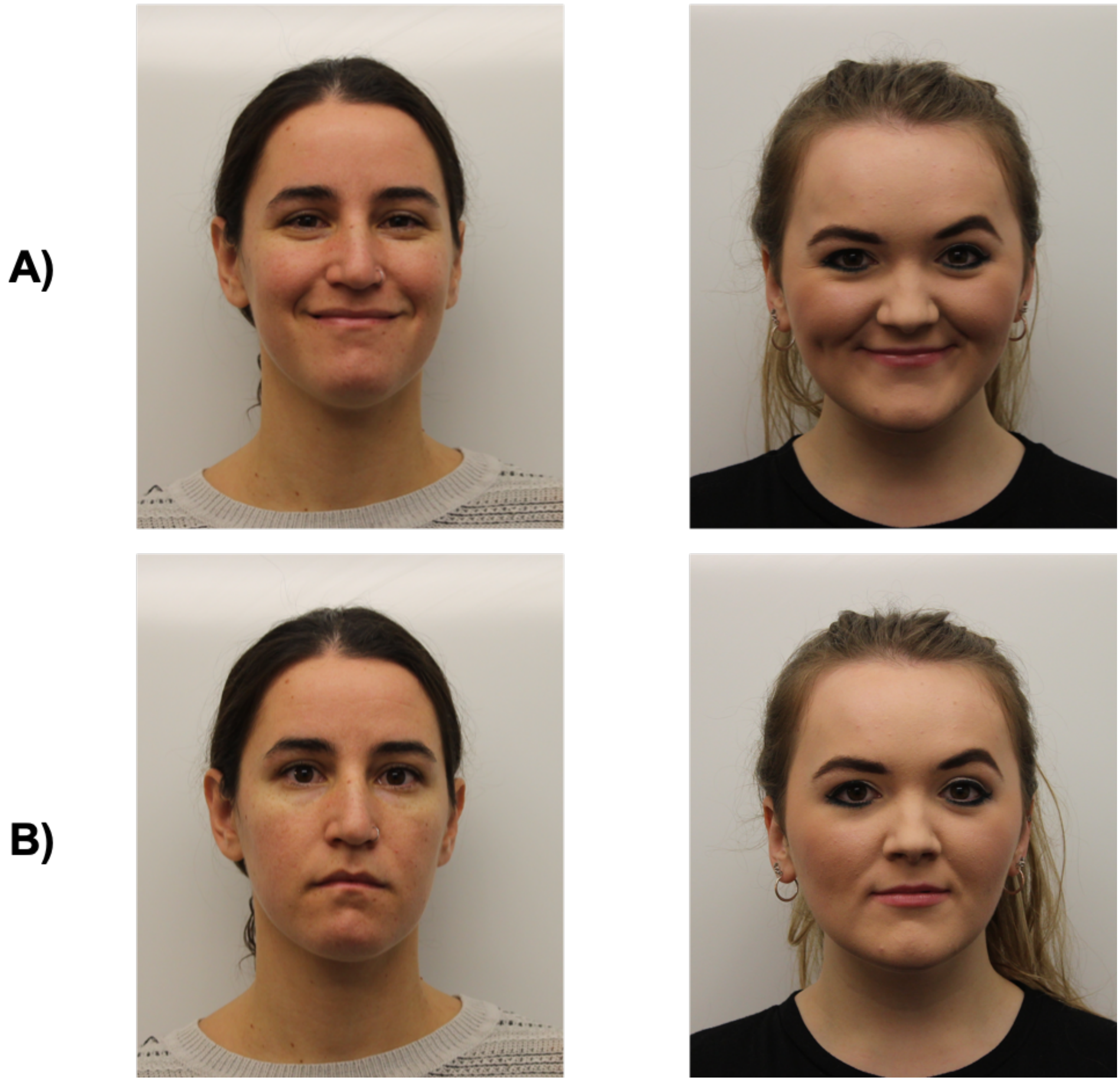

Figure 3.3: Examples of face stimuli used for Experiment 6 with both a Happy expression (A) and a Neutral expression (B)

Design

Experiment 6 followed the same design as Experiments 1 and 2, with the addition of one condition. An additional within-subjects manipulation investigated the number of faces looking towards an object across the trials (Face Number). Half of the objects were associated with a Single face, and half of the objects were associated with Multiple face identities. Objects that were associated with Single or Multiple identities were alternated between participants and identities were pseudo-randomised between conditions for all participants.

Procedures

Participants were recruited through Prolific Academic. Once they had confirmed they would like to take part participants were given a link to Gorilla. Here they were shown a full information sheet and were asked to consent to partake in the experiment. Once they had consented participants were displayed the instructions and participants began the experiment following the same procedure as in Experiment 5.

Data analysis

As in previous experiments, Accuracy was calculated as a proportion of trials that were responded to correctly. The Median RTs for all correct trials were calculated for each condition. The data were analysed using a 2X2X2 ANOVA with Expression (Happy vs Neutral) entered as a between-subjects factor, and Face Number (Multiple vs Single) and Congruency (Congruent vs Incongruent) as the within-subject factor. The same covariates as in Experiment 4 and 5 were entered into 2X2X2 ANCOVAs for AQ, SAS and Perceived Gaze Influence.

Outlier Exclusion

As in the previous experiments, participants were required to achieve a minimum accuracy score of 75%, but otherwise, no extreme data values were removed from the analysis.

Preregistration Statement

The design, procedure and data analysis plan for Experiment 6 were registered on OSF before data collection on 27/06/2017 (the registration document can be found at (https://osf.io/25kds/).

Results

Hypothesis 1: *Autistic and social anxiety traits as moderators of gaze cueing effects.*

Accuracy

Analysis of the effects of Gaze Congruency, Expression and Face Number on a proportion of correct responses was conducted using a Mixed effect ANOVA. It was found that there was no significant difference between Congruent or Incongruent gaze cues (M = .948, SD = .052 and M = .951, SD = .051 respectively). There was also no significant difference between the Happy (M = .949, SD = .055) and the Neutral (M = .948, SD = .051) facial Expressions ($F < 1$). Analysis of the main effects of Face Number found that there was also no significant difference between the Single (M = .949, SD = .045) and Multiple (M = .949, SD = .051) conditions .($F < 1$). There was a two-way interaction between Congruency and Expression (F (1,341) = 5.546, p =.019, η_p^2 = .016). Analysis of simple main effects demonstrated a significant difference between Congruent and Incongruently cued items in the Neutral condition (p = .011), but not in the Happy condition (p = .438). There were no other interactions between the factors ($F < 1$, for all interactions). A summary of the data can be found in Table 3.7.

Table 3.7: *Mean and Standard Deviation for Accuracy scores from Experiment 6 (N = 343). Scores are displayed by Expression and by Number of Faces, as well as aggregates.*

		Happy		Neutral		Total	
		M	SD	M	SD	M	SD
Single	Congruent	0.945	0.056	0.947	0.056	0.947	0.056
	Incongruent	0.954	0.051	0.952	0.054	0.952	0.054
Multiple	Congruent	0.948	0.053	0.952	0.046	0.950	0.050
	Incongruent	0.952	0.053	0.947	0.052	0.950	0.052
Total	Congruent	0.949	0.055	0.948	0.051	0.948	0.053
	Incongruent	0.951	0.055	0.951	0.052	0.951	0.053

Reaction Times

Analysis of the Median Reaction times found that there was a significant main effect of Gaze Congruency (F (1,341) = 50.466, $p < .001$, η_p^2 = .129) driven by faster reaction times for Congruently cued (M = 791.090 ms, SD = 152.292) than Incongruently cued objects (M = 805.023 ms, SD = 151.395). There were no significant differences between the Happy (M = 795.535 ms, SD = 147.875) and the Neutral Expressions (M = 800.562 ms, SD = 155.906; $F < 1$). There was also no significant main effect of Face Number (Single, M = 797.409 ms, SD = 155.491; Multiple M = 798.643 ms, SD = 148.172; $F < 1$). Analysis of the interactions found that there were no two-way interactions ($F < 1$ for all two-way interactions), or a three way interaction between all factors (F (1,341) = 3.464, p = .064, η_p^2 = .010). A summary of the data can be found in Table 3.8.

Table 3.8: *Mean and Standard Deviation for Median RT (ms) from Experiment 6 (N = 343). Scores are displayed by Expression and by Number of Faces, as well as aggregates.*

		Happy		Neutral		Total	
		M	*SD*	*M*	*SD*	*M*	*SD*
Single	Congruent	791.145	148.868	792.589	145.574	790.312	157.573
	Incongruent	804.147	151.091	805.107	156.212	804.629	153.457
Multiple	Congruent	791.145	148.868	792.589	145.574	791.869	147.011
	Incongruent	801.212	149.266	809.599	149.719	805.418	149.334
Total	Congruent	791.145	148.868	792.589	145.574	791.090	152.292
	Incongruent	802.679	150.179	807.353	152.965	805.023	151.395

Hypothesis 2: *Effects of Gaze Congruency on Object Ratings in both Photographic and Schematic gaze cues.*

Analysis of the mean Object Rating scores found that there was a main effect of Gaze Congruency for Object Ratings (F (1,341) = 12.600, $p < .001$, $\eta_p^2 = .036$) driven by higher ratings for Congruently cued objects (M = 5.246, SD = 1.118) compared to Incongruently cued objects (M = 5.147, SD = 1.168). There were no significant main effects for either Face Number (Single, M = 5.224, SD = 1.186; Multiple, M = 5.169, SD = 1.099; F (1,341) = 2.778, $p < .096$, $\eta_p^2 = .008$) or Face Expression (Happy, M = 5.224, SD = 1.065; $F < 1$). Analysis of the interactions found that there was a significant interaction between Face Number and Gaze Congruency (F (1,341) = 5.636, p = .018, $\eta_p^2 = .016$), driven by the significant difference between the Congruent and Incongruent condition for the Multiple ($p <$ .001) but not the Single (p = .511) condition. There was also a significant interaction between Face Number and the Expression (F (1,341) = 4.239, p = .040, $\eta_p^2 < .012$), driven by a significant difference between the Single and Multiple condition for the Happy (p = .002) but not the Neutral (p = .057) expression (See Figure 3.4). However, there was no significant interaction between the three conditions (F (1,341) = 1.280, p = .259, η_p^2 = 0.004). A summary of the data can be found in Table 3.9.

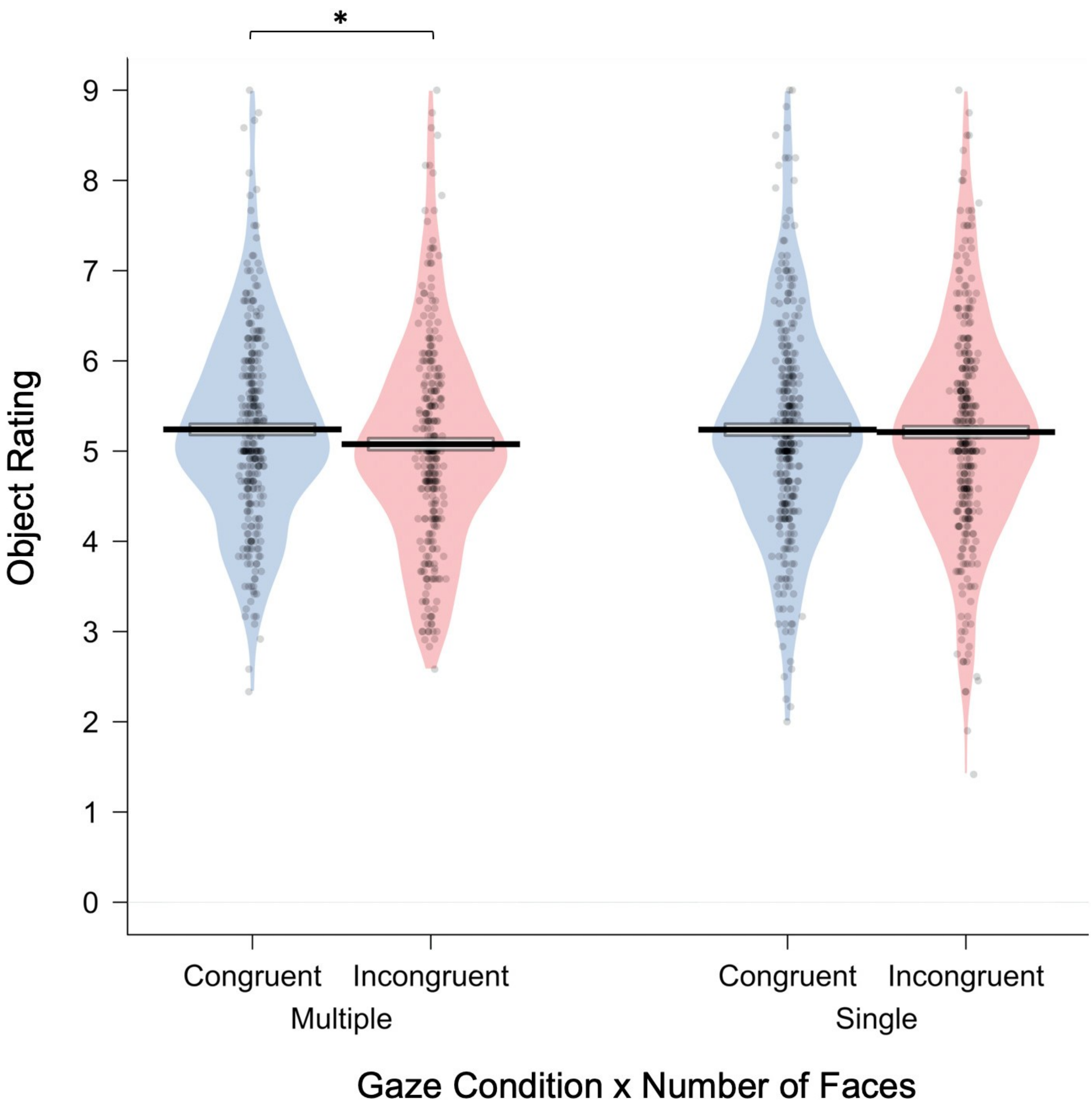

Figure 3.4: Mean object rating score by the Gaze condition and Number of Faces. The error box represents the *SD*. It was found that there was only a significant difference between Congruent and Incongruent conditions when Multiple faces attended to the objects, driven by higher scores for Congruently cued items. $p < .05$ indicated by *.

Table 3.9: *Mean and Standard Deviation for Rating scores from Experiment 6 (N = 343). Scores are displayed by Expression and by Number of Faces, as well as aggregates.*

		Happy		Neutral		Total	
		M	*SD*	*M*	*SD*	*M*	*SD*
Single	Congruent	5.295	1.135	5.181	1.221	5.238	1.179
	Incongruent	5.278	1.089	5.144	1.289	5.211	1.194
Multiple	Congruent	5.276	0.961	5.231	1.148	5.253	1.057
	Incongruent	5.050	1.076	5.120	1.206	5.085	1.142
Total	Congruent	5.286	1.048	5.206	1.184	5.246	1.118
	Incongruent	5.164	1.083	5.132	1.248	5.148	1.168

Hypothesis 3: ***Autistic and social anxiety traits as moderators of gaze cueing effects.***

Three covariates were entered (independently) into the model as in Experiment 5: The AQ, the SAS, and the Perceived Gaze Influence. It was found that there were no notable differences between the original model and the updated models, suggesting that the variance accounted for by the covariates was not influential on the outcome of the model.

Discussion

The findings of Experiment 6 replicate the results of Capozzi et al. (2015), as well as differences between positive and neutral expressions found by Bayliss et al. (2006). That is, for the liking effect to occur, the object must be directly gazed at by multiple faces with distinct identities, and the effect also appears to be boosted by a positive expression, therefore supporting our hypothesis. Our findings suggest that even though both expression conditions were able to elicit a bias in affective judgments, an ambiguous neutral expression may have led to an attenuated response to this effect. Possible explanations for why this condition influenced the effect will be discussed in more detail in the main discussion section.

The current experiment investigated, in relative detail, the optimum requirements for the liking effect to occur. There were some differences in the gaze cueing elements between prior studies, in particular in the neutral condition. The stimuli were collected specifically for this experiment, and unlike the Oslo Face database, there were no controls for visual saliency. The saliency of a cue has been found to influence the attention of an observer (Cerf, Harel, Einhäuser, in Neural, & 2008, 2007; Calvo & Nummenmaa, 2008). The ambiguity of the neutral face condition may have been exacerbated by the less visually salient stimuli; which may have ultimately interfered with the gaze cue reducing the effect on the RT. However, the bias seen in the affective judgments would indicate that participants were ultimately following the eye gaze of the face in some regards. Although this experiment does not account for why there may be differences between the two expressions, the results of this study confirm that the optimum design for this study is to use multiple faces with unique identities displaying a happy expression.

Experiment 7

The findings of Experiment 6 would suggest that the effect only occurs when the face that cues the object changes between exposures. Because this necessary condition requires a unique identity, the previously used schematic face may have been perceived as having a single identity and as a result, did not elicit the effect. Therefore, to specify that this effect does not occur in simple schematic representations of faces, it must be repeated schematic faces with unique identities. To do this a set of Schematic faces based on the Photographic faces used in Experiment 6 were developed to make eight unique schematic representations, showing a happy expression as this was found to be the optimal method of eliciting this effect. It was hypothesised that, as with photographic faces, only items that were congruently cued by multiple Schematic faces would be rated higher than objects that were incongruently cued. However, this effect would be attenuated compared to Photographic faces, as the Schematic faces would not elicit full facial processing that would optimise this effect.

Methods

Power analysis

The sample size of 173 participants for this experiment was determined based on the power analysis conducted in Experiment 6 as the design remained the same. Power analysis for the ANOVA gave an estimate of 0.85 for 173 participants and was therefore set as the target sample size.

Participants

As in Experiment 5, participants were recruited through Prolific (mean age = 36.5 (18-74); 82 female, 56 male). Participants were paid a base rate of £3.25 and a bonus of £1.25 if they achieved the minimum score of 75%. Participants were included in our analyses if they achieved over 75% with a target of 173 participants. However, due to technical difficulties, several participants had to be rejected after data collection had finished. In total, 138 participants' data were collected, which would achieve an estimated power of 0.75, slightly below the target minimum of .08.

Ethical considerations

Informed consent was received by all participants, and they were made aware of their right to confidentiality and to withdraw at any time. The experiment was given ethical approval by the University of Oxford Central University Research Ethics Committee (CUREC; R43226/RE004).

Apparatus

Data were collected using Gorilla between 20/6/2017 - 22/6/2017. Participants could only complete the experiment on a PC and were not able to complete the experiment on a tablet or mobile device. Participants used the keyboard to respond using either the H key or the Space key.

Measures

We collected AQ and SAS measures through Gorilla after the experiment was completed.

Stimuli

As in Experiment 6, a shortened procedure was used, reducing the number of objects to 24 Objects displayed in two colours for a total of 48 images.

For the schematic faces, the same face stimuli from Experiment 6 were used to make approximate schematic counterparts based on the outline of the face and marked locations to match the eyes, nose and mouth. Standardised shapes were fitted to represent these features and placed them to be the equivalent size then coloured to correspond with the model; ensuring that there were both featural and configural differences between the faces that would dictate a specific identity (see Figure 3.5)

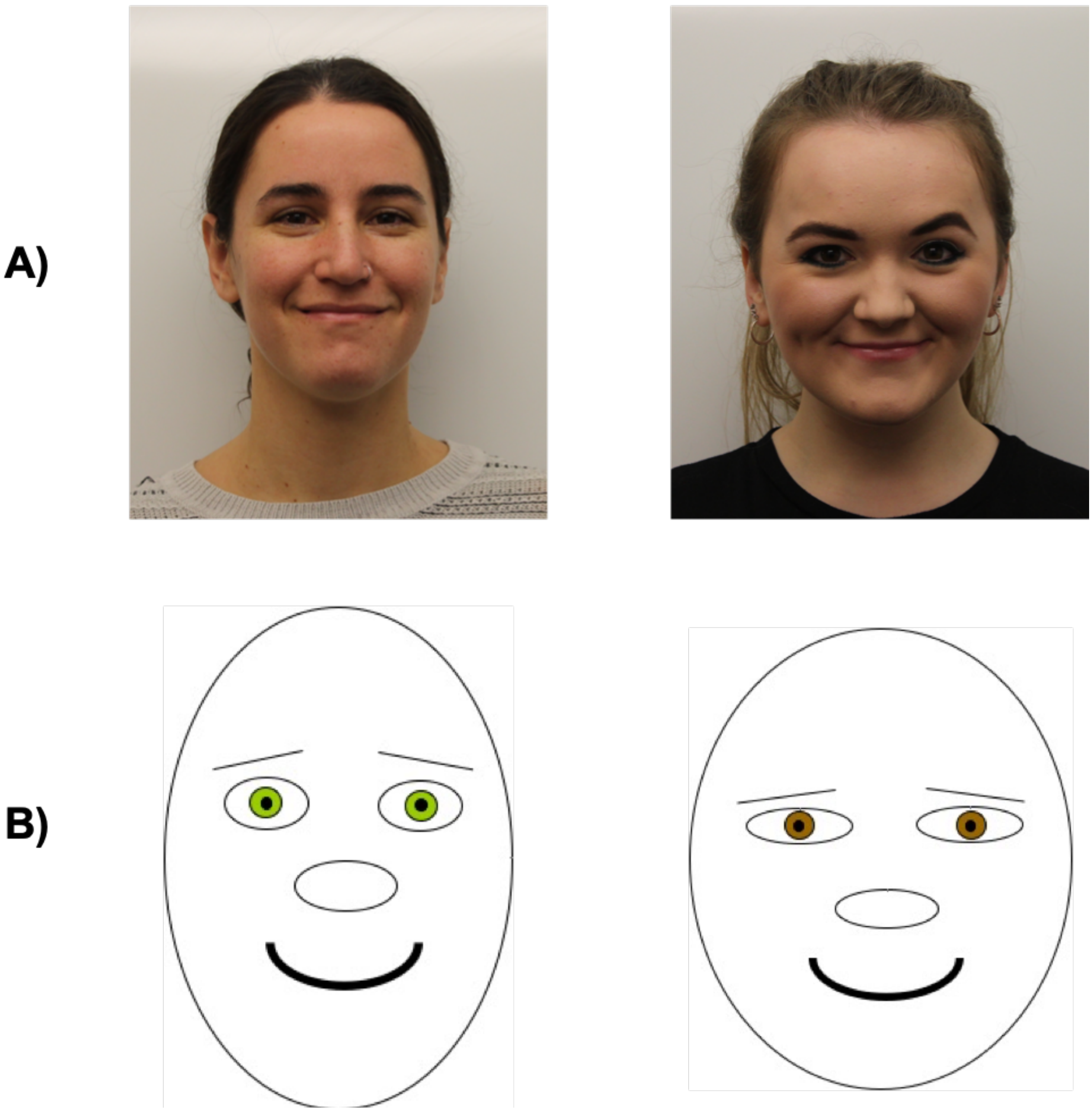

Figure 3.5: Examples of the schematic faces used in Experiment 7. By roughly matching the dimensions of the face and estimating similar eye colour, it was possible to make simple schematic faces with unique identities.

Procedure

The procedure was conducted in the same manner as in previous experiments. Participants were recruited through Prolific Academic and were given a full information sheet and consented to take part. Participants then began the experiment following the same procedure as referred to in Experiment 5.

Data Analysis

As before, the proportion of correct and incorrect scores and the Median of RTs for correct trials were calculated for the Accuracy scores and RT measures, respectively. The data were analysed using a 2x2 ANOVA within-subjects ANOVA with Face Type (Photo vs Schematic) and Congruency (Congruent vs Incongruent) as the within-subjects factors. As in the previous experiments, AQ, SAS and Perceived Gaze Influence were entered independently as Covariates into a series of 2X2 ANCOVAs.

Outlier Exclusion

As in previous chapters, participants were required to achieve 75% to be entered for analysis to ensure engagement in the experiment. Otherwise, there was no removal of outlying data from the analysis.

Results

The mean and standard deviation for the sample $N = 138$ can be found in Appendix G.

Hypothesis 1: *Autistic and social anxiety traits as moderators of gaze cueing effects.*

Response Accuracy

Analysis of the proportion of correct results found that there was no significant main effect of Gaze Congruency (Congruent, M = .951, SD = .057; Incongruent, M = .954, SD = .053; F(1,137) = 1.926, p = .167, η_p^2 = .014). There were also no significant differences between the Single (M = .951, SD = .057) or Multiple (M = .954, SD = .054) conditions for Face Number ($F < 1$). Analysis of the interactions found no significant effects ($F < 1$). A summary of the data can be found in Table 3.10.

Table 3.10: *Mean and Standard Deviations for Accuracy from Experiment 7 (N = 139). Scores are presented by Number of Faces and as aggregate scores.*

		M	*SD*
Single	Congruent	0.950	0.061
	Incongruent	0.954	0.053
Multiple	Congruent	0.953	0.054
	Incongruent	0.956	0.054
Total	Congruent	0.951	0.057
	Incongruent	0.955	0.054

Reaction Times

Analysis of Median RTs found that there was a significant main effect of Congruency (F(1,137) = 56.156, $p < .001$, η_p^2 = .255) driven by a faster reaction time in the Congruent condition(M = 797.031 ms, SD = 130.390) compared to the Incongruent condition (M = 816.947 ms, SD = 129.832). There was no significant effect of Face Number (Single, M = 807.151 ms, SD = 128.635; Multiple, M = 806.826 ms, SD = 131.287; $F < 1$). There was no significant interaction between the two conditions ($F < 1$). A summary of the data can be found in Table 3.11.

Table 3.11: *Mean and Standard Deviations for Median RT from Experiment 7 (N = 139). Scores are presented by Number of Faces and as aggregate scores.*

		M	*SD*
Single	Congruent	797.883	128.555
	Incongruent	816.420	128.715
Multiple	Congruent	796.179	131.626
	Incongruent	817.473	130.949
Total	Congruent	797.031	130.091
	Incongruent	816.947	129.832

Hypothesis 2: *Effects of Gaze Congruency on Object Ratings in both Photographic and Schematic gaze cues.*

For the Object Ratings there was a significant effect of Gaze Congruency (F(1,137) = 7.114, p = .009, η_p^2 = .050) driven by a higher rating for Congruent items (M = 5.411, SD = 1.159) than Incongruent items (M = 5.273, SD = 1.128). There were no significant differences between the Face Number conditions (Single, M = 5.341, SD = 1.138; Multiple, M = 5.343, SD = 1.149; $F < 1$). There was no significant interaction between the two factors of Face Number and Gaze Congruency ($F < 1$). A summary of the data can be found in Table 3.12.

Table 3.12: *Mean and Standard Deviations for Object Rating from Experiment 7 (N = 139). Scores are presented by Number of Faces and as aggregate scores.*

		M	*SD*
Single	Congruent	5.421	1.164
	Incongruent	5.286	1.143
Multiple	Congruent	5.402	1.155
	Incongruent	5.286	1.143
Total	Congruent	5.411	1.159
	Incongruent	5.286	1.143

Hypothesis 3: *Autistic and social anxiety traits as moderators of gaze cueing effects.*

The three covariates (AQ, SAS, and the Perceived Influence) were entered to independent ANCOVA's. It was found that there were no notable differences between the original model and the updated models; suggesting the variance accounted for by the covariates was not influential on the outcome of the model.

Discussion

It was found that there was a significant gaze cueing bias, but that there was no interaction between Congruent gaze cues and the Face Number condition, which would suggest that the multiple identities were not required to elicit the effect. However, this raises an interesting question: why did Congruent gaze result in a preference for items cued by schematic faces in this experiment, but not in previous experiments?

One possible explanation of the differences between the current experiment and Experiments 4 and 5 is the increase in details in the eyes of the schematic face. Eye colour was added to the schematic faces to add a featural element. The addition of more detailed featural information that was important for the perception of identity may have also impacted how the faces were perceived and may subsequently have resulted in a liking effect with Schematic faces where it was not seen previously. The relationship between facial realism and identity will be discussed further in the general discussion.

General Discussion

The main aim of the study was to replicate the findings of Bayliss et al. (2006), as well as to extend their previous findings to understand how this effect would occur in perceptually sparse face conditions. Throughout the experiment, it was found, with one exception, effects of gaze cueing on reaction times, with faster responses to objects that had been congruently cued, conforming with the previous literature. However, in our first two experiments, it was not possible to elicit a bias towards congruently cued objects in affective judgments (the "liking effect"), even in large samples. Further research found that the effect required multiple faces to gaze towards the object to elicit the liking effect, ultimately replicating more recent findings (Capozzi et al., 2015). However, in the schematic face condition, there were no differences between the multiple and single face conditions, rather a general preference for items that were congruently cued.

The Influence of Multiple Identities on Affective Judgments

Capozzi et al. (2015) suggested that multiple identities were able to elicit the effect as a result of the social consensus hypothesis. They suggested that seeing multiple people look towards the object leads to the conclusion that the object is liked by others and that therefore it is an efficient judgment to conform to the crowd. This explanation again would be in line with findings that perceived positive face attributes. However, this explanation would suggest that the consensus builds the effect up over trials. Therefore, it would be expected that there would be an attenuated effect even when using a single face condition. It is the case that this is seen in both Experiments 4 and 5. However, in these experiments, a positive difference is only seen in Experiment 7 with a negative difference in Experiment 5. These conflicting results make it difficult to make any claims about whether there is even a consistent trend towards the liking effect in single face conditions, although there are previous studies utilising a single face condition that have elicited an effect (King et al., 2011; Strick et al., 2008).

There may be alternative mechanistic explanations for why the effect is only consistently seen when objects are associated with Multiple faces using Photographic stimuli. One possible suggestion is that instead of multiple faces boosting the effect, repeated connections between an object and a single face identity may result in a habituation effect to the gaze connection, limiting the extent to which eye gaze will influence people's affective judgments of the targets. There are parallels in the fMRI literature of repetition suppression for repeated stimuli (Yi, Kelley, Marois, & Chun, 2006). There is no literature on habituation or repetition suppression on affective judgments. However, in more cognitive terms, the multiple/single differences could be explained through habituation. A habituating effect may also explain the original finding by Bayliss et al. (2006) that the arrow cue did not elicit a preference as a result of the cue. As the arrow cue does not change across trials, participants may habituate to the object and therefore may not show a preference.

This explanation of the liking effect mechanisms is complicated by the findings in Experiment 7. If it were the case that multiple identity associations with objects reduced habituation allowing the objects to benefit as a result of widespread attention to the item rather specific social attention related to eye gaze, it would be expected that this effect would also be seen in the multiple conditions with the schematic faces in Experiment 7. However, as it appeared that this effect was equal across both Single and Multiple conditions, this cannot be the case. Therefore, it is likely that whatever mechanisms that account for the social consensus explanation of the liking effect are more complex than simple visual habituation; raising the question: why do schematic faces appear to elicit a liking effect, irrelevant of expression or number of face/item connections? One possible reason is that the subcortical pathways that are triggered by a face shape begin activation of more extensive face processing networks that modulate how people associate the gaze with the item. However, in the schematic representations of faces, there is not enough information for full processing to occur. Therefore, the social consensus influence cannot act on participants' judgments. By raising the level of detail in the schematic condition, it may cross a threshold for the effect to occur that previous schematic faces did not. Any differences may have been identified in differences in the effect size, but the use of different tests (2 vs 3-factor ANOVAs) makes it difficult to make direct comparisons. The number of face/item connections will be addressed further in Chapter 4.

Issues with the Use of AQ as a Covariate for Investigating Functional Consequences of Eye Gaze

A further extension of the original study was the addition of measures of autism and social anxiety to investigate how they may have influenced the liking effect. ANCOVAs were used to remove the variance that was accounted for by these measures, and it was found that there were no notable differences. However, there were possible issues with the use of self-report measures that meant they might not have been indicative of the extreme levels of behavioural traits, in particular the AQ, that would impact face processing behaviours. For example, in Experiment 6, only 26 of 346 participants scored over the established diagnostic cut off of 32 for ASD, and of those 26, only 14 were in the multiple face condition. Therefore, the sample size is not adequate to allow useful conclusions. If the marked atypicalities in face perception and gaze following are only noted in individuals on the extreme end of this distribution, there may simply have not been enough variance within our participants to gain an adequate sample of extreme scores. Although there is evidence that there are differences between individuals with high vs low AQ scores in gaze cueing tasks (Bayliss & Tipper, 2005), limited variance in the recorded AQ meant that differences in social processes required for influencing affective judgments were possibly not represented. Therefore, to really investigate the extent of autistic and social anxiety traits on this effect, more sensitive measures of autistic behaviours or direct comparison between autistic participants and NT controls should be conducted. These concerns will be addressed in Chapters 4 and 5 by utilising an online study with specific high and low autistic trait cohorts, as well as collecting data with autistic participants.

Potential for Online Data Collection for Large Scale Replications

Since the findings of the reproducibility project were published (Open Science Collaboration, 2015), many psychological effects have been under scrutiny. Ensuring that testing for an effect is adequately powered has been recognised as one of the issues with experiments that were not able to replicate in the follow-up experiments.In this chapter, several large scale experiments were designed to replicate and extend the findings of Bayliss et al. (2006). Although Experiments 4 and 5 were unable to replicate the liking effect due to methodological differences, they were able to demonstrate that the Gorilla online experimental platform was useful in providing adequate results that aligned with a smaller lab-based sample; at least regarding the accuracy and RT (www.gorilla.sc; Anwyl-Irvine et al., 2019). The use of tools such as Prolific for participant recruitment and Gorilla for data collection will allow for much larger, and more diverse, samples to be collected quickly and at a relatively lower cost per participant, even for more extended, more complex behavioural studies. As long as adequate attentional checks are carried out to ensure that only data collected from participants who engages with the experiment are analysed, our research suggests that online data collection has the potential to improve how human data can be collected.

Conclusions

The main aim of this study was to replicate the findings of Bayliss et al. (2006). The effect was not replicated in experiments 4 and 5 due to minor differences in the methodology, but results of Experiment 6 did replicate the original findings of Bayliss et al. (2006, 2007). Furthermore, this confirms the findings by Cappozzi (2015) that the effect required connections between multiple face identities, as well as confirming that positive emotions expressed by the cue face would also influence the impact of the effect as demonstrated by Bayliss et al. (2007). However, although the liking effect was seen in the Schematic faces, there were no differences between the single and multiple face identity condition, raising questions about differences between face modalities. Although there were no differences based on AQ measures, it cannot be assumed that there will be no differences between people who have a diagnosis with ASD. Now that the parameters for the paradigm have been established, future chapters will now focus on investigating the liking effect in specific population samples. Finally, in this chapter, some successful data sets were collected online using Gorilla. Although further testing will be required to confirm the validity, this method does show promise as a solution for collecting large scale data sets of complex behavioural paradigms.

Chapter 4:

Functional Consequences of Gaze Cues on Affective Judgments in Individuals with High and Low Autistic Traits

Experiment 8

In the previous chapter, consistent with Capozzi et al. (2015), it was established that when photographic faces oriented their gaze towards items, there was a significant preference towards items that were congruently cued when cued by multiple face identities across trials. However, this liking effect was not seen when a single face consistently gazed towards an item.

Interestingly, the liking effect was seen in objects that were associated with both Multiple and Single face identities when schematic representations of faces were used. Although it was predicted that there would be no effect in schematic faces, these findings outline that there were distinguishable differences in the functional consequences of eye gaze in schematic vs photographic representations of faces. Further research is required to establish the mechanistic differences of this effect. However, in the context of the thesis aims, namely to disentangle differences between how faces are processed across a spectrum of realism in ASD, this effect may demonstrate potential.

In previous chapters, the Autistic Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, et al., 2001) has been used to account for variance in the broader autism phenotype that may account for behavioural differences. However, it has consistently been found that the AQ did not explain any such differences between individuals. Furthermore, using the AQ to account for a variance may indicate differences, but ultimately it may not provide us with insight into how processing of faces may be specifically impeded in autistic spectrum disorder. Possibly due to under-sampling among participants who demonstrate the most severe autistic traits.

For example, in Experiments 3 and 4 from Chapter 3, only 26 of 509 participants in total had an AQ score high enough to meet the criteria for potential diagnosis. One possible explanation is that the sample collected in the previous chapter failed to collect data from those demonstrating the most severe autistic traits.

Autism is a spectrum disorder, and as such diagnosis is made when behaviour exhibits at the extreme end of the spectrum in a way that impairs behaviour (Mandy, Charman, & Skuse, 2012; American Psychiatric Association, 2013). Therefore, by collecting samples from the general population and entering the AQ as a continuous covariate, it has not been possible to investigate those at the extreme of the spectrum. If gaze avoidance is a response to atypical processing at the end of the spectrum, it would be a plausible explanation as to why we have not seen any gaze related difficulties based on autistic traits. Therefore, by utilising the established liking effect paradigm with individuals who show high levels of autistic traits compared to individuals with below-average autistic traits, it will be possible to draw conclusions about how face processing is impeded in those with more severe autistic traits.

As previously discussed, there is a vast literature that shows atypical responses to gaze cueing among autistic people (Nation & Penny, 2008; Frischen, 2007). Autistic individuals are equally as good at detecting faces with direct and averted eye gaze, whereas neurotypical individuals show a significant bias towards faces with direct eye gaze (Senju, Yaguchi, Tojo, & Hasegawa, 2003). Difficulties in detecting eyes and interacting with others' eye contact have been shown to impact how individuals with ASD share attention (Caruana et al., 2018; Mundy, Sigman, & Kasari, 1994). Furthermore, it has been demonstrated that individuals with ASD are less effective at naturally sharing attention with others (Hobson & Hobson, 2007), impacting their ability to develop language skills (Charman, 2003; Morales et al., 2000; Tomasello & Farrar, 1986), as well as impacting their mentalising skills (Charman et al., 2000). However, regarding gaze cueing paradigms, the research has been inconclusive (for review see Nation & Penny). Although initial findings suggested that autistic individuals did not follow gaze cues (Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997; Leekam, López, & Moore, 2000), more recent findings have not supported these conclusions (Senju et al., 2005).

One proposal has been that many cueing experiments work by presenting a face whose eyes shift in a way that results in perceived movement. This perceived movement would in turn work as a directional cue, which could explain comparable results between ASD and NT participants (Farroni et al., 2000; Leekam et al., 2000;

Swettenham, Condie, Campbell, Milne, & Coleman, 2003). However, subsequent experiments with static gaze cues also demonstrated a gaze cueing bias with autistic participants; suggesting that autistic individuals were explicitly using the direction of the eyes rather than just motion to attend to the object. However, this raised a question: If autistic people can follow the direction of the eyes, why are there also reported deficits in gaze following and joint attention in day to day interactions? Research using gaze cueing paradigms found that although autistic individuals were utilising gaze direction, it was not reflexive as it appeared in NT individuals (Ristic et al., 2005). Ristic et al. (2005) found that when Autistic participants were told that gaze cues were 80% likely to be incorrect, the cueing effect was not seen, but it was in NT individuals. The effect also appears to be less nuanced than in neurotypical individuals. Whereas typical attention shifts are mediated by multiple factors, such as hemisphere presentation and expectations (Frischen & Tipper, 2006; Vlamings, Stauder, Van Son, & Mottron, 2005), as well as gender (Bayliss et al., 2005), these factors do not seem to have the same impact on ASD (Senju, Tojo, Dairoku, & Hasegawa, 2004).

It has also been proposed that the conflicting results in gaze cueing paradigms are associated with patterns of gaze avoidant behaviour (Dalton et al., 2005; Kliemann et al., 2012). Variance within such avoidant behaviours across autistic individuals may explain the heterogeneity in gaze cueing experiments with autistic participants as well as the wider population (Kliemann et al., 2012). If avoidance of direct eye contact limits how well autistic individuals interact with eye contact, variations in avoidance may explain why there have been differences between ASD and NT participants in individual experimental designs, but not others. By utilising the liking effect paradigm, it is possible to interpret how individuals interact with functional elements of eye gaze, not just the spatial aspects. Differentiating between visuospatial and functional consequences of eye gaze processing in this way will help to disentangle differences between NT and ASD participants when processing eye gaze.

If it were the case that there were differences between the ASD and NT individuals, there are several potential explanations. One possible explanation for an atypical liking effect may be the avoidance of the eyes of photographic representations of faces: gaze avoidance could impede gaze cueing for spatial tasks (accuracy, RT). In turn, reduced gaze cueing effects could relate to the subsequent processing of the object that results in the liking effect. Secondly, difficulties in discerning facial identity may impact social cueing (Lewis, Shakeshaft, & Plomin, 2018; Weigelt, Koldewyn, & Kanwisher, 2012), as the liking effect is reliant on distinguishing identity and building connections between multiple face identities and the items (Capozzi et al., 2015). Finally, if it was the case that complexity imposed by more detailed photographic representations of faces interferes with the ability to extract the information required to influence affective judgments, it may be the case that the liking effect will be present when low detail schematic face stimuli are used as cues.

Therefore, there were two primary goals of this study. Firstly, to establish whether the liking effect would occur among individuals with high and low traits of ASD, to move away from the limitations of testing the effects of AQ in population samples. Secondly, to try and disentangle whether differences in how individuals process low detail and high detail faces will also be present in both high and low autistic traits through the impact on the ensuing liking effects. As in Chapter 3, data collection leveraged online methods to allow us to meet the statistical power requirements of a large enough sample to investigate these effects. Participants were pre-screened for AQ scores so that we could recruit individuals who scored above the AQ at-risk cut off that would suggest behavioural traits equivalent to that of a diagnosis of ASD. Although this method limits to what extent conclusions can be drawn concerning a full diagnosis of ASD, this would give a good indication of differences compared to individuals who achieved below average AQ scores. For this experiment two hypotheses were proposed:

> *Hypothesis 1*: Individuals whose AQ score is above the clinical cut off for autistic traits (AQ-S score greater than 80; Hoekstra et al., 2010), would not be as influenced by the eye gaze of photographic faces when making affective judgments, compared to individuals with below-average AQ scores as they would have more difficulty processing eye gaze, limiting the effect.
>
> *Hypothesis 2*: There would be no significant differences between the High and Low AQ score groups when the same task was completed with Schematic representations of faces, as the High AQ group would not have the same difficulties processing eye gaze of schematic faces.

Methods

Participant Recruitment

As the goal of the study was to compare individuals who demonstrated both high and low levels of autistic traits, a pre-screen questionnaire was used to gauge autistic traits. Then individuals were designated to either a High or a Low group and were then invited to take part in the experiment.

Pre-screen

For the pre-screen, the shorter AQ-28 was used to measure ASD behavioural traits (Hoekstra et al., 2010). The shorter nature of the AQ-28 was ideal for the pre-screen questionnaire and has been demonstrated to correlate highly with the original AQ questionnaire. Hoekstra and colleagues established a score of >65 as a cut off for the presence of autistic traits. As the aim was to establish these findings in High autistic traits it was decided to set an AQ-s of 80, closer to the mean of the original autistic group (89; Hoekstra et al., 2010). As for the Low autistic traits group, an AQ-s score of 60 was established; below the threshold established by Hoekstra and colleagues for autistic traits and closer to the average AQ-s score for the original control groups (56; Hoekstra et al., 2010). The pre-screen questionnaire was made available through Prolific for up to 1000 participants and was completed on Gorilla. Of the 1000 participants who took part 101 participants met the threshold for the High group and 467 for the Low group.

Power analysis

A power analysis was conducted based on four independent variables. The effect size of the difference between object ratings in Experiment 6 was used to calculate the required sample size for a 2X2X2X2 mixed ANOVA. The resulting a priori power analysis based on the effect size (f = 0.142), the correlation between variables(0.7), and the Alpha level of 0.05 found that, for a power of 0.95, we required an overall sample of 156 participants. For the minimum power level of 0.8, a sample size of 102 was required. Seventy-eight participants were invited to participate for each of the High and Low ASD groups to achieve the upper power level of 0.95. The high and low group were then divided into two further groups of 35 for the Photographic and Schematic face groups.

Participants

Based on these targets, we invited 78 randomly selected participants who had scored an AQ-s score of 80 and 78 participants who scored below 65. Participants were given two weeks to complete the experiment. After the two weeks, a full sample was collected for both groups. To account for external influences that could not be controlled outside a lab setting, a minimum performance level requirement was set. Participants were expected to have over 75% accuracy. Once participants who had scored below the minimum accuracy threshold (75% accuracy) score was removed, a small number of extra participants were invited from the Low group to ensure the two groups were balanced and above the minimum participant number required. In total, there were 65 participants in the High group and 59 participants in the Low group. Based on the power calculations, this would give an overall power estimate of 0.88.

Participants mean age was 35.6 (range: 18 – 73) and had normal or corrected to normal vision. The experiment lasted approximately 45 minutes, and participants were reimbursed £3.75 for participation and a bonus £1.25 if they met a minimum accuracy score, as detailed in Chapter 3. Participants were recruited through Prolific Academic (Prolific Academic: Oxford UK; Website: www.Prolific.ac.uk) and were forwarded to Gorilla) Experiment Builder (www.gorilla.sc; Anwyl-Irvine et al., 2019) through which the study was conducted.

Ethical considerations

Informed consent was received by all participants, and they were made aware of their right to confidentiality and to withdraw at any time. The experiment was given ethical approval by the University of Oxford Central University Research Ethics Committee (CUREC; R57031/RE001).

Apparatus

Data were collected online via the participant's browser using Gorilla between the 13/7/2018 - 24/7/2018. Participants could only complete the experiment on a desktop or laptop computer, and the experiment was not available on a mobile or tablet. Participants were asked to put their browser in full screen, but there was no control over screen size and position. Participants responded using the H or Space key on their keyboard.

Measures

To measure autistic traits, the AQ-s was used to collect measures of autistic traits in the pre-screen (Hoekstra et al., 2010). This measure was selected over the original AQ-50 for two reasons. Firstly, the reduced size was ideal for short pre-screening questionnaires of a large number of participants. Secondly, it has been found that the AQ-50 may show significant overlap with individuals diagnosed with social anxiety disorder (Freeth, Bullock, & Milne, 2013). The AQ-s reduces the subcategories that show the most overlap and should, therefore, provide a more suitable measure of autistic behavioural traits.

Stimuli

For this experiment, the same stimulus set from Experiments 6 and 7 was used: 24 target items of which 12 would be typically found in the garage, and 12 which would typically be found in the garden. Each item was manipulated to be a shade of 1 of 4 colours (red, green, yellow or blue), for a total of 96 unique items. For photographic face cue, the same eight photographs of individuals from the prior experiment were used. These photographs consisted of direct gaze as well as a left and right directed gaze cue. The schematic faces were the same set from the prior study (Experiment 4) with eight unique identities. Each face was manipulated to have a direct as well as a left and right gaze cue (see Figure 3.5 in the previous chapter for an example of both Photographic and Schematic stimuli)

Design

The study was a 2x2x2x2 mixed design. The two within-subject factors were: Gaze Congruency with two levels defined by whether the gaze Congruently or Incongruently cued the item location; Number of faces, where the two levels were either Multiple or Single identity-item connections. There were also two between-subjects factors: Face Type -the effect of the modality of the face cues with the two levels defined as either Photographic or Schematic representations of faces; AQ Group - with two levels defined as either High AQ group or Low AQ group.

Procedure

For this study, participants completed one block fewer than in Experiments 6 and 7, completing four training blocks and one rating block of 48 trials per block to reduce the length of the experiment to facilitate the collection of the correct number of participants. First participants were displayed a fixation cross for 600 ms followed by a direct gaze face for 1500 ms. The face would then shift gaze to the left or right for 500 ms before the onset of the object. Objects were displayed for a maximum of 2500 ms while participants categorised the object as either a Kitchen or Garage item. The trial would end once participants had responded, and auditory feedback was provided before moving on to the next trial. The rating phase followed the same timeline as the categorising trials, but after participants had categorised each item, a blank screen was presented for 500 ms, followed by a rating screen which was presented asking participants to rate the item on a scale of 1-9 with one being the "do not like at all" and 9 "like very much". Subsequently, another blank screen was presented for 500 ms before the beginning of the next trial. After completing all trials, participants were asked if any features affected their ratings and gave a score between 0-100 on how they felt the eyes influenced their ratings.

Data analysis

These data were analysed across three measures: Response Accuracy, Median RT and Object ratings. Response Accuracy was calculated using the proportion of trials with correct to incorrect responses. The Median RT was calculated based on the trials with accurate responses, removing the impact of erroneous trials. To analyse the data, we used a series of 2x2x2x2 mixed ANOVA with Face type, and AQ group entered as between-subject factors and Congruency and Number of Faces entered as within-subject factors. Post hoc analysis of simple main effects with Bonferroni corrections was used to explore any significant interactions. All data were checked to ensure that they met the assumption of sphericity using Mauchly's Test of Sphericity. A Greenhouse-Geisser correction was applied if sphericity had been violated.

Assumptions of homogeneity of variances and normal distribution were checked using Levene's test and the Shapiro-Wilk test, respectively. The results of the Shapiro-Wilk test suggested that the residuals were not normally distributed. Visual inspection of the data suggested that the residuals did appear to be normally distributed (Histograms can be found in Appendix H. Failure of the test is likely due to the large sample sizes. Based on these observations, it was decided to continue with the ANOVA.

Outlier Exclusion

In line with the previous experiments, it was decided that participants must achieve a minimum accuracy score of 75% to ensure engagement with the experiment. However, there was no other removal of extreme values to ensure that trimming of data from one dependent variable did not unintentionally impact other dependent variables. As before median RTs were compared to account for extreme values without having to remove data entirely.

Results

Accuracy

Analysis of the proportion of correct responses found that there was a significant main effect of Face Number, with more correct responses for the Multiple face condition (M = .951, SD = .055) when compared to the Single face condition (M = .959, SD = .051 ;F(1,120) = 8.156, p = .005, η_p^2 = .064). However, there was no significant main effect of Congruency (Congruent, M = .955, SD = .054; Incongruent, M = .955, SD = .052; $F < 1$). There was also no significant main effect of Face Type on Accuracy (Photographic, M = .953, SD = .055; Incongruent, M = .957, SD = .050; $F < 1$). Analysis of the mean Response Accuracy between AQ groups was also non-significant (High, M = .954, SD = .055; Low, M = .957, SD = .052; $F < 1$).

Analysis of the interactions between factors found that there were interactions between three factors: AQ group, Face Type, and Congruency; F(1,120) = 4.224, p = .042, η_p^2 = .034), driven by significant differences between Congruent and Incongruent gaze cues in the High AQ-group, with the Schematic face only (p = .035). Finally, it was found that there was a significant interaction for the four way interaction (F(1,120) = 4.402, p = .038, η_p^2 = .035), and that significant differences between Congruent and Incongruent trials were found in the Photo and Schematic conditions, in the High AQ group, only when Multiple face interactions are made (High/Photo/Multiple: p = .019; High/Schematic/Multiple: p = .037). A summary of the data can be found in Table 4.1.

Table 4.1: *Mean and Standard Deviation for Accuracy for the High (N = 65) and Low AQ (N = 59) groups. The scores are presented by both Face Number and Face Type, as well as aggregate scores.*

			Photo		Schematic		Total	
			M	*SD*	*M*	*SD*	*M*	*SD*
High	Single	Congruent	0.953	0.064	0.949	0.057	0.951	0.060
		Incongruent	0.953	0.052	0.940	0.057	0.947	0.054
	Multiple	Congruent	0.954	0.052	0.964	0.050	0.959	0.051
		Incongruent	0.972	0.041	0.966	0.053	0.969	0.047
	Total	Congruent	0.953	0.058	0.957	0.054	0.955	0.056
		Incongruent	0.963	0.046	0.953	0.055	0.958	0.051
Low	Single	Congruent	0.941	0.061	0.969	0.034	0.955	0.047
		Incongruent	0.946	0.061	0.966	0.966	0.956	0.051
	Multiple	Congruent	0.954	0.052	0.957	0.044	0.956	0.048
		Incongruent	0.956	0.045	0.970	0.037	0.963	0.041
	Total	Congruent	0.948	0.057	0.963	0.039	0.955	0.048
		Incongruent	0.951	0.053	0.968	0.502	0.959	0.277

Median RT

Analysis of the median RTs was conducted with a mixed effects ANOVA. It was found that there was a signifcant main effect of RT, with Congruently cued items (M = 808.975 ms, SD = 124.041) being responded to faster than Incongruently cued items (M = 821.852 ms, SD = 127.121; F(1,120) = 12.592, p = .001, η_p^2 <.095). It was found that there was no significant difference between the AQ groups (High, M = 803.731 ms, SD = 118.199; Low, M = 827.183 ms, SD = 133.077; F < 1). There was also no significant main effect of Face Type (Photo, M = 831.725 ms, SD = 142.241; Low, M = 800.081 ms, SD = 102.187; F < 1), Face type (F(1,120) = 2.049, p = .155, η_p^2 <.017). There was also no significant difference between the Face Number conditions (Single, M = 814.091 ms, SD = 124.009; Multiple, M = 816.736 ms, SD = 127.712; F < 1). Analysis of the interactions found that there were no significant interactions between any of the factors (F < 1 for all interactions). A summary of the data can be found in Table 4.2.

Table 4.2: *Mean and Standard Deviation for Median RT for the High (N = 65) and Low (N = 59) AQ groups. The scores are presented by both Face Number and Face Type, as well as aggregate scores.*

			Photo		Schematic		Total	
			M	*SD*	*M*	*SD*	*M*	*SD*
High	Single	Congruent	806.652	104.048	791.176	112.600	798.914	108.324
		Incongruent	818.295	122.614	798.602	127.546	808.449	125.080
	Multiple	Congruent	818.518	121.633	782.068	108.726	800.293	115.179
		Incongruent	833.190	127.025	822.116	88.426	827.653	107.726
	Total	Congruent	812.585	112.841	786.622	110.663	799.604	111.752
		Incongruent	825.742	124.820	810.359	107.986	818.051	116.403
Low	Single	Congruent	840.018	167.108	796.368	82.765	818.193	124.937
		Incongruent	843.419	156.847	822.116	88.426	832.768	122.637
	Multiple	Congruent	818.518	121.633	796.772	84.365	807.645	102.999
		Incongruent	849.334	162.732	818.336	84.188	833.835	123.460
	Total	Congruent	829.268	144.370	796.570	83.565	812.919	113.968
		Incongruent	846.377	159.790	820.226	86.307	833.301	123.048

Object Rating

Analysis of the Object ratings found that there was no significant difference between the Congruent (M = 5.177, SD = 1.303) and Incongruent (M = 5.181, SD = 1.217) Gaze conditions ($F < 1$). There was also no significant difference between the AQ groups (High, M = 5.078, SD = 1.261; Low, M = 5.289, SD = 1.259; $F(1,120) = 1.107$, $p = .295$, $\eta_p^2 < .009$). Analysis of the Face Types found that there was no significant difference between the Photographic (M = 5.104, SD = 1.268) or Schematic faces (M = 5.255, SD = 1.258; $F < 1$). Finally, there was also no significant main effect of Face Number (Single, M = 5.174, SD = 1.303; M = 5.182, SD = 1.216; $F < 1$. There were no significant interactions between any of the factors ($F < 1$ for all interactions). A summary of the data can be found in Table 4.3.

Table 4.3: *Mean and Standard Deviation for Object Rating from the High (N = 65) and Low (N = 59) AQ groups. The scores are presented by both Face Number and Face Type, as well as aggregate scores.*

			Photo		Schematic		Total	
			M	*SD*	*M*	*SD*	*M*	*SD*
High	Single	Congruent	4.987	1.406	5.149	1.366	5.068	1.386
		Incongruent	4.964	1.154	5.162	1.402	5.063	1.278
	Multiple	Congruent	4.872	1.158	5.209	1.240	5.040	1.199
		Incongruent	5.058	1.242	5.351	1.188	5.204	1.215
	Total	Congruent	4.930	1.282	5.179	1.303	5.054	1.293
		Incongruent	5.011	1.198	5.256	1.295	5.134	1.246
Low	Single	Congruent	5.261	1.312	5.289	1.351	5.275	1.332
		Incongruent	5.257	1.282	5.351	1.187	5.304	1.235
	Multiple	Congruent	4.872	1.158	5.426	1.317	5.149	1.238
		Incongruent	5.191	1.279	5.323	1.081	5.257	1.180
	Total	Congruent	5.066	1.235	5.357	1.334	5.212	1.285
		Incongruent	5.224	1.280	5.337	1.134	5.281	1.207

Discussion

In this experiment, a liking effect paradigm was used to investigate differences between individuals with measured high and low levels of autistic traits. Regarding the visuospatial responses, there were some significant differences. For accuracy, there was a significant difference between congruently and incongruently cued items for the high AQ group when multiple identities were used. However, interestingly, the gaze direction of the gaze cueing bias was inverted for the photographic group. There was a significant main effect of congruency for the median RT, demonstrating the expected gaze cueing bias. For the measures of affective judgments, it was found that there were no significant differences in object ratings between congruent and incongruently cued objects. These results, although not in line with our previous finding, may be indicative of issues with the experimental design, which, as prior experience would suggest, may impact the reliability of the effect more than had been expected.

Relationship Between Experimental Design and Effect Size Estimates for Power Analysis

Although several experimental reasons may have impacted our results, it should be noted that the experimental parameters matched those of the last study, where replication of the effect was found, with one exception: the number of trials. Power calculations were based on the effect sizes from the previously successful study. These effect sizes should give an adequate estimate of the power required for the experiments. Power analysis calculations focus on the number of observations required to observe the effect. However, for research that utilises an average of scores across multiple observations, these calculations do not account for the number of trials. That is, effect size calculations are inherently affected by the number of trials (Brand & Bradley, 2016). Therefore, by reducing the number of trials, while overlooking how this may impact the observed effect sizes, the power analysis underestimated the power of the proposed sample. Furthermore, the analysis conducted here was more complex (two within-subject, two between-subject factors) than the previous study (two within-subject, one between-subject). Although these differences were inputted into the power analysis calculations, there was no accounting for how the change in design complexity would impact the results of the estimated effect sizes. By altering variables that would ultimately impact the observed effect sizes, the power level would have fallen below the suggested threshold, plausibly explaining the non-significant findings.

These findings outline the inherently problematic nature of estimating effect sizes for power analysis, particularly for paradigms such as this, where effect sizes are already small — making sample size estimates difficult, even when it is possible to collect large sample sizes. Even using online data collection methods can be restrictive, particularly in regard to the cost involved in collecting large data sets. Impacts of the effect sizes and methods for overcoming these issues will be considered in the main discussion of this thesis.

Influence of Face Identity and realism on Accuracy in Individuals with High AQ Scores

Even though the experiment in this chapter found null results for the measures of affective judgments, there was one interesting finding that stands out. The response accuracy data, a measure that participants consistently perform at ceiling in previous chapters, demonstrated some interesting interactions in individuals in the high group. These findings would suggest that there was an element of the multiple identities that impacted participants' responses. Visual inspection of the data shows that these differences appear to be in opposite directions, dependent on the face type. That is, in the photo condition, the Incongruent items are categorised more accurately, whereas, in the schematic condition, it is the congruent images that are categorised more accurately. It is possible these differences outline flaws in how autistic individuals integrate facial information. That is, if those with ASD have difficulties in integrating the facial information, it is possible that processing of highly detailed face stimuli may require more processing resources before a gaze cueing effect can occur, possibly inhibiting the response. However, as the demands are not so high in the low detailed faces, there is not the same impact. Due to the issues with the experimental design discussed above; it is not possible to draw full conclusions about these differences in accuracy measures. However, it outlines the future potential for investigating such differences in more detail with an appropriate sample.

Conclusion

Overall this study was unable to elicit the main liking effects that have previously been found. Although there may be other potential reasons for this, the overly complicated design which resulted in an underpowered sample makes it difficult to comment on how eye gaze has influenced individuals with high ASD traits, and subsequently the impact on the liking effect. However, some differences between the High and Low AQ groups found in the visuospatial measures suggest that there may be potential for utilising this paradigm in future with more appropriate samples sizes and more robust design.

Chapter 5: Eye Tracking Study of the Liking Effect in Neurotypical and Autistic Populations

Experiment 9

In Chapter 3, when multiple face identities cued an item, a liking effect was elicited consistent with Capozzi et al. (2015). In Chapter 4, the focus was on investigating differences across individuals with measured high and low levels of autistic traits in a large sample study, but, in Chapter 4, the liking effect was not replicated. The null findings could be attributed to a combination of an underpowered sample size for the more complex statistical design and reduced number of trials. In this chapter, a lab-based eye-tracking experiment with autistic individuals was conducted in tandem with Chapter 4 will be discussed.

Eye-tracking is utilised to measure where a participant is attending during a computerised task. Most conventional eye trackers produce a direct source of infrared light that is aimed at the participants' eyes. Infrared light is reflected off the cornea which presents as a 'glint' that is detected by infrared cameras. The location of the glint can then be calibrated to measure where a participant is looking on the screen. The location is typically recorded in the form of coordinates from the screen, indicating where a participant is looking in relation to timed events such as the onset of stimuli (Singh & Singh, 2012). By conducting this experiment in a lab-based environment alongside the eye tracker, it was possible to measure attention in ways that were not possible with the behavioural paradigm alone. There are two elements of this paradigm that were of interest: Firstly, how the individuals interacted with the face itself. Secondly, how individuals would utilise the eye gaze to attend to objects and how this may relate to preferential rating.

Some of the earliest methods of eye tracking by Yarbus (1967), demonstrated a preference for viewing faces within complex scenes (for a recent summary see Tatler, Wade, & Findlay, 2010). For an example of the initial results, see Figure 5.1.

Eye-tracking has also been used to distinguish how people view faces, showing a tendency in neurotypical individuals to direct and sustain gaze towards the eye region (Dalton et al., 2005; Klin, Jones, Schultz, Volkmar, & Cohen, 2002). In the current paradigm, participants were presented faces with both direct and averted eye gaze, providing the opportunity to investigate how people attend to eyes within a face in both NT and ASD populations. Is it the case that autistic individuals spend less time attending to the eyes? Does the time spent attending to the eyes impact either the visuospatial or the functional consequences of eye gaze?

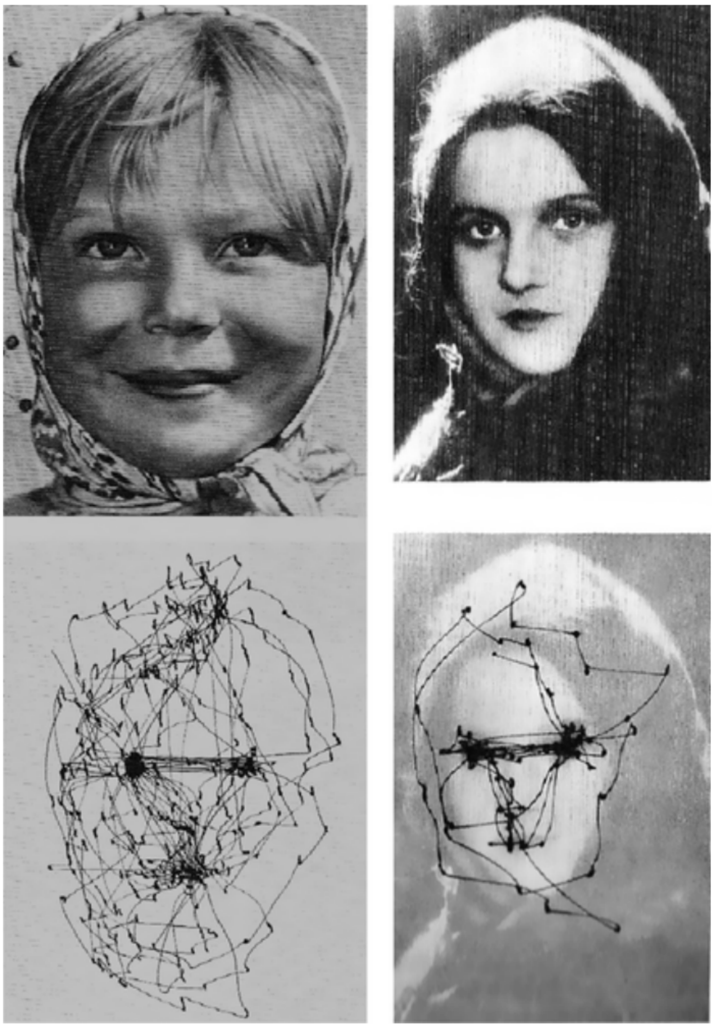

Figure 5.1: Original recordings of gaze fixation from Yarbus 1967 (Figure adapted from Tatler et al., 2010).

Eye-tracking has been used to distinguish differences in how autistic people attend to faces (Dalton et al., 2005; Kliemann et al., 2012, 2010). It has been noted that autistic individuals tend to spend less time attending to eyes (Dalton et al., 2005; Klin et al., 2002) and show spikes in amygdala activity when they engage with direct eye contact (Kliemann et al., 2012, 2010). One explanation is that autistic individuals show reduced attraction to direct eye contact (J. W. Tanaka & Sung, 2016), impeding effective gaze cueing. However, if instructed that the gaze cues are predictive, and therefore should be followed, Autistic participants would demonstrate a gaze cueing bias (Ristic et al., 2005). If these difficulties in making eye contact limit the impact of eye contact in ASD is it possible that the extra processing required for the liking effect may also be impacted, even if there is a gaze cueing bias for visual-spatial attention.

The use of eye-tracking has also been important in investigating how overt and covert attention shifts are important for a gaze cue bias. Mansfield, Farroni, and Johnson (2003), built on findings by Driver et al. (1999) and Friesen and Kingstone (1998), that eye gaze would cue visuospatial orienting. They found that, even though participants did make cue-driven saccades, these alone did not account for the gaze cueing bias (Mansfield et al., 2003); suggesting that although participants may make an overt attentional shift, the effects of the gaze cueing bias for visuospatial attention is driven by covert attention shifts. Although the goal of this experiment is not to investigate the explicit differences of covert versus overt orienting in regards to visuospatial measures, it may be able to indicate how overt orienting might influence the functional consequences of eye gaze, specifically, affective judgments.

Regarding the liking effect, there is currently no research utilising the current paradigm alongside eye tracking. However, other paradigms have been used to investigate how attention to an object can predict preferences towards an object (Schotter, Berry, McKenzie, & Rayner, 2010). It was demonstrated that people spend more time attending to an object that would subsequently be rated positively compared to objects rated negatively. Schotter et al. (2010) argued that it was selective encoding that drove the gaze bias which boosted the liking effect. In addition, similar paradigms have investigated how gaze leading can alter preferential ratings and viewing (Bayliss et al., 2012). In this gaze leading paradigm, the eye direction of a central face either shifted congruently or incongruently based on the gaze direction of a participant. It was found that participants demonstrated more consistent preferences for items in the congruent condition even though the gaze cue occurs after the initial saccade from the participant (Bayliss et al., 2012). The confirmatory gaze from the face may facilitate the selective encoding that was proposed by Schotter et al. (2010). Furthermore, it has been found that encoding as a result of an overt attention shift resulted in interference in memory tasks, a finding that was not replicated in covert orienting (Tas, Luck, & Hollingworth, 2016). As previously discussed, there appears to be some evidence that memory for cued objects facilitates the affective judgments (Bry et al., 2011). If it is the case that overt orienting is related to the selective encoding that facilitates both memory and the liking effect, it should be possible to detect these differences by looking at the dwell time on congruently versus incongruently cued items in the current liking effect paradigm.

Data reported in the current chapter were collected concurrently with the experiment in Chapter 4. Both individuals with a previous diagnosis of ASD and a NT control group were recruited. In contrast to the previous Chapters (exploring individual differences in population samples, in Chapters 2 and 3, sampling individuals with high AQ symptom scores in Chapter 4) the inclusion of individuals with an ASD diagnosis allows more direct conclusions about differences in ASD processing to be drawn, rather than having to extract this from measures of the broader autism phenotype. Therefore, in combination with Chapter 4, it would be possible to draw better conclusions about face and gaze processing in ASD. Based on the literature (Schotter et al., 2010; Bayliss et al., 2012; Dalton et al., 2005; Bayliss et al., 2006) as well as the prior results, several hypotheses were proposed:

> *Hypothesis 1*: In both the NT and ASD participants attention would be cued by the eye gaze of face stimuli which would result in a gaze cueing bias for visuospatial measures of attention (RT and Accuracy).
>
> *Hypothesis 2*: In the NT group, there would also be a gaze cueing bias for affective judgments as cued attention would facilitate selective encoding. However, in the ASD group, there would not be a gaze cueing bias for affective judgments as cued attention would not facilitate selective encoding.
>
> *Hypothesis 3*: The differences in affective judgment would also correspond to overt orienting measures. NT participants would dwell more on objects that are congruently cued, whereas there would be attenuated attention to target objects in the ASD group as they would be less likely to follow the gaze cue.
>
> *Hypothesis 4*: In the NT group, participants would attend more to the eyes of the face stimuli compared to the mouth. However, the aversive nature of eye

contact for autistic participants would result in less time attending to the eyes and more time attending to the mouth in the ASD group.

Methods

Participants

A total of 36 participants were recruited. Of these 36 participants, 7 participants (two female, five male) were recruited through the Oxford University Disability Advisory Service (”DAS”) and had received a previous diagnosis of ASD. The remaining 29 participants (16 female, 13 male) were recruited through the University participant recruitment system. Due to time constraints and difficulties with recruiting autistic participants at the time, it is noted that this experiment is significantly underpowered. Therefore, these findings are presented as a preliminary investigation and will require future replications to confirm – this will be discussed in more detail in the discussion of this chapter.

All participants were reimbursed £10 for their participation. The mean age of the NT group was 22 years (18-30 years) and for the ASD group it was22 years (18-28). All participants gave informed consent before the experiment commenced, they were made aware that they could withdraw at any time, and that all data collected would remain anonymous.

Ethical considerations

Informed consent was received by all participants, and they were made aware of their right to confidentiality and to withdraw at any time. The experiment was given ethical approval by the University of Oxford Central University Research Ethics Committee (CUREC; R57031/RE001).

Apparatus

Tasks were generated using Matlab (64-bit version R2015b; The MathWorks, Natick, MA), with images being displayed using Psychtoolbox (Kleiner et al., 2007; Pelli, 1997; Brainard, 1997). Responses were collected via keyboard input, and stimuli were presented on a 24-inch monitor with 1920 x 1080 pixel resolution. Eye-tracking data were collected using the Eye Tribe tracker (The Eye Tribe; Copenhagen, Denmark) collecting data at 60 Hz (binocular). The Eye Tribe eye tracker is a portable and low-cost eye tracker. Previous comparisons have found comparable results for the analysis of fixation data to other more expensive but gold-standard setups, for example, the Eye-link eye-tracker (Dalmaijer, 2014). Participants were positioned at the appropriate distance from the monitor using the default Eye Tribe software. The Eye Tribe was connected to Matlab using the Eye Tribe for Matlab toolbox (see https://github.com/esdalmaijer/EyeTribe-Toolbox-for-Matlab; Dalmaijer, 2014), and calibrations were performed using the default toolbox, 9-point calibration. Ten participants in the NT group were unable to provide adequate calibrations and were removed from the ET analysis, resulting in a total of 19 (16 female) in the NT group for the eye-tracking analysis.

Stimuli

Target stimuli were the same 24 items utilised in Chapter 4. Of these items, 12 items were typically found in the garage and 12 items typically found in the kitchen. These items were colourised to be either red, green, blue or yellow for a total of 96 items. Images were presented in a box measuring 300 x 300 pixels (approximately 83 mm x 83 mm). Face stimuli consisted of the same eight photographs of individuals with a direct, left or right-oriented gaze. The faces were presented with a happy expression to optimise the liking effect based on previous findings (Bayliss et al., 2007). Face images measured 534 x 716 pixels (147 mm x 198 mm).

Regions of Interest

Regions of interest (ROI) were defined using a custom script to present stimuli for a researcher to draw a 4-point-box around the target region. Two specific ROIs were defined for presentation stimuli: The mouth; and the eye region (a combined region of the individual eyes leaving space over the bridge of the nose; see 5.2). From these data, the proportion of time spent looking at each ROI during the stimuli presentation was calculated.

Figure 5.2: Example of Regions of Interests for one Face stimulus. The blue box represents the Mouth region. The Eye region consisted of a combination of the two red boxes. Division of the eye region ROI was to ensure explicit focus on the eyes rather than the bridge of the nose.

Design

For the behavioural analysis, a paired samples design was utilised to compare Congruent and Incongruent conditions data across three measures: Response Accuracy, Median RT, and Object Ratings. For the eye-tracking analysis, conditions were compared across three scenes during the trial sequence: First, the initial Gaze Fixation scene presented the central face alone with gaze directed at the participant. Second, in the Gaze Shift scene, during which the central face displayed gaze directed to the left or right of fixation. Finally, in the Object Onset scene, the central face, with gaze, averted, remained on scene with the target object. Across all scenes, a paired samples comparison was made between the proportional duration on the Eye ROI versus the Mouth ROI. In the Object onset scene, a comparison was also made between the proportional duration on the Congruent item ROIs versus the Incongruent item ROIs, resulting in a total of four comparisons made using the eye-tracking data.

Procedure

The current experiment followed the same procedure as in the previous chapter, with a slightly reduced number of trials (to see an example of the timeline, please refer to Figure 3.2 in Chapter 3). In blocks one to four, participants were advised that the gaze cue was uninformative and to respond by categorising the item as quickly as possible. After participants had read and confirmed consent, they were asked to read the instructions. Participants completed the eye-tracking calibration before starting the trials. They were then presented a fixation cross for 600 ms followed by a face with a direct gaze for 1500 ms. They were then presented with a face with averted gaze for 500 ms before an object was displayed. The object would stay on the screen until a response was made or until 2500 ms had passed. The screen went blank for 1500 ms before the beginning of the next trial signified by the fixation cross. Participants used the H key and the space bar to respond to the category of the target object. The key which corresponded to Garage or Kitchen objects was counterbalanced between participants. In the final block, the trials followed the same procedure, but after the participant had categorised the object, a blank screen was displayed for 500 ms, followed by the rating screen. Participants were asked to respond aloud to an experimenter facing away from the screen who entered the result using a second keyboard. Once the response was recorded, a blank screen was presented for 500 ms before the onset of the next trial.

Data analysis

For the Behavioural data, paired-samples t-tests were used to compare the Congruent versus Incongruent conditions for the three measures (Accuracy, Median RT, and Object Rating). As in previous experiments, the Accuracy score was calculated as the proportion of correct responses to trials and the RT was calculated from the median of accurate trials. Again, for the eye-tracking data, we used paired samples t-tests to compare the ROI pairings for each scene. T-tests were run independently for the NT and ASD groups, and no direct comparisons were made between the two groups due to the disproportionate sizes of the two samples.

Normal distribution of the samples was tested using the Kolmogorov-Smirnov test. However, due to the small sample, particularly for the ASD group, it may be possible that the test is not sensitive enough to detect normality. As a precaution, non-parametric equivalents (i.e., the Wilcoxon signed-rank tests; Wilcoxon, 1946) were run. There was no divergence between the parametric and non-parametric tests, and as such the parametric tests are reported here.

Outlier Exclusion

As in the previous chapters, there was no removal of extreme values from the data to limit unintended impacts on other measured variables. Therefore, as RT may be sensitive to extreme values, the median was used to analyse reaction times to limit the potential impact of extreme values.

Results

The mean and *SD* of the behavioural data for the NT group ($N = 29$) and the ASD group ($N = 7$) can be found in Appendix J.

Behavioural

Hypothesis 1: Effect of gaze cues on visuospatial attention

For the NT group there was no significant difference for Accuracy between Congruently (M = .938, SD = .042) and Incongruently (M = .924, SD = .061) cued items ($t(28) = 1.9137$, p = .066, d = .365), although there did appear to be a trend towards significance. In the ASD group there was again no significant differences between the Congruent (M = .968, SD = .035) and Incongruent (M = .950, SD = .026) conditions ($t(6) = 0.966$, p = .371, d = .365). A summary of the accuracy data can be found in Table 5.1.

Table 5.1: *Mean and Standard Deviation for Accuracy for the NT (N = 29) and the ASD groups (N = 7).*

	NT Group		ASD Group	
	M	*SD*	*M*	*SD*
Congruent	0.939	0.043	0.969	0.035
Incongruent	0.925	0.062	0.951	0.027

The Median RT was significantly different between the Congruent and Incongruent conditions for both the NT (Congruent M = 1313.299 ms, SD = 92.589; Incongruent M = 1338.843 ms, SD = 98.549; $t(28) = 3.494$, p = .002, d = -0.661) and the ASD groups (Congruent, M = 1330.120 ms, SD = 116.746; Incongruent, M = 1369.793 ms, SD = 134.258;$t(6)$ = -3.052, p = .022, d = -1.153). A summary of the data can be found in Table 5.2.

Table 5.2: *Mean and Standard Deviation for Median RT (ms) for the NT (N = 29) and the ASD groups (N = 7).*

	NT Group		ASD Group	
	M	*SD*	*M*	*SD*
Congruent	1313.300	92.589	1330.120	116.747
Incongruent	1338.843	98.549	1369.793	134.259

Hypothesis 2: Effect of gaze cues on affective judgments

Analysis of participants' Object Ratings found that in the NT group there was a significant preference for Congruently cued items (M = 5.011, SD = 0.840) over Incongruently cued items (M = 4.802, SD = 0.846; $t(28) = 2.746$, $p = .011$, d = 0.519; See Figure 5.3)). However, in the ASD group there was no significant difference between the two conditions (Congruent: M = 4.857, SD = 0.645; Incongruent: M = 5.047, SD = 0.851; $t(6) = -0.838$, $p = .434$, $d = -0.316$). A summary of the data can be found in Table 5.3.

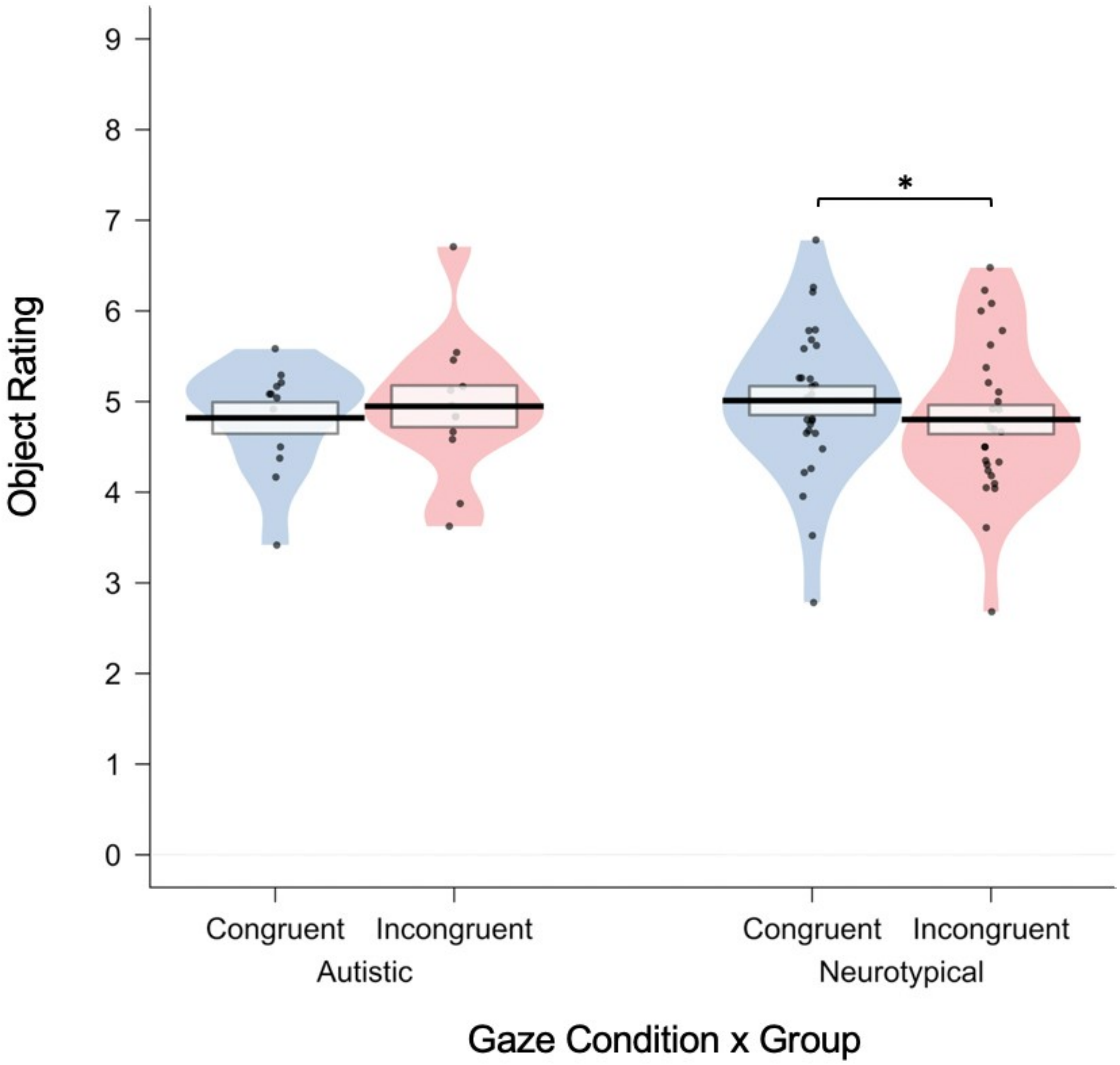

Figure 5.3: Participants' mean Rating score by the Gaze condition and Autism group. The error boxes represent the *SD*. It was found that there was only a significant difference between Congruent and Incongruent conditions among the Neurotypical group. $p < .05$ indicated by *.

Table 5.3: *Mean and Standard Deviation for Object Rating for the NT (N = 29) and the ASD groups (N = 7).*

	NT Group		ASD Group	
	M	*SD*	*M*	*SD*
Congruent	5.011	0.840	4.857	0.645
Incongruent	4.802	0.847	5.048	0.852

Eye tracking

Hypothesis 3 - Relationship between overt orienting gaze cues and overt orienting to target objects

As well as comparisons between ROIs of the face, ROIs between Congruently and Incongruently cued items object were compared. It was found that the NT group spent more time looking at Congruent (M = 0.0108 ms, SD = 0.0121) items than Incongruent items (M = 0.0003 ms, SD = 0.0006; $t(18)$ = -3.742, p = .001, d = -0.858). In the ASD there was no observable difference between either Congruent (M = .007 ms, SD = .013) or Incongruently (M = 0.0023 ms, SD = 0.0044) cued items ($t(6)$ = -1.309, p = .238, d = -0.494).

Hypothesis 4 - Comparison between attention to Eye and Mouth regions

During the direct Gaze Fixation stage, it was found that more samples identified participants' gaze as located on the Eye ROI (M = 0.178 ms, SD = 0.094) versus the Mouth ROI (M = 0.041 ms, SD = 0.037; $t(18)$ = 5.346, $p < .001$, d = 1.226) in the NT group. There was also a similar pattern in the ASD group (Eye ROI: M = 0.113 ms, SD = 0.052; Mouth ROI: M = 0.059, SD = 0.035; $t(6)$ = 2.944, p = .025, d = 1.112). In the Gaze Shift scene there was again a significant difference in the NT group (Eye ROI: M = 0.251 ms, SD = 0.13; Mouth ROI: M = 0.035 ms, SD = 0.167; $t(18)$ = 6.4, $p < .001$, d = 1.468). In the ASD group however, there was only a borderline difference between the Eye ROI (M = 0.147 ms, SD = 0.085) than the Mouth ROI (M = 0.062 ms, SD = 0.045; $t(6)$ = 2.42 p = .051, d = 0.914). In the Object Onset phase there was again a significant difference in the NT group with more time spent looking at the Eye ROI (M = 0.063 ms, SD = 0.053) than the Mouth (M = 0.009 ms, SD = 0.011; $t(18)$ = 4.586, $p < .001$, d = 1.052). However, this pattern was not observed in the ASD group with no significant difference found between the ROIs (Eye ROI: M = 0.037 ms, SD = 0.027; Mouth ROI: M = 0.032 ms, SD = 0.029; $t(6)$ = 0.0586, p = .578, d = 0.221).

Discussion

In this experiment, a gaze cueing bias was seen for both NT and ASD participants for measures of visuospatial attention. However, regarding the functional measures, it was found that there was a gaze cueing bias for affective judgments of objects in only the NT group and not the ASD group. These differences between the two groups suggest that even though attention is cued by eye gaze for autistic participants, they are not utilising the gaze in the same way to influence their judgments of objects. Furthermore, by comparing the outcome of this experiment to the findings in Experiment 6, these findings would lend support for the requirement of multiple identities cueing the object to elicit the effect.

The Results in the Context of Autistic Populations

The findings in the Autistic group, although very interesting, must be taken with a grain of salt. Potential interpretations of these findings are discussed, but it must be noted that the small sample was collected as a preliminary investigation into this effect in autistic people. As such, any conclusions discussed here based on these findings are speculative and will require replication with larger samples to confirm these differences. The main finding in the ASD group is that no liking effect was found, suggesting that autistic participants were not making interpretations about the object that were cued, unlike the NT participants who showed a preference for cued objects.

As discussed earlier, if it is the case that the liking effect is driven by selective encoding (Schotter et al., 2010), it may be the case that individuals with ASD are not encoding the cued objects. However, the eye-tracking results may suggest an alternative explanation. In the comparison between the proportion of dwell time on congruent and incongruent items, it was found that more time was spent attending to the congruently cued items and most NT participants spent no time attending to the incongruent item. It has been found that overt orienting may facilitate other functional measures of attention (for example, memory; Tas et al., 2016), so it is plausible that these results may demonstrate facilitation of selective encoding. However, in the ASD group, the proportional time looking towards the incongruent item was higher and equal to that of the congruently cued items. Therefore, if the eye-tracking measures are associated with selective encoding, it would be plausible that autistic participants were encoding both congruently and incongruently cued items equally, which would eliminate the liking effect.

Another interesting finding based on the eye-tracking results was the differences between how individuals attended to eyes versus the mouth of the face in the task. In the direct gaze condition, both NT and ASD groups attended to the eyes more than the mouth, although this was attenuated in the ASD group. The focus towards the eyes in the ASD group was an unexpected finding. However, visual inspection of the eye-tracking data shows a focus towards the bridge of the nose (see figure 5.4). It has been suggested that attending to this space is often used as a coping strategy for many individuals with ASD who have difficulties in making direct eye contact (Trevisan et al., 2017). Although in principle this space was excluded from the eye region of interest, it is likely that this was unsuccessful, which makes it hard to comment on possible differences between the current and previous findings.

As the eyes shift direction, the participants' preference for the eyes disappears in the ASD group, but not in the NT group, which may suggest that a reaction to shifting gaze results in an overt rather than covert shift of attention; possibly a consequence of the more voluntary response to attentional cues in ASD compared to the reflexive response in NT individuals. As the attention has then been averted from the face, NT participants may be inclined to spend more time exploring the face resulting in equal time spent attending to the eyes compared to the mouth.

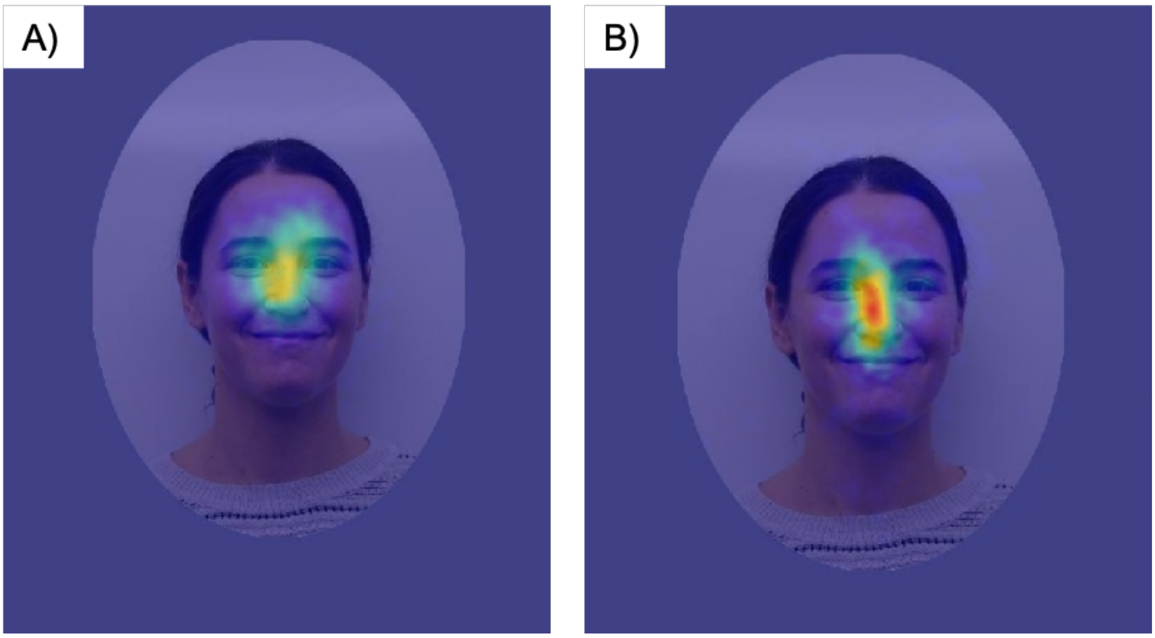

Figure 5.4: A simple representation of the average heat map of gaze fixation points for NT participants (A) and ASD participants (B).

Limitations and analysis of the methodology

Limited sample size

A notable limitation of this study is the limited sample size of the ASD group. Due to time constraints on the recruitment of individuals with a prior diagnosis, it was only possible to recruit a limited sample. In addition, there is also the question of the generalisability of the findings to all autistic individuals, because the participants in the current sample were all students at university. However, it could be argued that this may help to ensure a relative match with the control group to ensure a level of comparability to the control group, in terms of age and approximate intelligence. These data may give a good indication of how individuals with ASD interact with the liking effect, but caution must be taken when drawing conclusions from these data, in particular when making direct comparisons between the two groups.

Another interesting observation is the potential differences in the effect sizes of this experiment compared to our previous findings. In particular, it may raise questions about why we were unable to elicit the effect in Chapter 4 when the same, experiment was used, compared to this lab-based study. One possible explanation is that there is some level of demand characteristics that impact the effect. For example, the presence of an experimenter and direct reporting of likeability ratings to the experimenter may drive the participants to change their behaviour by rating gazed-towards items higher. However, if this were the case, there would have been no effect seen in the Experiments in Chapter 3. Alternatively, it is proposed that, although the presence of an experimenter may not have influenced the participants' ratings per se, it may have influenced how engaged participants were with the experiment, which in turn may have increased the encoding of the congruently cued items. In the online versions of the experiment, where engagement may be lower without the presence of an experimenter or in the presence of more external distractors, it would make sense that encoding of objects may have been impeded to some extent. The current findings do not inherently discredit the findings of the online experiments, but they do raise the question about whether future online experiments should account for differences in engagement with the target objects. In the meantime, until such engagement issues can be studied in greater detail, it would be advised to maintain the practice of collecting both online and lab-based data to confirm any findings.

Effectiveness of the Eye Tribe eye tracker

An ancillary objective of this study was to investigate the effectiveness of the Eye Tribe tracker in a research setting. The Eye Tribe is very compact and therefore portable, as well as being significantly cheaper (approximately $100 at the time of purchase) than other available eye trackers. The portability of the device means that it can be easily transported to participant's homes. For researchers who collect data from populations including ASD, who may prefer not to travel to an unfamiliar lab-based setting, portability would be particularly desirable — allowing data to be collected in the participants' homes for example.

As in previous validations of the tracker, we were able to get accurate tracking data suitable for fixation analysis. However, there were notable calibration difficulties with wearers of glasses and contact lens. Although fixation analysis was adequate for this study, saccade analysis would also have been a useful metric that was not possible with this device. In conclusion, this device is adequate for the data analysis conducted here and would be a useful tool for collecting data when portability of a device is paramount. However, issues with calibration as well as the relatively low frequency may suggest that it is not the best eye tracker to use in a lab setting if alternatives are available.

Wider impact of the findings

These results demonstrate the reproducibility of the liking effect, contributing to the literature further support for these findings. It also contributes to our understanding of differences in facial processing between NT and ASD participants. Specifically, although stated with caution, it suggests that autistic individuals do not encode information about the target of the gaze cue in the same way as NT individuals do, even if they do benefit from gaze cues. However, there appeared to be no conclusive differences between how autistic participants viewed the eyes compared to the mouth.

In the later stages of trials, the difference in gaze patterns, however, became more apparent. Once the gaze shifted, NT participants continued to focus on the eyes, whereas the ASD group spent an equal amount of time attending to the mouth and the eyes. Despite the fact the eyes are presenting dynamic signals, the ASD group was more focused on investigating other aspects of the face equally. Even though the ASD participants were primed to attend to the cued direction of the object onset, they were not engaging with targets of social gaze in the same way as NT participants, at least in a way that was conducive of the liking effect.

The eye versus mouth ROI data was also complemented by our findings that individuals with ASD spent an equal amount of time looking at congruent versus incongruently cued items. While the NT group showed significantly less time looking at incongruently cued items, the ASD group interacted with the congruently and incongruently cued objects equally. This raises the question of whether it was indeed the case that direct eye contact has an additive effect on object engagement (and ratings), or whether incongruent eye contact results in NT participants disregarding incongruently cued items, resulting in a reduced affective judgment score. In contrast, individuals in the ASD group who did not show a reduced time spent attending to the incongruently cued item rated them equally.

It has been proposed here that the measures of overt orienting towards the target are suggestive of selective encoding. However, it is not possible from this experiment to conclude whether the overt orienting is a result of, or the cause of, any selective encoding. Although participants were instructed in the original Bayliss et al. (2006) experiment not to make overt eye movements during the original experiment; there were no explicit controls for this. Also, this was not stipulated in the experiment outlined here, making it difficult to draw any conclusions. Therefore, future experiments that focus on disentangling overt movements from selective encoding would be beneficial for understanding the mechanisms of this effect. One potential method for investigating overt attentional shifts in future is to investigate the null findings related to single face – item connections across trials. If individuals are less likely to overtly shift attention once a repetitive connection to an object has been made, a reduction in attention to objects should be visible across trials, which may outline the effect of face identity on the relationship between the facilitating effect of overt attention shifts related to gaze cues and the impact on affective judgments. Relationships between selective encoding and the liking effect, as well as the potential for future research, are discussed in more detail in the general discussion of this thesis.

Conclusion

In conclusion, it was demonstrated that the liking effect replicates under the current conditions. Furthermore, eye tracking analysis showed that the preferences for items in the NT group were driven by more time spent looking at the congruently cued items compared to the incongruently cued items. However, this effect is not observed in a sample of individuals diagnosed with ASD, suggesting that they may not be influenced by eye gaze in the same way, although caution is required based on the limited sample of the ASD group.

Chapter 6: General Discussion

Throughout this thesis, a series of experiments have been presented that aim to establish the potential for using the liking effect paradigm to investigate how different faces are processed across a spectrum of realism. It had previously been established that both photographic and schematic face stimuli would elicit gaze cueing bias for spatial properties. However, it was unclear whether this gaze cueing bias would translate into more advanced social processing that may have been required for the liking effect. Additionally, experiments were presented that aimed to investigate inconsistent findings of deficits shown by autistic individuals in gaze cueing experiments and how high detail faces (such as photographic) may impede face processing in autism in ways that low detail faces (such as cartoons) may not. Summaries of each chapter's specific findings are presented below.

Synopsis of Chapter 2

In the first chapter, three experiments were run, which aimed to investigate the impact of memory on the liking effect in photographic faces. However, methodological issues in Experiment 1 resulted in a lack of gaze cueing bias for visuospatial measures (RT and Response Accuracy), which also appeared to impact functional measures (Memory and Affective Judgments). It was thought that these null findings might have been the result of an inhibition of return effect on attention that may have impacted attentional cueing due to a long SOA. Therefore, to overcome the potential that any faciliatory effect was decaying, two different SOA conditions were introduced in Experiment 2. Again, a null effect was found across both spatial and functional effects of eye gaze. Finally, in Experiment 3, it was proposed that task demands may impact gaze cueing. In Experiments 1 and 2, the task allowed a correct response to be distinguished from either a congruent or incongruent cue, and there was no need to reorient in a fashion that would elicit a gaze cueing bias. By increasing the task demands in a way that resulted in reorienting to the target image, a gaze cueing bias was found for spatial measures. This attentional bias to congruently cued targets resulted in increased memory, replicating previous findings. However, gaze cueing bias for visuospatial measures did not result in a gaze cueing bias for affective judgments in Experiment 3.

It was concluded that task demands play a significant impact in the gaze cueing bias for spatial as well as the functional consequences. Regarding the liking effect, it was concluded that although it was not found here, substantial difficulties in eliciting gaze cueing bias and differences from the original paradigm made it impossible to conclude a specific reason why the effect was not found.

Synopsis of Chapter 3

In Chapter 2, it was concluded that it was necessary to replicate the original effect to establish the potential for the liking effect as a method for researching facial realism. It was decided to replicate the original paradigm with the additional extension of using simple schematic stimuli rather than arrow cues. In Experiment 4, a direct replication of Bayliss et al. (2006) was conducted. Although it was found that there was a gaze cueing effect for visuospatial measures, this did not translate to a gaze cueing bias for affective judgments for either Photographic or Schematic face stimuli. It was concluded that this might have been the result of a low sample size resulting in a type two error. In Experiment 5, the paradigm was translated into an online data collection platform, allowing the collection of substantially more extensive data sets. However, again it was found that while there was a gaze cueing bias for RT, there was no difference between affective judgments based on the gaze cue. Subsequently, an investigation of the literature explained why this might have been. Capozzi et al. (2015) found that objects had to be associated with congruent eye gaze from multiple face identities across trials for participants' affective judgments to be influenced by eye gaze - a level of counterbalancing overlooked in the previous experiments. In Experiment 6 data were collected with the updated paradigm comparing objects that were cued by Single and Multiple Photographic faces across trials and it was found that the effect only occurred in the condition where multiple identities cued the object. Finally, in Experiment 7, the paradigm was used to investigate the liking effect with Schematic faces. It was found that congruently cued items were rated higher than incongruently cued items, although this happened consistently across both single and multiple face identities. Therefore, it was concluded that the effect was reliant on multiple identity associations with the object in Photographic condition, which would conform with a social cohesion account proposed by Capozzi et al. (2015). However, there were still questions remaining due to the null interaction between congruency and identity/object interactions in the Schematic condition.

Synopsis of Chapter 4

Experiment 8 aimed further to disentangle the differences between the Schematic and Photographic faces. Also, to get a better understanding of the impact of ASD, participants were prescreened and separated into High and Low autistic trait groups. However, this resulted in a complex design that attempted to contrast four factors (Gaze Congruency, Face type, Number of Identity - Object Interactions, and AQ group). A large sample was collected using online data collection, in line with a priori power analysis. In this experiment, it was found that there was a gaze cueing bias for RT, although this was less consistent than in prior experiments. However, there was no significant gaze cueing bias for any of the conditions. It was concluded that by raising the complexity of the design and reducing the number of trials in the experiment, the selected effect sizes used to calculate the sample size were invalid. Furthermore, the question was raised that differences in engagement with the experiment between online and lab-based paradigms may impact this effect, as well as online research in general. Therefore, although there was a null finding for the liking effect, it was not appropriate to draw conclusions on the effect based on the results of this experiment.

Synopsis of Chapter 5

In Experiment 9, a lab-based eye-tracking experiment was conducted concurrently with the Experiment 8 (in Chapter 4). This experiment focused purely on HSF faces and collected data from both NT participants, and participants who had a previous diagnosis of ASD. Eye-tracking technology gave an explicit measure of visual attention to the face cue stimuli and the target stimuli. It was found in this experiment that we were able to replicate a gaze cueing bias for both RT measures as well as affective judgments in the NT participants, but only the RT measures in the ASD group. These results were also supported by our eye-tracking results that demonstrated the significantly longer time spent looking at the congruent compared to the incongruent object in the NT group, but not in the ASD group. There was also a reduced focus towards the eyes compared to the mouth in autistic participants. Based on these findings it was concluded that although autistic individuals were able to elicit eye gaze cues from photographic face stimuli for visuospatial processes, this attentional bias did not translate to functional differences in eye gaze, outlining the possibility that they are unable to elicit the social information required for the liking effect to occur. Although these findings must be taken with caution due to the meagre sample size in the autistic group, these findings would suggest that there may be value in studying this in a large sample of autistic participants.

Summary of Findings in Relation to the Research Aims

This thesis aimed to answer three research questions: 1) To what extent can the liking effect be used to measure differences in how we process faces of different levels of realism? 2) Can autistic individuals utilise eye gaze in the same manner as neurotypical participants resulting in functional consequences of eye gaze? 3) Can this paradigm be used to distinguish differences in face processing in Autistic individuals and NT across a spectrum of facial realism? In answer to the first question regarding the liking effect, it was found that there were different patterns of behaviours associated with low detail and high detail faces. In particular, the liking effect was only observed in photographic faces when there were multiple identity to object interactions, a condition that was not required with schematic faces. However, the Schematic faces were able to elicit some sort of liking effect. Potential explanations for this will be discussed below. Regarding the second question, it was found that there were no differences in how individuals used eye gaze cues for visuospatial cueing, either based on measures of the broader autism phenotype or in a sample of autistic adults. For the liking effect, there was no significant impact of a gaze cue bias on affective judgments in our sample of autistic adults. However, our large-scale experiments with individuals who exhibited high and low autistic traits were inconclusive as a result of issues with the sampling technique. Finally, regarding the third research aim, it was not possible to extract conclusions of differences between the autistic and NT adults. The experiment optimised to answer precisely these questions suffered from design flaws (Chapter 4). Ideally, the research of Chapter 5 will be extended in future to collect a larger sample across both high and low detail faces to attempt to answer this question.

Impacts on Wider Research Area

Although it was not possible to provide answers for all the main research questions, our results have provided insight that will impact the existing literature. These impacts can be broken down into two broad areas: Firstly, the investigation of facial realism on the affective consequences of eye gaze in NT and ASD populations; Secondly, new lessons on the impact on the value, but also the limitations of online testing.

Facial Realism and the Functional Consequences of Eye Gaze

The original findings of the liking effect allowed us to investigate the functional consequences of eye gaze across both visuospatial and social processes. Although we did have issues in replication at times, when following the original paradigm of Bayliss et al. (2006) with correct counterbalancing, the result did appear to remain consistent. The three main findings that the current research has added to the research around the liking effect will be discussed, before proposing the direction of future research into facial realism.

Face Identity and Object Interactions

One significant finding was the stark difference in affective judgments once multiple identities were used to cue objects across trials in the Photographic face stimuli, providing a robust replication of Capozzi et al. (2015) findings in a large-scale online sample. It also provides a cautionary tale about how small methodological differences can have a significant impact on replication studies. A small oversight in counterbalancing procedures resulted in null results across multiple experiments. However, once the correct procedures were in place, it was possible to establish interesting differences between Photographic and Schematic face conditions.

The finding that there was a requirement for multiple identities in the Photographic face stimuli group was consistent with the social consensus explanation provided by Capozzi et al. (2015). However, in the Schematic face condition, both the Multiple and Single face conditions elicited a liking effect. As discussed in Chapter 4, it makes sense that subcortical activation as a result of LSF from both face stimuli would rapidly increase attention towards associated to the item, explaining an increase in affective judgment (Johnson et al., 2015). When more detailed information are available to make effective identity judgments, higher-order social effects such as the social consensus can have a more significant impact on the functional consequences of eye gaze (Capozzi et al., 2015).

In addition to the differences between photographic and schematic faces, there was also an interesting difference between the two types of schematic face stimuli used in different experiments in Chapter 3. In both Experiments 4 and 5, there was a condition that presented a single schematic face. However, the effect was only seen in the Single condition of Experiment 7. So why was this effect not found in these previous experiments? One potential explanation for these findings is that the presentation of repetitive stimuli in Experiments 4 and 5 resulted in suppressed attention to the object, thereby impeding the effect. Although comparisons of the RT suggest relatively similar patterns of RT responses, it is possible that without any variation of face stimuli the effects of repetition may have reduced full engagement with the stimuli, so there was no impact on preferences. Therefore, in Experiment 7, where novel features were added, there was less repetition which maintained the engagement that resulted in a liking effect. This does not explain why there were differences between the Multiple and Single conditions in the Photographic, but not the Schematic conditions. It was proposed that the social consensus hypothesis may only occur when more detail is present, but further research of the interactions between the four conditions is required to understand these differences fully.

Eye-tracking Measures of Visual Attention and Their Relationship with the Liking Effect

The addition of eye-tracking to study the liking effect offered a complimentary method to understand how gaze cues influence visual attention and the subsequent effects on affective judgments. In the experiment, it was seen that NT individuals spent significantly more time gazing towards the congruently cued items than the incongruently cued items, as would be expected because individuals would be required to reorient back to the target before they could respond (Posner, 1980). In contrast, in the ASD group, it was found that autistic participants did not attend to the congruently cued objects any more than the incongruently cued objects. Even though autistic participants showed a similar pattern of gaze cueing bias for the RT, they spent an equal amount of time attending to items irrelevant of the gaze cue condition. This provides a mechanism through which more time spent attending to an object facilitates engagement required to increase affective judgments for NT participants only. Therefore, in the ASD group where engagement was the same no effect was found. These findings are in line with the consequences of eye gaze on memory in Chapter 1. It was found that only when more prolonged engagement with targets was encouraged through higher task demands, was there any improvement in memory associated with cued attention. However, based only on the findings presented in this thesis, it is not possible to establish the extent to which memory may play on the liking effect. Potential methods for investigating potential cross over will be discussed later in this chapter.

The Liking Effect in Autistic Spectrum Disorder

The liking effect is of interest to research into autism for the same reason as it is for facial realism. Although there has been previous research investigating gaze cueing for visuospatial measures in ASD, there have been inconsistent results across methodologies. The liking effect offered a measure of functional consequences across both visuospatial as well as functional eye gaze.

In the presented experiments, three methods of analysing autistic traits were used. In the first instance, AQ scores were entered as a covariate into the analyses to remove variance account for as a result of AQ. However, at no point did the AQ account for any significant amount of the variance. In Chapter 4, participants were prescreened and assigned to either group based on their AQ score. Comparisons between the group were inconclusive as a result of limitations in the statistical design. Finally, in Chapter 5, participants who had a previous diagnosis were recruited directly through the university DAS. It was found that the liking effect was not found in the ASD group. However, the small sample size makes it difficult to conclude that the liking effect is atypical in ASD, though it raises the feasibility of investigating the effect in more detail.

The eye-tracking study in Experiment 9, compared the amount of time attending to the eyes and the mouth when viewing a face. It was found that in the ASD group, individuals tended to reduce their engagement with the eyes as they made dynamic gaze shifts, whereas the NT group maintained more time attending to the eyes across all phases of the trial.

It was found by Chuk et al. (2014) that a sustained focus towards the centre of the face and less time spent attending to individual features is indicative of holistic style processing. Although the ROIs were not designed to investigate this hypothesis, the eye regions detected such sustained focus in the NT group. In contrast, the movement to investigate features across phases may indicate disrupted holistic style processing in the ASD group. Although this does not demonstrate aversive attention towards the eyes (previously suggested by Kliemann et al., 2010, 2012), it does show different patterns of face processing between NT and ASD participants. Further research effort is warranted to investigate directly how facial processing styles impacts the liking effect.

Conclusions for Future Research

There is still much to be understood about face processing. Throughout this thesis, multiple experiments have been presented that have investigated the potential for using the liking effect paradigm as a method of processing differences in perception between facial realism in both ASD and NT adults. Several differences have been identified in how the liking effect operates, both in terms of face realism and ASD, that warrant further investigation. However, it may not be the optimal paradigm for investigating face processing for face realism in ASD. One reason the liking effect was chosen was that, even though the effect was small, it appeared consistent across participants. In our successful attempts, we did not see the same level of consistency as in the original experiment. Combining this with a small effect size would require large sample sizes, so it would not be the optimal paradigm to use. However, there are still several follow up experiments that could be proposed. For example, an extension of the eye-tracking study with a larger ASD sample using both photographic and schematic faces. By focusing on how attention is distributed to the target items, it is hoped that this would provide clarification of our current findings within the ASD group, while also providing further evidence about how differences in the processing of faces of different levels of realism may influence the liking effect in both ASD and NT participants.

General Research Methodology and the Future of Online Testing

This thesis has utilised and adapted several experimental paradigms and research methodologies. This section will discuss the impact of the findings for research methodologies in general: Firstly, the necessity of adequate power and issues that arise with selecting the correct effect sizes; secondly, the use of Gorilla as an online data collection tool; and finally, the use of open science and preregistration and its importance in effective research.

The Validity of Online Data Collection Through Gorilla

A major theme of this thesis has been the need for suitably sized data sets, and specifically the collection of data remotely from participants using the online data collection tool Gorilla (Anwyl-Irvine et al., 2019). In reaction to questions about replication in Psychology in 2015 (Open Science Collaboration, 2015), it has become clear that an important issue for behavioural data collection is the low number of participants in samples. The use of online data collection tools may be vital in helping to overcome this problem in future. In the experiments presented here, it was found that we were able to replicate the findings of lab-based experiments with significantly larger samples. Consistent replication of gaze cueing results persisted across every experiment conducted, and replication of the liking effect was also obtained in Chapter 3. However, it should be noted that there were differences between the effect sizes of our lab-based and online samples. One potential reason for differences between the lab-based and online experiments was the level of engagement of the participant. Despite the use of an accuracy threshold, there were probably direct differences between the two samples. In lab experiments, participants are less likely to be exposed to distractions, compare to online testing, which is reliant on participants maintaining a level of engagement that cannot be confirmed. Bonuses were offered to participants who were able to accurately respond to 75% of trials to encourage engagement. However, though this may have appeared to be successful, it is not possible to know the real impact of distractions without a non-bonus comparison group. Also, questions can be asked about how bonuses may influence participants beyond encouraging engagement.

Another issue with online data collection is the browser based software's ability to display stimuli and collect responses with high temporal resolution. In a recent article, researchers at Gorilla have conducted a series of experiments utilising a flanker task to replicate the 'conflict network' paradigm across a wide range of conditions (different browsers, multiple types of PC, a variety of internet conditions; Anwyl-Irvine et al., 2019). They found that they were able to reliably replicate this effect across a multitude of conditions, consistently showing a reliable level of RT measurements. Although their findings suggest that Gorilla can collect data at comparable levels to lab-based data collection methods, such as Psychtoolbox and E-prime, more evidence is probably required to instil confidence in researchers that the platform can provide the required precision, especially those who require very short stimulus presentation timings.

The reliable and simplistic nature of modern online platforms makes online testing more accessible than ever. A drive to collect more extensive and representative samples also makes online behavioural testing more attractive to researchers. In future, Gorilla is hoping to expand its data collection abilities, such as the addition of eye-tracking using conventional webcams. Utilising the javascript package webgazer.js (Papoutsaki et al., 2016), Gorilla has recently introduced beta testing of eye-tracking for online studies, offering a limited ability to measure fixation of target regions. The ability to measure gaze location during online experiments may have applications beyond attempting to replicate conventional eye-trackers. It may provide a measure of participants engagement to ensure attention to the screen before proceeding with the experiment.

Overall, online experimentation appears to be an effective way of collecting large amounts of data in a short period. It facilitates data collection from more diverse participant groups that are not typically seen in traditional university settings. However, researchers must be careful not to fall into the trap of recruiting an equally restrictive demographic by utilising the same recruitment websites and demographic settings. A further issue impacting the research conducted here is the expense associated with collecting large data sets. Currently, the cost is lower per participant, but the pure number of participants required can drive the cost up in a way that may make it restrictive for some researchers and must be considered in advance.

Furthermore, it is advisable that any data collection online is complemented with data collection in the laboratory environment to allow comparisons to be made from, as well as between, the large sample size of online data collection and the controlled setting of lab-based testing. Finally, when utilising online data collection to reach a specific sample size based on power calculations, it is vital to ensure that the effect size estimates are appropriate for the online paradigm.

Calculating Effect Sizes for Power Analysis

Statistical power refers to the probability that a statistical test will reject a false null hypothesis. That is, it is the likelihood that a test will detect an effect that is present. Power of a statistical test is calculated using four main parameters, the predefined alpha level, the effect size, the number of observations, and the power level itself. Therefore, using three of these parameters, it is possible to calculate the fourth. Power can be calculated as a post hoc statistic using the measured effect size, the number of participants and the alpha level. However, if the goal is to ensure that there is adequate power to detect an effect, the power analysis has to be conducted a priori. To calculate the number of observations required to meet the target power level, usually defined as a minimum of 0.8 with a goal of 0.95 based on the work of Cohen (1992), requires the alpha level (typically set at 0.05), and the estimated effect sizes of the desired effect.

It is this final parameter that makes a priori power analysis particularly tricky. There are ongoing debates about how to calculate effect sizes for future experiments best. One suggestion is to collect enough pilot data to get an idea of the size of an effect and then use this effect size to run a power analysis. A pilot method approach has the advantage of ensuring that the effect size is based on the exact experimental parameters. However, this method requires more time and expense to collect enough pilot data; and there is no way of guessing the size of a sufficient pilot sample to get an accurate estimation of the effect sizes. Alternatively, by taking the effect size from previous experiments, it is possible to estimate effect sizes in advance without having to collect enough pilot data. However, this method also has limitations that have impacted multiple experiments in this thesis.

In the experiments conducted in Chapter 3, an apriori analysis was calculated by using the effect sizes from the original Bayliss et al. (2006) study. The results of this power analysis resulted in a sample size that was prohibitive in lab-based testing. By moving experiments online, it was possible to collect samples that previously were not possible, resulting in successful replications of the effect once correct methodologies were put in place. However, in Chapter 4, when additional conditions were added, and the number of trials was reduced, the estimated effect sizes were not valid for our calculations, probably resulting in a sample that was underpowered and therefore was not adequate to test the effect of interest.

The example of Chapter 4 outlines a central issue with using prior experiments to estimate effect size. Altering the paradigm in any way will impact the effect sizes used for the calculations which may not invalidate this method, but it is something that must be accounted for. In this case, a more complex design was combined with reducing the number of trials in an experiment, which has been found to impact effect sizes directly (Brand & Bradley, 2016). For the experiments in Chapter 4, the decision to reduce the length of the experiment was made purely to ensure that it was possible to collect data from enough participants required from the power analysis calculations. Therefore, by focusing on collecting the calculated samples size, the importance of accurately calculating the effect sizes was overlooked at the expense of the quality of data collected. These findings provide a cautionary tale for those using power analysis to calculate sample size; accurate representation of effect sizes is paramount. One possible method to get a more representative sample is to collect an aggregate of effect sizes across experiments. However, in a similar way to how researchers may use p-hacking to find a significant result (Head, Holman, Lanfear, Kahn, & Jennions, 2015), researchers may be tempted to select effect sizes for an aggregate that provide the desired result.

There are also alternative statistical methods that may be more appropriate when dealing with experiments that yield small effect sizes. Throughout this thesis, frequentist statistics have been used to analyse the results of the experiments. However, alternative methods, such as Bayesian statistics, may offer more appropriate alternatives. In frequentist approaches the method of dealing with experiments with small effect sizes is to increase the number observations, but in a recent replication of the liking effect, Tipples and Pecchinenda (2018) used simulation modelling to estimate the required sample for a Bayesian statistical model. This analysis required only 90 participants using an equal number of trials as the original Bayliss et al. (2006) study. As with the approach presented here, they were able to replicate the effect, albeit with an attenuated effect. These findings would suggest that a more measured, analysis-driven approach to dealing with small effect sizes would be a better approach in future.

Importance of Preregistration and Open Science

For Chapters 3 and 4, the experiments were preregistered in advance on the Open Science Framework (OSF). OSF is a service provided by the Centre for Open Science, a non-profit technology organisation with a mission to increase the openness, integrity, and reproducibility of scientific research. They provide a framework for preregistration that allows separating hypothesis testing and hypothesis-generating research. By planning research design and methodology using the framework, it is possible to present methodologies transparently in advance of data collection. It also provides a facility for publicly storing data and analysis scripts for independent review. In this thesis, several hypothesis-driven experiments have been presented. However, in hindsight, there is an argument that exploratory analysis early in the research project could have been used to identify methods that would have been best suited to investigate the research questions, which would then have produced hypothesis-driven experiments in later chapters, rather than building hypothesis and testing the liking effect this way.

When using online experiments, the potential for dredging data, or sampling until a significant effect is found, is higher than in lab-based studies. Although exploratory analysis is encouraged when the methods are correctly preregistered, data that are presented as hypothesis-driven when they are not, undermine the scientific method and have the potential for adverse consequences, as was highlighted by the reproducibility crisis. Preregistration for online experiments allows for analyses to be easily defined as either exploratory or hypothesis-driven and requires the researcher to define the sample as such. By asking researchers to justify these decisions before data collection, Open Science will encourages open discussions about methodology before data collection. For example, opening up decisions of effect sizes for power analysis for examination before the data collection will allow people to question potential flaws in advance. Preregistration may not be perfect and will not fix methodological issues overnight, but by making research more open, it is a step in the correct direction.

Proposed Future Research

Memory as a Measure of Selective Encoding

In the early stages of this thesis, the goal was to utilise memory as a secondary measure of functional attention. In Chapter 2, tentative links were made to the impact of memory on the liking effect (Bry et al., 2011), but a strong connection was not made between selective encoding and the potential relationship between the two measures. As the primary focus became the development of the liking effect paradigm in Chapter 2, memory was dropped as a measure. In hindsight, potential to investigate the relationship between selective encoding and the subsequent effects on memory and affective judgments was missed. Although it was necessary at the time to focus on replicating the liking effect, it was an oversight not to return to measures of memory, and this merits more attention in future research.

Specifically, two adaptations of the paradigm are proposed that would hope to investigate the effects of selective encoding on memory and the relationship with affective judgments. The first adaptation would attempt to resolve the intentions of studying long-term memory in Chapter 2 using the liking effect paradigm from Chapter 3. By following the procedure of Chapter 3, with the repeated categorisation of items, but with an independent recall and rating block, it would be possible to test both the liking effect and memory as well as any relationship between the two. The addition of eye-tracking to measure how overt attention shifts facilitate both memory and effect would also be beneficial. Furthermore, the addition of item level analysis across the three measures (affective scores, memory recall and dwell time) would provide a more in-depth understanding of the relationship between selective encoding and functional consequences of eye gaze.

The second adaptation would seek to investigate the impact of the selective encoding in much longer-term memory traces. By using online experiments, with novel eye-tracking technology, large data sets could be collected using the same paradigm as Experiments 6 and 9. The use of eye-tracking will ensure engagement as well as indicating overt gaze shifts towards objects. However, the nature of the online experiment means that it is also relatively simple to do follow up memory recall experiments with participants across different time frames (from days to weeks and beyond). Using similar item-level analysis as proposed earlier, it may be possible to get an indication of the relationship between gaze cues and the impact on selective encoding.

Interactive Specialisation Account for Investigating Facial Realism Processing

Another topic of interested discussed throughout the thesis is the interactive specialisation account of face perception (Johnson et al., 2015, 2009). The interactive specialisation account can be used to explain how the very early attraction to LSF face shapes leads to the development of full face processing abilities across the full band of spatial frequencies. However, as this account was a relatively late addition to the process of designing the experimental paradigms, it had little influence on the experiments at hand. Therefore, there are several directions that this research could focus on in future. Firstly, investigating further differences between photographic and schematic faces in functional consequences, for which we had success in Experiments 6 and 7. Focusing on both affective judgments and memory, as well as utilising eye-tracking as in Chapter 5, it should be possible to get a greater understanding of how the eye gaze of schematic face shapes drive selective encoding. It may also be possible to investigate other functional consequences that were not considered here, such as the impact of accurate gaze cues in face ratings, as demonstrated by Bayliss et al. (2007). In addition to expanding research into functional consequences, it would also be essential to investigate differences in early visual processing across a spectrum of faces across different levels of facial realism. In the early stages of the thesis, it was planned that EEG would be utilised to investigate these effects. However, as there were no specific hypotheses, and again the focus shifted toward replicating the behavioural effects, these plans did not move forward. In future research, early visual properties, whether measured using EEG, eye-tracking or behavioural methods, will provide a better understanding of differences between faces of different levels of realism. The potential advantages of EEG are that hypotheses can be developed around network-based approaches that may provide better insight into how high detailed faces are processed differently.

General Conclusion

The work presented here investigated the potential of using the liking effect to further our understanding of the mechanisms of face perceptions in autistic individuals exploiting differences in how they process faces of differing levels of realism. However, methodological issues and limited sample sizes make it difficult to make substantive claims about the liking effect. Specific differences were found between how schematic face stimuli and photographic face stimuli elicited the liking effect. Also, it was found in small-scale eye-tracking experiments, that there was no explicit liking effect seen in autistic participants. Based on these findings, it is suggested that the future directions of research into the liking effect should highlight differences in face processing mechanisms in those with ASD. However, it appears that this effect is not the optimum paradigm for investigating developmental differences in face processing. Future experiments that focus on earlier visual processing of faces may be able to differentiate face processing differences across development more effectively. Finally, this thesis has outlined relatively novel approaches to collecting large data sets online, as well as discussing the limitations of such methods for future reference.

Appendix

Appendix A - Histograms of Residuals from Experiment 8

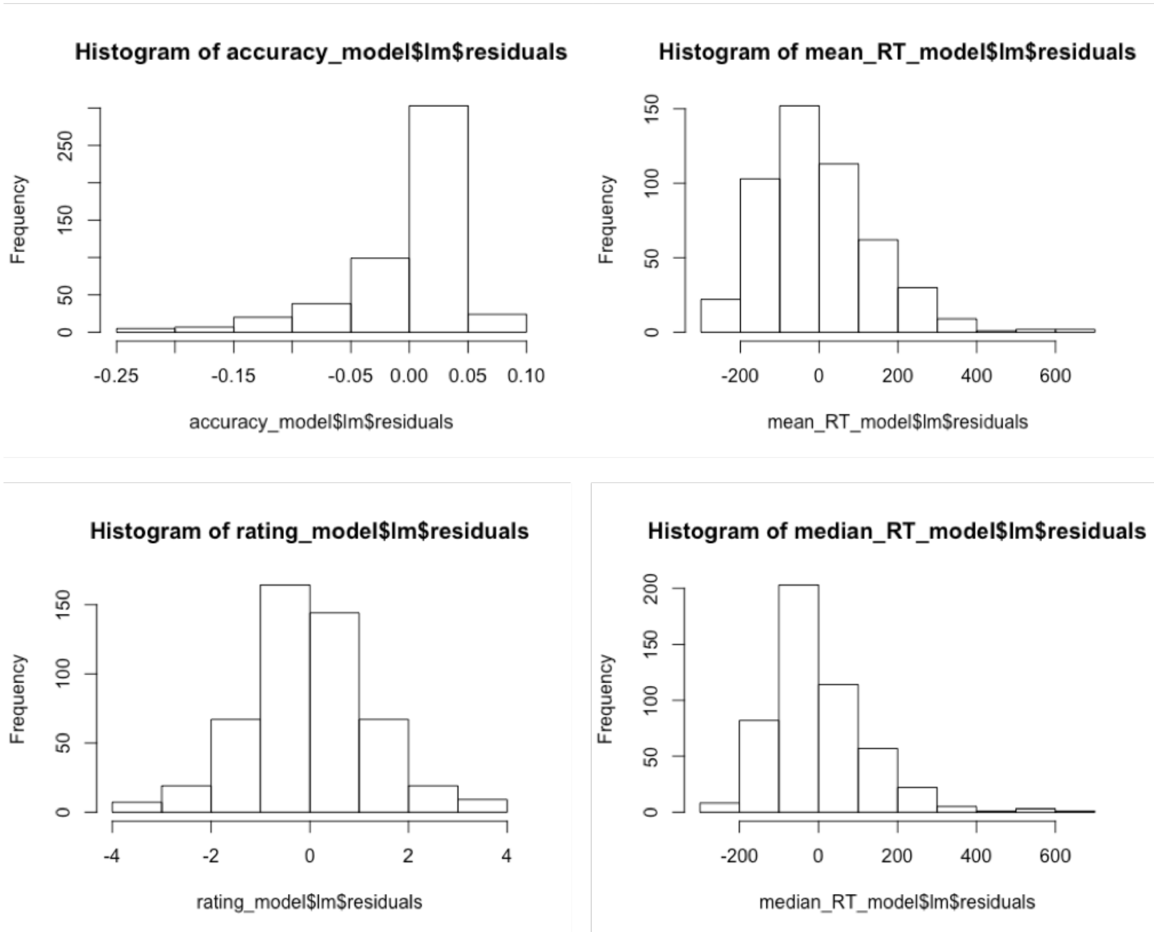

Figure A.1: Histograms of the residuals from Chapter 4. Although there is some skew in the Accuracy (tending towards ceiling effects) and the RT measures, they tend to follow a normal distribution.

References

Adams, R. B., Gordon, H. L., Baird, A. A., Ambady, N., & Kleck, R. E. (2003). Effects of gaze on amygdala sensitivity to anger and fear faces. *Science*, *300*(5625), 1536. doi: 10.1126/science.1082244

Adams, R. B., & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, *5*(1), 3–11. doi: 10.1037/1528-3542.5.1.3

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2004). Beautiful faces have variable reward value. *Neuron*, *32*(3), 537–551. doi: 10.1016/s0896-6273(01)00491-3

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Amso, D., & Casey, B. J. (2006). Beyond what develops when neuroimaging may inform how cognition changes with development. *Current Directions in Psychological Science*, *15*(1), 24–29. doi: 10.1111/j.0963-7214.2006.00400.x

Anderson, A. W., Gauthier, I., Skudlarski, P., & Gore, J. C. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience*, *3*, 191–197. Retrieved from `http://www.nature.com/?file=/neuro/journal/v3/n2/full/nn0200{\_}191.html` doi: 10.1038/72140

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2019). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*. Retrieved from `http://link.springer.com/10.3758/s13428-019-01237-x` doi: 10.3758/s13428-019-01237-x

Bailey, A. J., Braeutigam, S., Jousmäki, V., & Swithenby, S. J. (2005). Abnormal activation of face processing systems at early and intermediate latency in individuals with autism spectrum disorder: A magnetoencephalographic study. *European Journal of Neuroscience*, *21*(9), 2575–2585. doi: 10.1111/j.1460-9568.2005.04061.x

Balas, B., & Saville, A. (2016). N170 face specificity and face memory depend on hometown size. *Neuropsychologia*, *69*, 211–217. doi: 10.1016/j.neuropsychologia.2015.02.005.N170

Banks, S. J., Eddy, K. T., Angstadt, M., Nathan, P. J., & Luan Phan, K. (2007). Amygdala-frontal connectivity during emotion regulation. *Social Cognitive and Affective Neuroscience*, *2*(4), 303–312. doi: 10.1093/scan/nsm029

Barbeau, E. J., Taylor, M. J., Regis, J., Marquis, P., Chauvel, P., & Liégeois-Chauvel, C. (2008). Spatio temporal dynamics of face recognition. *Cerebral Cortex*, *18*(5), 997–1009. doi: 10.1093/cercor/bhm140

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` doi: 10.1111/1469-7610.00715

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning Autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17. doi: 10.1023/A:1005653411471

Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J., & Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, *23*(2), 223–229. doi: 10.1016/S0163-6383(01)00037-6

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104*(3), 644–653. doi: 10.1016/j.cognition.2006.07.012

Bayliss, A. P., Murphy, E., Naughtin, C. K., Kritikos, A., Schilbach, L., & Becker, S. I. (2012). "Gaze leading": Initiating simulated joint attention influences eye movements and choice behavior. *Journal of Experimental Psychology: General*, *142*(1), 76–92. doi: 10.1037/a0029286

Bayliss, A. P., Paul, M. a., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin and Review*, *13*(6), 1061–1066. doi: 10.3758/BF03213926

Bayliss, A. P., Pellegrino, G. D., & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *58*(4), 631–650. doi: 10.1080/02724980443000124

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*, *96*(1), 95–114. doi: 10.1348/000712604X15626

Bayliss, A. P., & Tipper, S. P. (2006). Predictive Gaze Cues and Personality Judgments. *Psychological Science*, *17*(6), 514–520. doi: 10.1111/j.1467-9280.2006.01737.x

Behrmann, M., Thomas, C., & Humphreys, K. (2006). Seeing it differently: visual processing in autism. *Trends in Cognitive Sciences*, *10*(6), 258–264. doi: 10.1016/j.tics.2006.05.001

Belin, P., Zatorre, R. J., Lafallie, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, *403*(6767), 309–312. doi: 10.1038/35002078

Bentin, S., Truett, A., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of Cognitive Neuroscience*, *8*(6), 551–565. doi: 10.1162/jocn.1996.8.6.551.Electrophysiological

Berger, A., Henik, A., & Rafal, R. (2005). Competition between endogenous and exogenous orienting of visual attention. *Journal of Experimental Psychology: General*, *134*(2), 207–221. doi: 10.1037/0096-3445.134.2.207

Bernstein, M., & Yovel, G. (2015). *Two neural pathways of face processing: A critical evaluation of current models* (Vol. 55). Elsevier Ltd. Retrieved from `http://dx.doi.org/10.1016/j.neubiorev.2015.06.010` doi: 10.1016/j.neubiorev.2015.06.010

Boutet, I., & Faubert, J. (2006). Recognition of faces and complex objects in younger and older adults. *Memory and Cognition*, *34*(4), 854–864. doi:

10.3758/BF03193432

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. doi: 10.1163/156856897X00357

Brand, A., & Bradley, M. T. (2016). The precision of effect size estimation from published psychological research: Surveying confidence intervals. *Psychological Reports*, *118*(1), 154–170. doi: 10.1177/0033294115625265

Brodeur, M. B., Guérard, K., & Bouras, M. (2014). Bank of Standardized Stimuli (BOSS) phase ii: 930 new normative photos. *PLoS ONE*, *9*(9). doi: 10.1371/journal.pone.0106953

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327. Retrieved from `http://onlinelibrary.wiley.com/doi/10.1111/j.2044-8295.1986.tb02199.x/full{\%}5Cnpapers3://publication/doi/10.1111/j.2044-8295.1986.tb02199.x/pdf` doi: 10.1111/j.2044-8295.1986.tb02199.x/pdf

Bry, C., Treinen, E., Corneille, O., & Yzerbyt, V. (2011). Eye'm lovin' it! The role of gazing awareness in mimetic desires. *Journal of Experimental Social Psychology*, *47*(5), 987–993. Retrieved from `http://dx.doi.org/10.1016/j.jesp.2011.03.023` doi: 10.1016/j.jesp.2011.03.023

Burke, D., & Sulikowski, D. (2013). The evolution of holistic processing of faces. *Frontiers in Psychology*, *4*(JAN), 1–10. doi: 10.3389/fpsyg.2013.00011

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*(8), 641–651. doi: 10.1038/nrn1724

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: Salient physical features guide effective visual search. *Journal of Experimental Psychology: General*, *137*(3), 471–494. doi: 10.1037/a0012771

Campbell, R., Heywood, C. A., Cowey, A., Regard, M., & Landis, T. (1990). Sensitivity to eye gaze in prosopagnosic patients and monkeys with superior temporal sulcus ablation. *Neuropsychologia*, *28*(11), 1123–1142. doi: 10.1016/0028-3932(90)90050-X

Capozzi, F., Bayliss, A. P., Elena, M. R., & Becchio, C. (2015). One is not enough: Group size modulates social gaze-induced object desirability effects. *Psychonomic Bulletin & Review*, *22*(3), 850–855. Retrieved from `http://link.springer.com/10.3758/s13423-014-0717-z` doi: 10.3758/s13423-014-0717-z

Carey, S., & Diamond, R. (1994). Are faces perceived as configurations more by adults than by children? *Visual Cognition*, *1*(2-3), 253–274. doi: 10.1080/13506289408402302

Carlin, J. D., & Calder, A. J. (2013). The neural basis of eye gaze processing. *Current Opinion in Neurobiology*, *23*(3), 450–455. doi: 10.1016/j.conb.2012.11.014

Caruana, N., Stieglitz Ham, H., Brock, J., Woolgar, A., Kloth, N., Palermo, R., & McArthur, G. (2018). Joint attention difficulties in autistic adults: An interactive eye-tracking study. *Autism*, *22*(4), 502–512. doi: 10.1177/1362361316676204

Cerf, M., Harel, J., Einhäuser, W., in Neural, C. K. A., & 2008, U. (2007).

Predicting human gaze using low-level saliency combined with face detection. *Advances in Neural Information Processing Systems*, 1–8. Retrieved from http://papers.nips.cc/paper/3169-predicting-human-gaze-using-low-level-saliency-combined-with-face-detection

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 315–324. doi: 10.1098/rstb.2002.1199

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation, and play as infancy precursors to language and theory of mind. *Cognitive Development*, *15*(4), 481–498. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0885201401000375 doi: 10.1016/S0885-2014(01)00037-5

Chelnokova, O., Laeng, B., Eikemo, M., Riegels, J., Løseth, G., Maurud, H., ... Leknes, S. (2014). Rewards of beauty: the opioid system mediates social motivation in humans. *Molecular Psychiatry*, *19*(12), 1342–1342. doi: 10.1038/mp.2014.149

Chien, S. H.-L. (2011). No more top-heavy bias: Infants and adults prefer upright faces but not top-heavy geometric or face-like patterns. *Journal of Vision*, *11*(6), 13–13. Retrieved from http://jov.arvojournals.org/Article.aspx?doi=10.1167/11.6.13 doi: 10.1167/11.6.13

Chua, K.-W., Richler, J. J., & Gauthier, I. (2015). Learned attention to parts results in holistic processing. *Journal of Experimental Psychology: General*, *144*(4), 723–729. doi: 10.1037/xge0000063.Holistic

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden Markov models. *Journal of Vision*, *14*(11), 8–8. doi: 10.1167/14.11.8

Cohen, J. (1992). Statistical power analysis. *Current directions in psychological science*, *1*(3), 98–101.

Collie, A., Maruff, P., Yucel, M., Danckert, J., & Currie, J. (2000). Spatiotemporal distribution of facilitation and inhibition of return arising from the reflexive orienting of covert attention. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1733.

Collins, C. (2016, jul). Are Neutral Faces Really Neutral ? *APS*, *29*(6). Retrieved from https://www.psychologicalscience.org/observer/are-neutral-faces-really-neutral{\#}.WmDF4HUgqKc.mendeley

Conty, L., N'Diaye, K., Tijus, C., & George, N. (2007). When eye creates the contact! ERP evidence for early dissociation between direct and averted gaze motion processing. *Neuropsychologia*, *45*(13), 3024–3037. Retrieved from http://dx.doi.org/10.1016/j.neuropsychologia.2007.05.017 doi: 10.1016/j.neuropsychologia.2007.05.017

Cooney, R. E., Atlas, L. Y., Joormann, J., Eugène, F., & Gotlib, I. H. (2006). Amygdala activation in the processing of neutral faces in social anxiety disorder: Is neutral really neutral? *Psychiatry Research - Neuroimaging*, *148*(1), 55–59. doi: 10.1016/j.pscychresns.2006.05.003

Dalmaijer, E. S. (2014). Is the low-cost EyeTribe eye tracker any good for research?

PeerJ PrePrints, *4*(606901), 1–35. doi: 10.7287/peerj.preprints.141v2

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., ... Davidson, R. J. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, *8*(4), 519–526. Retrieved from http://www.nature.com/doifinder/10.1038/nn1421 doi: 10.1038/nn1421

Damasio, H., Hichwa, R. D., Ponto, L. L. B., Grabowski, T. J., Parvizi, J., Damasio, A. R., & Bechara, A. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, *3*(10), 1049–1056. doi: 10.1038/79871

Davies, S., Bishop, D., Manstead, A. S., & Tantam, D. (1994). Face perception in children with Autism and Asperger's syndrome. *Journal of Child Psychology and Psychiatry*, *35*(6), 1033–1057. doi: 10.1111/j.1469-7610.1994.tb01808.x

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in Autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology*, *27*(3), 403–424. doi: 10.1207/s15326942dn2703_6

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, *22*(5), 545–559. doi: 10.1016/0042-6989(82)90113-4

De Valois, R. L., & De Valois, K. K. (1988). *Spatial vision [electronic resource].* Oxford: Oxford University Press.

de Gelder, B., Frissen, I., Barton, J., & Hadjikhani, N. (2003). A modulatory role for facial expressions in prosopagnosia. *Proceedings of the National Academy of Sciences*, *100*(22), 13105–13110. Retrieved from http://www.pnas.org/cgi/doi/10.1073/pnas.1735530100 doi: 10.1073/pnas.1735530100

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

Deruelle, C., Rondan, C., Gepner, B., & Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *34*(2), 199–210. doi: 10.1023/B:JADD.0000022610.09668.4c

Doherty, M. J., & Anderson, J. R. (2000). A new look at gaze: Preschool children's understanding of eye-direction. *Children*, *571*(1999), 549–571. doi: 10.1016/S0885-2014(99)00019-2

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). *Gaze perception triggers reflexive visuospatial orienting* (Vol. 6) (No. 5). doi: 10.1080/135062899394920

Dubey, I., Ropar, D., & Antonia, A. F. (2017). Brief report: A comparison of the preference for viewing social and non-social movies in typical and Autistic adolescents. *Journal of Autism and Developmental Disorders*, *47*(2), 514–519. doi: 10.1007/s10803-016-2974-3

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science*, *1*(1), 393–416. doi: 10.1146/annurev-vision

-082114-035518

Engell, A. D., Kim, N. Y., & McCarthy, G. (2018). Sensitivity to faces with typical and atypical part configurations within regions of the face-processing network: An fMRI study. *Journal of Cognitive Neuroscience*, *30*(7), 963–972. doi: 10 .1162/jocn

Ewing, L., Rhodes, G., & Pellicano, E. (2010). Have you got the look? gaze direction affects judgements of facial attractiveness. *Visual Cognition*, *18*(3), 321–330. doi: 10.1080/13506280902965599

Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition*, *7*(6), 705–718. doi: 10.1080/13506280050144399

Farroni, T., Massaccesi, S., Pividori, D., & Johnson, M. H. (2004). Gaze following in newborns. *Infancy*, *5*(1), 39–60. doi: 10.1207/s15327078in0501_2

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175–191.

Feng, S., Wang, X., Wang, Q., Fang, J., Wu, Y., Yi, L., & Wei, K. (2018). The uncanny valley effect in typically developing children and its absence in children with autism spectrum disorders. *PLoS ONE*, *13*(11), 1–14. doi: 10.1371/journal.pone.0206343

Freeth, M., Bullock, T., & Milne, E. (2013). The distribution of and relationship between autistic traits and social anxiety in a UK student population. *Autism*, *17*(5), 571–581. doi: 10.1177/1362361312445511

Freire, A., Lee, K., & Symons, L. A. (2000). The face-inversion effect as a deficit in the encoding of configurai information: Direct evidence. *Perception*, *29*(2), 159–170. doi: 10.1068/p3012

Freiwald, W., Duchaine, B., & Yovel, G. (2016). Face processing systems: From neurons to real-world social perception. *Annual Review of Neuroscience*, *39*(1), 325–346. Retrieved from `http://www.annualreviews.org/doi/10.1146/annurev-neuro-070815-013934` doi: 10.1146/annurev-neuro-070815-013934

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review*, *5*(3), 490–495. doi: 10.3758/BF03208827

Friesen, C. K., & Kingstone, A. (2003). Abrupt onsets and gaze direction cues trigger independent reflexive attentional effects. *Cognition*, *87*(1), B1 – B10. doi: https://doi.org/10.1016/S0010-0277(02)00181-6

Friesen, C. K., Moore, C., & Kingstone, A. (2005). Does gaze direction really trigger a reflexive shift of spatial attention? *Brain and Cognition*, *57*(1), 66–69. doi: 10.1016/j.bandc.2004.08.025

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), 319–329. doi: 10.1037/0096-1523.30.2.319

Frischen, A. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724. doi: 10.1037/0033-2909.133.4.694.Gaze

Frischen, A., & Tipper, S. P. (2004). Orienting attention via observed gaze shift evokes longer term inhibitory effects: Implications for social interactions, attention, and memory. *Journal of Experimental Psychology: General*, *133*(4), 516–533. doi: 10.1037/0096-3445.133.4.516

Frischen, A., & Tipper, S. P. (2006). Long-term gaze cueing effects: Evidence for retrieval of prior states of attention from memory. *Visual Cognition*, *14*(3), 351–364. doi: 10.1080/13506280544000192

Gauthier, I. (2018). Domain-specific and domain-general individual differences in visual object recognition. *Current Directions in Psychological Science*, *27*(2), 97–102. doi: 10.1177/0963721417737151

Gauthier, I., & Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, *11*, 219–224. doi: https://doi.org/10.1016/S0959-4388(00)00200-2

Gee, B. M., Thompson, K., & St John, H. (2014). Efficacy of a sound-based intervention with a child with an autism spectrum disorder and auditory sensory over-responsivity. *Occupational Therapy International*, *21*(1), 12–20. doi: 10.1002/oti.1359

Geisler, W. S., & Perry, J. S. (2013). High-order statistics for point prediction in natural images. *Journal of Vision*, *11*(2011), IMC1. doi: 10.1364/isa.2011.imc1

George, N., Driver, J., & Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *NeuroImage*, *13*(6), 1102–1112. doi: 10.1006/nimg.2001.0769

Gillespie-Lynch, K., Elias, R., Escudero, P., Hutman, T., & Johnson, S. P. (2013, dec). Atypical gaze following in Autism: A comparison of three potential mechanisms. *Journal of autism and developmental disorders*, *43*(12), 2779–2792. doi: 10.1007/s10803-013-1818-7

Gliga, T., Elsabbagh, M., Andravizou, A., & Johnson, M. (2009). Faces attract infants' attention in complex displays. *Infancy*, *14*(5), 550–562. doi: 10.1080/15250000903144199

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

Goffaux, V., & Rossion, B. (2006). Faces are "spatial" - Holistic face perception is supported by low spatial frequencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 1023–1039. doi: 10.1037/0096-1523.32.4.1023

Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., & Baron-Cohen, S. (2010). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders*, *40*(3), 269–279. doi: 10.1007/s10803-009-0862-9

Grelotti, D. J., Klin, A. J., Gauthier, I., Skudlarski, P., Cohen, D. J., Gore, J. C., ... Schultz, R. T. (2005). fMRI activation of the fusiform gyrus and amygdala to cartoon characters but not to faces in a boy with autism. *Neuropsychologia*, *43*(3), 373–385. doi: 10.1016/j.neuropsychologia.2004.06.015

Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research*, *41*(10-11), 1475–1482. doi: 10.1016/S0042-6989(00)00317-5

Haigh, S. M. (2018). Variable sensory perception in autism. *European Journal of Neuroscience*, *47*(6), 602–609. doi: 10.1111/ejn.13601

Hare, T. A., Tottenham, N., Davidson, M. C., Glover, G. H., & Casey, B. J. (2005). Contributions of amygdala and striatal activity in emotion regulation. *Biological Psychiatry*, *57*(6), 624–632. doi: 10.1016/j.biopsych.2004.12.038

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. doi: 10.1016/S1364-6613(00)01482-0

Hayden, A., Bhatt, R. S., Reed, A., Corbly, C. R., & Joseph, J. E. (2007). The development of expert face processing: Are infants sensitive to normal differences in second-order relational information? *Journal of Experimental Child Psychology*, *97*(2), 85–98. doi: 10.1016/j.jecp.2007.01.004

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of P-hacking in science. *PLoS Biology*, *13*(3), 1–15. doi: 10.1371/journal.pbio.1002106

Heimberg, R. G., Horner, K. J., Juster, H. R., Safren, S. A., Brown, E. J., Schneier, F. R., & Liebowitz, M. R. (1999). Psychometric properties of the Liebowitz Social Anxiety Scale. *Psychological Medicine*, *29*(1), 199–212. doi: 10.1017/S0033291798007879

Hernandez, N., Metzger, A., Magné, R., Bonnet-Brilhault, F., Roux, S., Barthelemy, C., & Martineau, J. (2009). Exploration of core features of a human face by healthy and autistic adults analyzed by visual scanning. *Neuropsychologia*, *47*(4), 1004–1012. doi: 10.1016/j.neuropsychologia.2008.10.023

Hileman, C. M., Henderson, H., Mundy, P., Newell, L., & Jaime, M. (2011). Developmental and individual differences on the P1 and N170 ERP components in children with and without autism. *Developmental Neuropsychology*, *36*(2), 214–236. doi: 10.1080/87565641.2010.549870

Hobson, & Hobson. (2007). Identification: The missing link between joint attention and imitation? *Development and Psychopathology*, *19*(2), 411–431. doi: 10.1017/S0954579407070204

Hobson, R. P., Ouston, J., & Lee, A. (1988). What's in a face? The case of autism. *British Journal of Psychology*, *79*(4), 441–453. doi: 10.1111/j.2044-8295.1988.tb02745.x

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., … van der Sluis, S. (2010). The construction and validation of an abridged version of the Autism-Spectrum Quotient (AQ-Short). *Journal of Autism and Developmental Disorders*, *41*(5), 589–596. doi: 10.1007/s10803-010-1073-0

Hood, B. M., Douglas Willen, J., & Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, *9*(2), 131–134. doi: 10.1111/1467-9280.00024

Hood, B. M., Macrae, C. N., Cole-Davies, V., & Dias, M. (2003). Eye remember you: The effects of gaze direction on face recognition in children and adults. *Developmental Science*, *6*(1), 67–71. Retrieved from `http://www.ingentaconnect.com/content/bpl/desc/2003/00000006/00000001/art00015` doi: 10.1111/1467-7687.00256

Hubel, S. D., & Livingstone, M. S. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*(4853), 740–749. doi: 10.1126/science.3283936

Iskra, A., & Gabrijelčič Tomc, H. (2017). Eye-tracking analysis of face observing and face recognition. *Journal of Graphic Engineering and Design*, *7*(1), 5–11. doi: 10.24867/jged-2016-1-005

Jeantet, C., Laprevote, V., Schwan, R., Schwitzer, T., Maillard, L., Lighezzolo-Alnot, J., & Caharel, S. (2019). Time course of spatial frequency integration in face perception: An ERP study. *International Journal of Psychophysiology*, *143*(July), 105–115. doi: 10.1016/j.ijpsycho.2019.07.001

Jemel, B., Mottron, L., & Dawson, M. (2006). Impaired face processing in autism: Fact or artifact? *Journal of Autism and Developmental Disorders*, *36*(1), 91–106. doi: 10.1007/s10803-005-0050-5

Johnson, M. H. (2000). Functional brain development in infants: Elements of an interactive specialization framework. *Child Development*, *71*(1), 75–81. Retrieved from `https://doi.org/10.1111/1467-8624.00120{\%}0Ahttps://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-8624.00120` doi: 10.1111/1467-8624.00120

Johnson, M. H. (2005). Subcortical face processing. *Nature Reviews Neuroscience*, *6*(10), 766–774. doi: 10.1038/nrn1766

Johnson, M. H., Grossmann, T., & Kadosh, K. C. (2009). Mapping functional brain development: Building a social brain through interactive specialization. *Developmental Psychology*, *45*(1), 151–159. doi: 10.1037/a0014548

Johnson, M. H., Senju, A., & Tomalski, P. (2015). The two-process theory of face processing: Modifications based on two decades of data from infants and adults. *Neuroscience and Biobehavioral Reviews*, *50*, 169–179. Retrieved from `http://dx.doi.org/10.1016/j.neubiorev.2014.10.009` doi: 10.1016/j.neubiorev.2014.10.009

Jones, W., & Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature*, *504*(7480), 427–431. Retrieved from `http://dx.doi.org/10.1038/nature12715` doi: 10.1038/nature12715

Kampe, Knut; Frith, Chris; Frith, U. (2003). "Hey John": Signals Conveying Communicative Intention toward the SelfActivate Brain Regions Associated with "Mentalizing," Regardless ofModality. *The Journal of Neuroscience*, *23*(12), 5258–5263. doi: 10.1002/hbm.21164

Kätsyri, J., Förger, K., Mäkäräinen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: Support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, *6*(MAR), 1–16. doi: 10.3389/fpsyg.2015.00390

Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., . . . Nakamura, K. (1999). The human amygdala plays an important role in gaze monitoring: A PET study. *Brain*, *122*(4), 779–783. doi: 10.1093/brain/122.4.779

Kéïta, L., Guy, J., Berthiaume, C., Mottron, L., & Bertone, A. (2014). An early origin for detailed perception in Autism Spectrum Disorder: Biased sensitivity for high-spatial frequency information. *Scientific Reports*, *4*, 17–19. doi: 10.1038/srep05475

King, D., Rowe, A., & Leonards, U. (2011). I trust you; hence I like the things you look at: gaze cueing and sender trustworthiness influence object evaluation. *Social Cognition*, *29*(4), 476–485. Retrieved from `http://guilfordjournals.com/doi/10.1521/soco.2011.29.4.476` doi: 10.1521/soco.2011.29.4.476

Kingstone, A., Tipper, C., Ristic, J., & Ngan, E. (2004). The eyes have it!: An fMRI investigation. *Brain and Cognition*, *55*(2), 269–271. doi: 10.1016/j.bandc.2004.02.037

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoobox-3? *Perception*, *36*(ECVP Abstract Supplement), 14.

Kleinhans, N. M., Richards, T., Sterling, L., Stegbauer, K. C., Mahurin, R., Johnson, L. C., ... Aylward, E. (2008). Abnormal functional connectivity in autism spectrum disorders during face processing. *Brain*, *131*(4), 1000–1012. doi: 10.1093/brain/awm334

Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., & Heekeren, H. R. (2012). The role of the amygdala in atypical gaze on emotional faces in Autism Spectrum Disorders. *Journal of Neuroscience*, *32*(28), 9469–9476. Retrieved from `http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5294-11.2012` doi: 10.1523/jneurosci.5294-11.2012

Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical reflexive gaze patterns on emotional faces in Autism Spectrum Disorders. *Journal of Neuroscience*, *30*(37), 12281–12287. Retrieved from `http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.0688-10.2010` doi: 10.1523/jneurosci.0688-10.2010

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. Retrieved from `http://archpsyc.jamanetwork.com/article.aspx?doi=10.1001/archpsyc.59.9.809` doi: 10.1001/archpsyc.59.9.809

Krolak-Salmon, P., Hénaff, M.-A. A., Vighetto, A., Bertrand, O., & Mauguière, F. (2004). Early amygdala reaction to fear spreading in occipital, temporal, and frontal cortex: A depth electrode ERP study in human. *Neuron*, *42*(4), 665–676. doi: 10.1016/S0896-6273(04)00264-8

Kühn, S., Brick, T. R., Müller, B. C. N., & Gallinat, J. (2014). Is this car looking at you? How anthropomorphism predicts fusiform face area activation when seeing cars. *PLoS ONE*, *9*(12), 1–14. doi: 10.1371/journal.pone.0113885

Ladouceur, M., Phillips, C., & Drevets, W. (2008). A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *molecular Psychiatry*, *13*(9), 829–857. doi: 10.1038/mp.2008.65.A

Langton, S., Watt, R., & Bruce, I. (2000). Do the eyes have it? Cues to the direction of towards expressive gaze manner in embodied virtual agentsocial attention. *Trends in cognitive sciences*, *4*(2), 50–59. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/10652522` doi: 10.1016/S1364-6613(99)01436-9

Lee, E., Kang, J. I., Park, I. H., Kim, J. J., & An, S. K. (2008). Is a neutral face really evaluated as being emotionally neutral? *Psychiatry Research*, *157*(1-3), 77–85. doi: 10.1016/j.psychres.2007.02.005

Lee, K., Anzures, G., Quinn, P. C., Pascalis, O., & Slater, A. M. (2012). Development of Face Processing Expertise. In *Oxford handbook of face perception* (pp. 1–15). doi: 10.1093/oxfordhb/9780199559053.013.0039

Leekam, S., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1997). Eye-direction detection: A dissociation between geometric and joint attention skills in autism. *British Journal of Developmental Psychology*, *15*(1), 77–95. Retrieved from `http://doi.wiley.com/10.1111/j.2044-835X.1997.tb00726.x` doi: 10.1111/j.2044-835X.1997.tb00726.x

Leekam, S., López, B., & Moore, C. (2000). Attention and joint attention in preschool children with autism. *Developmental psychology*, *36*(2), 261–273. doi: 10.1037/0012-1649.36.2.261

Leopold, D. a., & Rhodes, G. (2010). A comparative view of face perception. *Journal of Comparative Psychology*, *124*(3), 233–251. doi: 10.1162/jocn.2009.21401

Lewis, G. J., Shakeshaft, N. G., & Plomin, R. (2018). Face identity recognition and the social difficulties component of the Autism-like phenotype: Evidence for phenotypic and genetic links. *Journal of Autism and Developmental Disorders*, *48*(8), 2758–2765. Retrieved from `http://dx.doi.org/10.1007/s10803-018-3539-4` doi: 10.1007/s10803-018-3539-4

Liebowitz, M. R. (1987). Social phobia. *Modern Problems of Pharmacopsychiatry*, *22*, 141–173. doi: http://dx.doi.org/10.1159/000414022

Luck, S. J. (2014). *An introduction to the event-related potential technique.* MIT press.

Macchi Cassia, V., Valenza, E., Simion, F., & Leo, I. (2008). Congruency as a nonspecific perceptual property contributing to newborns' face preference. *Child Development*, *79*(4), 807–820. doi: 10.1111/j.1467-8624.2008.01160.x

Mandy, W. P. L., Charman, T., & Skuse, D. H. (2012). Testing the construct validity of proposed criteria for DSM-5 autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(1), 41–50. Retrieved from `http://dx.doi.org/10.1016/j.jaac.2011.10.013` doi: 10.1016/j.jaac.2011.10.013

Mansfield, E. M., Farroni, T., & Johnson, M. H. (2003). Does gaze perception facilitate overt orienting? *Visual Cognition*, *10*(1), 7–14. doi: 10.1080/713756671

Mari, M., Bennett, K. M., Scarpa, M., Brighetti, G., & Castiello, U. (1997). Processing efficiency of the orienting and the focusing of covert attention in relation to the level of disability in Parkinson's disease. *Parkinsonism and Related Disorders*, *3*(1), 27–36. doi: 10.1016/S1353-8020(96)00036-3

Materna, S., Dicke, P. W., & Thier, P. (2008). Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: A functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *20*(1), 108–119. doi: 10.1162/jocn.2008.20008

McCormick, E. M., van Hoorn, J., Cohen, J. R., & Telzer, E. H. (2018). Functional connectivity in the social brain across childhood and adolescence. *Social Cognitive and Affective Neuroscience*, *13*(8), 819–830. doi: 10.1093/scan/nsy064

McGugin, R. W., Richler, J. J., Herzmann, G., Speegle, M., & Gauthier, I. (2012). The Vanderbilt Expertise Test reveals domain-general and domain-specific sex effects in object recognition. *Vision Research*, *69*, 10–22. Retrieved from http://dx.doi.org/10.1016/j.visres.2012.07.014 doi: 10.1016/j.visres.2012.07.014

McKone, E., & Robbins, R. (2012). Are Faces Special? In *Oxford handbook of face perception* (pp. 1–45). doi: 10.1093/oxfordhb/9780199559053.013.0009

Mihaylova, M., Stomonyakov, V., & Vassilev, A. (1999). Peripheral and central delay in processing high spatial frequencies: Reaction time and VEP latency studies. *Vision Research*, *39*(4), 699–705. doi: 10.1016/S0042-6989(98)00165-5

Minnebusch, D. A., Suchan, B., & Daum, I. (2009). Losing your head: Behavioral and electrophysiological effects of body inversion. *Journal of Cognitive Neuroscience*, *21*(5), 865–874. doi: 10.1162/jocn.2009.21074

Mondloch, C. J., Le Grand, R., & Maurer, D. (2002). Configuration 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

Morales, M., Mundy, P., Delgado, C. E., Yale, M., Messinger, D., Neal, R., & Schwartz, H. K. (2000). Responding to joint attention across the 6 through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, *21*(3), 283–298. doi: 10.1016/S0193-3973(99)00040-4

Mori, M., MacDorman, K., & Kageki, N. (2012). The Uncanny Valley. *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. Retrieved from http://www.movingimages.info/digitalmedia/wp-content/uploads/2010/06/MorUnc.pdfhttp://ieeexplore.ieee.org/document/6213238/ doi: 10.1109/MRA.2012.2192811

Moriuchi, J. M., Klin, A., & Jones, W. (2017). Mechanisms of diminished attention to eyes in Autism. *American Journal of Psychiatry*, *174*(1), 26–35. doi: 10.1176/appi.ajp.2016.15091222

Morris, J. S. (2002). Differential extrageniculostriate and amygdala responses to presentation of emotional faces in a cortically blind field. *Brain*, *124*(6), 1241–1252. Retrieved from https://academic.oup.com/brain/article-lookup/doi/10.1093/brain/124.6.1241 doi: 10.1093/brain/124.6.1241

Morris, J. S., Friston, K. J., Büchel, C., Frith, C. D., Young, a. W., Calder, a. J., & Dolan, R. J. (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, *121*(1), 47–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9549487 doi: 10.1093/brain/

121.1.47

Morton, J., & Johnson, M. H. (1991). CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychological review*, *98*(2), 164–181.

Morton, J., Johnson, M. H., & Maurer, D. (1990). On the reasons for newborns' responses to faces. *Infant Behavior and Development*, *13*(1), 99–103. doi: 10.1016/0163-6383(90)90008-V

Mundy, P., Sigman, M., & Kasari, C. (1994). Joint attention, developmental level, and symptom presentation in autism. *Development and Psychopathology*, *6*(3), 389–401. doi: 10.1017/S0954579400006003

Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology*, *20*(1), 79–97. doi: 10.1017/S0954579408000047

Nichols, K. A., & Champness, B. G. (1971). Eye gaze and the GSR. *Journal of Experimental Social Psychology*, *7*(6), 623–626. doi: 10.1016/0022-1031(71)90024-2

Nishimura, M., Rutherford, M. D., & Maurer, D. (2008). Converging evidence of configural processing of faces in high-functioning adults with autism spectrum disorders. *Visual Cognition*, *16*(7), 859–891. doi: 10.1080/13506280701538514

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science (New York, N.Y.)*, *349*(6251), aac4716. doi: 10.1126/science.aac4716

Pantelis, P. C., & Kennedy, D. P. (2017). Deconstructing atypical eye gaze perception in autism spectrum disorder. *Scientific Reports*, *7*(1), 14990. Retrieved from `http://www.nature.com/articles/s41598-017-14919-3` doi: 10.1038/s41598-017-14919-3

Papoutsaki, A., Daskalova, N., Sangkloy, P., Huang, J., Laskey, J., & Hays, J. (2016). WebGazer: Scalable webcam eye tracking using user interactions. In *Ijcai international joint conference on artificial intelligence* (Vol. 2016-Janua, pp. 3839–3845).

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Pelphrey, K. a., Viola, R. J., & McCarthy, G. (2004). When strangers pass: Processing of mutual and averted social gaze in the superior temporal sulcus. *Psychological Science*, *15*(9), 598–603. doi: 10.1111/j.0956-7976.2004.00726.x

Pfaff, D., Ribeiro, A., Matthews, J., & Kow, L. M. (2008). Concepts and mechanisms of generalized central nervous system arousal. In *Annals of the new york academy of sciences* (Vol. 1129, pp. 11–25). doi: 10.1196/annals.1417.019

Pick, A. D., & Gibson, J. J. (1963). Perception of Another Person's Looking Behavior. *The American Journal of Psychology*, *76*(3), 386–394. Retrieved from `http://links.jstor.org/sici?sici=0002-9556{\%}28196309{\%}2976{\%}3A3{\%}3C386{\%}3APOAPLB{\%}3E2.0.CO{\%}3B2-8{\%}5Cnhttp://www.jstor.org/cgi-bin/jstor/printpage/00029556/ap050298/05a00050/0.pdf?dowhat=Acrobat` doi: 10.2307/1419779

Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Current Biology*, *19*(4), 319–324. Retrieved from `http://dx.doi.org/10.1016/j.cub.2009.01.007` doi: 10.1016/j.cub.2009.01.007

Posner, M. I. (1980, feb). Orienting of attention. *The Quarterly journal of experimental psychology*, *32*(1), 3–25. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/7367577` doi: 10.1080/00335558008248231

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Attention and Performance*, *32*, 531–556. Retrieved from `http://wknapp.com/class/cog/spring14/posner.pdf` doi: 10.1162/jocn.1991.3.4.335

Proverbio, A. M., Zani, A., & Avella, C. (1996). Differential activation of multiple current sources of foveal VEPs as a function of spatial frequency. *Brain Topography*, *9*(1), 59–68. doi: 10.1007/BF01191643

Quinn, P. C., & Slater, A. (2003). Face perception at birth and beyond. In *The development of face processing in infancy and early childhood: Current perspectives.* (pp. 3–11). Nova Science Publishers. Retrieved from `https://books.google.co.uk/books?id=pzbolUS6lIsC{\&}pg=PA3{\&}source=gbs{\_}toc{\_}r{\&}cad=4{\#}v=onepage{\&}q{\&}f=false`

Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., & Reissland, N. (2017). The human fetus preferentially engages with face-like visual stimuli. *Current Biology*, *27*(12), 1825–1828.e3. Retrieved from `http://dx.doi.org/10.1016/j.cub.2017.05.044` doi: 10.1016/j.cub.2017.05.044

Rhodes, G., Calder, A., Johnson, M., & Haxby, J. V. (2012). *Oxford Handbook of Face Perception.* Oxford University Press. doi: 10.1093/oxfordhb/9780199559053.001.0001

Richler, J. J., Wilmer, J. B., & Gauthier, I. (2017). General object recognition is specific: Evidence from novel and familiar objects. *Cognition*, *166*, 42–55. Retrieved from `http://dx.doi.org/10.1016/j.cognition.2017.05.019` doi: 10.1016/j.cognition.2017.05.019

Ristic, J., Mottron, L., Friesen, C. K., Iarocci, G., Burack, J. A., & Kingstone, A. (2005). Eyes are special but not for everyone: The case of autism. *Cognitive Brain Research*, *24*(3), 715–718. doi: 10.1016/j.cogbrainres.2005.02.007

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(2), 275–285. doi: 10.1037/0096-1523.7.2.275

Rosset, D. B., Rondan, C., Da Fonseca, D., Santos, A., Assouline, B., & Deruelle, C. (2008). Typical emotion processing for cartoon but not for real faces in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(5), 919–925. doi: 10.1007/s10803-007-0465-2

Rossion, B., Gauthier, I., Tarr, M. J., Despland, P., Bruyer, R., Linotte, S., & Crommelinck, M. (2003). The N170 occipito-temporal component is delayed and enhanced to inverted faces but not to inverted objects. *NeuroReport*, *11*(1), 69–72. doi: 10.1097/00001756-200001170-00014

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, *9*(2), 260–264. doi: 10.1037/a0014681

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266. doi: 10.1038/253265a0

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., ... Vogeley, K. (2010). Minds made for sharing: Initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, *22*(12), 2702–2715. doi: 10.1162/joc

Schneier, F. R., Rodebaugh, T. L., Blanco, C., Lewin, H., & Liebowitz, M. R. (2011). Fear and avoidance of eye contact in social anxiety disorder. *Comprehensive Psychiatry*, *52*(1), 81–87. Retrieved from `http://dx.doi.org/10.1016/j.comppsych.2010.04.006` doi: 10.1016/j.comppsych.2010.04.006

Schotter, E. R., Berry, R. W., McKenzie, C. R. M., & Rayner, K. (2010). Gaze bias: Selective encoding and liking effects. *Visual Cognition*, *18*(8), 1113–1132. doi: 10.1080/13506281003668900

Schultz, R. T., Grelotti, D. J., Klin, A., Kleinman, J., Van Der Gaag, C., Marois, R., & Skudlarski, P. (2003). The role of the fusiform face area in social cognition: Implications for the pathobiology of autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 415–427. doi: 10.1098/rstb.2002.1208

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition*, *12*(8), 1474–1496. doi: 10.1080/13506280444000797

Senju, A., & Johnson, M. H. (2009a). Atypical eye contact in autism: Models, mechanisms and development. *Neuroscience and Biobehavioral Reviews*, *33*(8), 1204–1214. doi: 10.1016/j.neubiorev.2009.06.001

Senju, A., & Johnson, M. H. (2009b). The eye contact effect: mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127–134. doi: 10.1016/j.tics.2008.11.009

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), 127–139. doi: 10.1016/j.bandc.2007.12.0001

Senju, A., Tojo, Y., Dairoku, D., & Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(3), 445–458. Retrieved from `http://ovidsp.ovid.com/ovidweb.cgi?T=JS{\&}PAGE=reference{\&}D=emed9{\&}NEWS=N{\&}AN=38380123` doi: http://dx.doi.org/10.1111/j.1469-7610.2004.00236.x

Senju, A., Yaguchi, K., Tojo, Y., & Hasegawa, T. (2003). Eye contact does not facilitate detection in children with autism. *Cognition*, *89*(1), 3–8. doi: 10.1016/S0010-0277(03)00081-7

Shakeshaft, N. G., & Plomin, R. (2015). Genetic specificity of face recognition. *Proceedings of the National Academy of Sciences*, *112*(41), 12887–12892. doi: 10.1073/pnas.1421881112

Shepherd, S. V., Klein, J. T., Deaner, R. O., & Platt, M. L. (2009). Mirroring of attention by neurons in macaque parietal cortex. *PNAS*, *106*(23), 9489–9494.

Retrieved from https://www.pnas.org/content/106/23/9489.short doi: 10.1073/pnas.0900419106

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317–1322. doi: 10.1038/nn1150

Siegal, M., & Varley, R. (2002). Neural systems involved in 'theory of mind'. *Nature Reviews Neuroscience*, *3*(6), 463–471. doi: 10.1038/nrn844

Simion, F., Macchi Cassia, V., Turati, C., & Valenza, E. (2003). Non-specific perceptual biases at the origins of face processing. In *The development of face processing in infancy and early childhood: Current perspectives* (pp. 13–25). Nova Science Publishers. Retrieved from https://books.google.co.uk/books?id=pzbolUS6lIsC{\&}pg=PA13{\&}source=gbs{_}toc{_}r{\&}cad=4{\#}v=onepage{\&}q{\&}f=false

Singh, H., & Singh, J. (2012). Human Eye Tracking and Related Issues: A Review. *International Journal of Scientific and Research Publications*, *2*(1), 2250–3153. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.375.5607{\&}rep=rep1{\&}type=pdf doi: 10.1.1.375.5607

Slater, A., Bremner, G., Johnson, S. P., Sherwood, P., Hayes, R., & Brown, E. (2000). Newborn infants' preference for attractive faces: The role of internal and external facial features. *Infancy*, *1*(2), 265–274. doi: 10.1207/S15327078IN0102_8

Slater, A., Von Der Schulenburg, C., Brown, E., Badenoch, M., Butterworth, G., Parsons, S., & Samuels, C. (1998). Newborn infants prefer attractive faces. *Infant Behavior and Development*, *21*(2), 345–354. doi: 10.1016/S0163-6383(98)90011-X

Sokolowski, K., & Corbin, J. G. (2012). Wired for behaviors: from development to function of innate limbic system circuitry. *Frontiers in Molecular Neuroscience*, *5*, 1–15. Retrieved from http://journal.frontiersin.org/article/10.3389/fnmol.2012.00055/abstract doi: 10.3389/fnmol.2012.00055

Strick, M., Holland, R. W., & van Knippenberg, A. (2008). Seductive eyes: Attractiveness and direct gaze increase desire for associated objects. *Cognition*, *106*(3), 1487–1496. doi: 10.1016/j.cognition.2007.05.008

Stuss, D. T., Gallup, G. G., & Alexander, M. P. (2001). The frontal lobes are necessary for 'theory of mind'. *Brain*, *124*, 279–286. doi: 10.1093/brain/124.2.279

Susan, C., & Rhea, D. (1977). From piecemeal to configurational representation of faces. *Science*, *195*(4275), 312–314. doi: 10.1126/science.831281

Swettenham, J., Condie, S., Campbell, R., Milne, E., & Coleman, M. (2003). Does the perception of moving eyes trigger reflexive visual orienting in autism? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 325–334. doi: 10.1098/rstb.2002.1203

Symons, L. A., Hains, S. M., & Muir, D. W. (1998). Look at me: five-month-old infants' sensitivity to very small deviations in eye-gaze during social interactions. *Infant Behavior and Development*, *21*, 531–536. doi: 10.1016/S0163-6383(98)90026-1

Tanaka, J., & Farah, M. (1995). Parts and wholes in face and object recognition. *The Quarterly Journal of Experimental Psychology*, *46A*(2), 225–245. doi: 10.1080/14640749308401045

Tanaka, J., & Gordon, I. (2012). Features, configuration, and holistic face processing. In *Oxford handbook of face perception* (pp. 177–194). Oxford University Press Oxford. doi: 10.1093/oxfordhb/9780199559053.013.0010

Tanaka, J. W., & Sung, A. (2016). The "Eye Avoidance" hypothesis of Autism face processing. *Journal of Autism and Developmental Disorders*, *46*(5), 1538–1552. doi: 10.1007/s10803-013-1976-7

Tas, C., Luck, S. J., & Hollingworth, A. (2016). The relationship between visual attention and visual working memory encoding: A dissociation between vovert and overt orienting. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1121–1138. doi: 10.1037/xhp0000212

Tatler, B. W., Wade, N. J., & Findlay, J. M. (2010). Yarbus, eye movements, and vision. *i-Perception*, *1*, 7–27. doi: 10.1068/i0382

Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders michael. *Journal of Speech Language and Hearing Research*, *52*(2), 336. Retrieved from `http://jslhr.pubs.asha.org/article.aspx?doi=10.1044/1092-4388(2009/07-0144)` doi: 10.1044/1092-4388(2009/07-0144)

Tipples, J., & Pecchinenda, A. (2018). A closer look at the size of the gaze-liking effect: a preregistered replication. *Cognition and Emotion*, *0*(0), 1–7. Retrieved from `https://doi.org/10.1080/02699931.2018.1468732` doi: 10.1080/02699931.2018.1468732

Tkačik, G., Garrigan, P., Ratliff, C., Milčinski, G., Klein, J. M., Seyfarth, L. H., ... Balasubramanian, V. (2011). Natural images from the birthplace of the human eye. *PLoS ONE*, *6*(6), 1–12. doi: 10.1371/journal.pone.0020409

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(1), 1454–1463. doi: 10.2307/1130423

Trevisan, D. A., Roberts, N., Lin, C., & Birmingham, E. (2017). How do adults and teens with self-declared Autism Spectrum Disorder experience eye contact? A qualitative analysis of first-hand accounts. *PLoS ONE*, *12*(11), 1–22. doi: 10.1371/journal.pone.0188446

Turk-Browne, N. B., Golomb, J. D., & Chun, M. M. (2013). Complementary attentional components of successful memory encoding. *NeuroImage*, *66*, 553–562. Retrieved from `http://dx.doi.org/10.1016/j.neuroimage.2012.10.053` doi: 10.1016/j.neuroimage.2012.10.053

Van Gulick, A. E., McGugin, R. W., & Gauthier, I. (2016). Measuring nonvisual knowledge about object categories: The Semantic Vanderbilt Expertise Test. *Behavior Research Methods*, *48*(3), 1178–1196. Retrieved from `http://dx.doi.org/10.3758/s13428-015-0637-5` doi: 10.3758/s13428-015-0637-5

Vlamings, P. H., Stauder, J. E., Van Son, I. A., & Mottron, L. (2005). Atypical visual orienting to gaze- and arrow-cues in adults with high functioning autism. *Journal of Autism and Developmental Disorders*, *35*(3), 267–277. doi: 10.1007/s10803-005-3289-y

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: Evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174–194. doi: 10.1016/j.neuropsychologia.2006.06.003

Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., & Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, *59*(6), 1037–1050. doi: 10.1016/j.neuron.2008.09.006

Walsh, D. G., & Hewitt, J. (2011). Giving men the come-on: Effect of eye contact and smiling in a bar environment. *Perceptual and Motor Skills*, *61*(3), 873–874. Retrieved from `http://www.amsciepub.com/doi/abs/10.2466/pms.1985.61.3.873` doi: 10.2466/pms.1985.61.3.873

Wang, Y., Newport, R., & Hamilton, A. F. d. C. (2010). Eye contact enhances mimicry of intransitive hand movements. *Biology Letters*, *7*(1), 7–10. Retrieved from `http://rsbl.royalsocietypublishing.org/content/early/2010/04/26/rsbl.2010.0279.short` doi: 10.1098/rsbl.2010.0279

Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: A review of behavioral studies. *Neuroscience and Biobehavioral Reviews*, *36*(3), 1060–1084. doi: 10.1016/j.neubiorev.2011.12.008

Wilcoxon, F. (1946). Individual comparisons of grouped data by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83. doi: 10.1007/978-1-4612-4380-9_16

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., ... Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *Proceedings of the National Academy of Sciences*, *107*(11), 5238–5241. doi: 10.1073/pnas.0913053107

Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, *45*(1), 195–206. doi: 10.1016/j.neuropsychologia.2006.05.009

Winston, J. S., Vuilleumier, P., & Dolan, R. J. (2003). Effects of low-spatial frequency components of fearful faces on fusiform cortex activity. *Current Biology*, *13*(20), 1824–1829. doi: 10.1016/j.cub.2003.09.038

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger Syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders*, *35*(3), 331–335. doi: 10.1007/s10803-005-3300-7

Yamasaki, T., Maekawa, T., Miyanaga, Y., Takahashi, K., Takamiya, N., Ogata, K., & Tobimatsu, S. (2017). Enhanced fine-form perception does not contribute to gestalt face perception in autism spectrum disorder. *PLoS ONE*, *12*(2), 1–18. doi: 10.1371/journal.pone.0170239

Yan, Y., Liu, C., Ye, L., & Liu, Y. (2018). Using animated vehicles with real emotional faces to improve emotion recognition in Chinese children with autism spectrum disorder. *PLoS ONE*, *13*(7), 1–6. doi: 10.1371/journal.pone.0200375

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In *Eye movements and vision.* Springer. doi: 10.1007/978-1-4899-5379-7_8

Yi, D. J., Kelley, T. A., Marois, R., & Chun, M. M. (2006). Attentional modulation of repetition attenuation is anatomically dissociable for scenes and faces. *Brain*

Research, *1080*(1), 53–62. doi: 10.1016/j.brainres.2006.01.090

Yin, R. K. (1969). Looking at upide-down faces. *Journal of Experimental Psychology*, *81*(1), 141–145. doi: 10.1037/h0027474

Young, A. W., & Bruce, V. (2011). Understanding person perception. *British Journal of Psychology*, *102*(4), 959–974. doi: 10.1111/j.2044-8295.2011.02045.x

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*, 1166–1178. doi: 10.1068/p160747n