Autism Spectrum Disorder and Emotion Recognition

Marie Antonia Georgopoulos

Bachelor of Psychology (Hons)

Thesis

Submitted to Flinders University

for the degree of

Doctor of Philosophy

College of Education, Psychology and Social Work

August 2021

Table of Contents

Abstract

While it has been previously suggested that adults with Autism Spectrum Disorder (ASD) have impairments with recognising and reacting to the emotions of others, existing research has been clouded by inconsistent findings. Given the conceivable connection of emotion recognition and reaction with effective social reciprocity and communication, clarity in this area is important for practical, research and clinical understanding.

Key questions remaining unanswered include how ASD and typically developing adults' emotional inferences are influenced by different stimulus types, response formats, and specific emotions or groups of emotions. Questions also exist regarding ASD adults' ability to react to the emotions of others in ways that are considered appropriate. These are issues which the study within this thesis attempted to address.

I examined emotion recognition abilities across three stimulus types (static, dynamic and social), two response formats (free-report and multiple-choice) and 12 emotions (6 basic and 6 complex) in samples of autistic and (*n*= 63) and typically developing (*n*= 67) adults. I also examined individuals' ability to provide appropriate reactions to the emotions of others across the 12 emotions. This study also examined emotion recognition latency and metacognitive monitoring of emotion recognition and reaction responses. Performance across these measures was examined while controlling for individuals' verbal and perspective taking abilities.

The findings showed that regardless of stimulus type, response format and emotion, ASD individuals were less accurate, slower and less confident than typically developing participants in the recognition of emotions. ASD individuals also reacted less appropriately than typically developing participants to others' emotions, and were slower and less confident when doing so, regardless of emotion type. Remarkably, ASD and typically developing groups did not differ at all in terms of their metacognitive awareness of their limitations.

Possible reasons for group differences and implications of any deficits found are discussed, along with limitations and future research directions.

Declaration

I certify that this thesis:

1. does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and

2. to the best of my knowledge and belief, does not contain any material previously published or written by another person except where due reference is made in the text.

This research was supported by an Australian Research Council Discovery grant to Emeritus Professor Neil Brewer and Professor Robyn Young.

....

Marie Antonia Georgopoulos

March 2021

Acknowledgements

It is with great pleasure that I acknowledge those who have been essential to my journey in completing the research contained within this PhD thesis.

I would like to express the deepest appreciation and gratitude to my primary supervisor Emeritus Professor Neil Brewer. Your persistence and perseverance to extract the best version of your students has been captivating. Thank you for your bottomless capacity to understand and have compassion for the many unpredictable speedbumps that have arisen along the way. Truly, you have made this an invaluable learning experience.

Thank you to my co-supervisor Professor Robyn Young for your input and feedback. Your approach has left a lasting impression. Thank you for always having the time to answer any questions that I have had no matter how big or small.

I would also like to thank those in the School of Psychology at Flinders University, particularly Carmen Lucas and Emily Barnett for your support through the years. I would also like to thank Pawel Skuza for your time and expertise in statistics and for taking the time to answer all my questions with complete detail.

Finally, I would also like to acknowledge my son Isaiah who entered our lives at the most perfect (and complicated) time. You have reminded me of why I embarked on this journey. It is with my deepest gratitude that I also thank my family, particularly Mum and Sam for your unconditional support and love. Thank you for being the inspiration and motivation behind all my achievements thus far. If not for you, none of this would be possible. The love and admiration I have for you is endless.

CHAPTER 1

Introduction

How are you feeling, right now, as you begin to read this thesis? Are you curious? Hopeful that you might learn something new? Perhaps excited or happy because it is an area you enjoy or daunted because it is such a big task. Emotions and states of mind such as these are an important part of human nature. They provide us with valuable insights into what we are experiencing and help guide reactions. From early on in life we feel and express our emotions with other people, perhaps even before we can name what those emotions might be. Typically, we become more skilled at recognising and understanding our own emotions and the emotions of others through the developmental trajectory. The smiles of others, the tears and the frowns are just some cues that start to form pieces of the emotional puzzles that reflect what others may be feeling. Being aware of, and able to, infer other people's emotions and mental states is thought to be critical within the social realm. However, what if development of this key social component does not follow the typical trajectory—and the cues that help us to infer what others may be feeling and expressing go unnoticed or are difficult to interpret?

It is commonly thought that those with Autism Spectrum Disorder (ASD) possess some degree of impairment in recognising, reacting to and expressing emotions (Harms et al., 2010; Uljarevic & Hamilton, 2013). ASD is the term used to describe the variable presentation of this pervasive developmental disorder. It captures many of the individuals previously diagnosed with Autistic Disorder, Asperger's Syndrome and Pervasive Developmental Disorder-Not Otherwise Specified in accordance with earlier versions of the Diagnostic and Statistical Manual of Mental Disorders (DSM). These neurodevelopmental disorders are known to be characterised by difficulties in social communication, social interaction and restricted and repetitive behaviour or interests (American Psychiatric

Association [APA], 2013; Faras et al., 2010). Individuals diagnosed with an ASD who do not display global intellectual impairments coupled with a history of language delay are sometimes referred to as high-functioning (Philip et al., 2010). It is not uncommon for those with high-functioning ASD to describe social communication and interaction as the most debilitating aspect of their condition (Philip et al., 2010). What underpins the lifelong social difficulties that characterise individuals with ASD is unknown, but atypical facial emotion recognition and reaction to others' expressions of emotions are thought to play a role.

Facial emotion recognition refers to the ability to infer emotional states or states of mind from facial configurations. Throughout the human lifespan faces are ubiquitously enmeshed in everyday life, whether it be through a passing smile or an in-depth conversation. In recent times, exposure to different online social media platforms (e.g., Facebook, Instagram and Snapchat), dating and geosocial networking platforms (e.g., Tinder, eharmony and Bumble), video sharing platforms (e.g., Youtube) and professional career platforms (e.g., LinkedIn) have not only diversified the way facial representations are delivered but have also increased exposure to a plethora of uniquely presented facial configurations. Faces have thus become the front cover of not only our face-to-face interactions but also our online personas. The face is arguably the most important source of information within social interactions, playing a significant role in the discernment of, for example, who is guilty, or perhaps who we should trust, love or help (Barrett et al., 2019; Todorov, 2017; Zebrowitz, 1997, 2017; Zhang et al., 2018). The ability to appropriately infer and react to the facial emotions of others is essential for participation in everyday social life. However, the potential for inaccurate interpretation of the emotions of others, and atypical reactions to those emotional expressions, makes adults with ASD vulnerable to negative social outcomes. These difficulties may have implications for an array of contexts in one's life, including social interactions at work and in one's private life.

It is not surprising, therefore, that facial emotion recognition has been an important focus in research exploring social difficulties in individuals with ASD. The primary objective of such research has been understanding how people with ASD process social cues, particularly from faces. However, despite the large body of literature that has accumulated in over 70 years since the first description of autism by Kanner (1943), empirical studies of face emotion processing by adults with ASD have yielded contradictory results, raising questions about whether purported difficulties persist into adulthood and, if so, the extent, nature and variability of any difficulties.

The possible links between difficulties in emotion recognition and ASD are foreshadowed by the diagnostic criteria for ASD. Current diagnostic criteria for ASD within the American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5; APA, 2013) include, in part, the following:

> A. Persistent deficits in social communication and social interaction across multiple contexts, as manifested by the following, currently or by history (examples are illustrative, not exhaustive, see text):
>
> 1. Deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation; to reduced sharing of interests, emotions, or affect; to failure to initiate or respond to social interactions.
> 2. Deficits in nonverbal communicative behaviors used for social interaction, ranging, for example, from poorly integrated verbal and nonverbal communication; to abnormalities in eye contact and body language or deficits in understanding and use of gestures; to a total lack of facial expressions and nonverbal communication.

3. Deficits in developing, maintaining, and understanding relationships, ranging, for example, from difficulties adjusting behavior to suit various social contexts; to difficulties in sharing imaginative play or in making friends; to absence of interest in peers (DSM-5; APA, 2013, p. 50).

Interestingly, although difficulties with inferring emotional states or identifying emotional cues are not explicitly referred to in the diagnostic criteria, they are implied. As explained by Herba and Phillips (2004), emotion processing can be described as consisting of processes within three broad domains: (1) the identification of emotion cues, (2) the production of an emotional state, and (3) the regulation of that emotional state and behaviour. To meet the diagnostic criteria for ASD as detailed in the DSM-5 and described above, evidence of difficulties in all three domains as described by Herba and Phillips (2004) are required[1]. In this context, it is commonly thought that those with ASD have difficulties in all three emotion processing domains, with impairment seen in the recognition of emotions, engagement in emotional reciprocity and the formulation of appropriate reactions. Recognising emotional cues, particularly from facial configurations, is a fundamental component of effective social cognition (Enticott et al., 2014) whereby the reciprocity of emotional inference and reacting to emotions plays a vital role in social interactions and interpersonal communication (Marinetti et al., 2011). Such deficits in emotional inference and reaction form, in part, the basis for diagnosis and constitute some of the core characteristics widely reported in people with ASD (Baron-Cohen, 1995), including the supposed lack of empathy conveyed in social interactions. As summarised here, and described in detail later, despite these well-known characteristics and stereotypical

[1] Not to be confused with the two main symptom domains described within the DSM-5 of deficits in 1) social communication and interaction and 2) restricted, repetitive patterns of behaviours (APA, 2013).

conclusions about ASD individuals' abilities in the recognition of and reaction to the emotions of others, researchers have not consistently demonstrated these difficulties within adults. The results from some studies suggest emotion recognition outcomes in persons with ASD that are worse than typically developing individuals (e.g., Gaigg, 2012; Harms et al., 2010; Uljarevic & Hamilton, 2013). Many other studies report opposing findings, with individuals with ASD performing in a similar manner to typically developing individuals (e.g., Castelli, 2005; Evers et al., 2014; Gepner et al., 2001; Grossman et al., 2000; Rosset et al., 2008; Rutherford & Towns, 2008), thus making it difficult to draw conclusions with certainty about the ability of adults with ASD to recognise and react to others' emotions.

It is unsurprising that face processing has been studied in individuals with ASD given their well-documented difficulties during childhood in processing or responding to social information (Cassidy et al., 2014). Facial emotional recognition is a key focus of such research because it is assumed to be the main aspect of face processing involved in social interactions. It is suggested that difficulties in using or understanding facial information are impaired in persons with ASD to a degree that may improve through the lifespan, without reaching typicality (Dziobek et al., 2006; Sucksmith et al., 2013). Elaborated further (briefly here, and in detail in later sections), the differences in accurately inferring basic emotions between those with ASD and typically developing persons show some evidence of narrowing in adult groups relative to child groups. However, difficulties again appear somewhat more marked in adults with ASD in recognising more complex and subtle emotions, and emotion blends, than is seen in age-matched typically developing individuals (Greimel et al., 2010; Humphreys et al., 2007; Law Smith et al., 2010). This pattern may suggest that adults with ASD understand and recognise basic emotions (i.e., happiness, sadness, anger, fear, disgust and surprise) but not more complex emotions (e.g., shame, pride and guilt).

Explanations of the reduction in, but persistence of, difficulties of emotion recognition into adulthood have originated from the suggestion of compensatory strategies that likely develop through the lifespan. For example, social processing is generally thought of as an implicit process that occurs automatically and rapidly. However, research with individuals with ASD has shown that emotion recognition is a more deliberative or effortful process and, thus, uses more cognitive resources (Harms et al., 2010; Kaland et al., 2011; Kliemann et al., 2013; Senju, 2012). This has been supported by behavioural and neuroimaging research which has demonstrated that individuals with ASD use more effortful compensatory strategies in decoding facial emotions, including using rule-based processing (Rutherford & McIntosh, 2007) and atypical fixation on faces (Sasson et al., 2007). Through the lifespan these compensatory processes are likely improved on, but their lack of efficiency hinders overall accuracy. To better understand underlying mechanisms for emotion recognition difficulties, it is imperative to assess what is currently known about these difficulties in ASD individuals.

Given the inconsistent research findings on emotion recognition difficulties in ASD individuals, efforts to disentangle findings have led to several meta-analyses and comprehensive reviews of research in the area (see, for example, Lozier et al., 2014; Nuske et al., 2013; Uljarevic & Hamilton, 2013). These meta-analyses and reviews have reached rather different conclusions, suggesting either poorer emotion recognition in those with ASD compared to typically developing controls (Uljarevic & Hamilton, 2013; Lozier et al., 2014) or that impairments are not universal in ASD as evidenced by variability across participants and testing methodologies (Nuske et al., 2013). Further to this, a literature review by Begeer et al. (2008) concluded no differences between ASD subjects and typically developing controls when matched for intelligence, leading to the conclusion that differences in previous studies are not attributable to ASD directly but rather to the intellectual disability that often accompanies ASD. Another literature review concerned with facial emotion recognition in persons with ASD was conducted by Harms et al. (2010). Their review included the assessment of 65 studies and reported that, overall, there were slightly more behavioural studies that demonstrated difficulties in facial emotion recognition than studies that did not. Interestingly, as described in the review, most psychophysiological studies (i.e., eye-tracking, neuroimaging and electrophysiological: event-related potentials) reported group differences with difficulties in facial emotion recognition more likely in those with ASD. Harms et al. (2010) proposed that perhaps mixed findings in past behavioural studies may have been a result of some methodologies being insensitive to subtle group differences in atypical emotion processing mechanisms.

Thus, these meta-analyses and reviews highlight the inconsistent research findings. However, they also highlight several potentially important determinants of the inconsistent findings. A primary methodological concern with previous research is sample size. Uljarevic and Hamilton (2013) noted that relatively small sample sizes (of studies included within their meta-analysis, in one case, as few as 5 participants with ASD) plague currently available research. Additionally, they reported that only 15 of the 48 studies that met their inclusion criteria included more than 20 participants. Thus, it is possible that many available studies are underpowered. Uljarevic and Hamilton (2013) noted that a power analysis indicates that 35 participants in each group (i.e., participants with ASD and typically developing controls) are required to obtain a power of 0.95 to detect even a large group difference (effect size=0.8).

Stimulus type and response format have been identified as two further potential influences on inconsistent results of past research (Uljarevic & Hamilton, 2013). For similar reasons, these two methodological factors have been discussed as potentially influential, primarily due to there being two main variants of each (i.e., static and dynamic stimulus types and multiple-choice and free-report response formats) which have been used in emotion recognition research with ASD individuals. The choice of which variant of each is used potentially speaks to ecological validity as well as providing opportunity for enhanced or hindered emotion recognition. For example, dynamic stimulus types have been suggested to provide participants with an opportunity to use context (as is possible in normal social interactions) to better recognise emotions. Similarly, multiple-choice response formats have been discussed in relation to assisting ASD individuals to use compensatory methods to decipher emotional inferences. The effects of stimulus type and response format are discussed in detail in subsequent sections.

Effects of Stimulus Type

It has been suggested that the heterogeneous findings in the current literature may in part be due to the variation in stimulus presentation (Harms et al., 2010; Nuske et al., 2013). The meta-analyses note the progression from the use of static face emotion stimuli to the use of dynamic stimuli designed to better match the demands of everyday life.

Static Tests of Emotion Recognition

Static images (i.e., drawings or photographs of faces) have been used to test facial emotion recognition in individuals with ASD. The "Reading the Minds in the Eyes" task is one of the most commonly used tests for facial emotion recognition used in ASD research. The test involves showing participants pictures of eye expressions and requiring them to make a choice from a menu of emotions of a response they believe represents the emotion the image is depicting (Baron-Cohen et al., 1997). The eye region has been of interest as it is the accepted location for high emotional information to be conveyed. The results of several studies demonstrate that individuals with ASD are less capable of accurately matching the correct emotion with eye expressions (e.g., Baron-Cohen et al., 2001; Franco et al., 2014). This test has also been adapted to include images of mouth expressions and full-face expressions, aiding exploration of the notion that persons with ASD display atypical gaze patterns, characterised by reduced fixation on the eye regions and increased fixation on the mouth region (Black et al., 2017, 2020; Harms et al., 2010). Some studies using full-face images (e.g., Wilson et al., 2012) showed children with ASD were less likely to accurately match the correct emotions when compared to typically developing participants. In the same study, it was noted that difficulties persisted when the mouth and eye regions were isolated. These difficulties are thought to contribute to communication and emotional intelligence impairments. Although the results of some studies that used a recognition paradigm comprising basic emotions (i.e., angry, afraid, disgusted, happy, sad, surprise) show differences between children with ASD and typically developing children in emotion recognition, there have been many similar studies which reported ceiling effects or failed to find differences (e.g., Adolphs et al., 2001; Kuusikko et al., 2009; Rump et al., 2009).

Similar trends have been reported in emotion recognition research involving adults with ASD. Emotion recognition studies comprising basic emotions have demonstrated deficits, as illustrated by Eack et al.'s (2015) finding that adult participants with ASD had difficulties distinguishing emotional from neutral facial expressions. Parallel to the trend in the child literature, although some studies showed adults with ASD were impaired in recognising basic emotions, other studies have failed to report such findings or have found ceiling effects (e.g., Ashwin et al., 2006; Castelli, 2005; Evers et al., 2014; Sucksmith, et al., 2013). Such discrepancies in the literature involving adults has led to postulations about the effects of age and developmental level on the recognition of basic emotions.

The ceiling effects reported within emotion recognition tasks involving basic emotions has led to a focus on recognition of more complex emotions. It has been suggested that, over their developmental trajectory, adults with ASD within the normal range of intelligence (i.e., IQ > 85) have developed skills required to pass emotion recognition tasks involving static presentations of basic emotions. It is important to highlight that some degree of inconsistency within the literature is expected given that ASD refers to a spectrum characterised by variable presentations. However, the use of more complex emotions may lead to the detection of more marked difficulties, at least in some ASD participants.

Static stimulus materials that include more complex emotions (Baron-Cohen et al., 1997) or emotions of lower intensities (Wong et al., 2012) have revealed some evidence for more marked impairments compared to basic emotions or emotions presented at higher intensities in those with ASD than in typically developing individuals. Specifically, studies with adolescents and adults with high functioning ASD have demonstrated that, although from about 12 years of age they can perform similarly to typically developing individuals in tests of basic emotions (e.g., Capps et al., 1992; Grossman et al., 2000), impairments persist in tests requiring recognition of complex emotions (e.g., Humphreys, et al., 2007). For example, a study conducted by Heerey et al. (2003), involving 8-15 year old children (25 children with high-functioning ASD and 21 typically developing children) found that participants with ASD showed difficulties in tests involving self-conscious emotions (e.g., shame and embarrassment) but did not differ on tests of basic emotions.

It is important to note, however, that Uljarevic and Hamilton's (2013) meta-analysis did not find any effects of age on emotion recognition abilities. This does not, of course, rule out the possibility that there may be person-specific patterns of accuracy improvement through the lifespan; rather, it indicates no significant changes in emotion recognition accuracy in ASD samples, at the group level. Therefore, the magnitude of any deficits in basic and complex emotions may fluctuate over the development trajectory, but the existence of any difficulties is likely to remain irrespective of age. Thus, even into adulthood, as the complexity of the emotion increases so should any difficulty in correctly inferring others' emotions.

Further, studies investigating emotion recognition have found deficits when tasks include stimuli that are more subtle (Humphrys et al., 2007), such as emotions expressed in just one part of the face (Gross, 2004) or only presented for a short amount of time (Kliemann et al., 2010). Consequently, it has been argued that, to assess subtle impairments in emotion recognition in adults with ASD, it is important to develop and use materials that mimic the demands of everyday life whilst maintaining experimental control (Cassidy et al., 2014).

Dynamic Tests of Emotion Recognition

Many studies of facial emotion processing in those with ASD have used static instead of dynamic (video clips and/or moving images of faces) material. The use of photographs and other still-displays may not capture the realistic nature of facial expressions encountered in everyday situations. Given that naturally occurring emotional expressions are dynamic in nature it has been argued that dynamic (i.e., moving) depictions of faces may have greater ecological validity than static stimuli because they are likely to integrate more life-like characteristics, such as body language, verbal cues and contextual cues (Enticott et al., 2014, Johnston et al., 2010). Thus, it is unsurprising that research evolved to include dynamic stimuli (Golan et al., 2008; Loveland et al., 2008; Loveland et al., 1997). One study that used dynamic stimuli to assess emotion recognition in adults with ASD involved the use of "The Awkward Moments Test" (Heavey et al., 2000). This test comprises seven short clips of social situations taken from television advertisements. Participants were required to predict the mental states of the characters within the scenes by being asked to state what the characters were feeling. As explained within the study, this task required the participants to take the perspective of others to accurately infer the emotion or mental states that the characters were expressing. The results showed that, at the group level, participants with ASD performed at significantly lower levels of accuracy compared to the comparison group.

This study suggests that group differences in emotion recognition in higher functioning adults with ASD persist even when using more naturalistic stimuli. Other studies have used materials adapted from static tests to assess any differences in emotion recognition in participants with ASD. Golan et al. (2006) adapted the static Reading the Mind in the Eyes task into a dynamic version, the “Reading the minds in films” task. The study included 22 adults aged 17-52 years who were diagnosed with Asperger syndrome or higher functioning autism, as well as 22 participants from the general population who formed the comparison group. Participants were required to view a series of 22 short scenes and match the characters’ emotions and mental states from a list. The emotions and mental states included within this study can all be categorised as complex emotions. Mirroring the results found with the static version, the results showed that those with an ASD were significantly less accurate than the comparison group. Interestingly, this study also reported positive correlations between verbal intelligence and task scores. The significant differences between groups in the study supports other studies’ findings (Heavey et al., 2000; Klin et al., 2002).

Of the studies conducted using adult samples of persons with ASD, some have compared the use of dynamic with static stimuli (e.g., Philip et al., 2010). Much like trends previously described in the above sections, there are inconsistencies regarding the effect of using dynamic stimuli, with some studies finding that dynamic stimuli facilitate recognition of facial emotion displays (e.g., Ambadar et al., 2005) whereas other studies have reported minimal, and even no, advantage of using dynamic over static stimuli (see, for example, Enticott et al., 2014). Nevertheless, for reasons of ecological validity, continued use of dynamic stimuli seems warranted.

In daily social interactions, multiple external cues are available to compensate for any difficulties with facial emotion recognition; these include body language, verbal information and contextual or situational information. Therefore, it seems plausible to hypothesise that movement and added contextual information (if dynamic displays include social scenes) would aid in using skills and cues learnt through the lifespan to infer emotions more accurately. Emotion recognition is undoubtedly complex given the diverse range of emotions that one can select to describe a facial configuration. It is further complicated by crossovers in cues or signals. For example, crying may be a prototypic cue for sadness, but it can also be a cue for considerable happiness or frustration. Context is essential to be able to distinguish which emotion would be most accurate in describing the facial configuration displaying crying. In real life situations, there are a variety of contextual cues which are available for individuals to employ to assist in deciphering facial configurations (Barrett et al., 2019). Therefore, not only simply dynamic depictions of facial emotions (i.e., moving faces) but also social scenes with context would be helpful in gaining more insights into emotion inference abilities in individuals with ASD.

It has been suggested that perhaps persons with ASD may have adapted and learnt individual means and compensatory strategies and unique cues to recognise emotions that are not adequately captured by static images (Cassidy et al., 2015; Chevallier et al., 2015). Thus, dynamic stimuli might provide more of an opportunity to infer others' emotions accurately. An interesting but opposing suggestion is that individuals with ASD find it more difficult to recognise emotions from dynamic stimuli. A few studies using dynamic stimuli have reported that those with ASD are better able to infer emotions when video stimuli are slowed down (e.g., Gepner et al., 2001; Tardif et al., 2007). In other words, dynamic stimuli which better represent the way emotional cues are presented in real-life are recognised less accurately unless the speed at which the emotion is represented is reduced, thus reducing the degree to which the stimuli mimic the demands of real-life. Furthermore, Roeyers et al. (2001) found that adults with ASD could more accurately infer complex emotions from static photos compared to dynamic video depictions of a social interaction. Difficulties in processing dynamic stimuli could be explained by persons with ASD missing important social cues. This could be attributed, in part, to a reduced tendency to focus on socially relevant information, as evidence by patterns of atypical gaze such as looking away from people in social situations or focusing on regions of the face such as the mouth instead of the eye region as typically expected (Fletcher-Watson et al., 2009; Klin et al., 2002; Speer et al., 2007). However, available research does not necessarily support either of the above suggestions. The trends that have emerged in research adopting dynamic stimuli have mimicked those found using static stimuli for individuals with ASD. There is little agreement about whether a deficit in emotion recognition is present and, more specifically, little is known of the magnitude of any difficulties. There are, however, parallels between research using static versus dynamic stimuli in that basic emotions appear to be more consistently recognised than more complex and subtle emotions by adults with ASD (Cassidy et al., 2015; Chevallier et al., 2015).

Summary

The preceding discussion regarding the effects of stimulus type highlights the inconsistencies in the empirical findings, to the extent that little can be confidently concluded about the impact of stimulus type on recognition performance in ASD and typically developing adult samples. Briefly, the precise effects of stimulus type are unresolved, with research involving both types of stimuli yielding similar results (i.e., some studies show difficulties and others fail to do so, as well as the recognition of complex emotions showing some evidence of more marked difficulties in those with ASD). There have been some important yet conflicting propositions regarding the use of dynamic stimuli which remain largely unanswered, such as whether it will yield more or less accurate emotion inferences. Therefore, it would be beneficial to have future research include both types of stimuli within the one study to allow comparisons. It would also be important to assess the role of context, comparing static stimuli with dynamic stimuli without context and dynamic stimuli with context (e.g., social scenes which I refer to subsequently as social stimuli).

Effects of Response Format

An important consideration when assessing available research is the nature of the emotion recognition response that was required of the participant: specifically, was the response format multiple-choice or free-report. Multiple-choice is a characteristic of a response paradigm where a small list of alternative emotions is provided along with a stimulus and participants are required to choose the response from that list that best describes the emotion (i.e., what is often referred to as multiple-choice). Conversely, free-report response paradigms require participants to freely label the facial configurations presented at each trial without being provided with a list of emotion words or other cues (Betz et al., 2019).

As noted in Lozier et al.'s (2014) meta-analysis of emotion recognition studies with ASD individuals, the majority of studies have employed multiple-choice paradigms. This begs the question as to whether difficulties that may characterise individuals in real-world contexts are captured by multiple-choice paradigms. Researchers (see, for example, Barrett et al., 2007; Gendron et al., 2013) have highlighted that the presentation of words (i.e., emotion label options), coupled with participants' conceptual knowledge of these words, can inadvertently provide a basis for achieving higher accuracy than is reflective of their actual abilities (Betz et al., 2019). For example, multiple-choice paradigms provide prompts to choose an answer that might not have spontaneously been thought of within a free-report paradigm condition. In further support of this are studies of emotion recognition demonstrating that free-report response paradigms markedly reduce accuracy scores of typically developing participants through the elimination of access to emotion concept cues and knowledge that word options provide (for a review, see Barrett et al., 2019).

Furthermore, multiple-choice paradigms provide perfect opportunities for participants to apply alternative methods for “working out” facial emotion configurations by employing strategies such as ‘process-of-elimination’ (Cassels & Birch, 2014). This is an important consideration, particularly in individuals suspected of having difficulties in emotion recognition. Using multiple-choice paradigms, no evidence may be found for a deficit, but it may be due to the reliance on compensatory processes that are able to circumvent deficits that would be otherwise present in free-report tasks or in real life (Cassels & Birch, 2014).

Given the possibility that response format may be an important influence on emotion recognition accuracy, it remains an important question for investigation within an ASD sample and may also help explain inconsistencies within the current literature. As noted earlier, the variable presentation of ASD will likely result in not all individuals employing or having developed effective compensatory strategies to use in a multiple-choice paradigm.

The use of free-report paradigms is also essential in working toward greater ecological validity within empirical research. In real life situations we are not confined in our choices when inferring the emotions of others, although it is likely that context will help narrow the potentially broad array of response options. Rather, there is a demand for spontaneous and rapid identification of emotions based predominantly on facial emotional configurations. The inclusion of both paradigms in one study comparing accuracy of emotion inference would be ideal within an ASD population to allow comparison between the two response types. This would help gauge whether observed differences (if any) are, in part, a result of strategy employed to decipher emotions or differences in accuracy.

Effects of Emotion

To better understand potential difficulties that individuals with ASD may experience in recognising the emotions of others, it is important to explore how typically developing individuals identify emotions from others’ facial configurations almost effortlessly. One of the main explanations suggests that emotions are recognised via 'bottom-up processing,' also known as stimulus-driven processing whereby emotions are recognised through visual stimuli and do not require previous knowledge or learning (Cook et al., 2012; Maekawa et al., 2011). It is suggested that a core set of facial expressions (i.e., basic emotions) are recognised instinctually, with this ability being innate (Ekman & Friesen, 1969, 1975). Studies investigating this explanation have identified six basic emotions: happiness, sadness, fear, anger, surprise, disgust (Ekman & Friesen, 1969, 1975), with all other emotions being labelled as complex emotions.

During establishment of the emotions that were to be categorised as 'basic', it was argued that the identified set of emotions could be accurately recognised across cultures, suggesting that, regardless of different social experiences, the recognition of this set of emotions appears to be biologically programmed (Ekman & Friesen, 1969, 1975; Frank & Stennett, 2001). In support of this explanation, Ekman et al. (1987) reported that individuals from ten different cultural backgrounds showed strong agreement on identifying the emotion that represented the most displayed facial expression in a set. They were required to provide a primary answer as well as identify a second most likely emotion. The participants showed strong agreement on both judgments, suggestive of innate recognition irrespective of social or learning experiences. However, this theory of innately driven emotion recognition has been challenged by the findings of other studies (e.g., Aviezer et al., 2008; Righart & Gelder, 2008) suggesting, for example, that recognition of basic emotions can be manipulated by pairings with incongruent contexts (e.g., a face expressing fear in the context of an anger invoking situation), where the emotions chosen would describe the situation not the facial configuration presented (Carroll & Russell, 1996). Despite this conflict in findings, emotion recognition research still widely promotes the idea that emotions can be categorised as either basic or complex. It is generally thought that complex emotions are more difficult to recognise as they are comprehended in more culturally specific ways, requiring abilities in perspective taking (i.e., an understanding of other people's thoughts, beliefs or mental states; Golan et al., 2008).

Considering emotion recognition involving individuals with ASD, emotion type (i.e., basic or complex) could be an important moderator of the relationship between emotion inference accuracy and ASD over the lifespan. While not directly tested through longitudinal studies, studies involving children and adults separately suggest that persons with ASD become more proficient at recognising basic emotions as they age, but even in adulthood continue to struggle with more complex and subtle expressions of emotion. This raises important questions about emotion recognition abilities of adults in more challenging tasks using complex emotions.

Focus on emotion type arose in an attempt to explain the inconsistent results of past research, guided by the possibility that facial emotion recognition in individuals with ASD might be emotion type specific rather than indicative of generalised difficulties (Song & Hakoda, 2017). Most emotion recognition studies involving individuals with ASD focus on the six basic emotions (i.e., happiness, sadness, fear, anger, surprise, disgust). Some studies have reported that persons with ASD have difficulties in recognising only specific basic emotions. For instance, the emotions of surprise (e.g., Bormann-Kischkel et al., 1995), sadness (e.g., Boraston et al., 2007) and fear (e.g., Ashwin et al., 2006; Corden et al., 2008; Howard et al., 2000) have been identified as the most difficult to recognise basic emotions in those with ASD. Further to this, some have suggested that ASD individuals appeared to ignore, or be less accurate at interpreting, negative basic emotions (but not positive emotions) than typically developing individuals (e.g., Ashwin et al., 2006; Enticott et al., 2014; Whitaker et al., 2017). However, other studies have failed to find impairments in negative emotions (e.g., Lacroix et al., 2009) or have also found deficits in the inference of positive basic emotions such as happiness (e.g., Humphreys et al., 2007). Other interesting patterns have emerged in emotion recognition studies involving ASD samples. For example, Eack et al. (2015) found that persons with ASD are more likely than typically developing individuals to have difficulty distinguishing emotional from neutral facial expressions. Furthermore, ASD subjects were reported to have misinterpreted happy facial configurations for non-emotional (neutral) expressions as well as being significantly more likely to attribute negative valence to non-emotional facial configurations.

Several studies investigating error patterns in the recognition of basic emotions in individuals with ASD did not include a comparison group of typically developing individuals (e.g., Humphreys et al., 2007; Jones et al., 2011; Wallace et al., 2011). Without the inclusion of a comparison group it is difficult to determine if reported impairments are associated with ASD or if extraneous experimental variables (e.g., stimulus types, response formats) can account for difficulties in emotion recognition abilities.

The few studies that used complex emotions with adults and children with ASD appeared to more consistently demonstrate difficulties with emotion recognition (e.g., Baron-Cohen et al., 2001; Golan et al., 2006; Rutherford et al., 2002), although inconsistencies in the reports of any difficulties remain. Additionally, it is unclear whether error patterns in the recognition of specific complex emotions reflect the variability in the complex emotions included due to there being no uniform specified set of complex emotions.

In sum, little can be confidently concluded about the impact of emotion type due to inconsistencies clouding empirical findings as well as a gap in research involving adults using complex emotion recognition tasks. Therefore, investigations into emotion recognition difficulties in adults with ASD would benefit from assessments of emotion types (i.e., basic and complex) to better understand error patterns. It is important for future studies to investigate ASD adults' abilities in recognising basic and complex emotions to gain a better understanding of any difficulties and to assist with disentangling the inconsistencies of past research. There are also potential practical applications as this information can help guide future programs teaching emotion recognition skills as part of social skills training. Should specific emotion types be found to be more challenging for adults with ASD to infer from facial stimuli, there may be merit in developing programs that focus on specific emotion types that are the most challenging to recognise.

Reacting to the Emotions of Others

What might the consequences be if there are impairments in how we understand how our treatment toward others is affecting them? Or how we should react to others' treatment of us? It is undisputed that emotion processing impairments or atypicalities are likely to be detrimental to social interactions as accurately inferring others' emotional states allows for the formulation of appropriate reactions (Brewer et al., 2016) Thus, not only is it important to assess emotion recognition but it is also important to evaluate abilities in reacting to the emotions of others to gain a comprehensive insight into the emotion processing of individuals with ASD. Understanding the processing of emotions within a social context will better assist in gauging where true deficits in emotion processing exist.

Symptoms or traits of ASD often include difficulties in areas such as the following: social-emotional reciprocity, abnormal social approach, difficulties with back-and-fourth conversations, reduced sharing of interests, emotions and affect, as well as difficulties with initiation of conversation and reacting to others within social interactions (APA, 2013; Kern Koegel et al., 2016). For those with ASD, a lack of production of these behaviours and difficulties with social communication generally may interfere with the development of social relationships. From what is known about emotion recognition difficulties and challenges in social conversation in individuals with ASD, it is possible that the inability to accurately recognise emotions may hinder abilities to react appropriately. Explained further, difficulties in emotion recognition could, for example, impact the expression of empathy (Baron-Cohen & Wheelwright, 2004; Golan et al., 2006), perhaps explaining some of the apparent lack of empathy often attributed to those with ASD. Difficulties in expressing empathy have reportedly been found to limit social interactions and lead to difficulties understanding and expressing interests to others as well as maintaining relationships (Baron-Cohen & Wheelwright, 2004; Kern Koegel et al., 2016; Laugeson et al., 2009). Empathy is a broadly defined and complex construct which is generally accepted to refer to one's reaction to the perceived experience of others (Davis, 1980; Kern Koegel et al., 2016). Therefore, to convey empathy to others, one must accurately recognise emotional cues, interpret them and then formulate appropriate reactions (Hill, 2009). Song et al.'s (2019) meta-analysis assessing empathy impairments in children with ASD revealed that, overall, children with ASD showed empathy impairments relevant to typically developing children.

Difficulties with expressing empathy have also been found in adults with ASD (e.g., Baron-Cohen & Wheelwright, 2004; Rogers et al., 2007; Shamay-Tsoory et al., 2002). However, there appears again to be less research involving adults. It may be that, given the evidence that difficulties in emotion processing appear to continue through the lifespan (as discussed previously), the findings of research involving children are being relied on to inform about difficulties in reacting appropriately to emotions in those with ASD generally (i.e., children and adults). Unfortunately, these studies say little about the potential for development or improvement of empathy-related behaviours over time (i.e., if differences in reactions to others' emotions between children and adults were found, this may suggest improvement with age). In a longitudinal study conducted by Dissanayake et al. (1996) which investigated the stability of reactions to examiners' distress, empathy ratings from the initial time-point (preschool) predicted empathic reacting to similar emotional displays 5 years later. This suggests that difficulties in reacting to others' emotions are likely to be maintained, suggesting the possibility that what is known from research investigating empathic reacting in children with ASD may be relevant to adults with ASD. However, this cannot be known for sure and is a limitation of the relevant literature.

Expressing empathy within social contexts is imperative for effective social functioning. Locke et al. (2013) reported that, although rates of reciprocal friendships for children with ASD and non-ASD disability were similar early in the school year to those of typically developing children, reciprocal friendship rates of the former samples decreased compared to typically developing children over the course of the school year. Locke et al. (2013) concluded that, despite typically developing children nominating their peers with ASD and non-ASD disabilities as friends at the beginning of the school year, they appeared less likely to do so as the year progressed and they became more aware of social differences between typically developing children and those with social difficulties (i.e., ASD and non-ASD disabilities). These patterns in negative social outcomes are likely to persist through the lifespan. For example, Howlin et al. (2004) conducted a longitudinal study exploring adult outcomes for children with Autism which included 68 individuals with ASD (children with an average age of 7 years, and followed up over a 29-year period with the follow-up average age being 29 years). They reported findings for a range of areas with particular focus on the following: educational attainments, employment, friendships, independent living, overall social outcome and Autism-related problems. Howlin et al. (2004) noted that the majority (i.e., 38 participants or 56%) reported having no friends or acquaintances, and only three of the participants had married (one of whom later divorced).

I could not locate any behavioural studies which have directly explored ASD adults' reactions to other people's emotions and mental states. However, Kern Koegel et al., (2016) described a study that provides some relevant information. They assessed the effectiveness of intervention techniques, involving video feedback, designed to improve ASD adults' abilities in empathic reacting, specifically targeting the production of empathic listening statements and questions that would convey empathy during conversations with others. Participants were required to participate in three phases (i.e., baseline, intervention, follow-up). A baseline session was followed by intervention sessions which occurred once per week for different lengths of time for each of the three participants (i.e., 5, 6 and 9 weeks). A follow-up session was completed one month after the intervention phase for all participants, with a subsequent follow-up session conducted with two of the three participants two years after the intervention phase. At baseline, it was found that participants had difficulty communicating empathic listening statements and they reacted with few empathic questions. They reported an improvement in the expression of empathy through statements and initiating questions in all three participants following intervention, with improvements maintained at both follow-up sessions. They also noted gains for participants in the general level of empathy expressed and confidence in their social communication abilities. These findings are at least consistent with stereotypical descriptions of characteristics of social communication in adults with ASD. However, the study leans more to highlighting the teachablity of reacting skills with an ASD sample rather than establishing data for any difficulties in reaction to emotions. Furthermore, there was no comparison group to assess differences in reacting compared to typically developing adults. And, there was only limited information provided regarding the types of reactions that were provided at the baseline phase, thus preventing assessment of abilities prior to intervention relative to those of typically developing adults.

In sum, there appears to be little empirical work that directly assesses the types of reactions ASD adults would provide within a social context and how these reactions compare to those of typically developing adults. In the absence of studies directly assessing how ASD adults react to the emotions of others, much of what is currently known is based on clinical observations and anecdotes. Thus, there is a clear need for empirical research involving adults with ASD to gain an insight into any deficits in reacting to the emotions of others, and the degree and variability of such deficits.

The Current Study

The research examined in the preceding sections leaves many key questions largely unanswered. These include how ASD and typically developing adults' emotional inferences are influenced by different stimulus types, response formats, and specific emotions or groups of emotions (i.e., basic and complex emotions). It is also clear that questions exist regarding ASD adults' ability to react to the emotions of others in ways that are considered appropriate. These are issues which the present study attempts to address.

To explore the effects of stimulus type (static, dynamic and social scenes), response format (multiple-choice and free-report) and emotion on facial emotion inferences and on reactions to others' emotions, samples of ASD and typically developing adult participants within the normal range of intelligence were exposed to multiple trials of basic and complex emotions, using static, dynamic and social scene stimuli presentations. Thus, I applied a 3 (Stimulus Type: static, dynamic, social) × 2 (Response Format: multiple-choice, free-report) × 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) design to examine the characteristics of emotion recognition performance. There were 3 within-group factors (stimulus type, response format and emotion) and one between-group factor (group).

The appropriateness of reactions to social stimuli was only examined for the social stimuli. ASD and typically developing individual's reactions to the emotions of others were examined using a 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) design.

Dependent Variables

Participants responded via both free-report and multiple-choice response formats for emotion recognition items. There were three key dependent "performance" variables: emotion recognition and reaction accuracy, response latency, and judgements of confidence in the recognition and reaction to the emotions of others; the latter variable is discussed in detail in a subsequent section.

The use of the term 'accuracy' when referring to facial emotion recognition has been challenged, particularly by Barrett et al. (2019) who mount a cogent argument that recognition accuracy is an inappropriate label for this variable. To summarise their argument briefly (see Rationale for Referring to Percent Agreement Rather than Accuracy within Chapter 2 for a detailed discussion), "accuracy" measures such as those used within this study rely on researcher judgments about what the normative recognition response would be—albeit informed by normative data— to classify whether a recognition response is accurate or correct (or incorrect). Thus, participants' recognition accuracy scores represent 'agreement' between the researcher and participants' inferences of emotion from the given stimuli. Thus, although accuracy is the commonly used label in the research literature, the term percent agreement (between the experimenters' targeted responses and the participants' responses) rather than percent correct will be used in the Results section when reporting on quantitative examinations of recognition performance. For the same reasons, accuracy/appropriateness of reactions to emotions will also be referred to as percent agreement rather than percent correct.

Response latency for emotion recognition decisions is an important index which has largely been undervalued. Few studies impose time restrictions on stimulus exposure or report the time taken for participants to respond (i.e., latency). Therefore, it is not known if, for example, ASD individuals required additional time to 'work out' answers and match the performance of typically developing individuals, are less accurate than typically developing individuals because they were less cautious or responded impulsively, or perhaps are characterised by some capacity limitation that undermined recognition while elevating latency. Without including time constraints on studying the stimuli or measuring response latency, it is difficult to assess if findings are likely to be representative of real-life situations where rapid decision-making is often necessary. Social exchanges require reasoning and decision-making skills to make judgments of what others may be feeling and to decide appropriate reaction to those emotions. Individuals with ASD have been previously associated with problematic decision making. For example, Luke et al. (2012) described three main areas of reasoning and decision making that are challenging for those with ASD, one of which is making decisions quickly. This is to be expected given persons with ASD are likely to rely on exhaustive emotion processing (e.g., rule-based processing) strategies that use more cognitive resources (Harms et al., 2010; Rutherford & McIntosh, 2007). A similar perspective is offered by Dual Process Theory. Dual Process Theory describes two types of processing Type 1: intuitive rapid decision-making/reasoning and Type 2: deliberate decision-making/reasoning (Brosnan et al., 2016). ASD individuals have been previously associated with the slower, more effortful, conscious processing of Type 2 (Brosnan et al., 2016, 2017). A propensity to engage in reflective rather than intuitive processing is likely to contribute to extended response times. Irrespective of whether a deficit in emotion recognition performance is found, differences between ASD and typically developing individuals in latency might have serious ramifications within real life situations. For example, if facial configurations for anger can be recognised, but only atypically slowly, this might lead to a range of negative outcomes and a delay in reaction may even be viewed as antagonising the situation further. Therefore, response latency is potentially an important index of emotion recognition.

Other Measures

Verbal Ability

Other variables that may differentiate adults with ASD and typically developing adults were also incorporated. Verbal intelligence was identified as one potentially important moderator of emotion recognition performance.

Emotion research has recognised that intelligence is a critical contributor to identifying emotions. Specifically, the Theory of Constructed Emotion (Barrett, 2017) suggests that language is central in developing normative emotional concepts. Emotional cues from the environment and facial configurations are organised by verbal labels. The role language has in supporting emotion recognition extends further than procurement of emotional vocabulary. It supports interactions with others whereby a conceptual alignment between emotional language and external cues (e.g., facial expressions, tone of voice, body language) is strengthened through experience (Berggren et al., 2018). Thus, a reduction in opportunities to learn about emotional concepts through language (i.e., social communication) compromises conceptual alignment, leading to less accurate emotion inferences. This has been demonstrated through research such as that conducted by Dunn et al. (1991) who reported that parent-child discussions about emotions predicted children's emotion identification accuracy. They found that children (baseline recorded at 36 months of age) who were exposed to more frequent discussion about emotion states were better at making inferences about the emotions of unfamiliar adults at follow-up (6 years of age) compared to children who were not exposed to as many conversation based emotion concept learning opportunities.

Intelligence may be associated with face emotion processing as described in previous studies (e.g., Dyck et al., 2006; Wright et al, 2008). Specifically, intelligence has been discussed as a potential compensatory mechanism for emotion processing in persons with

ASD (Rutherford & Troje, 2012). However, the meta-analysis by Uljarevic & Hamilton (2013) failed to find effects of intelligence on emotion recognition performance in individuals with ASD. A subsequent meta-analysis by Lozier et al., (2014) found limited evidence of intelligence effects on emotion recognition abilities. One of the analyses performed within their meta-analysis showed that intelligence scores moderated age-related deficits in those with ASD, with the largest advantage of higher intelligence on emotion recognition performance found in studies involving adults with ASD. It has been noted that few studies provide information about intelligence (Uljarevic & Hamilton, 2013) and the average intelligence reported in some studies was below average intelligence levels (Lozier et al., 2014). Therefore, it is unknown whether an association between verbal intelligence and emotion recognition exists in adults with ASD who have intelligence within the normal range.

Another important issue in considering intelligence as a possible moderator of emotion recognition abilities is the possibility that intelligence may affect abilities across task response types (i.e., multiple-choice and free-report). In light of the argument that verbal intelligence is a foundation for emotion concepts, coupled with the suggestion that intelligence can serve as a compensatory mechanism, it is thus important to consider the relationship between emotion recognition abilities in adults with ASD across both response types (i.e., multiple-choice and free-report) when verbal intelligence is controlled. Consequently, although I targeted ASD and typically developing samples within the normal range of intelligence (IQ > 85) and, hopefully, at least roughly matched in terms of verbal ability, verbal ability was measured to use as a control in statistical analyses.

Perspective Taking

The second potential moderator that was examined in the current study was ASD individuals' ability to take the perspective of others. This ability is also known as "Theory of

Mind". Theory of Mind (ToM) is a theoretical construct that represents an individual's ability to recognise the mental states of others or take others' perspective (i.e., recognising other people's thoughts, desires, beliefs, intentions). A well-developed ToM allows the attribution of mental states to oneself and others to explain and predict behaviour (Blair, 2005). It has been argued that perspective taking deficits are in part responsible for multiple social and communication impairments witnessed in ASD (Buitelaar et al., 1999). And, it has been shown that many individuals with ASD, though certainly not all, perform worse than typically developing individuals of similar intelligence on perspective taking or ToM tasks (e.g., Brewer et al., 2017; Chevallier et al., 2015).

Not surprisingly, it has been suggested that a deficit in taking others' perspectives may be associated with difficulties in emotional recognition (LaCava et al., 2007). Without the ability to accurately take the perspective of others, recognising the emotions of others might be expected to be impaired. Some studies have used emotion recognition paradigms as measures of ToM (e.g., Reading the Minds in Eyes task). However, the validation of these tasks as a measure of ToM with an ASD population has been criticised, in part due to the presence of comorbidities that sometimes exist alongside a diagnosis of ASD (Oakley et al., 2016) that may impede facial emotion recognition. Considering the potential social and emotion processing implications that a deficit in ToM abilities may have, a dedicated measure of ToM was included in the current study. This was used to examine whether any group differences in emotion recognition are moderated by ToM abilities.

Metacognitive Monitoring

ASD individuals' awareness of their ability to draw appropriate facial emotional inferences was also examined. The ability to plan, monitor and assess one's own understanding and performance is generally referred to as metacognition. Flavell (1979) referred to it as "thinking about thinking". Accurate metacognition is thought to foster effective self-regulation of cognition and behaviour (Nelson & Narens, 1990).

Metacognitive monitoring informs the individual of issues likely to affect their judgments and decision making which can have practical implications in many contexts. As described by Hacker et al. (2008), for many professionals an inability to make accurate, realistic predictions can have dismal consequences, such as the case for example, with doctors who are overconfident in their diagnoses, or an airline pilot who overestimates their ability to fly safely during challenging weather conditions. For students, the ability to monitor what they have learned can help them regulate their study patterns. For example, an underconfident student may unduly spend time revising content they already know, whereas an overconfident student is likely to believe they already know the content, and fail to study sufficiently (e.g., Hacker et al., 2008; Maki et al., 2005). Thus, when judgments of confidence do not accurately reflect abilities, it can have implications for academic achievement. Within the judicial system, for a witness to a crime, the ability to monitor the likely accuracy of what they are retrieving from memory can shape their decisions about whether to report or withhold information (e.g., Maras et al., 2020).

With regard to emotion recognition, accurate metacognition facilitates monitoring of abilities in accurately inferring emotions of others and, based on judgments of confidence in their level of accuracy, adjust reactions to others' emotions accordingly (Kelly & Metcalfe, 2011). For example, if someone accurately infers an emotional expression of sadness in another person, it should elicit an appropriate reaction which is likely to enhance social connectivity. Similarly, if an individual is unable to infer the emotion being displayed, but metacognitive monitoring abilities are efficient, they will be aware of this uncertainty and have the opportunity to adjust reactions to compensate for their uncertainty in efforts to react in a more socially appropriate manner. Conversely, social difficulties arise when there is unsatisfactory awareness of actual abilities (i.e., poor metacognitive monitoring), perhaps increasing the likelihood that the individual will not react appropriately.

The current study's methodology provided a perfect opportunity to examine metacognitive monitoring of emotion recognition judgments. Measuring metacognitive monitoring of emotion response judgments is essentially attempting to reveal whether individuals with ASD are aware of their strengths and limitations of their abilities in inferring emotions from other people's facial configurations. The ability to monitor the likely accuracy of emotion recognition judgments can thus have a range of important implications for the way individuals approach and regulate social communicative behaviours.

A standard way of assessing metacognitive monitoring is by the use of confidence judgments. These tasks generally involve participants reporting their level of confidence in the responses they provide. Thus, this concept is concerned with the relationship between task performance (accuracy) and confidence. If people are aware of their strengths and limitations in drawing emotional inferences then their confidence judgments should accurately discriminate between correct and incorrect answers (Grainger et al., 2016). Further, the confidence-accuracy relationship should ideally demonstrate that confidence and accuracy are calibrated. Thus, responses or decisions made with very high confidence should be highly likely to be accurate, and responses made with lesser confidence should be less likely to be accurate. Perfect calibration would be indicated if all responses made with 100% confidence were accurate, 90% of responses made with 90% confidence were accurate, 80% of responses made with 80% confidence were accurate, and so on. Thus, examining the calibration of confidence and accuracy provides additional informative information compared with that obtained when simply assessing whether confidence discriminated correct from incorrect responses.

Some studies have examined confidence judgments in samples involving individuals with ASD, with mixed results regarding the existence of impairments having been reported (see Bebko & Ricciuti, 2000; Grainger et al., 2016; Maras et al., 2020; Sawyer et al., 2014; Wilkinson et al., 2010; Wojcik et al., 2011; Wojcik et al., 2013). Wilkinson et al. (2010) investigated children (*n*= 18, age= 9-17 years) and adults (*n*= 16, age= 18-45 years) with ASD and typically developing children (*n*=13) and adults (*n*= 15) matched for age and intelligence, and compared their metacognitive awareness during a face recognition task. Participants were asked to specify whether a face in a photograph had previously been shown to them. After each response during the face recognition task participants were required to make a confidence judgment about their answer, reporting that they were either 'certain', 'somewhat certain' or 'guessing'. They reported that confidence judgments of children with ASD did not align with their actual level of accuracy as well as those of typically developing children. Adults with ASD and typically developing adults distinguished when they were 'guessing' from when they were 'somewhat certain'. Conversely, adults with ASD did not meaningfully distinguish between when they were 'certain' and when they were 'somewhat certain' whereas typically developing adults demonstrated significantly better accuracy for items rated at 'certain' compared to 'somewhat certain'. Typically developing adults were correct on 85% of the answers they were 'certain' of, while adults with ASD were correct on only 72% of answers they were 'certain' of. This difference was reported as a moderate effect size (Cohen's *d*= 0.53). This study provides tentative evidence that supports the findings of other studies reporting impaired metacognitive monitoring in ASD individuals (e.g., Brosnan et al., 2016; Grainger et al., 2014, 2016; Williams et al., 2018). Moreover, it highlights the potential importance of determining the degree to which confidence reliably predicted accuracy of responses at each confidence level (i.e., certain, somewhat certain, guessing) as is possible through the use of confidence-accuracy calibration analyses.

I identified only one previous study specifically assessing confidence in emotion inferences involving individuals with ASD. Sawyer et al. (2014) required adult participants to complete a facial emotion recognition task. After each response, participants were asked to rate how confident they were that they selected the correct response. Although both groups' confidence estimates indicated comparable discrimination between accurate from inaccurate responses (i.e., on average, accurate responses were associated with higher confidence judgments than inaccurate responses), other aspects of the confidence-accuracy relationship were not examined. Of particular interest would be the investigation of comparative accuracy across the various levels of confidence. For instance, comparing the ASD participants and typically developing participants at, for example, 90% confidence might reveal that one group is 80% accurate and another 60% accurate. Thus, the latter group would be displaying greater overconfidence in their recognition abilities, something that would not be revealed by simply showing that confidence was significantly higher for accurate than inaccurate responses. Impairments in making realistic judgements of confidence can have negative consequences in real-world contexts and be particularly damaging to social relationships (Bègue et al., 2019). Ideally, emotion recognition involves quick and efficient mechanisms to recognise facial configurations. Even though this is important for social communication, our beliefs about the degree we trust (confidence) our recognition of what a face is conveying are vital for self-regulation of verbal and behavioural reactions (Bègue et al., 2019). Effective social functioning thus involves not only accurate and timely inferences about others' emotions but also realistic appraisals of the likely accuracy of those inferences. Therefore, judgments of confidence are an important area that should be addressed in research aiming to further investigate emotion recognition abilities in adults with ASD.

Considering the above, further investigation into the confidence accuracy relationship is warranted. I assessed metacognitive monitoring through judgements of confidence about individual facial emotion recognition responses as well as about reactions to others' emotions. No prior study appears to have assessed confidence with regards to responses to others' emotions. I also examined metacognition of emotion recognition through confidence-accuracy calibration analyses.

Hypotheses and Research Objectives

To reiterate, the current study employed a 3 (Stimulus Type: static, dynamic, social) × 2 (Response Format: multiple-choice, free-report) × 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) mixed design (with stimulus type, response format and emotion as within-subjects' factors) to examine emotion recognition percent agreement scores, response latency and confidence. Additionally, reactions to emotions' agreement, latency and confidence measures were examined using a 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) mixed design.

Given the inconsistent results in several of the areas reviewed, it is difficult to make directional predictions about many of the effects that will be examined. However, existing research suggests some directional hypotheses specifically regarding response format and emotion. The literature provides a sound basis for expecting higher levels of recognition performance with multiple-choice than free-report response formats. This suggests that a main effect of response format on percent agreement with normative recognition responses should be expected. Furthermore, given that the free-report format required typing of responses whereas the multiple-choice format simply required the selection of responses via mouse click, an (uninteresting) main effect of response format on latency was obviously expected. A main effect of emotion on agreement with normative recognition responses was also predicted, with basic emotions better recognised than complex emotions.

Additionally, I attempted to resolve a number of other issues regarding emotion recognition:

- whether the context provided by dynamic and social stimuli assists accuracy (i.e., increases agreement with normative recognition responses), reduces latency and increases confidence, and whether any such effects differ for ASD and typically developing adults,
- whether adults with ASD respond less accurately (i.e., lower agreement with normative recognition responses) compared to typically developing adults across both complex and basic emotions, or whether group interacts with emotion type, and whether emotion type affects latency and confidence,
- whether group interacts with response format whereby adults with ASD respond less accurately than typically developing adults in the free-report response format compared to multiple-choice response format, and whether response format (i.e., free-report or multiple-choice) affects confidence,
- whether the confidence-accuracy relationship differs for the two groups across stimulus type, response format and emotion.

Finally, I explored several issues regarding reacting to the emotions of other people:

- whether adults with ASD react less appropriately to others' emotions (i.e., lower percent agreement) compared to typically developing adults,
- whether percent agreement, latency and confidence in the reactions to others' emotions are differentially affected for the two groups across basic and complex emotions.

CHAPTER 2

Method

Participants

One hundred and thirty-four participants were recruited, 130 (66 males and 64 females) of whom met the inclusion criteria and fully completed the study. The ASD participant group included 63 participants who were diagnosed with an Autism Spectrum Disorder (46 males and 17 females) and aged 18 years and older (range 18-66 years, *M*= 31.1 *SD*= 13). The comparison group of typically developing adults included 67 subjects (20 males and 47 females) who were aged 18 years or older (range 18-65 years, *M*= 23.8 *SD*= 8.9). The results from an independent samples t-test revealed that there was a significant group difference in age, t (109.26) = 3.73, $p < .001$, but a non-significant difference in verbal IQ, t (128) = 1.00, $p = .318$.

Given the sample size issues highlighted in reviews of emotion recognition and ASD research (e.g., Uljarevic and Hamilton, 2013), I targeted as large a sample as manageable given the extensive testing sessions (4-5 hours and 3-3.5 hours for ASD and typically developing participants, respectively[2]), the associated coding of responses (414 test items per participant across all tasks equating to approximately 2.5 hours of coding per participant), and the anticipated training and coding of reliability observers (two hours of training for the free-report coding across the three tasks was required. In total, 15%, or 25 test items per participant, of free-report emotion recognition responses across the three tasks and 30%, or 22 test items per participant, of the reactions to others' emotions were required for reliability checks. Therefore, approximately 12 minutes per participant was required for reliability

[2] Despite typically developing participants being required to complete extra components (i.e., A-ToM and WASI-II verbal subtests), on average they completed the study within a shorter testing session compared to ASD participants. ASD participants also took advantage of breaks offered between tasks, whereas typically developing participants did not.

checks). Although there were these time constraints, the aim was to secure a sample of at least 50 participants in each group.

All but one participant resided in South Australia (one ASD participant resided in Victoria). Typically developing adult subjects were recruited from the general public and ASD participants were recruited from a database of individuals who had voluntarily registered to be contacted to participate in research pertaining to ASD. This database was developed by Flinders University staff and students and reached people through a variety of organisations that specialised in working with those diagnosed with ASD (e.g., psychological practices, employment services, support groups) within South Australia and Victoria. To be eligible to enrol as part of the database, individuals needed to have received a diagnosis of ASD from an external multidisciplinary team or medical professional that had determined that individuals met the criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR or DSM-5) for ASD. The majority of participants recruited within the ASD group are likely to have received a diagnosis from the revised fourth edition of the diagnostic and statistical manual for mental disorders (DSM-IV-TR; APA, 2000). Unless reassessed, those who had been diagnosed under the DSM-IV-TR criteria for an ASD would retain diagnosis, despite a more recent edition of the manual with major revisions relating to ASD (DSM-5; APA, 2013)[3]. Date of diagnosis was not recorded on the database.

To ensure that performance was not confounded by low intelligence, only ASD participants with Wechsler Abbreviated Scale of Intelligence, second edition (WASI-II; Wechsler, 2011) full-Scale Intelligence Quotient (FSIQ) scores of 85 and above were included (Spinks et al., 2009). Additional exclusion criteria were an inability to speak English, a lack of appropriate reading and writing abilities, and being below the age of 18

---

[3] According to the DSM-5 (APA, 2013) “Individuals with a well-established DSM-IV diagnosis of autistic disorder, Asperger’s disorder, or pervasive developmental disorder not otherwise specified should be given the diagnosis of autism spectrum disorder”.

years. ASD participants had previously been vetted for these abilities as part of the requirements to be eligible to be accepted onto the database. Typically developing participants must have never received a diagnosis of ASD, scored at least 85 on the Verbal Comprehension Index (VCI) of the WASI-II (Wechsler, 2011), and been 18 years or older.

Measures

Demographic Information

All participants provided basic demographic information, including age, gender and primary language spoken. Although this information was already available for ASD participants as part of the database inclusion requirements, it was collected again for purposes of cross-checking given the large scale of the database. All typically developing comparison participants were asked if they had ever received a diagnosis of ASD during recruitment.

Intelligence Quotient (IQ)

The WASI-II (Wechsler, 2011) was developed for individuals aged 6-89 years and consists of four subtests: vocabulary, similarities, block design and matrix reasoning. It provides an estimate of an individual's full-scale IQ as well as estimates of verbal and performance intelligence. The four subtests are used to derive Full Scale (FSIQ). The vocabulary and similarities subtests are amalgamated to form the Verbal Comprehension Index (VCI), an estimate of verbal IQ and the block design and matrix reasoning subtests form the Perceptual Reasoning Index (PRI), an estimate of nonverbal abilities and co-coordination skills. McCrimmon and Smith (2013) demonstrated that the test-retest reliabilities for all subtests were high (*r*s = .83 - .94). The same study also demonstrated high concurrent validity (*r*s = .71 - .92) between the WASI-II and the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), the Wechsler Intelligence Scale for Children, Fourth edition (WISC-IV; Wechsler, 2003) and the Wechsler Adult Intelligence Scale, Fourth edition (WAIS-IV; Wechsler, 2008).

IQ scores for the ASD participants were available on the database. Participants from the typically developing comparison group were required to complete the vocabulary and similarities subtests to provide an estimate of their VCI scores and enable comparisons of the two groups with VCI controlled.

Theory of Mind

Theory of Mind was assessed for all participants using the Adult Theory of Mind (A-ToM; Brewer et al., 2017) test. The A-ToM is comprised of two separate subtests: Social and Physical. The Social subtest items include six videos that depict different social behaviours (e.g., telling white lies, sarcasm) and require the participants to infer the mental state or intent of the protagonist in each video. The Physical subtest items include six videos that serve as control items, as comparisons to the responses from the social videos, and do not require participants to infer the mental states of the protagonists. Typically, it is expected that individuals with ASD and typically developing persons will differ in their responses for the social videos to a much greater degree than for the physical videos.

Participants watch the series of 12 videos (ranging between 24 and 61 seconds in length) of actors portraying interpersonal social interactions. For counterbalancing purposes, participants were randomly assigned to one of 4 different randomised orders of the 12 scenarios. At the end of each video, participants were presented with a question on the screen (e.g., “Why does the burglar give himself up?”) and they were required to provide a written response on the answer sheet supplied to them (see Brewer et al., 2017, Supplemental Materials). Additionally, for three of the videos, there was an additional response component (e.g., “When the mother said, ‘That meal must have really filled you up’, did she mean it? If not, why did she say it?”), requiring participants to indicate their answer by circling yes/no on the response sheet as well as reporting their written answer in the space provided (as required with all other questions). To more closely approximate the demands of real-life interactions, the A-ToM task provides participants with a 60 second response time limit to finalise each answer; participants were informed of this time limit at the beginning of the task. Should the participant have not finished before the time ended, they were instructed to stop writing and move on to the next item.

Responses for each item were scored on a scale ranging from 0-2: 0 (incorrect), 1 (partially correct) or 2 (correct), where a higher score (on the social subtest) was reflective of a greater ToM ability. A score of two was assigned to answers that demonstrated understanding of the protagonist's mental state or intent in the Social items and the logic of the behaviour in the Physical items. For responses which were fundamentally correct but lacked depth of understanding (i.e., lacked explanation or explanation was incorrect but the yes/no response was correct) a score of one was awarded. A score of zero was given to all incorrect responses. Detailed scoring criteria are presented in Brewer et al. (2017).

As reported in the original validation study by Brewer et al. (2017), the A-ToM task was developed using a large sample (n = 163) of adults with a diagnosis of ASD and a comparison group comprising of 80 typically developing participants. The psychometric evaluation reported by Brewer et al. (2017) revealed test-retest reliability coefficients for the Social and Physical subscales of .82 and .64, respectively. It also highlighted the instrument's construct validity. First, principal components analysis confirmed two underlying factors, one representing social items and the other physical items. Within the Brewer at al. (2017) study, some overlap between individuals with ASD and typically developing participants was reported for the Social subscale; however, at a group level the A-ToM task Social subscale differentiated individuals with ASD and participants in the comparison group who were matched for perceptual reasoning ability. Specifically, the performance differential between the two groups was greater for the A-ToM Social subscale (Social: Cohen's d [95% CIs]= 0.64 [0.37, 0.92]) than for the Physical subscale (Physical: Cohen's d [95% CIs]= 0.22 [-

0.05, 0.48]). These differences persisted even after controlling for verbal comprehension ability. Second, concurrent validity was illustrated by correlations between A-ToM Social and the Strange Stories task Social subscale (r = .60), A-ToM Social and the Frith- Happé animations (White et al., 2011) 'Mental' (r = .25) and 'Feelings' (r = .33) subscales, and A-ToM Physical and Strange Stories task Physical subscale (r = .56). Third, there was evidence of discriminant validity of the A-ToM task, as group differences were shown in tools which measured different, but related, constructs (e.g., social anxiety) but were not meaningfully related to the A-ToM Social scores.

Emotion Recognition and Reactions to Emotions Tasks

To assess the ability of individuals to recognise and react to the emotional expressions of others, three tasks (static stimuli emotion recognition task, dynamic stimuli emotion recognition task, and social stimuli emotion recognition and reaction task) were administered.

All stimuli used to create these tasks were adopted from the images and videos from the EU-Emotion Stimulus Set database (O'Reilly et al., 2016) after appropriate permissions were granted by the Autism Research Centre, University of Cambridge database manager. The dataset contained 20 emotions/mental states and one neutral state. As described within the validation study for the dataset by O'Reilly et al., (2016), these 20 emotions/mental states were selected from an initially evaluated set of 27 (Lundqvist et al., 2014, as cited in O'Reilly et al., 2016). ASD clinical experts (n = 47) and the parents of children with ASD (n = 88) rated these 20 emotions/mental states as being the most important for social interactions, out of the potential 27. The emotion set of 20 was portrayed by a multi-ethnic group of child and adult actors (age range 10- 70 years; Females= 10, Males= 9) through facial emotional expressions (static images and dynamic clips), and contextual social scenes.[4] Actors were provided with scripts (to guide the intensity of emotion portrayals) and guided through their performances by an experienced theatre and film production director. Actors were first instructed and filmed performing the set of facial expressions, followed by the social scenes. The actors performed each of the basic emotions (i.e., anger, afraid, disgust, happy, sad and surprise) twice, once at high intensity and once at low intensity. All other emotions were portrayed at high intensity. Contextual social scenes were all filmed at high intensity for all emotions. There were no vocalisations while the actors performed and filming was shot using an infinite white background (O'Reilly et al., 2016). Available within the EU-Emotion Stimulus Set database were static images. Although descriptions of these images are not available within the validation study, it is understood that they were created via freeze-frame at the apex of each facial emotional expression (see O'Reilly et al. 2016 Supplementary Materials for validation data for static images).

Given the large quantity of footage produced, it was filtered by three of the original research team to include what they deemed as the best portrayals of each emotion. Their decisions on which portrayals to put forward for validation were based on their agreement with each other; where they did not agree they put forward two versions of the same emotion. As described within the validation study by O'Reilly et al. (2016), there were 418 multimodal emotion and mental state representations with the face, body gesture and contextual social scene stimuli which were prepared for distribution to collect recognition data for validation. The emotions/mental state representations were divided into 14 separate modality specific online surveys (i.e., six face surveys, two body gesture surveys and six social scene surveys) with emotions/mental states evenly distributed. The 14 surveys were

[4] Available as part of the stimulus set were body gesture and vocal stimuli, however, these were not used in the current study.

first developed in English and then translated into Swedish and Hebrew, and were distributed in the UK, Sweden, and Israel.

The validation study conducted by O'Reilly et al. (2016) consisted of 1,231 participants spread across the 14 surveys; a minimum of 54 participants completed each survey. The study included only multiple-choice responses where participants were asked to select labels (there was a choice of six options; the target emotion/mental state, four control emotions/mental states, and a 'none of the above') which best describe what the person was expressing within the clip. The inclusion of a 'none of the above' option was designed to assist in the prevention of artificially forced agreement (O'Reilly et al. 2016) and followed Frank & Stennett (2001) who observed that using the 'none of these terms are correct' option resulted in greater than chance agreement on the emotion label for facial emotion expressions and artificial agreement on incorrect emotion labels was obviated. Additionally, as the EU-Emotion Stimulus Set validation study included multiple-choice responses, O'Reilly et al. (2016) used data from another study conducted by Lundqvist et al. (2014, as cited in O'Reilly et al., 2016) to ensure that the response options were consistent in difficulty across all emotion/mental states. Lundqvist et al. (2014, as cited in O'Reilly et al., 2016) investigated the similarity and dissimilarity of the same 20 emotions/mental states using ratings from 700 participants, forming a 20 × 20 emotion similarity and dissimilarity matrix. O'Reilly et al. (2016) specified similarity ranges which corresponded with the results from the Lundqvist et al. (2014, as cited in O'Reilly et al., 2016) similarity/dissimilarity matrix. The resulting ranges of similarity were very similar, quite similar, quite dissimilar, and very dissimilar. (see O'Reilly et al. 2016 "Emotions Matrix" within supplementary materials for a list of target and control emotions/mental states with their corresponding ranges). Based on these ranges, O'Reilly et al (2016) selected corresponding emotion/mental states from the similarity/dissimilarity matrix from the Lundqvist et al. (2014, as cited in

O'Reilly et al., 2016) study. This method of selecting response options was used to assist with maintaining consistency in difficulty across all emotions/mental states for the different modalities.

Overall, the results of the validation study showed that participants recognised the emotions/mental states presented across the three modalities with the following chance-corrected recognition scores: facial expression (*M* = 63%, *SD* = 16%) and contextual social scenes (*M* = 72%, *SD* = 17%). Thus, across the modalities there were moderate to high recognition scores. Further, contextual social scenes produced higher recognition scores than facial expressions (see O'Reilly et al. 2016 supplementary data).

Within the current study, 12 emotions were used rather than the full set available in the EU-Emotion Stimulus Set. Because the current study examined multiple different trials for stimulus types and response formats, as well as different emotions, the time demands of the testing session(s) for participants were large. Consequently, a narrower range of emotions than the 20 described in the O'Reilly et al. (2016) study was included within the current study. The same emotions were tested in all three tasks (i.e., static stimuli task, dynamic stimuli task and social stimuli task) and were comprised of the 6 basic emotions (afraid, angry, disgusted, happy, sad, surprised) and 6 complex emotions (ashamed, disappointed, frustrated, hurt, jealous, worried). All emotion stimuli were used in their high intensity format for the purpose of maintaining consistency in stimulus intensity.

Static and Dynamic Stimuli Emotion Recognition Tasks. The static and dynamic emotion recognition tasks followed the same format but differed in the stimuli that were presented. The static task presented participants with still (i.e., static) images of a facial configuration (from the shoulders up) whereas the dynamic task presented participants with short clips of a person (from the shoulders up) moving their face into configurations depicting target emotions. The static and dynamic stimuli tasks were embedded in separate questionnaires. At the beginning of each task the participants were asked to rate their confidence in their ability to accurately recognise the emotions of others based on the stimulus type (i.e. "Please rate your level of confidence in recognising emotions/states of mind in other people from images") for each task[5]. The confidence scale ranged from 0-100% (100% being completely confident in their abilities) going up in increments of 10%. Participants were then presented with two practice questions that followed the format of test items as described below but were not included as part of their final scores, serving to ensure participants understood the task requirements.

The stimuli within the static task and dynamic tasks were presented in randomised order, separately for each task and participant. There were four trials for each emotion (n= 48 test trials) for the static stimuli task and four trials for each emotion (n= 48 test trials) for the dynamic stimuli task. Participants viewed the image (in the static task) or short clip (in the dynamic task) and were first asked to provide a free-report response by typing into the box provided what they thought the person presented in the image or clip was feeling (e.g., "Please name one emotion/mental state of mind which best describes what the person in the image was expressing"). Subsequently, they were asked to rate their confidence (from 0% to 100% on an 11-point scale) by selecting the percentage value that reflected their level of confidence that their previous answer was accurate (e.g., "Please rate your level of confidence in the previous answer"). Then for the same stimulus, participants were required to choose from 4 possible multiple-choice options the emotion/mental state that best described what the person was feeling.

The response format order was always free-report followed by multiple- choice for all participants across all three tasks. During the study design phase, I considered that ideally

[5] Dynamic stimuli showed the highest mean confidence rating followed by social stimuli and static stimuli. The means and standard deviations for these confidence ratings are presented in Appendix A. No further analyses of these ratings were conducted.

half of each participant group would complete the multiple-choice response format first and the other half of the participants in each group complete the free-report response format first. However, there was a low prospect of securing sufficient ASD participants to provide an adequately powered study using such a design. There was also a concern that completing the multiple-choice response format first might so strongly 'bias' a subsequent free-report response that the free-report data might be extremely difficult to interpret. In this context, it was of primary importance to have as much uncompromised free-report data as possible given that free-report responding has largely been neglected in previous research and it more closely aligns with the demands of real-world responding. However, the limitations of having free-report always followed by multiple-choice (e.g., unknown influence of free-report on subsequent multiple-choice response) was acknowledged and contemplated; however, ultimately having an appropriately powered measure of free-report responding took precedence.

The 4 multiple-choice options that were presented for each item included the target emotion and three foils of randomly selected emotions/mental states from a larger pool (n= 61, see Appendix B). This pool of emotions included the 12 emotions focused on within this study, plus a compilation of 49 other emotions/mental states. These other emotions were selected from a larger pool of 70 emotions and mental states that I generated which were then reviewed by two independent judges to ensure that they were not too similar to the 12 target emotions. This was to ensure that there was only one "correct" answer available for each of the facial configurations depicted within the stimuli. The feedback from both judges was considered and any emotions/mental states that were deemed to be too similar to the target emotions were removed from the list of potential foils. The three foils used on each trial were randomly selected for each participant. Immediately following the participant's response, they were asked to rate their confidence (on the same 11-point scale) by selecting the value that reflected their level of confidence that their previous answer was accurate. Participants were then asked to choose from the remaining three options a response that would be their next best suggestion regarding the facial configuration they were previously presented. This was not a compulsory question, and participants were able to withhold a selection and skip to the next item if they believed the remaining responses were not suitable to describe the given stimuli. They were not required to rate their confidence for this response. All response sections for both the static and dynamic tasks used an inbuilt timing system (via Qualtrics which delivered the task components and recorded responses using their online platform), allowing the recording of response latency (how long it took for participants to answer and progress to the next question).

Taking advantage of this feature available within the Qualtrics platform allowed for preliminary insights into latency. However, it is acknowledged that its sensitivity is not as reliable as dedicated reaction time software. Specifically, Qualtrics records the time answers were submitted, potentially leading to a slight delay from the time the decision is made until the time Qualtrics records submission. An individual who engages in more deliberative or effortful processing may, of course, not respond immediately when their decision is made as they might, for example, engage in some form of double-checking of responses. In the Discussion section of this thesis this possibility was foreshadowed in the context of discussion about dual process theory.

The following URL links provide examples of these tasks:[6]

- Static stimuli task example:
  https://qualtrics.flinders.edu.au/jfe/form/SV_9pqp0YDttUmdts1

[6] Due to the high resolution of the images and video clips, Google Chrome internet browser is recommended when viewing the tasks in any of the URL links provided.

- Dynamic stimuli task example:

  https://qualtrics.flinders.edu.au/jfe/form/SV_8vafpqESU40Iyu9

The full versions of the tasks can be accessed via the following URL links:

- Full version of the static stimuli task:

  https://qualtrics.flinders.edu.au/jfe/form/SV_eVuWSejT8SEWFVj

- Full version of the dynamic stimuli task:

  https://qualtrics.flinders.edu.au/jfe/form/SV_6sQAYZOh4hJQ5yl

Social Stimuli Task (Emotion Recognition and Reactions to Emotions). The social task was created to assess both emotion recognition and reaction to others' emotions from viewing a short social scene which provides potentially important contextual visual information. There were 74 test trials within the social scene task. To assess participants on as many social stimuli trials as possible, the author used all appropriate clips within the EU-Emotion Stimulus Set database that corresponded with the 12 target emotions. In doing so, it resulted in uneven trial numbers across the 12 target emotions. Table 1 shows the number of trials for each emotion.

Initially, the social task followed the same format as the static and dynamic tasks, but with stimuli that comprised short clips depicting social interactions (there was no audio of verbal exchanges and any music in the clips was muted). Thus, participants first rated their confidence in their ability to recognise emotions based on short clips of social interactions, followed by two full practice questions in the same format as described previously. Participants then progressed to the test items. After presentation of each clip of the depicted social interaction, participants were required to provide a free-report response by typing a response that best described the emotion/state of mind of the specified protagonist within the clip just viewed. Following this, participants were required to rate their level of confidence (using the same 11-point scale) that their previous answer was accurate. Subsequently, participants were required to choose the most suitable multiple-choice option from a choice of four and rate their level of confidence that their previous answer was accurate. Participants were then given the option to select a second response that would be their next best suggestion about what the protagonist was feeling from the remaining three options (not mandatory). The section of the social stimuli task which captures insight into the reactions to others' emotions then followed. Participants were asked to type a short reaction that they would give to the specified protagonist in response to their emotional expression. For example, one clip depicted that, during an interaction involving two people, one of the people sneezed without shielding or directing their sneeze away from the other person. As a result, the other person appeared visibly disgusted. A reaction deemed appropriate within this situation would be one that makes reference to an apology and some type of explanation (e.g., "I am very sorry that I sneezed on you. I did not feel that sneeze coming"). An inappropriate reaction would be dismissal of having sneezed on the other person and no apology and explanation. See below for URL links to examples and the full task. A confidence rating was collected for this answer following the same format as previously described. As with the static stimuli and dynamic stimuli tasks, all response sections for the social stimuli task used an inbuilt timing system (via Qualtrics which delivered the task components and recorded responses using their online platform) allowing the recording of latency.

The following URL link provides an example of the social stimuli task:

- Social stimuli task example: https://qualtrics.flinders.edu.au/jfe/form/SV_9oad9qroRmS3yrr

The full versions of the tasks can be accessed via the following link:

- Full version of the social stimuli task: https://qualtrics.flinders.edu.au/jfe/form/SV_etedxYCn8FLXbOR

Table 1

Social Scenes Trial Count

Emotion	Trial Count
Afraid	6
Angry	7
Ashamed	7
Disappointed	7
Disgusted	6
Frustrated	4
Happy	7
Hurt	7
Jealous	4
Sad	7
Surprised	7
Worried	5

Scoring Protocols and Inter-Rater Reliability

Multiple-Choice Responses for Emotion Recognition

Across all three tasks (static, dynamic, social response), multiple-choice responses were scored manually as either as 0 for incorrect answers or 1 for correct answers. There was only one correct answer: namely, the target emotion. These scores represent agreement (as opposed to 'accuracy') between the normative recognition responses judgments for each stimulus (informed by the data of the EU-Emotion Stimulus Set validation study; O'Reilly et al., 2016) and participants' inferences about each stimulus. Given the number (n= 22,100) of data points that were required to be manually scored for multiple-choice response across the three tasks, I decided to consider only the first multiple-choice answer provided).

Rationale for Referring to Percent Agreement Rather than Accuracy. To investigate facial emotion perception, objective and subjective measures have been used to identify facial movements and gestures associated with different emotional expressions, with subjective measures being employed most commonly. Objective measures do not require inferences from researchers or judges to make decisions on emotional expressions. Rather they use technology such as facial electromyography which detects electrical activity from muscular contractions that are not necessarily visible by the naked eye (Barrett et al., 2019). However, these methods are seen in only a limited number of studies, probably largely due to the practical difficulties of placing electrodes on the face. Studies that have used this method have been noted to contain electrical signal measurements from relatively few muscles, with no studies including naturalistic facial moments (Feldman Barrett et al., 2019). For these reasons, subjective measures (measures reliant on judgments of researchers or other judges) are more commonly used within research. As Barrett et al. (2019) have pointed out, these measures are perceiver-dependent in the sense that they are reliant on the experimenter's inferences about what the 'correct' observable emotion is. Barrett et al. (2019) noted that, given that the derivation of these measures is perceiver-dependent, there is a problem with using the term "accuracy" when classifying performance. Because subjective measures are perceiver-dependent, it is difficult for a researcher to be certain that the emotional state reported was present. This is especially challenging due to there being no fixed or unitary interpretation of facial emotional configuration related to any specific emotion, something that is highlighted by the fact that cultural differences influence emotional expression and interpretation (Golan et al., 2008). This interpretative issue limits the observer-independent validity of a study design that is assuming particular facial configurations validly express a specific emotion. Therefore, Barrett et al. (2019) argue that face-emotion configurations observed in research reflect human consensus. Thus, the dependent variable should be correctly referred to as 'agreement,' which represents the degree of agreement between the participants and the researcher or judges (who have been asked to infer emotions from stimuli materials).

As described under the Emotion Recognition and Reaction Tasks within the Measures section of this Chapter, the facial emotion configurations used as stimuli within the current study were taken from the EU-Emotion Stimulus Set database. The validation study by

O'Reilly at al. (2016) described that stimuli were validated using many levels of perceiver-dependent methods. First, video footage depicting actors who were instructed to portray facial emotion configurations based on the information provided in scripts had been filtered by three of the researchers to include what they deemed as the best portrayals of each emotion. Their decisions on which portrayals to put forward for validation were based on their agreement with each other and, where there was no agreement, they put forward two versions of the same emotion for validation by participants. Second, the participants within the validation study were asked to select labels (there was a choice of six options) which best described what the person was expressing within the clip. Therefore, it can be clearly seen that this validation process is dependent on perceiver-dependent inferences. The present study also relied on scoring methods (as described in the Scoring Protocols and Inter-Rater Reliability section within the current chapter) that followed similar protocols by using judges to infer appropriateness of free-report emotion recognition responses and reactions to emotions. For these reasons, each participant's performance is classified in terms of 'percent agreement' between the participant's responses and the judges' inferences.

Free-Report Responses for Emotion Recognition

For each of the 12 emotions, synonyms (see Table 2 for summary statistics) were identified using four popular online thesaurus and/or dictionary platforms: thesaurus.com, Merriam-Webster.com, OxfordDictionaries.com and CollinsDictionary.com. The participant free-report responses across the three emotion recognition tasks were sorted and responses that were different to the synonyms that had been identified were compiled into a list (see Table 2 for summary statistics). This list was given to three independent judges who were asked to imagine that they were looking at someone's face and they had to describe what emotion was shown on their face. For each of the 12 emotions, they were asked to look at the corresponding list of responses and classify each of the alternative labels in terms of how close it was in meaning to the emotion. For each alternative, a score of 3 (strict: strictly the same, the label means exactly the same as the target emotion), 2 (lax: similar type of meaning, but not exactly the same), 1 (boundary: plausible alternative but not really a synonym) or 0 (incorrect: not at all like the emotion, a different meaning). Judges were asked to provide their ratings on a prepared spreadsheet.

Once the three judges provided their ratings these were collated and assessed for discrepancies. When one of the judge's ratings differed from the others, a majority rule principle was used. When all judges differed, the average (rounded) of the three ratings was accepted as the final score. Table 3 provides examples of free-report emotion recognition responses and the final allocated scores. The full set of scoring codes for the emotion recognition free-report responses is available within the Supplementary Materials Part A.

Table 2

Free-Report Emotion Recognition Coding Summary

Emotion	Number of Synonyms Coded	Number of Non-Synonym Responses Coded
Afraid	29	115
Angry	61	135
Ashamed	40	211
Disappointed	42	179
Disgusted	36	95
Frustrated	45	135
Happy	52	113
Hurt	30	173
Jealous	25	197
Sad	70	180
Surprised	34	127
Worried	80	166

Initially, a coding reliability check was conducted which entailed the scoring protocol being provided to another independent observer, along with a database that contained a subset of the free-report responses (n= 2,200) across the three emotion recognition tasks. The observer used the protocol to assign each response a score from 0-3 in accordance with the scoring criteria detailed in the protocol. For each of the participants' answers, the observer searched in the relevant emotion column in the database to find the corresponding answer and its awarded score. This score was entered into the datafile in the column adjacent to the response and labelled to contain the scores for the particular emotion. To assess inter-rater reliability for the scoring of the free-report emotion recognition responses, I independently scored the same subset of responses. Agreement between the observer's and my scoring indicated good inter-rater reliability (k = .89; McHugh, 2012).

The same observer was then provided with a datafile which contained the remaining participant responses (n= 19,900) and scored the responses according to the same protocol as previously described. To ensure that consistency was maintained throughout, I scored another random subset of approximately 5% (n= 1,115) of the participant responses. Inter-rater reliability was again high (k = .83; McHugh, 2012).

Table 3

Free-Report Emotion Recognition Scoring Codes Examples

Emotion	Strict (score of 3)	Lax (score of 2)	Boundary (score of 1)	Incorrect (score of 0)
Afraid	Fearful	Alarmed	Discomfort	Amazed
	Scared	Horrified	Cautious	Happy
Hurt	Offended	Unhappy	Sorrow	Bored
	Miserable	Glum	Uneasy	Alone
Worried	Concerned	Fearful	Melancholic	Disgusted
	Stressed	Uncertain	Ashamed	Curious

Free-Report Reactions to Emotions

Two independent judges and I individually stated what would be considered a socially appropriate reaction for each of the 74 scenarios. Once independently completed, responses were collated and assessed for discrepancies in the themes of reactions to the emotions. A 'majority rules' approach was applied when assessing these reactions. There were no cases where a common theme was not identified and thus there was no further discussion needed to establish what would be considered a socially appropriate answer.

The judges and I collaborated to write detailed scoring criteria for each scenario based on the agreed appropriate answers. Agreement scores were assigned whereby an appropriate answer was assigned a score of 1 and an incorrect answer was assigned 0. For some of the scenarios, the judges and I agreed that there may be some answers that lacked depth but were not entirely incorrect and thus would attract a borderline score of 0.5. For example, within one of the social scenes' clips, it depicts one person using a fake spider to play a trick on another. The other person becomes visibly afraid of the spider. An appropriate reaction to the person who was afraid of the spider, attracting a score of 1, would characterise a response that includes an apology and reassurance/explanation that the spider was fake and thus a prop in an attempted prank/joke (e.g., "I am sorry. Don't worry it is only a fake spider") whereas just an apology (e.g., "I am sorry") without further explanation would attract a score of 0.5. In this example, although an apology was appropriate, the absence of an explanation that the spider was fake does little to help diffuse the inflicted fear of the person who was visibly afraid of the spider. A reaction that only addresses the spider as being fake or that it was an attempted prank/joke (e.g., "It is a fake spider" or "It was a joke") was assigned a 0 agreement score as the absence of an apology does not show acknowledgment that the prank caused the other person to feel afraid. See Supplementary Materials Part B for the full scoring protocol.

This scoring protocol was then provided to one of the judges along with a database that contained a subset of approximately 20% (*n*= 1,930) of randomly selected participant reactions for the social scenes task to conduct a coding reliability check. The judge used the protocol to assign each reaction a score from 0-1 in accordance with the scoring criteria detailed in the protocol. This score was entered into the datafile in the column adjacent to the reaction and labelled to contain the scores for the particular scenario. To assess inter-rater reliability for the scoring of the free-report responses, I independently scored the same subset of reactions. Inter-rater reliability was again high (*k* = .87; McHugh, 2012).

The same judge was then provided with a datafile which contained the remaining participant reactions (*n*= 7,960) and scored the responses according to the same protocol as previously described. To ensure that consistency was maintained throughout, I scored another random subset of approximately 10% (*n*= 800) of the participant reactions; inter-rater reliability was maintained (*k* = .93; McHugh, 2012).

Design and Procedure

Approval from Flinders University Social and Behavioural Research Ethics Committee was sought prior to commencement of participant recruitment. The study took place predominantly at Flinders University, Bedford Park campus (*n* = 114) or at participant's place of residence (*n* = 20). Typically developing comparison participants were not given the option to complete the study off-campus. When the study was completed off-campus, the author was accompanied by a research assistant as a chaperone in accordance with the University's safety guidelines. Those in the ASD group who chose to participate in their place of residence were located in South Australian suburbs and towns that were up to 79 kilometres from Flinders University. All locations required a quiet and non-distracting environment with access to comfortable seating and a flat surface to rest the laptop. An honorarium of $75 was awarded for a full session for comparison group participants taking part at Flinders University, ASD participants tested in their place of residence and ASD participants who travelled less than 10 kilometres from their place of residence to Flinders University. A $90 honorarium was given to ASD participants who travelled over 10 kilometres from their place of residence to Flinders University to take part in the study.

Participants could complete the study in either one session (approximately 3.5-4 hours) or over two shorter sessions. Of the 136 participants, 8 participants (ASD $n = 2$, typically developing $n = 6$) chose to complete the study over two sessions.

Typically developing participants were required to complete the WASI-II (vocabulary and similarities subtests) and the A-ToM task prior to the commencement of the study; these measures were already available for the ASD database participants. The A-ToM task was presented to participants on a desktop computer (Dell, model D11S) with a 21.5-inch sized screen.

All three tasks were delivered via an online survey package (Qualtrics) allowing for the amalgamation of the stimuli and tasks, distribution and data collection. The study sections were presented to participants in randomised order over six possible combinations of presentations.

All participants completed the emotion recognition components of the study on a 15-inch Apple Macbook Pro (connected to a wireless mouse) and sat approximately 50cm from the laptop. Participants began by reading the letter of introduction and information sheets and signed the consent form prior to commencing the study. At the beginning of each session, I ensured that the laptop audio was muted (to ensure that any music overlaid in the clips could not be heard) and I read out the instructions that were presented on the laptop screen and ensured participants understood what was required. Participants were reminded at the commencement of each task to respond as quickly and accurately as they could for each response element. Additionally, throughout the study (i.e., for all trials across all three tasks) a reminder note was presented at the top of the screen requesting that participants answer as quickly and accurately as possible. At the beginning of each task, the participants were required to enter their demographic information and where given two practice questions to ensure they understood what was required before they proceeded with the experimental trials.

CHAPTER 3

Results

Participant Characteristics

Table 4 displays the demographic information for all participants who fully completed the study, met the inclusion criteria and had WASI-II VCI scores of 85 and above[7]. The ASD group was comprised of 63 participants and the typically developing group was comprised of 67 participants.

Table 4

Participant Demographic Information

Group	Gender	*n*	Age		VCI	
			Range	*M (SD)*	*Range*	*M(SD)*
ASD	Male	46	18-63	31.0 (12.6)	85-133	104.0 (12.0)
	Female	17	21-66	31.6 (14.4)	85-143	105.6 (14.7)
	Overall	63	18-66	31.1 (13.0)	85-143	104.4 (12.7)
TD	Male	20	18-34	22.7 (3.9)	98-130	112.0 (10.0)
	Female	47	18-65	24.3 (10.3)	85-136	104.3 (11.3)
	Overall	67	18-65	23.8 (8.9)	85-136	106.6 (11.4)

Note. Excluded participants are not represented in the table above. 'TD' is typically developing

[7] Table 4 includes only VCI scores as this was used to 'match' ASD and typically developing participants. ASD participants' WASI-II FSIQ ranged from 85-138 (M = 107.0, SD = 12.4) for males and 87-132 (M = 107.5, SD = 12.3) for females.

Data Preparation

Excluded Participants

During preliminary analysis, it was identified that four ASD participants had VCI composite scores which were lower than 85. Although they had FSIQs of 85 and above, they were excluded to optimise VCI matching with the typically developing sample.

Missing Data

All data were analysed using IBM SPSS version 25 statistical package software. No missing data values were identified for the A-ToM or WASI-II tasks. Missing data values across stimulus types for free-report (recognition and reaction) responses ranged from 0-13, totalling 38 or 0.12% of the total 31,720 responses. Across multiple-choice responses, missing data values ranged from 18-61 with a total of 231 or 1.05% of the total 22,100 responses. See Appendix C for breakdowns.

As the percent agreement data were dealt with by calculating agreement for each emotion within each task, any missing participant responses were counted as incorrect, attracting a score of 0. However, it is important to note that subsequent analyses of latency and confidence data omitted responses that were missing.

Assessing and Dealing with Univariate Outliers

Outliers were assessed and found within the latency data for both emotion recognition and reactions. The approach described by Field (2013) which involves adjusting the data containing outliers was applied: that is, an outlier was any value that fell 2.5 standard deviations above or below the mean of each variable. Possible outliers were identified within the emotion reaction and emotion recognition mean latency scores for each emotion for each participant, across the three stimulus types and for both response formats. The number of mean latency scores identified outside of these boundaries ranged from 1 to 7 across emotions, stimulus types and response formats. See Appendix D for a record of the number of latencies adjusted.

These outliers were adjusted to 2.5 standard deviations above and below the mean (calculated for each emotion for each stimulus type within both response formats) for the mean latency scores.

Emotion Recognition: Percent Agreement

Percent agreement scores were calculated for each participant for each of the cells in the design. To derive percent agreement scores for each emotion, the three stimulus types and both response formats, an agreement score was calculated for each participant's performance across the relevant trials and converted into a percentage score. For the static and dynamic stimulus type conditions, this meant the percent agreement score for each of the different emotions in each response format was based on four trials; for the social stimulus type the percent agreement was based on 4-7 trials. As already noted, uneven trials across emotions in the social tasks were the result of attempting to maximise trial numbers by using all corresponding stimuli available in the EU-Emotion Stimulus Set database. Using a percent agreement score for each emotion within each condition allowed for inferential statistical comparisons across stimulus types that included different numbers of trials.

A 3 (Stimulus Type: static, dynamic, social scenes) × 2 (Response Format: multiple-choice, free-report) × 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) ANOVA was conducted on percent agreement scores. There were 3 within-group factors (stimulus type, response format and emotion) and one between-group factor (group). See Tables 5 and 8 for outcomes of statistical analyses on the percent agreement data. First, I examined the impact (i.e., the main effects) of the manipulations of stimulus type, response format and emotion. Then, I considered the main effect of group, followed by its interactions with the manipulated variables.

The results were read from the Greenhouse-Geisser correction to account for violations of sphericity as indicated by Mauchly's test for stimulus type, $\chi^2(2) = 6.50$, $p=.039$, emotion, $\chi^2(65) = 185.22$, $p<.001$, Response Format × Emotion, $\chi^2(65) = 137.94$, $p<.001$, Stimulus Type × Emotion, $\chi^2(252) = 433.36$, $p<.001$, and Response Format × Stimulus Type × Emotion, $\chi^2(252) = 404.24$, $p<.001$.

Throughout, the alpha level was 0.05 and effect sizes were reported as partial eta squared (Partial η^2) and interpreted within the following guidelines of 0.01, 0.06, and 0.14, suggesting small, medium and large effects respectively (Field; 2013).

Stimulus Type: Static Versus Dynamic Versus Social

The first objective was to test whether the context provided within dynamic and social stimuli enhanced emotion recognition accuracy. There was a significant and strong main effect of stimulus type (see Table 5) on percent agreement scores. Percent agreement was significantly higher overall for the social scenes stimuli (M = 77.85%, SE = .75, 95% CI [76.39, 79.30], followed by the dynamic stimuli (M = 73.24%, SE = .80, 95% CI [71.66, 74.82] and then the static stimuli task (M = 66.95%, SE = .86, 95% CI [65.24, 68.65].

Response Format: Free-Report Versus Multiple-Choice

A second objective was to assess whether the two different response formats (free-report and multiple-choice) affected emotion recognition percent agreement scores. Table 6 shows the means and standard deviations for the two participant groups for each of the twelve emotions across the three stimulus types and response formats. There was a significant main effect of response format on percent agreement scores (see Table 5). Agreement scores were significantly and markedly higher overall for multiple-choice responses (M = 81.81%, SE =

.76, 95% CI [80.31, 83.30] than for free-report (*M* = 63.55%, *SE* = .62, 95% CI [62.32, 64.77].

Table 5

Mixed ANOVA Outcomes for the Effects of Stimulus Type, Response Format and Emotion on Recognition Percent Agreement, and their Interaction with Group

Source	*df*	Mean Square	*F*	*p*	Partial η^2
Stimulus Type	1.91	98000.77	100.26	<.001	.44
Error (stimulus Type)	243.83	977.51			
Response Format	1	779192.74	2298.24	<.001	.95
Error (response format)	128	339.04			
Emotion	8.66	120662.36	160.35	<.001	.56
Error (Emotion)	11.08.17	752.50			
Stimulus Type×Group	1.91	1140.30	1.17	.312	.01
Response Format×Group	1	345.91	1.02	.314	.01
Emotion×Group	8.66	581.32	0.77	.637	.01
Response Format×Stimulus Type	1.99	630.75	2.09	.126	.02
Error (Response Format×Stimulus Type)	254.43	301.26			
Response Format×Emotion	9.31	13664.74	47.33	<.001	.27
Error (Response Format×Emotion)	1191.82	288.74			
Stimulus Type×Emotion	17.34	32210.49	72.92	<.001	.36
Error (stimulus Type×Emotion)	2206.05	441.75			
Response Format×Stimulus Type×Group	1.99	169.32	.56	.570	.00
Response Format×Emotion×Group	9.31	331.68	1.15	.324	.01
Stimulus Type×Emotion×Group	17.24	513.50	1.16	.287	.01
Response Format×Stimulus Type×Emotion	17.11	5575.22	21.76	<.001	.15
Response Format×Stimulus Type×Emotion×Group	17.11	259.85	1.01	.439	.01
Error (Response Format×Stimulus Type×Emotion)	2190.45	256.16			

Table 6

Means and Standard Deviations of Recognition Percent Agreement for ASD and Typically Developing Groups Across Emotion, Stimulus Type and Response Format

Emotion	Response Format	Static		Dynamic		Social	
		ASD *M% (SD)*	TD *M% (SD)*	ASD *M% (SD)*	TD *M% (SD)*	ASD *M% (SD)*	TD *M% (SD)*
				Basic Emotions			
Afraid	FR	60.98 (24.63)	60.70 (24.65)	64.95 (23.34)	71.27 (20.12)	75.75 (18.97)	81.92 (13.82)
	MC	76.98 (27.79)	77.24 (26.38)	78.97 (22.54)	83.21 (20.59)	87.30 (17.64)	91.79 (12.44)
Angry	FR	62.57 (19.22)	66.29 (20.84)	68.39 (17.97)	73.38 (15.30)	76.87 (13.19)	79.82 (12.19)
	MC	69.44 (23.95)	75.37 (24.42)	80.95 (22.77)	85.45 (16.94)	89.12 (14.45)	91.47 (10.27)
Disgusted	FR	78.31 (26.54)	82.21 (18.46)	82.54 (19.38)	92.04 (13.33)	78.66 (17.23)	83.75 (11.58)
	MC	88.89 (19.97)	89.93 (17.44)	94.84 (12.83)	96.27 (10.88)	92.06 (15.51)	96.27 (8.11)
Happy	FR	96.03 (10.02)	98.26 (6.57)	85.85 (17.04)	85.45 (16.50)	72.11 (13.89)	73.28 (14.36)
	MC	92.46 (15.32)	95.52 (12.25)	94.84 (12.83)	93.66 (11.79)	87.53 (14.40)	87.63 (14.69)
Sad	FR	57.01 (24.33)	58.08 (24.32)	64.02 (24.99)	64.30 (18.85)	60.54 (15.80)	69.51 (16.03)
	MC	74.21 (26.17)	75.00 (26.11)	75.79 (26.93)	82.09 (18.87)	79.59 (16.96)	86.14 (15.12)
Surprised	FR	93.65 (16.17)	97.14 (8.28)	82.80 (16.39)	87.19 (15.65)	51.55 (22.57)	55.44 (13.89)
	MC	97.22 (7.92)	98.13 (6.62)	91.27 (15.66)	94.03 (11.59)	88.21 (12.18)	90.19 (10.61)
				Complex Emotions			
Ashamed	FR	43.52 (22.42)	44.78 (20.86)	44.05 (26.41)	54.48 (22.95)	64.55 (23.78)	71.14 (20.94)
	MC	65.87 (29.89)	66.42 (26.31)	72.22 (30.49)	79.85 (23.93)	83.45 (21.19)	86.99 (17.89)
Disappointed	FR	27.78 (20.52)	36.57 (20.87)	50.79 (22.34)	62.44 (24.16)	46.26 (18.13)	48.54 (14.07)
	MC	73.41 (25.74)	74.25 (26.10)	81.35 (22.44)	86.94 (19.15)	78.91 (18.48)	82.30 (12.94)
Frustrated	FR	29.50 (16.86)	31.84 (18.69)	52.25 (23.29)	59.20 (19.63)	86.11 (15.77)	93.41 (11.19)
	MC	61.90 (26.13)	69.78 (23.65)	79.76 (22.39)	87.69 (16.50)	98.02 (8.16)	96.27 (11.72)
Hurt	FR	50.79 (13.78)	56.34 (15.83)	54.50 (16.45)	57.46 (13.54)	60.39 (18.62)	68.30 (15.12)
	MC	65.48 (28.90)	74.63 (27.00)	72.62 (29.00)	77.61 (23.49)	81.18 (21.44)	87.42 (15.83)

Table 6 continued

Emotion	Response Format	Static		Dynamic		Social	
		ASD *M% (SD)*	TD *M% (SD)*	ASD *M% (SD)*	TD *M% (SD)*	ASD *M% (SD)*	TD *M% (SD)*
Jealous	FR	26.46 (17.03)	29.35 (18.99)	28.17 (16.36)	33.58 (15.55)	56.75 (23.37)	59.95 (18.70)
	MC	48.02 (30.88)	54.48 (30.44)	50.40 (31.27)	64.18 (31.45)	80.56 (21.27)	85.07 (17.44)
Worried	FR	48.28 (23.10)	53.86 (23.76)	61.77 (23.41)	66.92 (20.72)	61.38 (21.29)	61.39 (19.82)
	MC	78.57 (20.01)	79.85 (22.71)	78.57 (24.53)	85.07 (17.97)	86.98 (16.13)	84.78 (16.36)

Note. M% is mean percent agreement, 'MC' is multiple-choice, 'FR' is free-report, 'TD' is typically developing

Emotion

To confirm the findings of previous studies, that basic emotions can be better recognised than complex emotions, the main effect of emotion on percent agreement scores was assessed. There was a significant main effect of emotion (see Table 5), with percent agreement varying considerably across emotions (see Table 7). The emotions with the highest percent agreement were five of the six basic emotions; sad was the only basic emotion to not be included in the top six.

Table 7

Recognition Percent Agreement for Each Emotion in Descending Order

Emotion	*M* (%)	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Happy*	88.55	.62	87.32	89.78
Disgusted*	87.98	.95	86.10	89.86
Surprised*	85.57	.72	84.15	86.99
Angry*	76.59	.95	74.71	78.48
Afraid*	75.92	1.20	73.55	78.29
Worried	70.62	1.11	68.42	72.82
Sad*	70.52	1.13	68.28	72.77
Frustrated	70.48	0.91	68.67	72.28
Hurt	67.23	1.01	65.23	69.22
Ashamed	64.78	1.41	61.98	67.57
Disappointed	62.46	1.14	60.21	64.71
Jealous	51.41	1.36	48.73	54.10

Note. * is a basic emotion

Group: ASD Versus Typically Developing

The possibility that ASD individuals may be poorer at emotion recognition than typically developing individuals as indicated by lower percent agreement scores across response format, stimulus type, and emotion was evaluated. As shown in Table 8, there was a significant main effect of group, with, across all conditions, the typically developing group showing higher percent agreement (*M* = 74.76%, *SE* = .93, 95% CI [72.93, 76.59] compared to the ASD group (*M* = 70.60%, *SE* = .95, 95% CI [68.71, 72.48]. Controlling for VCI, A-ToM-social or both measures had minimal impact on the group main effect. Examination of the inferential statistics in Table 8 reveals that the effect size for group was only minimally affected by considering either or VCI or A-ToM.

Given the age differences between the two groups previously highlighted, I examined the correlations between age and percent agreement for each emotion under the various response format and stimulus type conditions for the two groups combined. Although the correlations were generally weak, most were negative: for example, 61 of the 72 coefficients ranged from -0.17 to -0.37, indicating higher percent agreement scores for younger participants. Consequently, age was also used as a covariate in the analysis of percent agreement. As shown in Table 8, the significant effect for group remained but the effect size was reduced quite substantially.

Interaction Effects Involving Group

To test whether ASD individuals' emotion recognition performance was disproportionately enhanced by increased context, or perhaps whether typically developing individuals would use context more effectively, the interaction between group and stimulus type was examined, revealing no significant interaction (see Table 5). Similarly, no significant interaction of group and response format was found (Table 5). Additionally, to test whether recognition performance for the two groups was differentially affected by complex and basic emotions, the interaction between group and emotion was examined, also revealing no significant interaction (Table 5).

There were some significant interactions not involving group for percent agreement. For full examinations of significant interactions that do not involve group refer to Appendices E, F and G.

Table 8

Mixed ANOVA Outcomes for Main Between-Subjects Effect of Group (with and without Covariates of A-ToM, VCI and Age) on Recognition Percent Agreement

Source	*df*	Mean Square	*F*	*p*	Effect Size (partial η^2)
		No covariates			
Group	1	40517.39	9.81	.002	.07
Error	128	4129.36			
		Covariate of A-ToM			
A-ToM	1	51425.25	13.69	<.001	.10
Group	1	35420.47	9.43	.003	.07
Error	127	3756.95			
		Covariate of VCI			
VCI	1	28558.35	7.25	.008	.05
Group	1	34445.19	8.75	.004	.06
Error	127	3937.00			
		Covariates of A-ToM and VCI			
A-ToM	1	32990.26	8.90	.003	.07
VCI	1	10123.36	2.73	.101	.02
Group	1	32489.67	8.77	.004	.07
Error	126	3706.42			
		Covariate of Age			
Age	1	37027.09	9.57	.002	.07
Group	1	16914.54	4.37	.039	.03
Error	127	3870.32			

Summary of Percent Agreement ANOVA Findings

There were five main findings within the percent agreement results. First, percent agreement was higher for the typically developing than the ASD group. Second, both groups showed much higher percent agreement under the multiple-choice than the free-report response format. Third, percent agreement was highest for social scenes stimuli, followed by dynamic and then static stimuli. Fourth, basic emotions generally showed higher percent agreement compared to complex emotions. Fifth, although there were some significant interaction effects involving the three within-subjects variables, none of the interactions with group even approached significance.

Emotion Recognition: Latency

Mean latency (to the nearest hundredth of a second) was calculated for each participant for the relevant number of trials in each of the cells in the design and any outliers were excluded (as per the earlier discussion). Thus, as previously described for percent agreement scores, for the static and dynamic stimulus type conditions, mean latency for each of the different emotions in each response format was based on four trials; for the social stimulus type the mean was based on 4-7 trials. A 3 (Stimulus Type: static, dynamic, social scenes) × 2 (Response Format: multiple-choice, free-report) × 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) ANOVA was again conducted on individual participants' mean latencies. See Tables 9 and 12 for statistical analysis outcomes. Again, I consider the main effects of stimulus type, response format and emotion, followed by the effect of group and its interactions with the manipulated experimental variables.

As for percent agreement data, the results were read from the Greenhouse-Geisser correction to account for violations of sphericity as indicated by Mauchly's test for emotion, $\chi^2(65) = 206.01$, $p<.001$, Response format × Emotion, $\chi^2(65) = 230.68$, $p<.001$, Stimulus type × Emotion, $\chi^2(252) = 717.63$, $p< .001$, and Response Format × Stimulus Type × Emotion, $\chi^2(252) = 703.42$, $p<.001$.

Stimulus Type: Static Versus Dynamic Versus Social

The main effect of stimulus type on latency was evaluated to determine if the added context provided by dynamic and social stimuli affected latency of emotion recognition responses. There was a significant main effect of stimulus type (see Table 9). Mean latency was significantly shorter overall for the dynamic stimuli task (M = 2.37, SE = .06, 95% CI [2.25, 2.49], followed by the social stimuli (M = 2.51, SE = .06, 95% CI [2.39, 2.63], and then the static stimuli (M = 2.67, SE = .07, 95% CI [2.54, 2.80].

Response Format: Free-Report Versus Multiple-Choice

To confirm whether the multiple-choice response format reduced latency given the reduced response demands, or perhaps as a result of the added contextual clues provided by the multiple-choice response options, the main effect of response format on latency was explored. Table 10 shows the means and standard deviations for the two participant groups for each of the twelve emotions across the three stimulus types and response formats. There was a significant main effect of response format on latency, with latencies shorter for multiple-choice responses (M = 1.92, SE = .04, 95% CI [1.83, 2] compared with free-report (M = 3.12, SE = .06, 95% CI [2.99, 3.25]. Inspection of Table 10 shows that latencies for the multiple-choice response format were consistently shorter for both ASD and typically developing groups across all emotions.

Emotion

There was a significant main effect of emotion on latency (see Table 9). As shown in Table 11, the relatively neat distinction between basic and complex emotions observed for percent agreement was not replicated for latency. The confidence intervals reported in Table 11 suggest that the most notable differences are perhaps between the top three emotions (hurt, ashamed, happy) with the shortest latency and the bottom three emotions with the longest latency (disgusted, frustrated, jealous).

Table 9

Mixed ANOVA Outcomes for the Effects of Stimulus Type, Response Format and Emotion on Latency, and their Interaction with Group

Source	*df*	Mean Square	*F*	*p*	Partial η^2
Response Format	1	3388.09	980.79	<.001	.89
Error (response format)	128	3.45			
Stimulus Type	1.91	72.18	11.30	<.001	.08
Error (stimulus Type)	245	6.39			
Emotion	8.41	13.46	9.89	<.001	.07
Error (Emotion)	1075.99	1.36			
Response Format×Group	1	106.37	30.79	<.001	.19
Stimulus Type×Group	1.91	6.98	1.09	.335	.01
Emotion×Group	8.41	1.06	.78	.630	.01
Response Format×Stimulus Type	1.95	47.83	30.86	<.001	.19
Error (Response Format×Stimulus Type)	249.26	1.55			
Response Format×Emotion	8.28	14.50	12.71	<.001	.09
Error (Response Format×Emotion)	1060.14	1.14			
Stimulus Type×Emotion	14.25	15.24	10.19	<.001	.07
Error (Stimulus Type×Emotion)	1824.48	1.50			
Response Format×Stimulus Type×Group	1.95	9.29	6.00	.003	.05
Response Format×Emotion×Group	8.28	2.01	1.76	.077	.01
Stimulus Type×Emotion×Group	14.25	2.34	1.56	.081	.01
Response Format×Stimulus Type×Emotion	14.25	14.13	11.14	<.001	.08
Response Format×Stimulus Type×Emotion×Group	14.25	2.46	1.94	.018	.02
Error (Response Format×Stimulus Type×Emotion)	1823.79	1.27			

Table 10

Means and Standard Deviations of Recognition Latency for ASD and Typically Developing Groups Across Emotion, Stimulus Type and Response Format

Emotion	Response Format	Static		Dynamic		Social	
		ASD *M* (*SD*)	TD *M* (*SD*)	ASD *M*(*SD*)	TD *M* (*SD*)	ASD *M* (*SD*)	TD *M* (*SD*)
Basic Emotions							
Afraid	FR	4.01 (2.41)	2.72 (.78)	3.59 (1.61)	2.57 (.84)	3.15 (1.04)	2.55 (0.79)
	MC	2.19 (1.11)	1.45 (.48)	1.98 (.91)	1.50 (.65)	2.30 (1.21)	1.62 (.58)
Angry	FR	3.77 (1.68)	2.75 (.79)	3.46 (1.57)	2.49 (.80)	3.36 (1.25)	2.64 (1.17)
	MC	2.19 (1.07)	1.70 (1.08)	2.36 (2.14)	1.50 (.58)	2.00 (.77)	1.62 (.76)
Disgusted	FR	4.77 (3.15)	3.01 (1.00)	3.42 (1.46)	2.46 (.73)	3.50 (1.53)	2.57 (.72)
	MC	2.91 (1.37)	2.28 (.92)	1.92 (.86)	1.53 (.98)	2.07 (.82)	1.93 (1.80)
Happy	FR	4.00 (1.69)	2.71 (.76)	3.24 (1.27)	2.47 (.78)	3.46 (1.47)	2.72 (1.08)
	MC	2.14 (1.28)	1.67 (1.26)	1.90 (.79)	1.39 (.42)	2.07 (.83)	1.43 (.42)
Sad	FR	4.11 (2.12)	2.82 (.94)	3.65 (1.94)	2.68 (1.23)	3.41 (1.51)	3.20 (2.88)
	MC	2.09 (.84)	1.75 (1.10)	2.09 (.84)	1.75 (1.10)	2.29 (1.29)	1.55 (.55)
Surprised	FR	3.73 (1.64)	2.90 (1.37)	3.49 (1.37)	2.74 (1.07)	3.48 (1.46)	2.65 (.79)
	MC	2.13 (1.22)	1.73 (1.29)	2.02 (1.18)	1.58 (.91)	2.16 (.87)	1.57 (.56)
Complex Emotions							
Ashamed	FR	4.05 (2.03)	2.66 (.73)	3.40 (1.44)	2.37 (.61)	3.26 (1.24)	2.50 (.74)
	MC	1.99 (.70)	1.63 (.71)	1.98 (.89)	1.48 (.59)	2.14 (.87)	1.58 (.75)
Disappointed	FR	3.88 (1.80)	2.70 (.81)	3.59 (1.70)	2.54 (.91)	3.27 (1.19)	2.55 (.62)
	MC	2.28 (1.28)	1.65 (.63)	1.88 (.79)	1.46 (.54)	2.16 (.93)	1.63 (.71)
Frustrated	FR	3.86 (1.72)	2.92 (1.18)	3.40 (1.23)	2.50 (.86)	3.65 (1.62)	2.44 (.79)
	MC	2.27 (1.06)	1.75 (1.15)	2.04 (1.12)	1.34 (.39)	2.06 (1.07)	1.52 (.63)
Hurt	FR	3.68 (1.96)	2.82 (1.42)	3.61 (1.57)	2.41 (.77)	3.43 (1.49)	2.45 (.70)
	MC	2.04 (.83)	1.47 (.46)	2.04 (.94)	1.48 (.74)	2.01 (.83)	1.58 (.76)

Table 10 continued

Emotion	Response Format	Static		Dynamic		Social	
		ASD *M* (*SD*)	TD *M* (*SD*)	ASD *M*(*SD*)	TD *M* (*SD*)	ASD *M* (*SD*)	TD *M* (*SD*)
Jealous	FR	4.02 (2.25)	2.80 (.86)	3.18 (1.18)	2.42 (.63)	3.18 (1.14)	2.44 (.82)
	MC	2.25 (1.03)	1.55 (.57)	2.06 (1.03)	1.59 (.63)	4.52 (2.98)	3.24 (.99)
Worried	FR	3.76 (1.51)	2.80 (.89)	3.58 (1.85)	2.43 (.65)	3.29 (1.32)	2.52 (0.93)
	MC	2.16 (1.12)	1.60 (.73)	1.93 (.84)	1.36 (.43)	2.08 (.81)	1.80 (.81)

Note. 'MC' is multiple-choice, 'FR' is free-report, 'TD' is typically developing

Table 11

Recognition Latency for Each Emotion in Ascending Order

Emotion	*M* (sec.)	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Hurt	2.42	.06	2.30	2.53
Ashamed	2.42	.05	2.32	2.53
Happy*	2.43	.06	2.32	2.55
Worried	2.44	.05	2.40	2.63
Afraid*	2.45	.06	2.35	2.59
Disappointed	2.47	.06	2.35	2.59
Angry*	2.49	.06	2.35	2.62
Surprised*	2.51	.06	2.40	2.63
Sad*	2.61	.07	2.47	2.76
Disgusted*	2.70	.07	2.57	2.83
Frustrated	2.70	.07	2.57	2.83
Jealous	2.77	.07	2.64	2.90

Note. * is a basic emotion

Group: ASD Versus Typically Developing

There was a significant main effect of group (see Table 12) with, across all conditions, the typically developing group showing shorter mean latency (M = 2.14, SE = .07, 95% CI [2.00, 2.28] than the ASD group (M = 2.90, SE = .07, 95% CI [2.75, 3.04]. Examination of the inferential statistics in Table 12 reveals that the effect size for group was only minimally affected by considering either or both of the A-ToM and VCI covariates.

As noted previously, age differences between the groups exist, thus correlations between age and latency for each emotion under the various response format and stimulus type conditions for the two groups combined were assessed. For 63 of the 72 age-latency correlations the coefficients were significant and ranged from .18 to .45, indicating longer latencies for older participants. Therefore, age was included as a covariate in the analysis of latency. The outcomes of this analysis are shown in Table 12, noting that although the effect size for group was reduced, the effect size remained large.

Table 12

Mixed ANOVA Outcomes for Main Between-Subjects Effect of Group (with and without Covariates of A-ToM, VCI and Age) on Recognition Latency

Source	*df*	Mean Square	*F*	*p*	Effect Size (partial η^2)
		No covariates			
Group	1	1328.09	55.15	<.001	.30
Error	128	24.08			
		Covariate of A-ToM			
A-ToM	1	35.28	1.47	.228	.01
Group	1	1299.67	54.16	<.001	.30
Error	127	24.00			
		Covariate of VCI			
VCI	1	17.43	.72	.397	.01
Group	1	1291.15	53.50	<.001	.30
Error	127	24.14			
		Covariates of A-ToM and VCI			
A-ToM	1	23.44	.97	.326	.01
VCI	1	5.62	.23	.630	.01
Group	1	1280.00	53.02	<.001	.30
Error	126	24.14			
		Covariate of Age			
Age	1	360.39	16.81	<.001	.12
Group	1	815.85	38.06	<.001	.23
Error	127	21.44			

Interaction Effects Involving Group

The absence of significant interactions involving Group with Emotion and Stimulus Type illustrate that no disproportionate effects of stimulus type or emotion were present for ASD participants.

As shown in Table 9, there were significant two-way, three-way and four-way interactions. The two-way interactions (with large and medium effect sizes) were between response format and group, response format and stimulus type, response format and emotion and stimulus type and emotion. The significant three-way interactions were between response format, stimulus type and group (small effect size) and response format, stimulus type and emotion (medium effect size). The significant four-way interaction (with a small effect size) was between response format, stimulus type, emotion and group; I make no attempt to interpret this interaction.

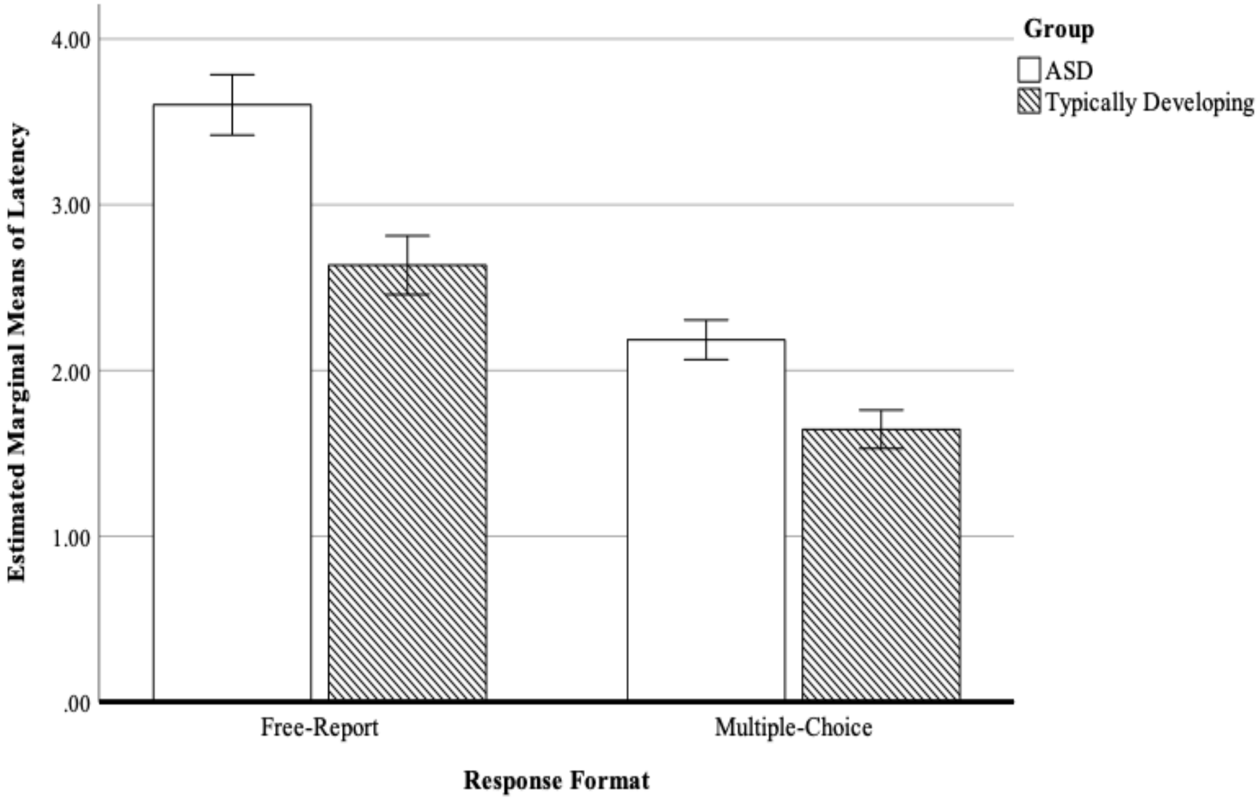
In the succeeding sub-sections I examine each of interactions involving group. For full examinations of significant interactions that do not involve group refer to Appendices H, I, J and K.

Response Format × Group. To explore whether either group's emotion recognition latencies were disproportionately reduced under multiple-choice responding, the interaction between response format and group was assessed. Figure 1 illustrates this interaction, indicating that the group difference under the two response formats was more marked for free-report than multiple-choice.

Response Format × Stimulus Type × Group. The significant three-way interaction between response format, stimulus type and group noted in Table 9 is difficult to interpret. Visual inspection of Figure 2 suggests some possible trends. For the ASD group, the static-dynamic and static-social latency differences are more marked in the free-report than the multiple-choice conditions. This was not the case for the typically developing group. This observation is confirmed by inspection of the pairwise comparisons in Table 13.

Figure 1

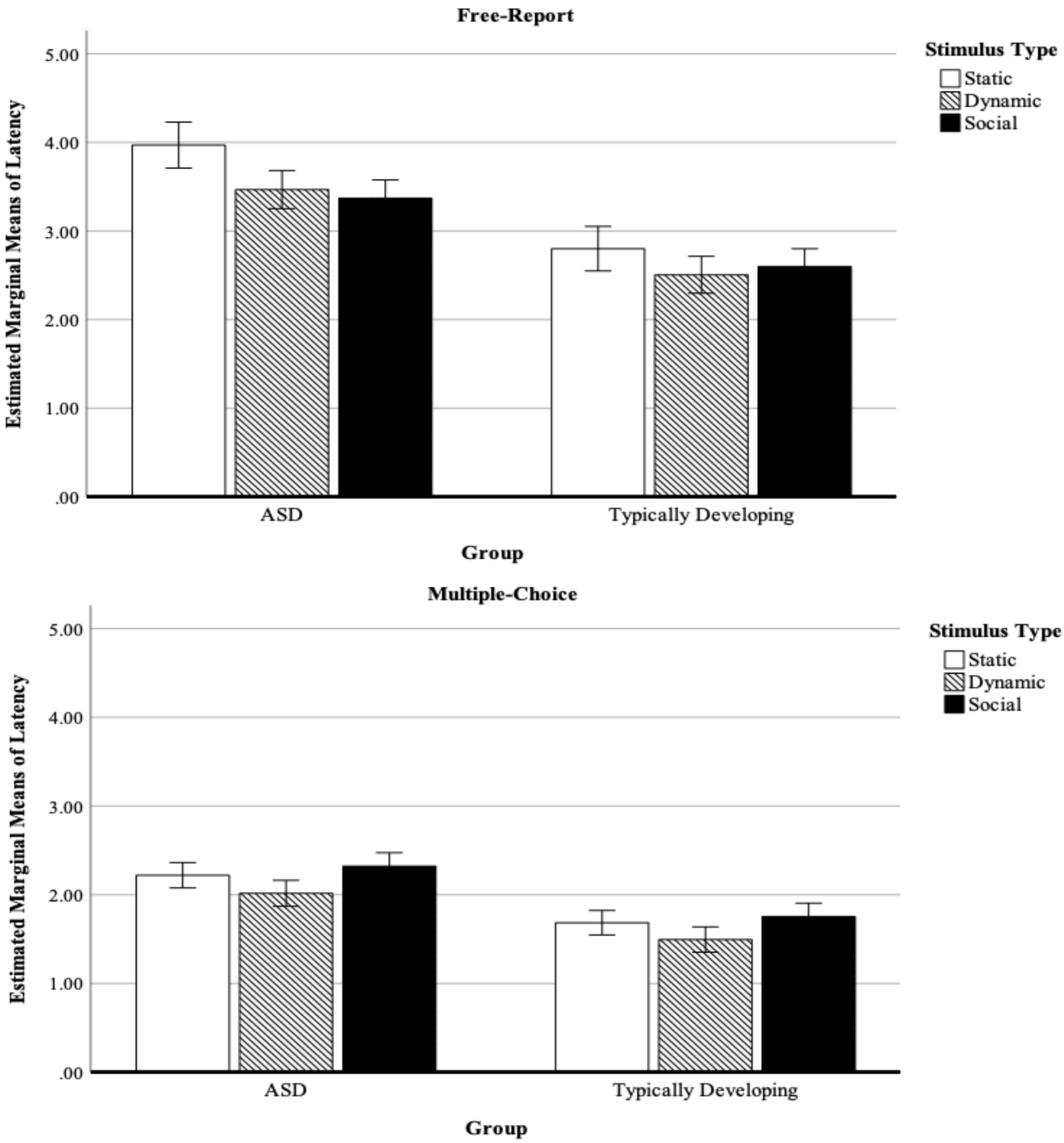
Two-Way Interaction of Group and Response Format on Latency



Note. Error bars represent 95% confidence interval

Figure 2

Three-Way Interaction of Response Format, Stimulus Type and Group on Latency



Note. Error bars represent 95% confidence interval

Table 13

Pairwise Comparison Outcomes Examining the Response Format × Stimulus Type × Group Interaction on Latency

Group	Response Format	Stimulus Type	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
ASD	FR	Static-Dynamic	.50	.14	.001	.18	.83
		Static-Social	.60	.12	<.001	.31	.89
		Dynamic-Social	.10	.11	1.000	-.16	.35
	MC	Static-Dynamic	.20	.07	.021	.02	.38
		Static-Social	-.10	.08	.612	-.30	.09
		Dynamic-Social	-.31	.07	<.001	-.48	-.13
TD	FR	Static-Dynamic	.29	.13	.078	-.02	.61
		Static-Social	.20	.11	.244	-.08	.48
		Dynamic-Social	-.09	.10	1.000	-.34	-.16
	MC	Static-Dynamic	.19	.07	.028	.02	.36
		Static-Social	-.07	.08	1.000	-.26	.12
		Dynamic-Social	-.26	.07	.001	-.43	-.09

Note. 'TD' is typically developing, 'MC' is multiple-choice, 'FR' is free-report, ** is based on estimated marginal means

Summary of Latency ANOVA Findings

Overall, typically developing participants had consistently shorter latencies than the ASD group. Both the ASD and typically developing groups had shorter latencies in the multiple-choice than the free-report response format across all emotions. These patterns are similar to those found for percent agreement. As will be discussed later, the longer latencies for ASD individuals could reflect specific processing difficulties or a greater tendency towards reflective (rather than intuitive) processing, the latter typically conceptualised within a speed-accuracy operating framework as greater caution. For both groups, latencies were significantly shorter for the dynamic stimuli, followed by the social stimuli and the static stimuli, but stimulus type did not interact with group. This does not follow the pattern found within the percent agreement results where a higher percent agreement was found for social stimuli, dynamic stimuli and then static stimuli, respectively. Additionally, despite there being a significant main effect of emotion for latency, there was no obvious pattern observed for latency for basic versus complex emotions.

Emotion Recognition: Confidence

Confidence data were collected using rating scales after each emotion inference question (multiple-choice and free-report) where participants rated their level of confidence (from 0% to 100% on an 11-point scale) in recognising the emotional expression. To derive mean confidence ratings for each emotion, for each of the three stimuli types and for both response formats, mean confidence ratings were calculated using the recorded confidence ratings for each participant across the relevant trials. As previously described for percent agreement and latency for the static and dynamic stimulus type conditions, mean confidence for each of the different emotions in each response format was based on four trials; for the social stimulus the mean was based on 4-7 trials.

Again a 3 (Stimulus Type: static, dynamic, social scenes) × 2 (Response Format: multiple-choice, free-report) × 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) ANOVA was conducted on each participant's mean confidence ratings for the trials they completed in each cell of the design. As with previous sections, the main effects of stimulus type, response format and emotion were examined followed by considerations of the main effects of group and then interactions with manipulated variables. See Tables 14 and 17 for analysis outcomes.

As for percent agreement and latency data, to account for violations of sphericity as indicated by Mauchly's test for Stimulus Type, $\chi^2(2) = 11.12$, $p=.004$, Emotion, $\chi^2(65) = 411.89$, $p<.001$, Response Format × Stimulus Type, $\chi^2(2) = 6.74$ $p=.034$, Response Format × Emotion, $\chi^2(65) = 171.97$ $p <.001$, Stimulus Type × Emotion, $\chi^2(252) = 675.68$, $p<.001$, and Response Format × Stimulus Type × Emotion, $\chi^2(252) = 521.74$, $p<.001$, the results were read from the Greenhouse-Geisser correction.

Stimulus Type: Static Versus Dynamic Versus Social

The possibility that added context provided within dynamic and social stimuli could affect confidence ratings was explored though the evaluation of the main effect of stimulus type on confidence. There was a significant main effect of stimulus type (see Table 14). Mean confidence ratings were significantly greater overall for the social stimuli task (*M* = 85.52, *SE* = 1.06, 95% CI [83.41, 87.62], followed by the dynamic stimuli (*M* = 78.83, *SE* = 1.21, 95% CI [76.45, 81.22] and then the static stimuli (*M* = 74.93, *SE* = 1.25, 95% CI [72.46, 77.39].

Response format: Free-Report Versus Multiple-Choice

To explore the possibility that multiple-choice response format could increase confidence in emotion recognition through the added contextual clues compared to the free-report format, the main effect of response format on latency was explored. Table 15 shows the means and standard deviations for the two participant groups for each of the twelve emotions across the three stimulus types and response formats. There was a significant main effect of response format on confidence with ratings significantly greater overall for multiple-choice responses (*M* = 81.46, *SE* = 1.08, 95% CI [79.32, 83.61] compared with free-report (*M* = 78.05, *SE* = 1.10, 95% CI [75.89, 80.22]. Inspection of Table 15 shows that the confidence ratings for multiple-choice responses were consistently greater for both ASD and typically developing groups across all emotions.

Table 14

Mixed ANOVA Outcomes for the Effects of Stimulus Type, Response Format and Emotion on Recognition Confidence Ratings, and their Interaction with Group

Source	*df*	Mean Square	*F*	*p*	Partial η^2
Response Format	1	27141.95	59.30	<.001	.32
Error (response format)	128	457.68			
Stimulus Type	1.85	96880.35	78.99	<.001	.38
Error (stimulus Type)	236.20	1226.55			
Emotion	5.26	22439.65	102.72	<.001	.45
Error (Emotion)	673.71	218.45			
Response Format×Group	1	190.62	0.42	0.520	.00
Stimulus Type×Group	1.85	3755.76	3.06	.053	.02
Emotion×Group	5.26	1199.26	5.49	<.001	.04
Response Format×Stimulus Type	1.90	482.85	6.35	.002	.05
Error (Response Format×Stimulus Type)	243.42	76.00			
Response Format×Emotion	8.83	459.42	12.66	<.001	.09
Error (Response Format×Emotion)	1130.62	36.30			
Stimulus Type×Emotion	13.25	4443.48	43.14	<.001	.25
Error (stimulus Type×Emotion)	1695.90	103.00			
Response Format×Stimulus Type×Group	1.90	41.25	0.54	0.573	.00
Response Format×Emotion×Group	8.83	45.00	1.24	.267	.01
Stimulus Type×Emotion×Group	13.25	217.95	2.12	0.010	.02
Response Format×Stimulus Type×Emotion	15.96	117.18	3.43	<.001	.03
Response Format×Stimulus Type×Emotion×Group	15.96	53.16	1.55	0.073	.01
Error (Response Format×Stimulus Type×Emotion)	2043.37	34.18			

Table 15

Means and Standard Deviations of Recognition Confidence Ratings (ranging from 0-100) for the ASD and Typically Developing Groups Across Emotion, Stimulus Type and Response Format

Emotion	Response Format	Static		Dynamic		Social	
		ASD *M* (*SD*)	TD *M* (*SD*)	ASD *M*(*SD*)	TD *M* (*SD*)	ASD *M* (*SD*)	TD *M* (*SD*)
				Basic Emotions			
Afraid	FR	69.52 (19.54)	80.93 (12.20)	70.95 (18.07)	83.21 (11.96)	80.98 (14.84)	88.08 (11.74)
	MC	73.33 (20.59)	83.81 (12.70)	76.47 (17.81)	86.46 (11.86)	83.94 (15.88)	90.95 (10.90)
Angry	FR	67.82 (20.89)	78.25 (12.34)	71.71 (17.39)	82.39 (12.36)	82.77 (15.12)	89.87 (10.50)
	MC	69.33 (19.31)	83.17 (12.10)	74.21 (19.45)	84.93 (13.37)	85.90 (13.77)	91.19 (9.21)
Disgusted	FR	70.52 (14.64)	77.76 (9.07)	77.46 (17.72)	85.63 (12.09)	83.04 (14.58)	89.70 (11.30)
	MC	75.95 (19.51)	85.49 (12.71)	83.37 (16.22)	91.49 (10.22)	85.42 (15.23)	92.09 (10.60)
Happy	FR	82.86 (18.13)	90.00 (10.62)	82.46 (15.53)	87.09 (12.03)	81.54 (14.74)	88.51 (11.28)
	MC	87.14 (15.46)	92.05 (10.15)	86.23 (14.96)	91.34 (10.59)	85.69 (14.88)	91.45 (10.75)
Sad	FR	66.35 (19.41)	77.91 (13.88)	71.51 (17.86)	81.94 (13.14)	77.66 (15.57)	85.97 (12.23)
	MC	71.63 (20.32)	80.90 (13.40)	74.80 (18.66)	86.53 (12.06)	79.59 (14.98)	89.17 (10.15)
Surprised	FR	77.90 (18.47)	85.78 (11.98)	77.78 (15.77)	85.41 (11.57)	82.95 (14.15)	89.53 (11.14)
	MC	83.27 (17.64)	90.37 (11.74)	82.86 (15.56)	89.18 (10.94)	85.80 (13.87)	90.21 (10.22)
				Complex Emotions			
Ashamed	FR	63.69 (21.87)	75.34 (12.81)	66.55 (18.72)	77.65 (15.25)	77.69 (15.81)	86.80 (11.71)
	MC	65.48 (22.29)	79.07 (13.41)	71.98 (19.61)	83.13 (13.77)	80.75 (16.22)	88.29 (11.59)
Disappointed	FR	62.94 (20.78)	74.22 (14.10)	69.68 (18.14)	82.09 (13.38)	75.96 (16.93)	84.69 (11.06)
	MC	66.23 (21.42)	78.25 (14.92)	75.16 (19.43)	86.75 (12.80)	80.05 (16.05)	86.82 (10.54)
Frustrated	FR	57.90 (22.11)	71.83 (14.49)	68.49 (18.86)	80.49 (13.06)	85.63 (13.40)	90.34 (12.51)
	MC	62.14 (21.97)	74.66 (14.83)	74.68 (19.74)	85.67 (12.59)	89.01 (12.62)	93.47 (10.81)
Hurt	FR	65.71 (19.46)	79.10 (12.03)	67.94 (19.13)	82.99 (11.62)	79.21 (15.20)	87.08 (12.32)

Table 15 continued

Emotion	Response Format	Static		Dynamic		Social	
		ASD *M* (*SD*)	TD *M* (*SD*)	ASD *M*(*SD*)	TD *M* (*SD*)	ASD *M* (*SD*)	TD *M* (*SD*)
	MC	69.40 (20.01)	82.39 (12.83)	71.47 (20.29)	83.81 (11.39)	82.27 (15.41)	90.19 (11.75)
Jealous	FR	63.97 (20.24)	74.22 (14.43)	66.43 (19.81)	78.40 (14.82)	79.84 (14.38)	86.98 (13.36)
	MC	62.58 (22.08)	74.25 (17.06)	64.76 (23.97)	77.01 (16.74)	81.79 (16.99)	86.87 (12.20)
Worried	FR	63.69 (20.47)	76.16 (13.54)	66.79 (18.44)	78.66 (13.86)	79.27 (15.79)	87.70 (12.05)
	MC	70.87 (18.53)	80.22 (12.86)	73.41 (17.91)	84.44 (12.37)	83.02 (15.18)	88.99 (11.94)

Note. 'MC' is multiple-choice, 'FR' is free-report, 'TD' is typically developing

Emotion

To explore confidence rating across basic and complex emotions and to investigate whether basic emotions can perhaps be more confidently recognised by participants, the main effect of emotion on latency were assessed. There was a significant main effect of emotion (see Table 14), with mean confidence ratings varying across emotions (see Table 16). The mean confidence rating for all 6 basic emotions showed greater confidence compared to the complex emotions.

Table 16

Recognition Confidence Ratings for Each Emotion in Descending Order

Emotion	*M*	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Happy*	87.20	1.03	85.17	89.23
Surprised*	85.10	1.05	83.02	87.17
Disgusted*	83.16	1.04	81.10	85.23
Afraid*	80.72	1.11	78.52	82.92
Angry*	80.13	1.10	77.94	82.31
Sad*	78.66	1.15	76.40	80.93
Hurt	78.46	1.15	76.18	80.74
Frustrated	77.86	1.11	75.66	80.06
Worried	77.77	1.11	75.57	79.97
Disappointed	76.90	1.16	74.60	79.21
Ashamed	76.37	1.19	74.01	78.73
Jealous	74.76	1.25	72.29	77.22

Note. * is a basic emotion

Group: ASD Versus Typically Developing

Given ASD individuals had demonstrated poorer emotion recognition than typically developing individuals, the possibility that ASD individuals would also demonstrate reduced confidence was explored through assessment of the main effect of group on confidence. There was a significant main effect of group (Table 17), with, across all conditions, the typically developing group showing greater mean confidence (*M* = 84.41, *SE* = 1.49, 95% CI [81.47, 87.35] than the ASD group (*M* = 75.10, *SE* = 1.53, 95% CI [72.07, 78.13]. Examination of the inferential statistics in Table 17 reveals that the effect size for group was only minimally affected by controlling for VCI, A-ToM-social or both measures.

Again, due to the group differences regarding age, I investigated the correlations between age and confidence for each emotion under each response format and stimulus type for the two groups combined. Similar to percent agreement, the correlation outcomes for the confidence data revealed generally weak but mostly negative correlations. For example, 70 of the 72 coefficients ranged from -0.17 to -0.33, indicating greater confidence for younger participants. Age was thus used as a covariate in the analysis of confidence, with results displayed in Table 17. The effect size for group was not meaningfully affected when age was considered.

Interaction Effects Involving Group

Table 14 showed there were no significant interactions between response format and group and stimulus type and group. There were, however, significant two-way and three-way interactions. The two-way interactions were between emotion and group (small effect size), response format and stimulus type (small effect size), response format and emotion (medium effect size), and stimulus type and emotion (large effect size). The significant three-way interactions were between stimulus type, emotion and group (small effect size), and response format, stimulus type and emotion (small effect size). In the following sub-sections I examine the interactions involving group. For full examinations of significant interactions not involving group, refer to Appendices L, M, N and O.

Table 17

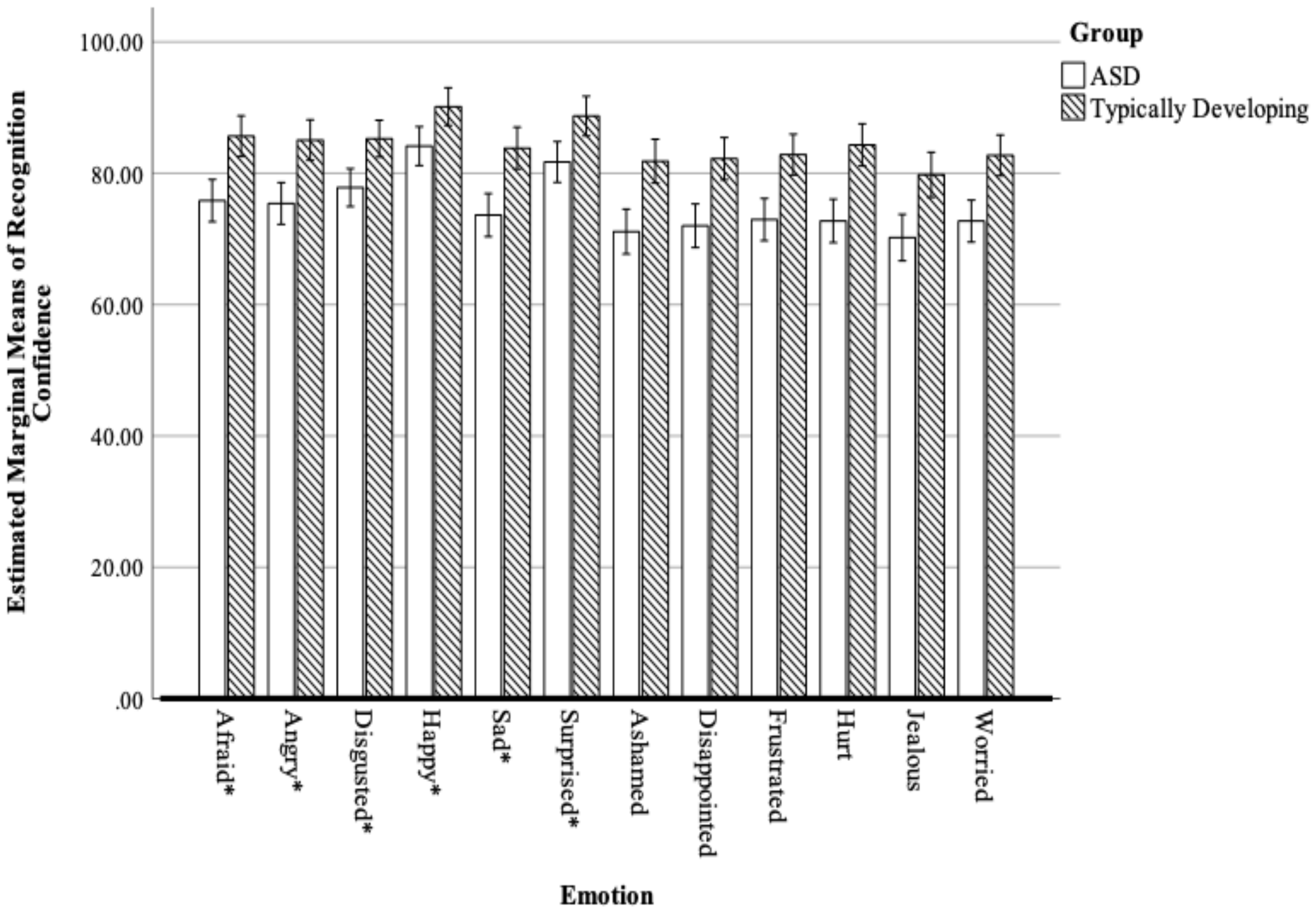
Mixed ANOVA Outcomes for Main Between-Subjects Effect of Group (with and without Covariates of A-ToM, VCI and Age) on Recognition Confidence Ratings

Source	*df*	Mean Square	*F*	*p*	Effect Size (partial η^2)
No covariates					
Group	1	202749.03	19.05	<.001	.13
Error	128	10643.59			
Covariate of A-ToM					
A-ToM	1	33923.31	3.24	.074	.03
Group	1	192890.78	18.44	<.001	.13
Error	127	10460.28			
Covariate of VCI					
VCI	1	5555.06	.52	.472	.00
Group	1	195315.05	18.28	<.001	.13
Error	127	10683.66			
Covariates of A-ToM and VCI					
A-ToM	1	28604.75	2.71	.102	.02
VCI	1	236.51	.02	.881	.00
Group	1	190841.54	18.10	<.001	.13
Error	126	10541.43			
Covariate of Age					
Age	1	8069.01	.76	.385	.01
Group	1	158972.81	14.91	<.001	.11
Error	127				

Emotion × Group. To investigate group differences between individuals' emotion recognition confidence ratings on basic and complex emotions, the interaction between emotion and group was assessed. As Figure 3 shows, although the ASD group was less confident than the typically developing group across all emotions (confirmed by the main effect for emotion) the difference generally appears to be more marked for complex than basic emotions. Although the mean difference in confidence for the two groups was significant for all emotions, Table 18 shows that the mean difference was generally larger for complex than basic emotions.

Figure 3

Two-Way Interaction of Emotion and Group on Recognition Confidence Ratings



Note. Error bars represent 95% confidence interval, * is a basic emotion

Table 18

Pairwise Comparison Outcomes Examining the Group × Emotion Interaction on Recognition Confidence Ratings

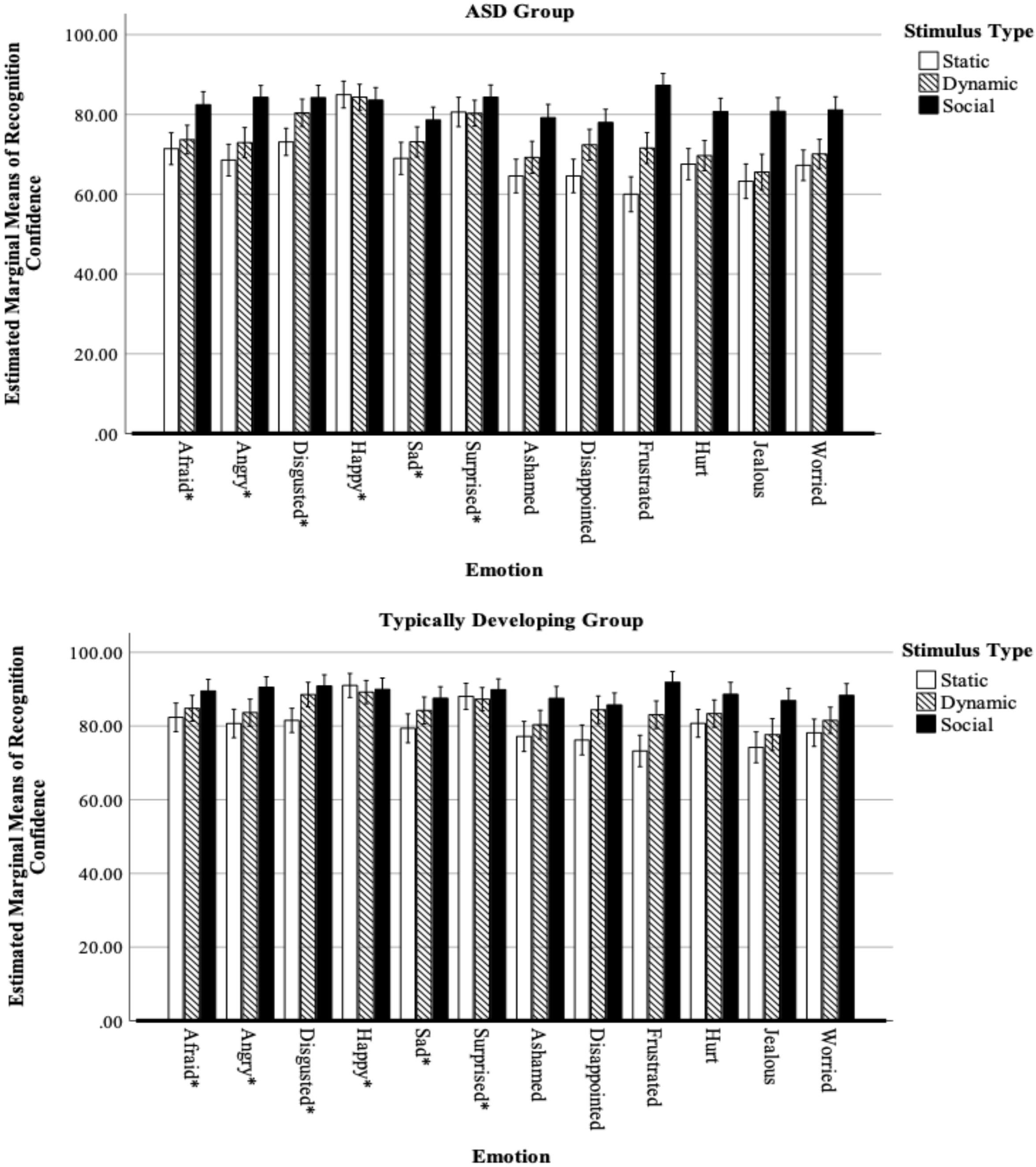
Group Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
TD-ASD	Afraid*	9.71	2.22	<.001	5.31	14.11
	Angry*	9.68	2.21	<.001	5.31	14.05
	Disgusted*	7.73	2.09	<.001	3.60	11.86
	Happy*	5.75	2.05	.006	1.70	9.81
	Sad*	10.15	2.29	<.001	5.61	14.68
	Surprised*	6.64	2.10	.002	2.48	10.79
	Ashamed	10.69	2.38	<.001	5.97	15.41
	Disappointed	10.47	2.33	<.001	5.86	15.08
	Frustrated	9.77	2.22	<.001	5.37	14.16
	Hurt	11.59	2.30	<.001	7.04	16.15
	Jealous	9.73	2.49	<.001	4.79	14.66
	Worried	9.85	2.22	<.001	5.46	14.25

Note. * is a basic emotion, ** is based on estimated marginal means, 'TD' is typically developing

Stimulus Type × Emotion × Group. The significant three-way interaction between stimulus type, emotion and group is not easy to interpret. Inspection of Figure 4 tentatively suggests that the source of the interaction effect may be that the disparity between social confidence and dynamic and static confidence for complex, though not for basic, emotions appears to be larger for the ASD group than for the typically developing group.

Figure 4

Three-Way Interaction of Stimulus Type, Emotion and Group for Confidence



Note. Error bars represent 95% confidence interval, * is a basic emotion

Summary of Confidence ANOVA Findings

Overall, typically developing participants demonstrated greater confidence compared to the ASD group. Participants were most confident in recognising emotions presented through social stimuli followed by dynamic and then static stimuli, thereby following a similar pattern to the percent agreement data. Also following the percent agreement patterns, confidence was higher for multiple-choice than free-report responses, and for basic than for complex emotions. In other words, when the percent agreement suggest that participants were experiencing greater difficulty in emotion recognition, they (quite appropriately) reported lower confidence.

Emotion Recognition: Confidence-Accuracy Calibration

The calibration of recognition confidence and accuracy was explored to investigate metacognitive monitoring of emotion recognition abilities in adults with ASD. Calibration curves were derived by plotting accuracy against confidence. As indicated in Chapter 1, perfect calibration is indicated by a perfect alignment between confidence and accuracy (e.g., 90% of responses judged with 90% confidence are accurate, 50% of responses judged with 50% confidence are accurate, etc.). Underconfidence is identified by calibration curves which lie above the 'perfect' calibration line (denoted by the dotted line in Figures 5, 6 and 7). Underconfidence occurs when confidence judgments are lower than actual performance (i.e., lack of confidence in correct responses, Dentakos et al., 2019). Overconfidence is identified by calibration curves which lie below the 'perfect' calibration line. Overconfidence occurs when judgements of confidence are greater than actual performance (i.e., too confident in incorrect responses, Dentakos et al., 2019).

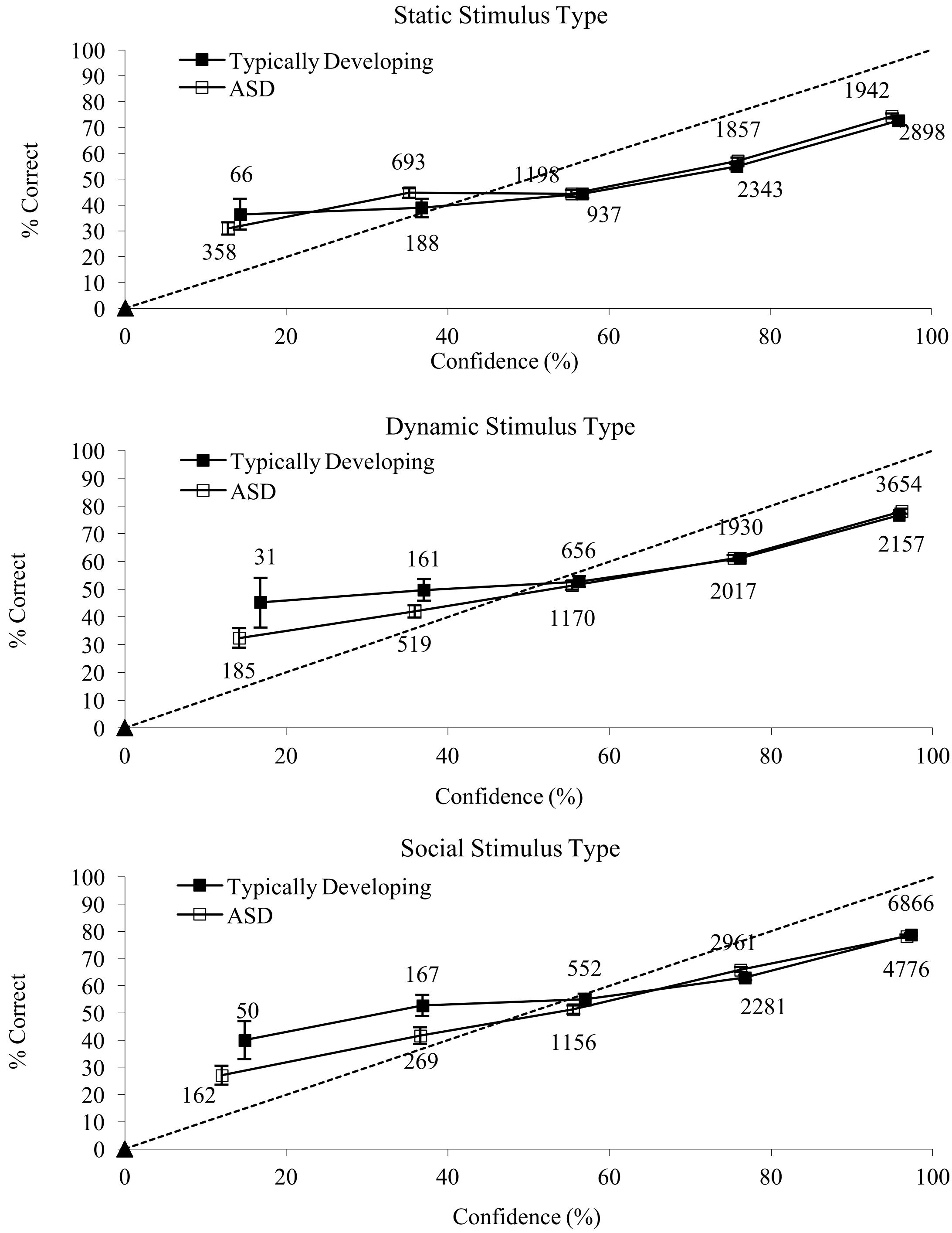
For the purposes of these calibration analyses, free-report percent agreement scores which were originally scored as 3 were recoded as 1 (correct) and scores 0-2 were recoded as 0 (incorrect). Following methods used in prior confidence- accuracy calibration research (e.g., Brewer & Wells, 2006; Juslin et al., 1996; Palmer et al., 2013), confidence data were collapsed from 11 categories (i.e., 0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%) to five categories (i.e., 0-20%, 30-40%, 50-60%, 70-80%, 90-100%) to provide more stable estimates in each confidence category. In all associated figures, proportion correct in each of the five categories was compared against the weighted mean confidence for that category.

The confidence-accuracy calibration curves for ASD and typically developing participants are displayed in Figures 5, 6 and 7 which represent the overall match between perceived and actual performance in inferring emotions for the main experimental conditions (stimulus type, response format and emotion).

Across Figures 5, 6 and 7, it is clear that, regardless of how the data are sectioned into experimental conditions, there are similar patterns observed in all three figures. Unsurprisingly, given the large number of data points, the patterns are very stable as reflected by the narrow error bars within each panel of each figure. All three figures reveal that higher levels of agreement (or accuracy) were associated with higher confidence. However, none of the calibration curves indicated perfect calibration. In all the comparisons, at the upper end of the scale the curves are characterised by overconfidence. This indicates that participants did not lower their confidence estimates sufficiently to align with the lower percent agreement. Additionally, in most (though not all) curves, at the lower end of the scale, participants lowered their confidence estimates more than was necessary given the associated levels of agreement. These patterns are consistent with the hard-easy effect which refers to the tendency to exhibit overconfidence during hard tasks and underconfidence during easy tasks (Juslin et al., 2000). Examples of this within the calibration figures are seen particularly in Figure 6, comparing response formats, and Figure 7, comparing complex and basic emotions. The exceptions to overconfidence at the top end of the scale occur for the multiple-choice response format (Figure 6) and for basic emotions (Figure 7). The exceptions to underconfidence at the lower ends of the scale are under free-report response format (Figure 6) and for complex emotions (Figure 7), the conditions where participants generally perform worse.

Figure 5

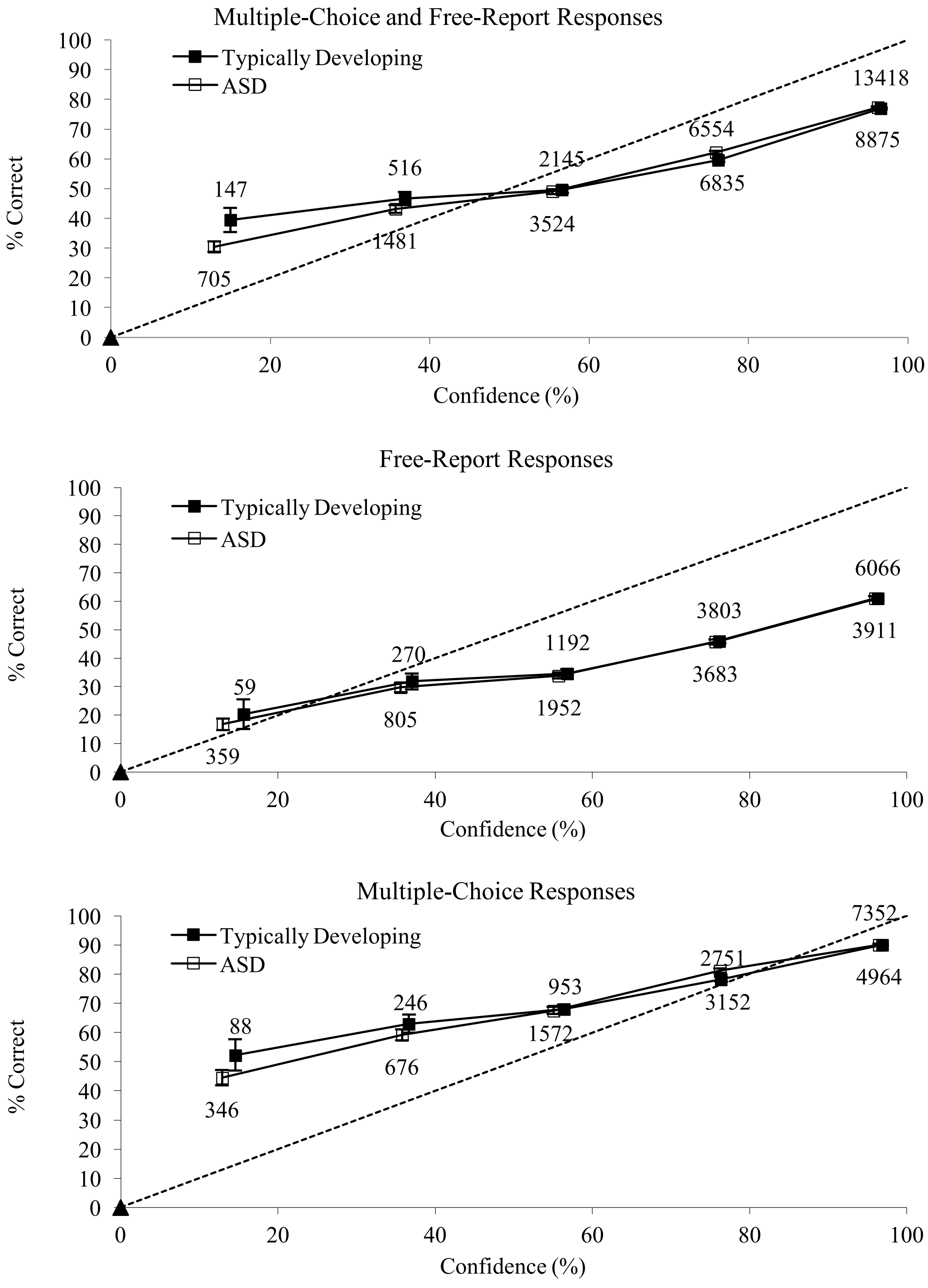
Confidence- Accuracy Calibration Curves for ASD and Typically Developing Groups for All Three Stimulus Types (i.e., Static, Dynamic and Social)



Note. The frequency of judgements of confidence are presented at each data point. Dotted line represents perfect calibration. Error bars represent standard error.

Figure 6

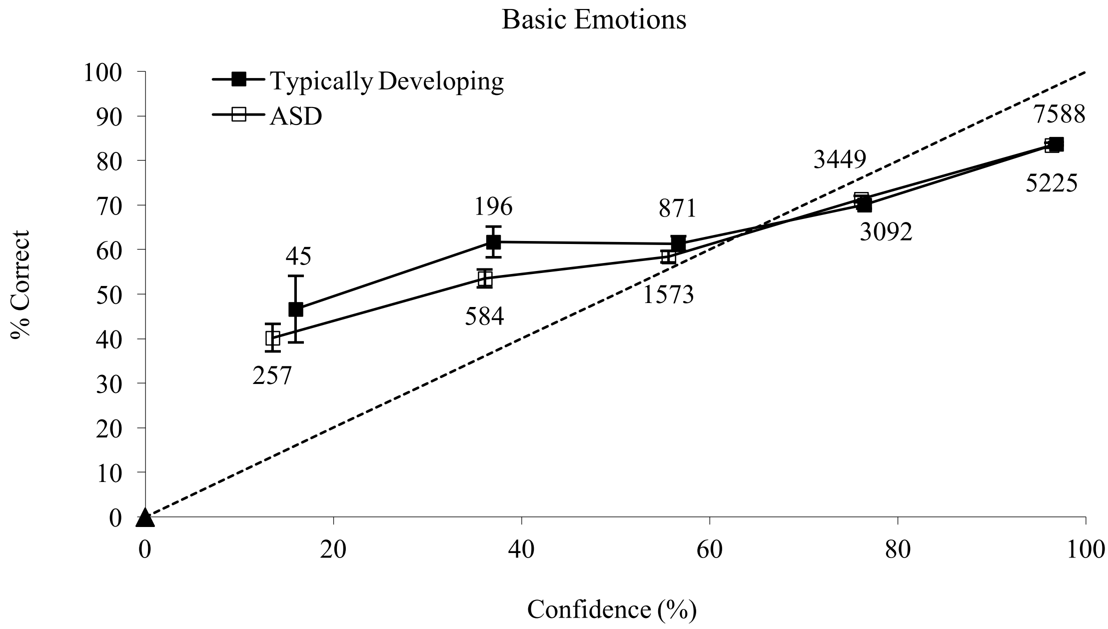
Confidence- Accuracy Calibration Curves for ASD and Typically Developing Groups Overall and for Both Response Formats Individually

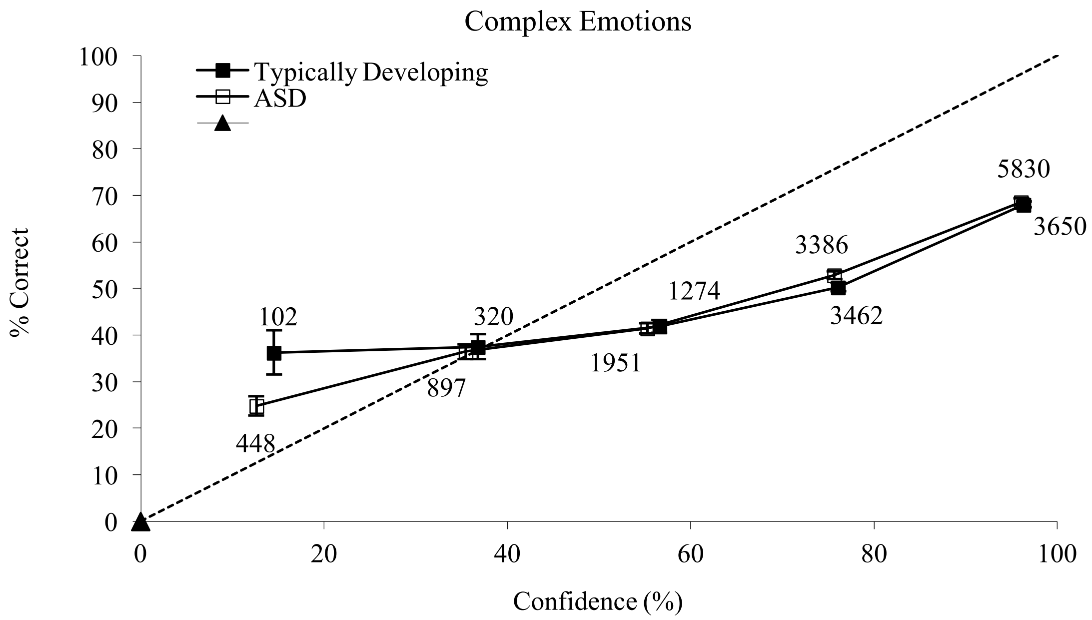


Note. The frequency of judgements of confidence are presented at each data point. Dotted line represents perfect calibration. Error bars represent standard error.

Figure 7

Confidence- Accuracy Calibration Curves for ASD and Typically Developing Groups for Basic and Complex Emotions





Note. The frequency of judgements of confidence are presented at each data point. Dotted line represents perfect calibration. Error bars represent standard error.

Perhaps the most striking observation is that (apart from the typically developing sometimes being a little more underconfident in the lower section of the scale) the calibration curves do not differ in any meaningful way for the two groups, despite the ASD group being characterised by lower percent agreement (i.e., accuracy), longer mean latencies and lower mean confidence compared to typically developing individuals. In other words, despite the clear performance differences between the two groups, the groups were not distinguished in terms of their metacognitive awareness in relation to their emotion recognition judgments.

Reactions to Emotions: Percent Agreement

As described within the emotion recognition section, percent agreement scores were calculated for each participant for each emotion. However, percent agreement within this section refers to agreement with the normative appropriate reaction to each emotion, not with the normative recognition response. To derive percent agreement scores for each emotion, a mean agreement score was calculated for each participant's performance across the relevant trials (4-7 trials) within the social task and converted into a percentage score allowing for comparison across emotions with varying trial numbers.

A 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) ANOVA was conducted on percent agreement scores. Emotion was a within-group factor and Group was a between-group factor. I assessed the main effect of emotion first followed by main effect of group and then its interactions with emotion. The ANOVA statistics appear in Tables 19 and 21. I subsequently explored reactions to emotions when recognition performance was high.

Consistent with previous sections, the results were read from the Greenhouse-Geisser correction to account for violations of sphericity as indicated by Mauchly's test for emotion, $\chi^2(65) = 201.70, p<.001$.

Emotion

To explore reactions to others' emotions, the main effect of emotion on percent agreement was assessed. There was a significant main effect of emotion (see Table 19), with mean percent agreement varying quite markedly across emotions as shown in Table 20. It was notable that the three emotions on which participants performed best were the basic emotions sad, happy and surprised, and the three worst were the complex emotions jealous, hurt and ashamed.

Table 19

Mixed ANOVA Outcomes for the Effects of Emotion on Reaction Percent Agreement and its Interaction with Group

Source	*df*	Mean Square	*F*	*p*	Partial η^2
Emotion	8.22	41352.48	105.26	<.001	.45
Error (Emotion)	1051.68	392.85			
Emotion×Group	8.22	1227.25	3.12	.002	.02

Table 20

Percent Agreement for Reactions to Emotion for Each Emotion in Descending Order

Emotion	*M* (%)	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Sad*	63.46	1.80	59.90	67.02
Happy*	62.65	1.96	58.77	66.52
Surprised*	51.73	1.77	48.23	55.23
Frustrated	45.16	1.77	41.67	48.66
Afraid*	43.15	2.00	39.19	47.11
Worried	41.03	2.06	36.95	45.12
Disappointed	40.03	1.55	36.96	43.10
Angry*	37.45	1.51	34.47	40.44
Disgusted*	35.43	1.59	32.29	38.57
Jealous	32.89	2.71	27.53	38.25
Hurt	25.36	1.55	22.30	28.43
Ashamed	7.02	.95	5.15	8.89

Note. * is a basic emotion

Group: ASD Versus Typically Developing

Given that ASD individuals demonstrated poorer emotion recognition than typically developing individuals, it is possible that ASD individuals would also be less likely to demonstrate appropriate reactions to others' emotions. The typically developing group showed greater mean percent agreement with the normative reactions to emotions (M = 42.99, SE = 1.53, 95% CI [39.95, 46.02] than the ASD group (M = 37.90, SE = 1.58, 95% CI [34.77, 41.04] and there was a statistically significant main effect of group. The effect size index, however, indicated a relatively weak effect (see Table 21).

Controlling for VCI, A-ToM-social or both measures did not have meaningful impacts on the main effect of group. Examination of the inferential statistics in Table 21 reveals that the effect size for group was not affected by considering either or both of the covariates. Consistent with previous sections, due to the group differences in age, correlations between age and reaction to emotions percent agreement for each emotion was explored for both groups. The correlations across all emotions were weak and, for 5 of the 12 emotions were negative, ranging from -0.08 to 0.00 indicating that younger participants had higher percent agreement for those 5 emotions. For the remaining 7 emotions (with correlations ranging from 0.00 to 0.12) older participants recorded higher percent agreement. Therefore, age was also used as a covariate and as displayed in Table 21, the effect size for group was not affected.

Interaction Effects Involving Group: Emotion × Group

As shown in Table 19 there was a significant two-way interaction (with a small effect size) between emotion and group. Although Figure 8 shows that means for the typically developing group were generally higher than for the ASD group, the error bars indicate large overlap for many emotions. This is confirmed by the pairwise comparisons shown in Table

22 which indicate that typically developing and ASD participants differed on only 5 of the 12 emotions, 4 of which were basic emotions (i.e., angry, happy, sad, surprised).

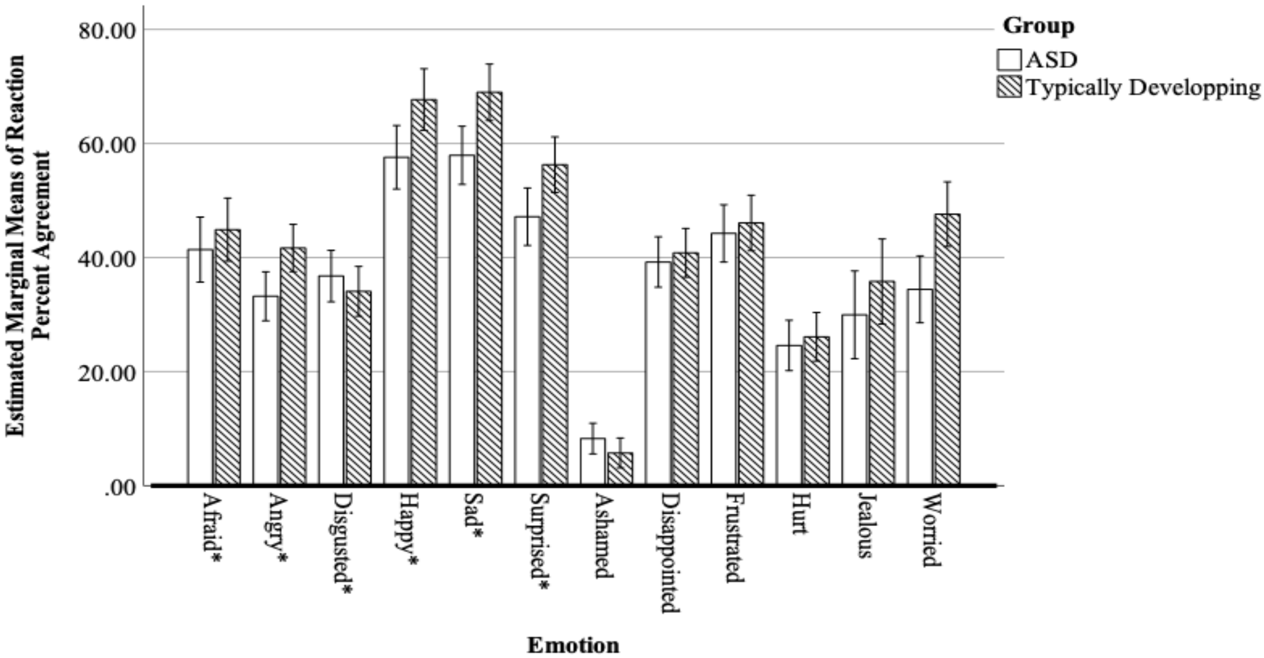
Table 21

Mixed ANOVA Outcomes for Main Between-Subjects Effect of Group (with and without Covariates of A-ToM, VCI and Age) on Percent Agreement for Reaction to Emotions

Source	*df*	Mean Square	*F*	*p*	Effect Size (partial η^2)
		No covariates			
Group	1	10067.67	5.32	.023	.04
Error	128	1892.14			
		Covariate of A-ToM			
A-ToM	1	2420.56	1.28	.260	.01
Group	1	9488.48	5.03	.027	.04
Error	127	1887.97			
		Covariate of VCI			
VCI	1	38.66	.02	.887	.00
Group	1	9879.94	5.18	.025	.04
Error	127	1906.73			
		Covariates of A-ToM and VCI			
A-ToM	1	2489.49	1.31	.255	.01
VCI	1	107.60	.06	.812	.00
Group	1	9586.20	5.04	.027	.04
Error	126	1902.10			
		Covariate of Age			
Age	1	1681.75	.89	.348	.01
Group	1	11698.28	6.18	.014	.05
Error	127	1893.79			

Figure 8

Two-Way Interaction of Emotion and Group on Percent Agreement for Reaction to Emotions



Note. Error bars represent 95% confidence interval, * is a basic emotion

Table 22

Pairwise Comparison Outcomes Examining the Emotion × Group Interaction on Percent Agreement for Reactions to Emotions

Group Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
TD-ASD	Afraid*	3.50	4.01	.384	-4.43	11.42
	Angry*	8.46	3.02	.006	2.50	14.43
	Disgusted*	-2.69	3.17	.398	-8.97	3.59
	Happy*	10.10	3.92	.011	2.35	17.85
	Sad*	11.04	3.60	.003	3.92	18.16
	Surprised*	9.12	3.54	.011	2.12	16.13
	Ashamed	-2.52	1.89	.185	-6.26	1.22
	Disappointed	1.60	3.11	.607	-4.54	7.75
	Frustrated	1.84	3.53	.604	-5.16	8.83
	Hurt	1.52	3.10	.626	-4.62	7.65
	Jealous	5.86	5.42	.281	-4.86	16.58
	Worried	13.17	4.12	.002	5.01	21.32

Note. * is a basic emotion, ** is based on estimated marginal means, 'TD' is typically developing

Reactions to Emotions When Recognition Performance was High

A failure to react appropriately to others' emotions is not surprising if the individual has not correctly recognised the emotion being displayed. Of particular interest, therefore, is whether individuals reacted appropriately to the emotions of others when they had in fact correctly recognised the other individual's emotion. Consequently, I examined participants' reactions to emotions when their emotion recognition performance was reasonably good. Reasonably good emotion recognition performance was defined separately for each of the 12 emotions and required the individual achieving a percent agreement score for any particular emotion of 75% or higher. I examined performance at the level of the individual emotion because individuals' performance varied across emotions, as did the number of individuals performing well. Setting the criterion at 75% agreement was done in an effort to secure a reasonable number of participants who recognised any particular emotion with reasonable accuracy. It is important to note that, irrespective of these efforts, two emotions (surprised and disappointed) contained very small participant numbers ranging from 3-7 participants whose exceeded 75% agreement. For the other emotions, 15-65 individuals exceeded the 75% agreement criterion.

From Table 23, angry was the only emotion with a significant difference between groups in reactions to others' emotions. However, the effect size indices suggest a medium effect for two emotions (hurt and worried) and a relatively weak effect for 4 other emotions. Thus, for those who scored relatively well in emotion recognition, it is possible that for at least 7 emotions (i.e., Hedge's g effect sizes >.20) the differences between groups for emotion reaction would have been statistically significant with more power. Note, however, that in 2 cases (ashamed and hurt), the ASD individuals were superior. Thus, when emotion recognition was relatively good, ASD individuals performed worse than typically developing on 4 of the 6 basic emotions but only on one of the complex emotions.

Table 23

Independent Samples T-Test Outcomes for the Differences Between Groups for Reactions to Emotions When Emotion Recognition is High

Emotion	Group	*n*	*M(SD)*	*t(df)*	*p*	Hedge's g
Afraid*	ASD	37	41.22(24.61)	1.29(85)	.202	.276
	TD	50	47.50(20.91)			
Angry*	ASD	36	33.93(17.86)	2.18(81)	.032	.478
	TD	47	42.55(17.87)			
Disgusted*	ASD	44	38.26(21.52)	.81(77.76)	.422	.168
	TD	52	35.10(15.86)			
Happy*	ASD	31	61.29(24.52)	1.71(55.20)	.093	.426
	TD	33	70.56(1819)			
Sad*	ASD	16	62.05(23.72)	1.11(22.62)	.280	.380
	TD	28	69.39(15.66)			
Surprised*	ASD	7	57.14(14.29)	.00(8)	1.000	.000
	TD	3	57.14(14.29)			
Ashamed	ASD	26	8.24(12.24)	1.49(55)	.142	.391
	TD	31	4.15(8.41)			
Disappointed	ASD	5	30(12.78)	.13(6)	.915	.082
	TD	3	28.57(18.90)			
Frustrated	ASD	57	45.18(22.13)	.32(107.73)	.751	.058
	TD	65	46.35(17.91)			
Hurt	ASD	15	35.71(25.03)	1.59(40)	.120	.502
	TD	17	25.66(16.00)			
Jealous	ASD	21	36.31(31.60)	.36(46)	.721	.103
	TD	27	39.81(34.85)			
Worried	ASD	16	30.63(25.16)	2.02(34)	.051	.662
	TD	20	45.50(19.05)			

Note. * is a basic emotion, 'TD' is typically developing

Reactions to Emotions: Latency

A 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) ANOVA was conducted on latency, with emotion and group as within- and between-group factors, respectively. The ANOVA statistics appear in Tables 24 and 26. As for percent agreement for reactions to emotions, I examined the main effect of emotion first followed by main effect of group and then its interactions with emotion.

The results were read from the Greenhouse-Geisser correction to account for violations of sphericity as indicated by Mauchly's test for emotion, $\chi^2(65) = 989.13$, $p<.001$.

Emotion

To explore latency of reaction to other people's emotions across basic and complex emotions, the main effect of emotion on latency was assessed. There was a significant but weak effect of emotion (see Table 24), with mean latency varying across emotions (see Table 25). Table 25 shows that 5 (i.e., surprised, disgusted, afraid, sad, happy) of the 6 basic emotions were among the 6 fastest emotion reactions while 5 (i.e., jealous, ashamed, frustrated, hurt, worried) of the 6 complex emotions were among the 6 slowest reactions.

Table 24

Mixed ANOVA Outcomes for the Effects of Emotion on Reaction Latency and its Interaction with Group

Source	*df*	Mean Square	*F*	*p*	Partial η^2
Emotion	3.06	98.62	5.96	<.001	.04
Error (Emotion)	391.83	16.56			
Emotion×Group	3.06	11.77	.711	.549	.01

Table 25

Latency for Reactions to Emotion for Each Emotion in Ascending Order

Emotion	*M* (sec.)	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Surprised*	3.36	0.12	3.12	3.60
Disgusted*	3.46	0.14	3.18	3.73
Disappointed	3.47	0.12	3.23	3.70
Afraid*	3.47	0.15	3.18	3.77
Sad*	3.48	0.14	3.19	3.76
Happy*	3.54	0.16	3.21	3.86
Jealous	3.55	0.16	3.23	3.87
Ashamed	3.55	0.14	3.27	3.83
Angry*	3.70	0.19	3.32	4.08
Frustrated	3.77	0.18	3.42	4.13
Hurt	4.08	0.28	3.53	4.64
Worried	5.03	0.50	4.05	6.01

Note. * is a basic emotion

Group: ASD Versus Typically Developing

The possibility that ASD individuals may be slower at reacting to the emotions of others compared to typically developing individuals was evaluated. The main effect of group was significant (see Table 26), with shorter mean latencies for the typically developing group (M = 3.13, SE = .16, 95% CI [2.80, 3.45] compared with the ASD group (M = 4.28, SE = .17, 95% CI [3.95, 4.62]). The effect size index indicated a strong effect (see Table 26).

Controlling for VCI, A-ToM-social or both measures did not have meaningful impacts on the main effect of group. Examination of the inferential statistics in Table 26 reveals that the effect size for group was not affected by considering either or both of the covariates. Consistent with previous sections, due to the group differences in age, correlations between age and reaction to emotions latency for each emotion was explored for both groups. The correlations across all emotions were weak for 10 of the 12 emotions with correlations ranging from 0.10. to 0.29; for two emotions (sad and surprised) correlations were moderate ranging from 0.31 to 0.33. The correlations indicate that older participants recorded longer latencies across all emotions. Therefore, age was also used as a covariate in the analysis of reaction to emotions latency. As displayed in Table 26, the effect size for group was reduced.

Table 26

Mixed ANOVA Outcomes for Main Between-Subjects Effect of Group (with and without Covariates of A-ToM, VCI and Age) on Latency for Reactions to Emotions

Source	*df*	Mean Square	*F*	*p*	Effect Size (partial η^2)
		No Covariates			
Group	1	520.92	23.95	<.001	.16
Error	128	21.75			
		Covariate of A-ToM			
A-ToM	1	15.17	.70	.406	.01
Group	1	509.33	23.36	<.001	.16
Error	127	21.81			
		Covariate of VCI			
VCI	1	.24	.01	.917	.00
Group	1	514.92	23.49	<.001	.16
Error	127	21.92			
		Covariates of A-ToM and VCI			
A-ToM	1	15.62	.71	.401	.01
VCI	1	.69	.03	.860	.00
Group	1	509.30	23.18	<.001	.16
Error	126	21.97			
		Covariate of Age			
Age	1	149.08	7.18	.008	.05
Group	1	316.39	15.25	<.001	.11
Error	127	20.75			

Interaction Effects Involving Group

Whether ASD individuals' latencies in their reactions to the emotions of other people were affected differently by emotion type compared to those of typically developing individuals was evaluated. However, as shown in Table 24 there was a non-significant two-way interaction between emotion and group.

Reactions to Emotions: Confidence

A 12 (Emotion: afraid, angry, ashamed, disappointed, disgusted, frustrated, happy, hurt, jealous, sad, surprised, worried) × 2 (Group: ASD, typically developing) ANOVA was conducted on confidence. As with percent agreement and latency, there was one within-group factor (emotion) and one between-group factor (group). Refer to Tables 27 and 29 for analysis outcomes.

Mauchly's test indicated violations of sphericity for emotion, $\chi^2(65) = 222.19$, $p<.001$, therefore, the results were read from the Greenhouse-Geisser correction.

Emotion

To examine confidence in providing appropriate reactions to others' emotions across basic and complex emotions, the main effect of emotion on confidence was assessed. There was a significant main effect of emotion with a medium effect size (see Table 27). The mean confidence varying across emotions (see Table 28), However, there were no obvious pattern of differences between basic versus complex emotions.

Table 27

Mixed ANOVA Outcomes for the Effects of Emotion on Reaction confidence and its Interaction with Group

Source	*df*	Mean Square	*F*	*p*	Partial η^2
Emotion	7.78	975.61	13.15	<.001	.09
Error (Emotion)	996.09	74.22			
Emotion×Group	7.78	205.51	2.77	.005	.02

Table 28

Confidence for Reactions to Emotion for Each Emotion in Descending Order

Emotion	*M*	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Frustrated	81.66	1.52	78.65	84.67
Sad*	79.82	1.43	76.99	82.66
Disgusted*	79.16	1.47	76.26	82.06
Worried	78.08	1.58	74.96	81.20
Surprised*	77.95	1.61	74.76	81.14
Ashamed	77.91	1.53	74.89	80.93
Angry*	77.35	1.58	74.22	80.48
Happy*	77.16	1.65	73.90	80.41
Afraid*	77.04	1.61	73.86	80.22
Disappointed	75.98	1.58	72.85	79.11
Hurt	74.56	1.72	71.16	77.96
Jealous	72.92	1.83	69.31	76.53

Note. * is a basic emotion

Group: ASD Versus Typically Developing

Confidence in the appropriateness of reactions to other people's emotions was explored across ASD and typically developing participants. As demonstrated in Table 29, there was a significant main effect of group with mean confidence higher for the typically developing group (M = 83.03, SE = 2.05, 95% CI [78.97, 87.09] compared to the ASD group (M = 71.91, SE = 2.12, 95% CI [67.72, 76.09]. The effect size index indicated a medium effect (see Table 29).

Controlling for VCI, A-ToM-social or both measures did not have meaningful impacts on the main effect of group. Examination of the inferential statistics in Table 29 reveals that the effect size for group was not affected by considering either or both of the covariates. Correlations between age and reaction to emotions confidence ratings for each emotion were explored for both groups. The correlations across all emotions were weak with correlations ranging from -.01 to .18. Generally, however, the correlations were positive, indicating higher confidence for older participants ratings. Therefore, age was also used as a covariate in the analysis of reaction to emotions confidence. As displayed in Table 29, the effect size for group actually increased a little.

Table 29

Mixed ANOVA Outcomes for Main Between-Subjects Effect of Group (with and without Covariates of A-ToM, VCI and Age) on Reaction Confidence Ratings

Source	*df*	Mean Square	*F*	*p*	Effect Size (partial η^2)
		No Covariates			
Group	1	48161.63	14.23	<.001	.10
Error	128	3385.46			
		Covariate of A-ToM			
A-ToM	1	3588.64	1.06	.305	.01
Group	1	49498.28	14.63	<.001	.10
Error	127	3383.86			
		Covariate of VCI			
VCI	1	459.02	.14	.714	.00
Group	1	46963.85	13.78	<.001	.10
Error	127	3408.51			
		Covariates of A-ToM and VCI			
A-ToM	1	5004.37	1.47	.227	.01
VCI	1	1874.74	.55	.459	.00
Group	1	47817.34	14.08	<.001	.10
Error	126	3395.84			
		Covariate of Age			
Age	1	13636.65	4.13	.044	.03
Group	1	60090.02	18.18	<.001	.13
Error	127	3304.75			

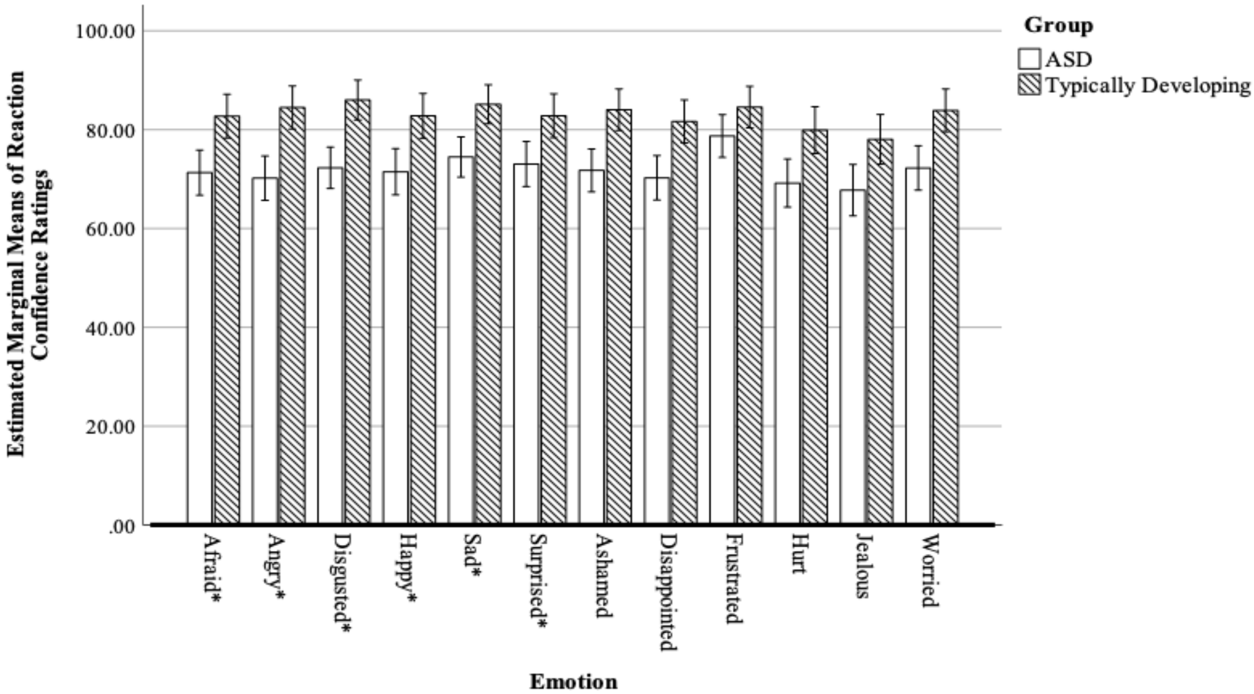
Interaction Effects Involving Group

As shown in Table 27 there was a significant two-way interaction (with a small effect size) between emotion and group, indicating that there were disproportionate effects of emotion type on confidence for ASD participants compared with typically developing.

Figure 9 shows that means for the typically developing group were generally higher than for the ASD group for 11 of the 12 emotions. The error bars indicate overlap for the emotion frustrated. The pairwise comparisons shown in Table 30 confirmed these observations, indicating that typically developing and ASD participants significantly differed on all emotions except for frustrated. There did not appear to be any patterns reflecting different effects for basic versus complex emotions.

Figure 9

Two-Way Interaction of Emotion and Group on Confidence Ratings for Reactions to Emotions



Note. Error bars represent 95% confidence interval, * is a basic emotion

Table 30

Pairwise Comparison Outcomes Examining the Emotion × Group Interaction on Confidence Ratings for Reactions to Emotions

Group Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
TD-ASD	Afraid*	11.44	3.22	.001	5.08	17.80
	Angry*	14.30	3.16	<.001	8.04	20.55
	Disgusted*	13.72	2.93	<.001	7.91	19.52
	Happy*	11.32	3.29	.001	4.81	17.83
	Sad*	10.67	2.86	<.001	5.00	16.34
	Surprised*	9.78	3.22	.003	3.40	16.15
	Ashamed	12.28	3.05	<.001	6.24	18.33
	Disappointed	11.41	3.17	<.001	5.15	17.68
	Frustrated	5.86	3.04	.057	-.17	11.88
	Hurt	10.71	3.44	.002	3.91	17.51
	Jealous	10.28	3.65	.006	3.06	17.51
	Worried	11.66	3.15	<.001	5.42	17.90

Note. * is a basic emotion, ** is based on estimated marginal means, 'TD' is typically developing

CHAPTER 4

Discussion

The study discussed within this thesis had a primary focus of exploring emotion recognition and reaction to the emotions of others in adults with ASD. Participants with ASD showed lower percent agreement—usually referred to as emotion recognition accuracy—than typically developing participants, regardless of stimulus type, response format and emotion. ASD participants were also slower than typically developing participants at emotion recognition, and significantly less confident, regardless of stimulus type, response format and emotion. There were also some tentative signs that they reacted less appropriately than typically developing participants to many of the basic emotions, and were less confident and slower at reacting regardless of emotion type. But strikingly, ASD and typically developing groups displayed similar levels of metacognitive awareness of their limitations. Previous research has said little about latency of emotion recognition, awareness of performance and reactions to emotions. Thus, this study provided a more comprehensive picture of emotion recognition difficulties of adults with ASD.

Importantly, these findings were obtained with reasonably sized samples, with (a) multiple observations for each participant within each cell of the design, (b) FSIQ for ASD people in the typical range (85-143), (c) verbal ability, perspective taking (ToM) ability and age controlled, (d) stimulus types varying in contextual information (static, dynamic, social) and (e) free-report of recognition responses. In sum, the study addressed many of the issues raised in Chapter 1 regarding previous research.

The group differences detected did not appear to reflect IQ, emotion type, the ability to use cues provided by multiple-choice response format or by stimuli that were more or less inclusive of context. Within everyday life, facial expressions are rarely interpreted without the presence of context and other pertinent social information. Therefore, it is unsurprising that social stimuli are better recognised by both groups given the greater experience in interpreting emotions by using readily available information to make inferences. Facial expressions have been previously referred to as inherently ambiguous with interpretation being strongly reliant on the context in which they are presented (Hassin et al., 2013). This is a particularly important notion for emotional facial configurations that appear similar: for example, when discriminating emotions that can have crossover presentations such as the facial configurations for surprise versus fear.

As recognition became enhanced through the stimulus layers from static to dynamic to social, it is likely (although beyond the scope of this study) that more than just facial muscle movement strengthened recognition, with the possibility that the coupling of both facial and body movement as well as context contributed to these recognition patterns. Explained further, the progression from static to dynamic saw improvements in recognition based on movement alone, with recognition further strengthened when context and body movement was introduced (as in the social stimuli). Many studies have demonstrated that with body movement alone (i.e., without facial expressions) emotions can be successfully identified (Atkinson et al., 2004; Crane & Gross, 2007, 2013; De Meijer, 1989). Taken together, the advantage of social stimuli become obvious as participants are offered important information that is not available in the presentation of static stimuli.

Latency and Metacognitive Awareness

The current study also provided hitherto neglected examinations of latency and metacognitive awareness across a large number of trials. Typically, social interactions are likely to benefit from rapid emotion recognition and reaction to others' emotions. In absence of this rapid recognition and reaction the functional outcome for social interactions is likely to be hindered. Within this study, ASD participants were not only less accurate but also were slower in recognising emotions than typically developing subjects. This is consistent with previous research showing that ASD participants demonstrate more accurate emotion recognition when stimulus presentations were slowed down (e.g., Gepner et al., 2001; Tardif et al., 2007). However, even if recognition was found to be just as accurate as that shown by typically developing participants, social interactions would still be negatively impacted by slow recognition when rapid recognition is demanded. Knowing what someone is feeling may be of little value when it takes too long to decide or infer what that emotion might be even before providing a reaction, likely resulting in missed information (e.g., subsequent emotional expressions) and missed opportunities (e.g., to react, to keep up with changing emotional expressions) all of which contribute to awkward social exchanges.

Finding that both accuracy and latency are impaired for ASD individuals could well diminish effective social functioning within the fast-paced nature of real-world interactions. This conceivably indicates processing difficulties and the possibility of a greater tendency towards reflective processing. It is important to highlight that the present study cannot distinguish whether slower emotion recognition reflects processing difficulties made worse by extra caution or whether extra caution alone can account for differences.

The Dual Process Theory (Evans, 2008; Wason & Evans, 1974) offers insight into general differences in processing styles that may help explain, at least to some degree, the latency results found. As described in Chapter 1, Type 1 processing involves intuitive rapid decision-making/reasoning whereas Type 2 relies on deliberate, more conscious decision-making/reasoning. The association between Type 2 processing and autism has been established in previous research (e.g., Brosnan et al., 2016, 2017). Emotion processing is argued to be an intuitive process that requires rapid and automatic extraction and interpretation of information from the social environment (i.e., Type 1 processing). Thus, individuals with preferences for Type 2 processing may encounter problems relating to delayed recognition and, in turn, impaired social-emotional functioning. There may be situations where deliberate processing is beneficial such as within certain academic or professional contexts. However, within social contexts the outcomes are likely detrimental.

Interestingly, the present study found significant group differences in latency for both emotion recognition and reaction to emotions. Put simply, before you can respond to someone's emotions you must first recognise them. If recognition is impaired and slow, this will cause challenges prior to even beginning to formulate a reaction. There must be coherent inferences of emotion, formulation of appropriate reactions and execution of those reactions, all within an appropriate timeframe, otherwise social interactions within a naturalistic setting are likely to be awkward and problematic.

The examination of metacognitive awareness via the calibration analyses revealed patterns in both groups consistent with what has been described as the hard-easy effect in the confidence literature: that is, over-confidence under conditions when discriminations are more difficult and under-confidence when they are relatively easy. Despite any challenges in metacognitive monitoring of emotion recognition accuracy, as reflected in less than perfect confidence-accuracy calibration (for both groups), the absence of meaningful group differences in calibration affirms that awareness of the accuracy of one's emotion recognition was not a contributor to group differences in emotion recognition. Intact self-awareness is important for the regulation of social-communicative behaviours. However, equivalent awareness of one's abilities in the two groups does not necessarily mean that those with ASD will actively regulate social interactions based on their understanding of the likely accuracy of their inferences. This skill is referred to within relevant literature as metacognitive control. Previous research by Sawyer et al. (2014) found that those with ASD continued to struggle with metacognitive control even when metacognitive monitoring was shown to be equal to that of typically developing individuals. ASD participants failed to act on their metacognitive monitoring skills and filter incorrect responses. Therefore, it is uncertain if, in this instance, intact monitoring improves real-world social outcomes.

Reactions to Other's Emotions

The current study also provided information on reactions to emotions, an area previously overlooked in much of the existing research. ASD participants reacted less appropriately to others' emotions than typically developing participants. Interestingly though, when participants who scored less than 75% correct across the trials for each emotion were excluded, significant group differences for reaction to emotions were not found for 11 of 12 emotions. However, the effect size statistics tentatively suggest the possibility that ASD individuals may be less effective at reacting appropriately to basic emotions. These patterns suggest that even when recognition of others' emotions is good, group differences in reactions to emotions may exist. This finding is relevant for understanding how social interactions involving those with ASD have previously been characterised as being plagued by an array of negative outcomes such as lacking empathy (e.g., Baron-Cohen & Wheelwright 2004; Mul et al., 2018). The possibility that these negative outcomes in part stem from emotion recognition inaccuracies, extended recognition timeframes, and generally poor reactions to others (irrespective of accurately inferring emotions) seems plausible, albeit, beyond the scope of what can be concluded from the current study.

Limitations and Future Directions

The current study is not without limitations. An important caveat to consider when interpreting the results of the current study relates to the response format data. The result showed that free-report responses were more challenging for both groups, suggesting that there was a level of strategy employed beyond actual recognition abilities for multiple-choice responses which was not possible for free-report. However, to obtain sufficient power to assess free-report performance, something that has been often neglected in previous research, the presentation order of response format was fixed (i.e., free-report followed by multiple-choice). Therefore, even though ASD participants performed worse than typically developing participants at recognising emotions in both response formats, exactly what would have happened under the multiple-choice response format if free-report had not preceded it is unknown. Specifically, questions around whether free-report responding enhanced performance on multiple-choice responses (to an unknown degree) and whether group differences under multiple-choice responding depended on whether it was preceded by free-report responding (e.g., if one group would have benefited more than the other from completing free-report before multiple-choice) remain. Nevertheless, real-world interactions almost exclusively rely on free-report responding, with available sources of information assisting with 'working out' answers from conventional social cues (e.g., facial movement, body movement, situational context, verbal tone etc.). Therefore, the current study arguably highlights potential difficulties encountered by ASD individuals when confronted with providing more realistic emotion recognition responses.

Both groups were relatively evenly matched on VCI. Moreover, VCI did not moderate results despite participants having to describe what they saw (i.e., provide an emotion label in free-report responding) and sort between emotion label options (i.e., in multiple-choice responding). However, the potential moderating effects of other aspects of intelligence cannot be ruled out. For example, the Perceptual Reasoning Index (PRI) of the WASI-II might well be related to the ability to dissect and interpret subtle aspects of facial configurations and contextual information. Nevertheless, the consistency of the results of the current study and the strength of the group differences observed suggest it seems unlikely that variations in PRI would have underpinned the substantial differences observed between groups.

The experimental conditions of the study unfortunately had the unavoidable consequence of not representing real-world contexts in a manner that allows a more realistic estimation of emotion recognition difficulties. I noted that, during the testing phase, three ASD participants described that they completed the emotion reaction tasks by writing what they ‘know’ they should say, not what they would ‘actually’ say in reaction to someone’s emotion in real-life situations. Thus, there is a possibility that greater difficulties than observed here might be observed in real social situations. Therefore, future research should aim to address this issue, perhaps initially through stricter instructions given to participants to answer with what they would say in real-life situations, not what they know they should say, and subsequently by assessing responses in more life-like settings.

Representation of everyday challenges is imperative to draw conclusions which accurately replicate the confronts in real-life settings. As described in the key diagnostic criteria, individuals with ASD have been characterised to struggle with eye contact in social interactions. However, it is not known whether this translates when viewing an artificial situation through technological means (e.g., computer). One possibility might be that an increase in motivational drive coupled with a potential reduction in cognitive load (O’Neil & Jones, 1997) may increase the likelihood that those with ASD attend to more details than they would in person. An individual with ASD may recognise more emotions of others in these specific artificial situations due to the reduced likelihood to produce overstimulation. Overstimulation is often reported in people with ASD and refers to sensory atypicalities such as hyposensitivity to impending stimuli (e.g., lighting and sounds) (Pellicano & Burr, 2012). Therefore, it has been argued that the ability to focus on micro-expressions (particularly eye contact) becomes cognitively and emotionally overstimulating and overwhelming during real-life social interactions (Baron-Cohen, 1997; O’Neil & Jones, 1997) but not in artificial situations. The current study does not explore whether participants would perform the same if perceiving emotions embedded in an interaction rather than observing images and clips. An avenue for future research is to explore or compare abilities of those with ASD and typically developing persons to recognise and react to emotions presented in more naturalistic ways.

Stimulus intensity may also contribute to the difficulties experienced by individuals with ASD. The current study used stimuli with high intensity for all 12 emotions for continuity. Research has previously highlighted that stimuli presented with lower intensities more consistently demonstrate difficulties (e.g., Law Smith et al., 2010; Wong et al., 2012). Thus, difficulties identified in the present study have the potential to be much greater within real-world contexts where emotional displays vary in presentation intensity. Additionally, the current study is limited to posed expressions (i.e., produced at the request of another person). Real-life interactions almost exclusively consist of spontaneous expressions. Spontaneous expressions are more subtle than posed expressions and can represent more than one emotion (Matsumoto & Willingham 2009). The current study speaks little to the abilities of those with ASD to recognise emotion blends or emotions presented at low intensity (i.e., subtle emotion presentations). The inclusion of both low and high intensity stimuli as well as emotion blends is important to explore in the future.

The possibility that comorbidities exist alongside a diagnosis of ASD that might account for at least some emotion recognition difficulties also constitutes a limitation within the current study. For example, it has been suggested that co-occurring alexithymia, which presents as difficulties in emotional awareness such as the inability to identify and describe one's own emotions, might account for emotion recognition difficulties rather than ASD per se. Growing attention to the intersection between alexithymia and ASD has led to reports of alexithymia traits being present in at least 50% of clinical populations diagnosed with ASD (e.g., Hill et al., 2004; Bird & Cook, 2013). An interesting study by Cook et al. (2013) compared performance in facial identity and emotional expression over two experiments with

ASD and typically developing participant groups in which both also included a subgroup of participants that also met the criteria for alexithymia. Among the findings of these experiments, it was suggested that alexithymia, not autism, predicted abilities in emotion expression recognition when asked to assign expressions to one of two label choices. This is an important finding as reports of group differences in emotion recognition between ASD and typically developing participants could be explained in part due to higher rates of alexithymia within ASD samples rather than a true reflection of difficulties represented by ASD alone. Given that there are greater reports of alexithymia within ASD populations, it would be justifiable to suggest that the ASD samples used in related research might have a higher representation of alexithymia traits than in typically developing samples. Thus, examining and controlling for alexithymia should be strongly considered.

Given the variability in the presentation and severity of ASD symptoms, the question of a core deficit in emotion processing in ASD becomes difficult to answer confidently. However, the heterogeneity of performance of some individuals within the ASD group and the overlap with the typically developing group is inconsistent with the notion of core deficits, particularly given that emotion recognition was relatively strong (albeit not as strong as the typically developing group) with added contextual information (such as within the social stimulus task) which comprised information likely available within everyday social interactions. The current study is constrained in terms of its ability to offer insights into potential explanations of the performance heterogeneity in some of the ASD participants under certain conditions. However, from what is known about the heterogeneity of ASD symptoms and comorbidities, it is possible that subgroups of ASD participants are impaired more than others with regards to recognising and reacting to emotions. Recent research has attempted to identify subgroups within child samples with ASD, by identifying varying performance in recognition of social signals, expressive aspects of social communication and the motivation to interact (Uljarević et al., 2020). Furthermore, it is often reported that rates of depression and anxiety related symptoms are higher in some adults with ASD (e.g., Nah et al., 2018). These conditions have the potential to impair processing of face emotion stimuli, thereby affecting performance. Therefore, more research aiming to identify potential subgroups related to social performance is essential in understanding various subtleties of emotion performance for a range of adults with ASD presentations.

Similarly, the question as to whether certain traits associated with ASD that might also be seen to some degree in typically developing adults may also account for impairments in emotion processing remains unanswered. Explained further, it is not known whether a possible "threshold effect" exists, whereby, despite adults demonstrating some ASD characteristics or traits, their emotion processing is only impaired if they demonstrate ASD symptoms to a level that satisfies a clinical diagnosis. These questions will remain unanswered until sufficiently large sample studies are conducted to permit meaningful comparisons of performance on emotion recognition tasks across the range of ASD characteristics.

The current study did not attempt to evaluate underlying mechanisms that may have contributed to group differences generally. Therefore, explanations for the ASD group underperforming in comparison to the typically developing group can merely be hypothesised by drawing on existing literature. Challenges in emotion recognition and reaction for the ASD participants are conceivably the result of many contributing factors that result in abnormal emotion processing. Below I briefly discuss three main areas (i.e., motivation, perception and attention) for consideration as possible mechanisms underpinning group differences in emotion recognition and reaction.

Motivation

Emotion processing is said to be, in part, driven by motivational factors (Skelly & Decety, 2012). The "Social motivational hypothesis" has been discussed and explored in relation to individuals with ASD to explain why deficits in emotion processing might occur (see Chevallier et al., 2012, for a review). This hypothesis suggests that deficits in emotion processing are a consequence of abnormalities in social reward mechanisms (Dawson et al., 2005; Nuske et al., 2013). It is based on research showing that those with ASD demonstrate more positive emotions when engaged in self-absorbed activities compared to activities involving a partner. This suggests that there do not appear to be the same difficulties expressing emotions when the situation is found to be emotionally motivating. Additionally, ASD participants have previously demonstrated a preference towards non-social objects over people, faces and eyes when presented with images or video clips (Klin et al., 2002; Riby & Hancock, 2008 Trevisan & Birmingham, 2016) as well as reduced interest in recorded speech but not non-social sounds (Čeponienė et al., 2003; Klin, 1991). Not surprisingly, it has also been suggested that those with ASD report less enjoyment and interest in friendships (Baron-Cohen & Wheelwright, 2003; Trevisan & Birmingham, 2016). Taken together, it appears that people with ASD tend to have a preference for non-social activities and find these more emotionally motivating than social interactions. Given that emotion processing is developed through experience, decreased social motivation can potentially hinder the proper development of emotion processing abilities in those with ASD. Therefore, in the context of the current study, emotion recognition and reaction deficits found for ASD participants may have a motivational basis impacting in two ways: firstly, by reduced motivation to engage in emotion recognition and/or react to emotions and secondly, by a lack of motivation to engage in such exchanges, they are less well versed in recognition and reaction through a reduction in practice. Future research may wish to consider exploring participant motivation throughout the tasks, by perhaps having ‘motivational checks’. These motivational checks can be used to explore participant insights into their motivations to, perhaps decipher the emotional states of others generally and within the task stimuli. The possible inclusion of non-emotion related stimuli within these checks will be interesting to compare differences in motivation with emotion and non-emotion stimuli. It has been hypothesised that autistic individuals may be less motivated to attend to face stimuli than non-face stimuli. If so, processing difficulties would be exhibited on face but not non-face stimuli. Given Sawyer et al. (2014) found no evidence of such a pattern, and the testing session for participants was already several hours long, such a control condition was not included.

Perception

Abnormalities in perceptual processes have often been reported in people with ASD (e.g., Carther-Krone et al., 2016; Deruelle et al., 2004; Happé & Frith, 2006; Hobson et al., 1988). Atypical perceptual processing may not entirely explain or account for the social difficulties in those with ASD but it seems plausible that detail focus (attention to detail) may interfere with already abnormal social functioning. In particular, featural processing (analysing facial characteristics) can impair facial emotion recognition and reduce context-sensitive interpretation of social behaviour (Happé & Frith, 2006). For example, Ozonoff et al. (1991) reported that participants with ASD demonstrated difficulties in distinguishing emotions that shared perceptual features. For example, difficulty was reported in distinguishing between fear and surprise as both exhibited the same facial feature of an open mouth. This attention to featural details rather than the entirety of the face can help explain deficits in facial emotion recognition. This is supported through research which further suggests that those with ASD rely on fine-grained details rather than gestalt methods of processing (Happé & Frith, 2006; Joseph & Tanaka, 2003). But how they use features to decipher emotions is particularly useful to help explain reasons for differences. As touched on in Chapter 1, individuals with ASD tend to rely on explicit rule-based strategies to identify emotional expressions. Over time individuals with ASD may have learnt to recognise facial emotions based on consistencies in emotional expressions, referred to as 'rules'. For example, the 'rules' for the emotion 'sad' may include down-turned mouth, lowered eyebrows and narrowing of the eyes (Rutherford & McIntosh, 2007). Considering that consistency between facial configurations and emotion labels has been previously acknowledged (e.g., acceptance in literature of basic and complex emotions; see Chapter 1 for discussion), an exhaustive rule-based strategy would be functional for emotion recognition, although, not as accurate. Given the many ways to express emotions with many emotions overlapping in facial characteristics, focusing on one aspect of available information (e.g., one part of the face) as an indicator is likely to reveal inaccurate recognition of emotions. Due to this overlap, recognising and representing emotions becomes a challenging process for those with ASD. Currently it is not clear to what degree perceptual difficulties may affect emotion recognition abilities and, of course, the results of the current study do not address this question. However, these questions remain an important direction for future research. Some research possibilities are discussed in the next section on possible attentional difficulties.

Attention

A central element in emotion processing is the proper distribution of attentional resources (Wong et al., 2005). The social communication impairments that underpin ASD emerge gradually over the first two years of life. In typical development, basic attentional processes may provide a critical foundation for social communication abilities. Therefore, early attentional dysfunction may result in atypical development of important social communicative abilities (Keehn et al., 2013).

It has been well documented that those with ASD are likely to have atypical attention patterns marked by reduced attention to emotional stimuli (Klin et al., 2009). It is suggested that those with ASD demonstrate reduced attention to faces, abstaining from facial eye contact (Grice et al., 2005). Available research in the area largely focuses on children, nevertheless, it is valuable to assess observable patterns present during critical learning periods (such as in childhood). In children with ASD, reduced facial attention results in the decreased opportunity to become ‘face experts’ (Dawson et al., 2005; Faja et al., 2012). The majority of studies concerned with the influence of these altered attention patterns on emotion recognition have examined the effects of attention to the mouth or eye regions. Within available research, there is evidence to suggest that those with ASD demonstrate diminished attention to the eye region. Attention to the eye region is regarded as important for accuracy to infer emotions due to previous positive associations with more accurate facial emotion recognition in typically developing individuals (Hall et al., 2010). Those with ASD are reported to direct their attention to the mouth region rather than the eyes (Black et al., 2020; Tang et al., 2019). It remains unclear why there may be reduced attention to emotive information conveyed through eye expressions. Motivational explanations have been used to suggest why the eyes are neglected. Regardless, a reduced attention to key informational features of faces and abnormal processing of facial stimuli and emotional cues may together contribute to the difficulties with emotion recognition and understanding in those with ASD. Again, however, the current study did not address the possible roles of such mechanisms and this remains an important and interesting direction for future research.

A variety of approaches might be used to explore contributing attentional and perceptual factors. For example, eye-tracking measures could be used to compare attentional gaze on emotional stimuli which vary in presentation of context similar to what the current study employed. This would help investigate whether attentional gaze is altered depending on the level of context present. In a similar vein, examinations of attentional gaze across emotions or under different presentation conditions might shed light on differences in decision latency observed across groups. Additionally, experimental manipulations of the perceptual features of faces will likely continue to be a focus of research designed to elucidate perceptual processing issues. For example, research on featural versus holistic processing has largely focused on processing of static images. How such research generalizes to the processing of perceptual features of faces under more ecologically valid dynamic presentations would appear to be one obvious research direction.

Conclusion

The current study was a necessary contribution to the vast body of research on emotion recognition in ASD individuals. The focus on adults, the reasonable sample size, the range of emotions examined, as well as the examination of different stimulus types and response formats helps delineate areas of difficulty and under what conditions those difficulties may be more pronounced. Previously, there were only limited grounds within available research to declare confidently that difficulties in emotion recognition and reaction exist in adults with ASD. The current study demonstrates that regardless of stimulus type, response format or emotion, ASD individuals are less accurate, slower and less confident than typically developing participants at recognising and (quite likely) reacting to the emotions of others. Precisely how the emotion processing difficulties identified in the current study might play out in real interactions and shape social-communicative competence remain major issues for ongoing investigation.

References

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience*, *13*(2), 232–240. https://doi.org/10.1162/089892901564289

Ambadar, Z., Cohn, J. F., & Reed, L. I. (2009). All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of Nonverbal Behavior*, *33*(1), 17–34. https://doi.org/10.1007/s10919-008-0059-5

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). American Psychiatric Publishing.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Ashwin, C., Chapman, E., Colle, L., & Baron-Cohen, S. (2006). Impaired recognition of negative basic emotions in autism: A test of the amygdala theory. *Social Neuroscience, 1*(3–4), 349–363. https://doi.org/10.1080/17470910601040772

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6)*,* 717–746. http://doi.org/10.1068/p5096

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., Moscovitch, M., & Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724–732. https://doi.org/10.1111/j.1467-9280.2008.02148.x

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* MIT Press/Bradford Books.

Baron-Cohen, S., & Wheelwright, S. (2003). The friendship questionnaire: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *33*(5), 509–517. https://doi.org/10.1023/a:1025879411971

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175. https://doi.org/10.1023/b:jadd.0000022607.19833.00

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind. *Journal of Child Psychology and Psychiatry, 38*(7), 813–822. https://doi.org/10.1111/j.1469-7610.1997.tb01599.x.

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, 42*(2), 241–251. https://doi.org/10.1111/1469-7610.00715

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*(11)*,* 1833–1855. http://doi.org/10.1093/scan/nsx060

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1–68. https://doi.org/10.1177/1529100619832930

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. *Trends in Cognitive Sciences, 11*(8)*,* 327–332. http://dx.doi.org/10.1016/j.tics.2007.06.003

Barrett, L., F., Adolphs, R., Marsella, S., Martinez, A. M, & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1-68. https://doi.org/10.1177/1529100619832930930

Bebko, J. M., & Ricciuti, C. (2000). Executive functioning and memory strategy use in children with autism: The influence of task constraints on spontaneous rehearsal. *Autism, 4*(3), 299–320. https://doi.org/10.1177/1362361300004003006

Begeer, S., Koot, H. M., Rieffe, C., Meerum Terwogt, M., & Stegge, H. (2008). Emotional competence in children with autism: Diagnostic criteria and empirical evidence. *Developmental Review, 28*(3), 342–369. https://doi.org/10.1016/j.dr.2007.09.001

Bègue, I., Vaessen, M., Hofmeister, J., Pereira, M., Schwartz, S., & Vuilleumier, P. (2019). Confidence of emotion expression recognition recruits brain regions outside the face perception network. *Social Cognitive and Affective Neuroscience, 14*(1), 81–95. https://doi.org/10.1093/scan/nsy102

Berggren, S., Fletcher-Watson, S., Milenkovic, N., Marschik, P. B., Bölte, S., & Jonsson, U. (2018). Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability. *Developmental Neurorehabilitation*, *21*(3), 141–154. https://doi.org/10.1080/17518423.2017.1305004

Betz, N., Hoemann, K., & Barrett, L. F. (2019). Words are a context for mental inference. *Emotion*, *19*(8), 1463-1477. http://doi.org/10.1037/emo0000510

Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. *Translational psychiatry*, *3*(7), e285. https://doi.org/10.1038/tp.2013.61

Black, M. H., Chen, N. T. M., Iyer, K. K., Lipp, O. V., Bölte, S., Falkmer, M., Tan, T., & Girdler S. (2017). Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. *Neuroscience and Biobehavioral Reviews, 80,* 488–515. https://doi.org/10.1016/j.neubiorev.2017.06.016.

Black, M. H., Chen, N. T., Lipp, O. V., Bölte, S., & Girdler, S. (2020). Complex facial emotion recognition and atypical gaze patterns in autistic adults. *Autism*, *24*(1), 258–262. https://doi.org/10.1177/1362361319856969

Blair R. J. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, *14*(4), 698–718. https://doi.org/10.1016/j.concog.2005.06.004

Boraston, Z., Blakemore, S. J., Chilvers, R., & Skuse, D. (2007). Impaired sadness recognition is linked to social interaction deficit in autism. *Neuropsychologia*, *45*(7), 1501–1510. https://doi.org/10.1016/j.neuropsychologia.2006.11.010

Bormann-Kischkel, C., Vilsmeier, M., & Baude, B. (1995). The development of emotional concepts in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *36*(7), 1243–1259. https://doi.org/10.1111/j.1469-7610.1995.tb01368.x

Brewer, N., & Wells, G. L. (2006). The confidence-accuracy relationship in eyewitness identification: Effects of lineup instructions, foil similarity, and target-absent base rates. *Journal of Experimental Psychology: Applied, 12*(1), 11–30. https://doi.org/10.1037/1076-898X.12.1.11

Brewer, N., Young, R. L., & Barnett, E. (2017). Measuring theory of mind in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(7), 1927–1941. https://doi.org/10.1007/s10803-017-3080-x

Brewer, R., Biotti, F., Catmur, C., Press, C., Happé, F., Cook, R., & Bird, G. (2016). Can neurotypical individuals read autistic facial expressions? Atypical production of emotional facial expressions in autism spectrum disorders. *Autism Research*, *9*(2), 262–271. https://doi.org/10.1002/aur.1508

Brosnan, M., Ashwin, C., & Lewton, M. (2017). Brief report: Intuitive and reflective reasoning in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(8), 2595–2601. https://doi.org/10.1007/s10803-017-3131-3

Brosnan, M., Lewton, M., & Ashwin, C. (2016). Reasoning on the autism spectrum: A dual process theory account. *Journal of Autism and Developmental Disorders*, *46*(6), 2115–2125. https://doi.org/10.1007/s10803-016-2742-4

Buitelaar, J. K., Van Der Wees, M., Swaab-Barneveld, H., & Van Der Gaag, R. J. (1999). Theory of mind and emotion-recognition functioning in autistic spectrum disorders and in psychiatric control and normal children. *Development and Psychopathology, 11*(1), 39–58. https://doi.org/10.1017/ S0954579499001947

Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of simple and complex emotions in non-retarded children with autism. Journal of *Child Psychology & Psychiatry & Allied Disciplines, 33*(7), 1169–1182. https://doi.org/10.1111/j.14697610.1992.tb00936.x

Carroll, J., & Russell, J. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205-18. https://doi.org/10.1037/0022-3514.70.2.205

Carther-Krone, T. A., Shomstein, S., & Marotta, J. J. (2016). Looking without perceiving: impaired preattentive perceptual grouping in autism spectrum disorder. *PloS One*, *11*(6), 158566. https://doi.org/10.1371/journal.pone.0158566

Cassels, T. G., & Birch, S. A. (2014). Comparisons of an open-ended vs. forced-choice 'mind reading' task: Implications for measuring perspective-taking and emotion recognition. *PLoS One, 9*(12), e93653. http://doi.org/10.1371/journal.pone.0093653

Cassidy, S., Mitchell, P., Chapman, P., & Ropar, D. (2015). Processing of spontaneous emotional responses in adolescents and adults with autism spectrum disorders: Effect of stimulus type. *Autism Research*, *8*(5), 534–544. https://doi.org/10.1002/aur.1468

Cassidy, S., Ropar, D., Mitchell, P., & Chapman, P. (2014). Can adults with autism spectrum disorders infer what happened to someone from their emotional response? *Autism Research*, *7*(1), 112–123. https://doi.org/10.1002/aur.1351

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism, 9*(4), 428–449. https://doi.org/10.1177/1362361305056082

Ceponiene, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech-sound-selective auditory impairment in children with autism: They can perceive but do not attend. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5567–5572. https://doi.org/10.1073/pnas.0835631100

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Chevallier, C., Parish-Morris, J., McVey, A., Rump, K. M., Sasson, N. J., Herrington, J. D., & Schultz, R. (2015). Measuring social attention and motivation in autism spectrum disorder using eye-tracking: Stimulus type matters. *Autism Research, 8*(5), 620–628. https://doi.org/10.1002/aur.1479

Cook, J., Barbalat, G., & Blakemore, S. J. (2012). Top-down modulation of the perception of other people in schizophrenia and autism. *Frontiers in Human Neuroscience*, *6*, Article 175. https://doi.org/10.3389/fnhum.2012.00175

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not Autism, predicts poor recognition of emotional facial expressions. *Psychological Science*, *24*(5), 723–732. https://doi.org/10.1177/0956797612463582

Corden, B., Chilvers, R., & Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social–perceptual impairment in Asperger's syndrome. *Neuropsychologia, 46*(1), 137–147. https://doi.org/10.1016/j.neuropsychologia.2007.08.005

Crane, E. A., and Gross, M. M. (2007). Motion capture and emotion: Affect detection in whole body movement. In A. C. R. Paiva, R. Prada, & R. W. Picard (Eds.) *Affective Computing and Intelligent Interaction ACII.* Lecture Notes in Computer Science, vol 4738 (pp.95–101). Springer. https://doi.org/10.1007/978-3-540-74889-2_9

Crane, E. A., & Gross, M. M. (2013). Effort-shape characteristics of emotion-related body movement. *Journal of Nonverbal Behavior, 37*(2), 91–105. https://doi.org/10.1007/s10919-013-0144-2

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. https://doi.org/10.1207/s15326942dn2703_6

De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior, 13*(4), 247–268. https://doi.org/10.1007/BF00990296

Dentakos, S., Saoud, W., Ackerman, R., & Toplak, M. (2019). Does domain matter? Monitoring accuracy across domains. *Metacognition Learning 14*,(3) 413–436. https://doi.org/10.1007/s11409-019-09198-4

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. https://doi.org/10.1023/b:jadd.0000022610.09668.4c

Dissanayake, C., Sigman, M., & Kasari, C. (1996). Long-term stability of individual differences in the emotional responsiveness of children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *37*(4), 461–467. https://doi.org/10.1111/j.1469-7610.1996.tb01427.x

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development, 62*(6), 1352-1366. https://doi.org/10.2307/1130811

Dyck, M. J., Piek, J. P., Hay, D., Smith, L., & Hallmayer, J. (2006). Are abilities abnormally interdependent in children with autism? *Journal of Clinical Child and Adolescent Psychology, 35*(1), 20–33. https://doi.org/10.1207/ s15374424jccp3501_3

Dziobek, I., Fleck, S., Rogers, K., Wolf, O., & Convit, A. (2006). The ‘amygdala theory of autism’ revisited: Linking structure to behavior. *Neuropsychologia, 44*(10), 1891–99. https://doi.org/10.1016/j.neuropsychologia.2006.02.005

Eack, S. M., Mazefsky, C. A., & Minshew, N. J. (2015). Misinterpretation of facial expressions of emotion in verbal adults with autism spectrum disorder. *Autism*, *19*(3), 308–315. https://doi.org/10.1177/1362361314520755

Ekman, P., & Friesan, W. V. (1969). The repertoire of nonverbal behaviour: Categories, origins, usage, and coding. *Semiotica, 1*(1), 49–98. https://doi.org/10.1515/9783110880021.57

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues.* Prentice-Hall.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., & Ricci-Bitti, P. E. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, *53*(4), 712–717. https://doi.org/10.1037//0022-3514.53.4.712

Enticott, P. G., Kennedy, H. A., Johnston, P. J., Rinehart, N. J., Tonge, B. J, Taffe, J. R., & Fitzgerald, P. B. (2014). Emotion recognition of static and dynamic faces in autism spectrum disorder. *Cognition and Emotion, 28*(6), 1110-1118. https://doi.org/10.1080/02699931.2013.867832

Evans J. S. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evers, K., Kerkhof, I., Steyaert, J., Noens, I., & Wagemans, J. (2014). No differences in emotion recognition strategies in children with autism spectrum disorder: Evidence from hybrid faces. *Autism Research and Treatment*, *2014*(2) 1–8. https://doi.org/10.1155/2014/345878

Faja, S., Webb, S. J., Jones, E., Merkle, K., Kamara, D., Bavaro, J., Aylward, E., & Dawson, G. (2012). The effects of face expertise training on the behavioral performance and brain activity of adults with high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(2), 278–293. https://doi.org/10.1007/s10803-011-1243-8

Faras, H., Al Ateeqi, N., & Tidmarsh, L. (2010). Autism spectrum disorders. *Annals of Saudi Medicine*, *30*(4), 295–300. https://doi.org/10.4103/0256-4947.65261

Field, A. (2013) *Discovering Statistics Using IBM SPSS Statistics: And Sex and Drugs and Rock "N" Roll*, (4th ed.). Sage.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist, 34*(10), 906-911. https://doi.org/10.1037/0003-066X.34.10.906

Fletcher-Watson, S., Leekam, S., Benson, V., Frank, M.C., & Findlay, J.M. (2009). Eye-movements reveal attention to social information in autism spectrum disorder. *Neuropsychologia, 47*(1), 248-257. https://10.1016/j.neuropsychologia.2008.07.016

Franco, F., Itakura, S., Pomorska, K., Abramowski, A., Nikaido, K., & Dimitriou, D. (2014). Can children with autism read emotions from the eyes? The eyes test revisited. *Research in Developmental Disabilities*, *35*(5), 1015–1026. https://doi.org/10.1016/j.ridd.2014.01.037

Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology*, *80*(1), 75–85. https://doi.org/10.1037//0022-3514.80.1.75

Gaigg, S.B. (2012). The interplay between emotion and cognition in autism spectrum disorder: Implications for developmental theory. *Frontiers in Integrative Neuroscience, 6*, Article 113. https://doi.org/10.3389/fnint.2012.00113

Gendron, M., Mesquita, B., & Barrett, L. F. (2013). Emotion perception: Putting the face in context. In D. Reisberg (Ed.), *Oxford handbook of cognitive psychology* (pp. 379–389). Oxford University Press.

Gepner, B., Deruelle, C., & Grynfeltt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders, 31*(1), 37–45. https://doi.org/10.1023/a:1005609629218

Golan, O., Baron-Cohen, S., & Golan, Y. (2008). The 'reading the mind in films' task (child version): Complex emotion and mental state recognition in children with and without autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *38*(8), 1534–1541. https://doi.org/10.1007/s10803-007-0533-7

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). The "reading the mind in films" task: Complex emotion recognition in adults with and without autism spectrum conditions. *Social Neuroscience*, *1*(2), 111–123. https://doi.org/10.1080/17470910600980986

Grainger, C., Williams, D. M., & Lind, S. E. (2014). Online action monitoring and memory for self-performed actions in autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*(5), 1–14. https://doi.org/doi:10.1007/s10803-013-1987-4

Grainger, C., Williams, D. M., & Lind, S. E. (2016). Metacognitive monitoring and control processes in children with autism spectrum disorder: Diminished judgement of confidence accuracy. *Consciousness and Cognition*, *42*, 65–74. https://doi.org/10.1016/j.concog.2016.03.003

Greimel, E., Schulte-Rüther, M., Kircher, T., Kamp- Becker, I., Remschmidt, H., Fink, G., Herpertz- Dahlmann, B., & Konrad, K. (2010). Neural mechanisms of empathy in adolescents with autism spectrum disorder and their fathers. *NeuroImage, 49*(1), 1055-1065. https://doi.org/10.1016/j.neuroimage.2009.07.057

Grice, S. J., Halit, H., Farroni, T., Baron-Cohen, S., Bolton, P., & Johnson, M. H. (2005). Neural correlates of eye-gaze detection in young children with autism. *Cortex*, *41*(3), 342–353. https://doi.org/10.1016/s0010-9452(08)70271-5

Gross, T. (2004). The perception of four basic emotions in human and nonhuman faces by children with autism and other developmental disabilities. *Journal of Abnormal Child Psychology, 32*(5), 469-480. https://doi.org/10.1023/B:JACP.0000037777.17698.01

Grossman, J. B., Klin, A., Carter, A. S., & Volkmar, F. R. (2000). Verbal bias in recognition of facial emotions in children with Asperger syndrome. *Journal of Child Psychology and Psychiatry, 41*(3), 369–379.

Hacker, D. J., Bol, L., & Keener, M. C. (2008). *Metacognition in education: A focus on calibration.* In J. Dunlosky, & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429–455). Psychology Press.

Hall, G., & Rodríguez, G. (2019). Attention to perceive, to learn and to respond. *Quarterly Journal of Experimental Psychology, 72*(2), 335–345. https://doi.org/10.1080/17470218.2017.1339719

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Harms, M. B., Martin, A., & Wallace, G. L. (2010). Facial emotion recognition in Autism Spectrum Disorders: A review of behavioral and neuroimaging studies. *Neuropsychology Review*, *20*(3), 290-322. https://doi.org/10.1007/s11065-010-9138-6

Hassin, R. R., Aviezer, H., & Bentin, S. (2013). Inherently ambiguous: Facial expressions of emotions, in context. *Emotion Review, 5*(1)*,* 60–65. http://doi.org/10.1177/1754073912451331

Heavey, L., Phillips, W., Baron-Cohen, S., & Rutter, M. (2000). The awkward moments test: A naturalistic measure of social understanding in autism. *Journal of Autism and Developmental Disorders, 30*(3), 225–236. https://doi.org/10.1023/A:1005544518785

Heerey, E. A., Keltner, D., & Capps, L. M. (2003). Making sense of self-conscious emotion: Linking theory of mind and emotion in children with autism. *Emotion, 3*(4), 394–400. https://doi.org/10.1037/1528-3542.3.4.394

Herba, C., & Phillips, M. (2004). Development of facial expression recognition from childhood to adolescence: Behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry, 45*(7), 1185–1198. https://doi.org/10.1111/j.1469-7610.2004.00316.x

Hill, E., Berthoz, S., & Frith, U. (2004). Brief report: cognitive processing of own emotions in individuals with Autistic Spectrum Disorder and in their relatives. *Journal of Autism and Developmental Disorders*, *34*(2), 229–235. https://doi.org/10.1023/b:jadd.0000022613.41399.14

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. https://doi.org/10.1111/j.2044-8295.1988.tb02745.x

Howard, M. A., Cowell, P. E., Boucher, J., Broks, P., Mayes, A., Farrant, A., & Roberts, N. (2000). Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. *Neuroreport*, *11*(13), 2931–2935. https://doi.org/10.1097/00001756-200009110-00020

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *45*(2), 212–229. https://doi.org/10.1111/j.1469-7610.2004.00215.x

Humphreys, K., Minshew, N., Leonard, G. L., & Behrmann, M. (2007). A fine-grained analysis of facial expression processing in high-functioning adults with autism. *Neuropsychologia, 45*(4), 685–695. https://doi.org/10.1016/j.neuropsychologia.2006.08.003

Johnston, P. J., Enticott, P. G., Mayes, A. K., Hoy, K. E., Herring, S. E., & Fitzgerald, P. B. (2010). Symptom correlates of static and dynamic facial affect processing in schizophrenia: Evidence of a double dissociation? *Schizophrenia Bulletin*, *36*(4), 680–687. https://doi.org/10.1093/schbul/sbn136

Jones, C. R., Pickles, A., Falcaro, M., Marsden, A. J., Happé, F., Scott, S. K., Sauter, D., Tregay, J., Phillips, R. J., Baird, G., Simonoff, E., & Charman, T. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *Journal of Child Psychology and Psychiatry, 52*(3), 275–285. https://doi.org/10.1111/j.1469-7610.2010.02328.x

Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *44*(4), 529–542. https://doi.org/10.1111/1469-7610.00142

Juslin, P., Olsson, N., & Winman, A. (1996). Calibration and diagnosticity of confidence in eyewitness identification: Comments on what can be inferred from the low confidence–accuracy correlation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1304–1316. https://doi.org/10.1037/0278-7393.22.5.1304

Juslin, P., Winman, A., & Olsson, H. (2000). Naive empiricism and dogmatism in confidence research: A critical examination of the hard–easy effect. *Psychological Review, 107*(2), 384–396. https://doi.org/10.1037/0033-295X.107.2.384

Kaland, N., Mortensen, E. L., & Smith, L. (2011). Social communication impairments in children and adolescents with Asperger syndrome: Slow response time and the impact of prompting. *Research in Autism Spectrum Disorders, 5*(3), 1129–1137. https://doi.org/10.1016/j.rasd.2010.12.009

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2,* 217–250.

Keating, C. T., Fraser, D. S., Sowden, S., & Cook, J. L. (2021, online ahead of print). Differences between autistic and non-autistic adults in the recognition of anger from facial motion remain after controlling for alexithymia. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-021-05083-9

Keehn, B., Müller, R. A., & Townsend, J. (2013). Atypical attentional networks and the emergence of autism. *Neuroscience and Biobehavioral Reviews*, *37*(2), 164–183. https://doi.org/10.1016/j.neubiorev.2012.11.014

Kelly, K. J., & Metcalfe, J. (2011). Metacognition of emotional face recognition. *Emotion, 11*(4), 896-906. https://doi.org/10.1037/a0023746

Kern Koegel, L., Ashbaugh, K., Navab, A., & Koegel, R. L. (2016). Improving empathic communication skills in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 921–933. https://doi.org/10.1007/s10803-015-2633-0

Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. (2010). Atypical reflexive gaze patterns on emotional faces in autism spectrum disorders. *The Journal of Neuroscience, 30*(37), 12281-12287. https://doi.org/10.1523/JNEUROSCI.0688-10.2010

Kliemann, D., Rosenblau, G., Bölte, S., Heekeren, H. R., & Dziobek, I. (2013). Face puzzle-two new video-based tasks for measuring explicit and implicit aspects of facial emotion recognition. *Frontiers in Psychology, 4*, Article 376. https://doi.org/10.3389/fpsyg.2013.00376

Klin A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders*, *21*(1), 29–42. https://doi.org/10.1007/BF02206995

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), 809–816. https://doi.org/10.1001/archpsyc.59.9.809

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, *459*(7244), 257–261. https://doi.org/10.1038/nature07868

Kuusikko, S., Haapsamo, H., Jansson-Verkasalo, E., Hurtig, T., Mattila, M., Ebeling, H., Jussila, K., Bölte, S., & Moilanen, I. (2009). Emotion recognition in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(6), 938-45. http://doi.org/10.1007/s10803-009-0700-0

Lacava, P. G., Golan, O., Baron-Cohen, S., & Smith Myles, B. (2007). Using assistive technology to teach emotion recognition to students with Asperger syndrome: A pilot study. *Remedial and Special Education, 28*(3), 174–181. https://doi.org/10.1177/07419325070280030601

Lacroix, A., Guidetti, M., Rogé, B., & Reilly, J. (2009). Recognition of emotional and nonemotional facial expressions: A comparison between Williams syndrome and autism. *Research in Developmental Disabilities, 30*(5), 976-985. https://doi.org/10.1016/j.ridd. 2009.02.002

Law Smith, M., Montagne, B., Perrett, D., Gill, M., & Gallagher, L. (2010). Detecting subtle facial emotion recognition deficits in high-functioning autism using dynamic stimuli of varying intensities. *Neuropsychologia, 48*(9), 2777-2781. https://doi.org/10.1016/j.neu- ropsychologia.2010.03.008

Locke, J., Kasari, C., Rotheram-Fuller, E., Kretzmann, M., & Jacobs, J. (2013). Social network changes over the school year among elementary school-aged children with and without an autism spectrum disorder. *School Mental Health: A Multidisciplinary Research and Practice Journal, 5*(1), 38–47. https://doi.org/10.1007/s12310-012-9092-y

Loveland, K. A., Bachevalier, J., Pearson, D. A., & Lane, D. M. (2008). Fronto-limbic functioning in children and adolescents with and without autism. *Neuropsychologia*, *46*(1), 49–62. https://doi.org/10.1016/j.neuropsychologia.2007.08.017

Loveland, K., Tunali-Kotoski, B., Chen, Y., Ortegon, J., Pearson, D., Brelsford, K., & Gibbs, M. (1997). Emotion recognition in autism: Verbal and non-verbal information. *Development and Psychopathology, 9*(3), 579-593. https://doi.org/10.1017/S095457949 7001351

Lozier, L. M., Vanmeter, J. W., & Marsh, A. A. (2014). Impairments in facial affect recognition associated with autism spectrum disorders: A meta-analysis. *Development and Psychopathology, 26*(4), 933–945. https://doi.org/10.1017/S0954579414000479

Luke, L., Clare, I. C., Ring, H., Redley, M., & Watson, P. (2012). Decision-making difficulties experienced by adults with autism spectrum conditions. *Autism*, *16*(6), 612–621. https://doi.org/10.1177/1362361311415876

Maekawa, T., Tobimatsu, S., Inada, N., Oribe, N., Onitsuka, T., Kanba, S., & Kamio, Y. (2011). Top-down and bottom-up visual information processing of non-social stimuli in high-functioning autism spectrum disorder. *Research in Autism Spectrum Disorders. 5*(1), 201-209. https://doi.org/10.1016/j.rasd.2010.03.012

Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology, 97*(4), 723–731. https://doi.org/10.1037/0022-0663.97.4.723

Maras, K., Norris, J. E., & Brewer, N. (2020). Metacognitive monitoring and control of eyewitness memory reports in autism. *Autism Research, 13*(11), 2017-2029. https://doi.org/10.1002/aur.2278

Marinetti C., Moore P., Lucas P., & Parkinson B. (2011). Emotions in social interactions: Unfolding emotional experience. In R. Cowie, C. Pelachaud, & P. Petta (Eds), *Emotion-oriented systems. Cognitive technologies* (pp. 31-46). Springer. https://doi.org/10.1007/978-3-642-15184-2_3

Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology, 96*(1), 1–10. https://doi.org/10.1037/a0014037

McCrimmon, A. W., & Smith, A. D. (2013). Review of the Wechsler abbreviated scale of intelligence, second edition (WASI-II). *Journal of Psychoeducational Assessment*, *31*(3), 337-341. https://doi.org/10.1177/0734282912467756

McHugh M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282.

Mul, C. L., Stagg, S. D., Herbelin, B., & Aspell, J. E. (2018). The feeling of me feeling for you: Interoception, alexithymia and empathy in autism. *Journal of Autism and Developmental Disorders*, *48*(9), 2953–2967. https://doi.org/10.1007/s10803-018-3564-3

Nah, Y. H., Brewer, N., Young, R. L., & Flower, R. (2018). Brief Report: Screening Adults with Autism Spectrum Disorder for Anxiety and Depression. *Journal of Autism and Developmental Disorders*, *48*(5), 1841–1846. https://doi.org/10.1007/s10803-017-3427-3

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *The Psychology of Learning and Motivation, 26*, 125-174. https://doi.org/10.1016/S0079-7421(08)60053-5

Nuske, H., Vivanti, G., & Dissanayake, C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. *Cognition & Emotion, 27*(6), 1042–1061. https://doi.org/10.1080/02699931.2012.762900

Oakley, B., Brewer, R., Bird, G., & Catmur, C. (2016). Theory of mind is not theory of emotion: A cautionary note on the Reading the Mind in the Eyes Test. *Journal of abnormal psychology*, *125*(6), 818–823. https://doi.org/10.1037/abn0000182

O'Neill, M., & Jones, R. S. (1997). Sensory-perceptual abnormalities in autism: A case for more research? *Journal of Autism and Developmental Disorders*, *27*(3), 283–293. https://doi.org/10.1023/a:1025850431170

O'Reilly, H., Pigat, D., Fridenson, S., Berggren, S., Tal, S., Golan, O., Bölte, S., Baron-Cohen, S., & Lundqvist, D. (2016). The EU-Emotion Stimulus Set: A validation study. *Behavior Research Methods*, *48*(2), 567–576. https://doi.org/10.3758/s13428-015-0601-4

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive function deficits in high-functioning autistic individuals: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32*(7), 1081-1105. https://doi.org/10.1111/j.14697610.1991.tb00351.

Palmer, M. A., Brewer, N., Weber, N., & Nagesh, A. (2013). The confidence-accuracy relationship for eyewitness identification decisions: Effects of exposure duration, retention interval, and divided attention. *Journal of Experimental Psychology: Applied*, *19*(1), 55–71. https://doi.org/10.1037/a0031602

Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': A bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, *16*(10), 504–510. https://doi.org/10.1016/j.tics.2012.08.009

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological Medicine*, *40*(11), 1919-1929. https://doi.org/10.1017/S0033291709992364

Riby, D. M., & Hancock, P. J. (2009). Do faces capture the attention of individuals with Williams syndrome or autism? Evidence from tracking eye movements. *Journal of Autism and Developmental Disorders*, *39*(3), 421–431. https://doi.org/10.1007/s10803-008-0641-z

Righart, R. G. R., & De Gelder, B. (2008). Recognition of facial expressions is influenced by emotional scene gist. *Cognitive, Affective and Behavioral Neuroscience*, *8*(3), 264-272. https://doi.org/10.3758/cabn.8.3.264

Roeyers, H., Buysse, A., Ponnet, K., & Pichal, B. (2001). Advancing advanced mind-reading tests: Empathic accuracy in adults with a pervasive developmental disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 271–278.

Rogers, K., Dziobek, I., Hassenstab, J., Wolf, O. T., & Convit, A. (2007). Who cares? Revisiting empathy in Asperger syndrome. *Journal of Autism and Developmental Disorders*, *37*(4), 709–715. https://doi.org/10.1007/s10803-006-0197-8

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. https://doi.org/10.1007/s10803-007-0465-2

Rump, K. M., Giovannelli, J. L., Minshew, N. J., & Strauss, M. S. (2009). The development of emotion recognition in individuals with autism. *Child Development*, *80*(5), 1434-1447. https://doi.org/10.1111/j.1467-8624.2009.01343.x

Rutherford, M. D., & Troje, N. F. (2012). IQ predicts biological motion perception in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(4), 557–565. https://doi.org/10.1007/s10803-011-1267-0

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of Autism and Developmental Disorders*, *32*(3), 189–194. https://doi.org/10.1023/a:1015497629971

Rutherford, M., & McIntosh, D. (2007). Rules versus prototype matching: Strategies of perception of emotional facial expressions in the autism spectrum. *Journal of Autism and Developmental Disorders, 37*(2), 187-196. https://doi.org/10.1007/s10803-006-0151-9

Rutherford, M., & Towns, A. (2008). Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*(7), 1371-1381. https://doi.org/10.1007/s10803- 007-0525-7

Sasson, N., Tsuchiya, N., Hurley, R., Couture, S. M., Penn, D. L., Adolphs, R., & Piven, J. (2007). Orienting to social stimuli differentiates social cognitive impairment in autism and schizophrenia. *Neuropsychologia, 45*(11), 2580–2588. https://doi.org/10.1016/j.neuropsychologia.2007.03.009

Sawyer, A. C., Williamson, P., & Young, R. (2014). Metacognitive processes in emotion recognition: are they different in adults with Asperger's disorder? *Journal of Autism and Developmental Disorders*, *44*(6), 1373–1382. https://doi.org/10.1007/s10803-013-1999-0

Senju, A. (2012). Spontaneous theory of mind and its absence in autism spectrum disorders. *Neuroscientist*, *18*(2), 108–113. https://doi.org/10.1177%2F1073858410397208

Shamay-Tsoory, S. G., Tomer, R., Yaniv, S., & Aharon-Peretz, J. (2002). Empathy deficits in Asperger syndrome: A cognitive profile. *Neurocase*, *8*(3), 245–252. https://doi.org/10.1093/neucas/8.3.245

Skelly, L. R., & Decety, J. (2012). Passive and motivated perception of emotional faces: Qualitative and quantitative changes in the face processing network. *PloS One*, *7*(6), e40371. https://doi.org/10.1371/journal.pone.0040371

Skelly, L. R., & Decety, J. (2012). Passive and motivated perception of emotional faces: Qualitative and quantitative changes in the face processing network. *PloS One*, *7*(6), e40371. https://doi.org/10.1371/journal.pone.0040371

Song, Y., & Hakoda, Y. (2018). Selective impairment of basic emotion recognition in people with autism: Discrimination thresholds for recognition of facial expressions of varying intensities. *Journal of Autism and Developmental Disorders*, *48*(6), 1886–1894. https://doi.org/10.1007/s10803-017-3428-2

Speer, L. L., Cook, A. E., McMahon, W. M., & Clark, E. (2007). Face processing in children with autism: Effects of stimulus contents and type. *Autism*, *11*(3), 265–277. https://doi.org/10.1177/1362361307076925

Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. A. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia, 51*(1), 98–105. https://doi.org/10.1016/j.neuropsychologia.2012.11.013

Tang, J., Chen, N., Falkmer, M., Bölte, S., & Girdler, S. (2019). Atypical visual processing but comparable levels of emotion recognition in adults with autism during the processing of social scenes. *Journal of Autism and Developmental Disorders*, *49*(10), 4009–4018. https://doi.org/10.1007/s10803-019-04104-y

Tardif, C., Lainé, F., Rodriguez, M., & Gepner, B. (2007). Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial–vocal imitation in children with autism. *Journal of Autism and Developmental Disorders, 37*(8), 1469–1484. https://doi.org/10.1007/s10803-006-0223-x.

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton University Press. https://doi.org/10.1515/9781400885725

Trevisan, D.A., & Birmingham, E. (2016). Are emotion recognition abilities related to everyday social functioning in ASD? A meta-analysis. *Research in Autism Spectrum Disorders, 32*, 24-42. https://doi.org/10.1016/J.RASD.2016.08.004

Uljarevic, M., & Hamilton, A. (2013). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, *43*(7), 1517–1526. https://doi.org/10.1007/s10803-012-1695-5

Uljarević, M., Phillips, J. M., Schuck, R. K., Schapp, S., Solomon, E. M., Salzman, E., Allerhand, L., Libove, R. A., Frazier, T. W., & Hardan, A. Y. (2020). Exploring social subtypes in Autism Spectrum Disorder: A preliminary study. *Autism Research*, *13*(8), 1335–1342. https://doi: 10.1002/aur.2294

Wallace, G. L., Case, L. K., Harms, M. B., Silvers, J. A., Kenworthy, L., & Martin, A. (2011). Diminished sensitivity to sad facial expressions in high functioning autism spectrum disorders is associated with symptomatology and adaptive functioning. *Journal of Autism and Developmental Disorders*, *41*(11), 1475–1486. https://doi.org/10.1007/s10803-010-1170-0

Wason P. C., Evans J.St B. T. (1975). Dual processes in reasoning? *Cognition, 3*(2), 141–154. https://doi.org/10.1016/0010-0277(74)90017-1

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI).* Harcourt Assessment.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children—4th Edition (WISC-IV)*. Harcourt Assessment.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II)*. Pearson.

Wechsler, D., Coalson, D. L., & Raiford, S. E. (2008). *WAIS-IV: Wechsler Adult Intelligence Scale*. Pearson.

Whitaker, L. R., Simpson, A., & Roberson, D. (2017). Brief report: Is impaired classification of subtle facial expressions in children with autism spectrum disorders related to atypical emotion category boundaries? *Journal of Autism and Developmental Disorders*, *47*(8), 2628–2634. https://doi.org/10.1007/s10803-017-3174-5

White, S. J., Coniston, D., Rogers, R., & Frith, U. (2011). Developing the Frith- Happé animations: A quick and objective test of theory of mind for adults with autism. *Autism Research*, *4*(2), 149-154. https://doi.org/10.1002/aur.174

Wilkinson, D. A., Best, C. A., Minshew, N. J., & Strauss, M. S. (2010). Memory awareness for faces in individuals with autism. *Journal of Autism and Developmental Disorders*, *40*(11), 1371–1377. https://doi.org/10.1007/s10803-010-0995-x

Williams, D. M., Bergström, Z., & Grainger, C. (2018). Meta-cognitive monitoring and the hypercorrection effect in autism and the general population: Relation to autism(-like) traits and mindreading. *Autism, 22*(3), 259–270. https://doi.org/10.1177/1362361316680178

Wilson, C. E., Palermo, R., & Brock, J. (2012). Visual scan paths and recognition of facial identity in autism spectrum disorder and typical development. *Plos One, 7*, 1-9. https://doi.org/10.1371/journal.pone.0037681

Wojcik, D. Z., Allen, R. J., Brown, C., & Souchay, C. (2011). Memory for actions in autism spectrum disorder. *Memory*, *19*(6), 549–558. https://doi.org/10.1080/09658211.2011.590506

Wojcik, D. Z., Moulin, C. J., & Souchay, C. (2013). Metamemory in children with autism: Exploring "feeling-of-knowing" in episodic and semantic memory. *Neuropsychology*, *27*(1), 19–27. https://doi.org/10.1037/a0030526

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal aging. *Neuropsychology, 19*(6), 739–749. https://doi.org/10.1037/0894-4105.19.6.739

Wong, N., Beidel, D.C., Sarver, D.E., & Sims, V. (2012). Facial emotion recognition in children with high functioning autism and children with social phobia. *Child Psychiatry & Human Development, 43*(5), 775–794. https://doi.org/10.1007/s10578-012-0296-z.

Wright, B., Clarke, N., Jordan, J. O., Young, A. W., Clarke, P., Miles, J., Nation, K., Clarke, L., & Williams, C. (2008). Emotion recognition in faces and the use of visual context in young people with high-functioning autism spectrum disorders. *Autism, 12*(6), 607–626. https://doi.org/10.1177/1362361308097118

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Westview Press.

Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science*, *26*(3), 237–242. https://doi.org/10.1177/0963721416683996

Zhang, Q., Chen, L., & Yang, Q. (2018). The effects of facial features on judicial decision making. *Advances in Psychological Science*, *26*(4), 698–709. https://doi.org/10.3724/SP.J.1042.2018.00698

Appendix A

Projected Confidence in Emotion Recognition for Each Stimulus Type

Table A1

Means and Standard Deviations for Projected Confidence Ratings

Stimulus Type	Group	*M (SD)*
Static	ASD	65.08 (21.47)
	Typically Developing	74.78 (18.86)
	Overall	70.08 (20.67)
Dynamic	ASD	69.84 (19.88)
	Typically Developing	76.87 (16.63)
	Overall	73.46 (18.54)
Social	ASD	70.48 (21.13)
	Typically Developing	75.37 (17.87)
	Overall	73.00 (19.59)

Appendix B

Pool of Emotions Used for Multiple-Choice Format

Target Emotions/States of Mind

Afraid	Happy
Angry	Hurt
Ashamed	Jealous
Disappointed	Sad
Disgusted	Surprised
Frustrated	Worried

Filler Emotions/States of Mind

Appreciated	Doubtful	Hopeful	Optimistic
Amazed	Dreading	Heartbroken	Pained
Awful	Desperate	Helpless	Passionate
Belligerent	Dismay	Hostile	Panicked
Bored	Eager	Inspired	Proud
Brave	Excited	Interested	Relieved
Bitter	Exhausted	Joking	Sympathetic
Certain	Exhilarated	Kind	Satisfied
Complacent	Flattered	Lonely	Sneaky
Confident	Grateful	Liberated	Suspicious
Dejected	Hatred	Nervous	Uncertain
Despair	Honoured	Offended	Weary

Appendix C

Missing Data Examinations

Table C1

Free-Report Format (Recognition and Reaction) Missing Responses Analysis

Stimulus type		Number of participants with missing data	Number of missing data points across emotions	Number of responses missing overall (percent missing for subcategory)
Static	ASD	4	Afraid=2 Sad=1 Jealous=2	5 (0.17%)
	TD	4	Angry=3 Afraid=2 Worried=1 Frustrated= 1	7 (0.22%)
Dynamic	ASD	2	Sad=1 Hurt=1	2 (0.07%)
	TD	4	Happy=2 Frustrated=2 Disgusted=2 Disappointed=2	8 (0.25%)
Social (emotion recognition)	ASD	3	Sad= 3 Happy=3 Disappointed=3 Afraid=4	13 (0.28%)
	TD	0		0 (0%)
Social (reactions to emotions)	ASD	3	Worried=2 Happy=1	3 (.08%)
	TD	0		0 (0%)
				Total responses missing across all tasks= 38 (.12%) of 31,720 responses

Note. 'TD' is typically developing

Table C2

Multiple-Choice Format (Recognition) Missing Responses Analysis

Stimulus Type		Number of participants with missing data (percent of all participants)	Number of missing data points across emotions		Number of responses missing overall (percent of all responses for each section missing)
Static	ASD	15 (23.81%)	Jealous=10 Ashamed=11 Afraid=7 Frustrated=3 Disappointed=3 Worried=2	Hurt=2 Angry=3 Sad=4 Happy=1 Disgust=1	47 (1.55%)
	TD	13 (19.40%)	Jealous=12 Ashamed=3 Afraid=6 Frustrated=7 Disappointed=7 Worried=5	Hurt=8 Angry=4 Sad=3 Happy=2 Disgust=3 Surprised=1	61 (1.90%)
Dynamic	ASD	11 (17.46%)	Jealous=9 Disappointed=1 Angry=2 Hurt=2 Ashamed=1	Afraid=2 Sad=1	18 (0.60%)
	TD	7 (10.45%)	Jealous=9 Disappointed=3 Angry=3 Hurt=6 Ashamed=2	Afraid=5 Frustrated=2 Disgusted=3	33 (1.03%)
Social	ASD	12 (19.05%)	Disappointed=6 Sad=7 Jealous=4 Disgusted=4 Ashamed=8 Surprised=1	Happy=4 Angry=1 Afraid=5 Worried=1 Frustrated=2 Hurt=2	45 (0.97%)
	TD	16 (23.88%)	Disappointed=4 Sad=5 Jealous=4 Disgusted=2 Ashamed=3 Surprised=2	Angry=3 Afraid=3 Worried=1	27 (0.54%)
				Total responses missing across all tasks= 231 (1.05%) of 22,100	

Note. ‘TD’ is typically developing

Appendix D

Latency Data Outliers

Table D1

Number of Individual Mean Latencies Adjusted to 2.5 Standard Deviations from Mean

Emotion	Multiple-Choice			Free-Report		
	Static	Dynamic	Social	Static	Dynamic	Social
Emotion Recognition						
Afraid*	3	2	2	6	4	3
Angry*	3	6	6	4	6	6
Disgusted*	6	6	3	4	5	3
Happy*	3	3	4	5	3	3
Sad*	3	2	3	3	3	5
Surprised*	4	3	4	3	3	4
Ashamed	4	3	4	2	5	3
Disappointed	5	4	1	7	5	6
Frustrated	5	4	4	5	2	4
Hurt	3	4	3	1	2	2
Jealous	3	3	4	5	2	2
Worried	3	3	5	3	2	3
Emotion Reaction						
Afraid*	-	-	-	-	-	1
Angry*	-	-	-	-	-	1
Disgusted*	-	-	-	-	-	6
Happy*	-	-	-	-	-	5
Sad*	-	-	-	-	-	4
Surprised*	-	-	-	-	-	3
Ashamed	-	-	-	-	-	5
Disappointed	-	-	-	-	-	5
Frustrated	-	-	-	-	-	7
Hurt	-	-	-	-	-	4
Jealous	-	-	-	-	-	1
Worried	-	-	-	-	-	3

Note. * is a basic emotion

Appendix E

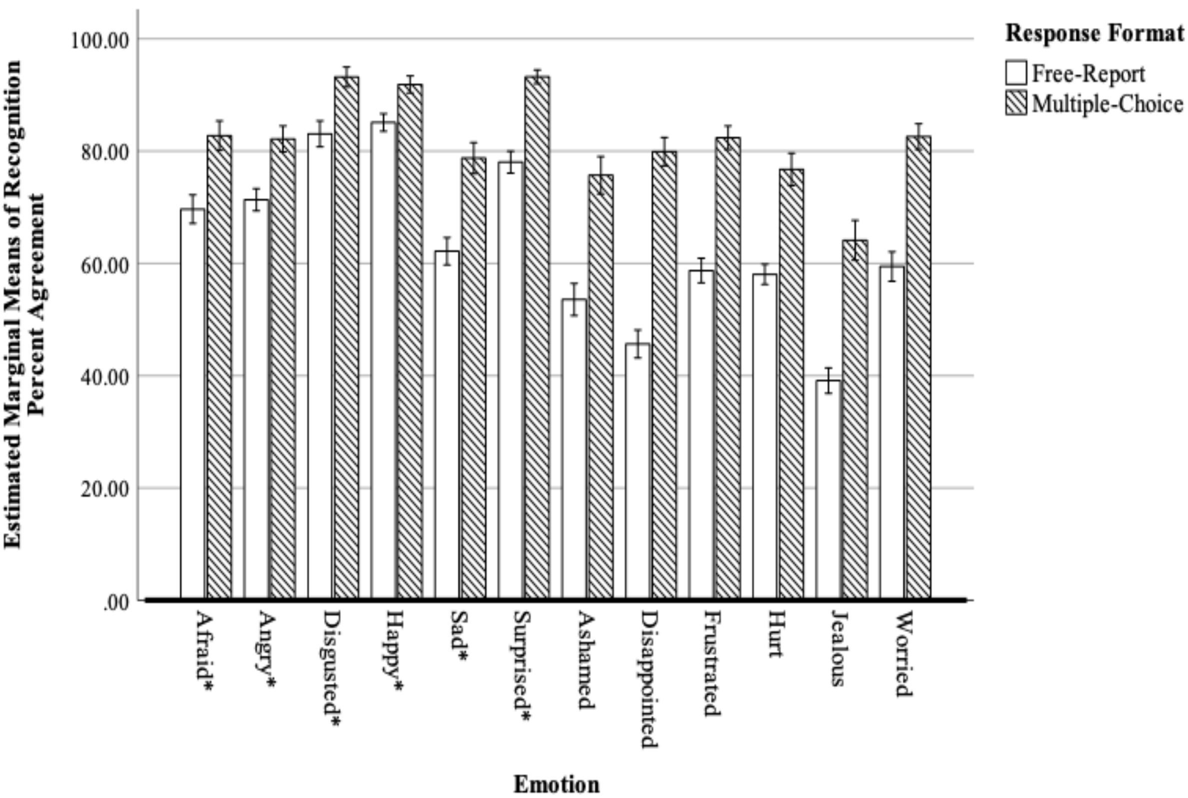
Emotion Recognition Percent Agreement Interaction Effects Not Involving Group

Response Format × Emotion

There was a significant two-way interaction between response format and emotion. Although Figure E1 shows higher percent agreement for multiple-choice than free-report (i.e., the difference in performance between the two response formats) across all emotions (indicated by the main effect for response format), this pattern is generally more marked for complex than basic emotions. Confirming this observation, inspection of the 95% CIs for the mean difference between percent agreement for each emotion shown in Table E1 reveals that the only overlap between the basic and complex emotions' difference scores involves the basic emotion, sad, and the complex emotion, hurt.

Figure E1

Two-Way Interaction of Response Format and Emotion on Recognition Percent Agreement



Note. Error bars represent 95% confidence interval, * is a basic emotion

Table E1

Pairwise Comparison Outcomes Examining the Response Format × Emotion Interaction on Recognition Percent Agreement

Response Format Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
MC-FR	Afraid*	13.32	1.15	<.001	11.05	15.59
	Angry*	10.75	1.00	<.001	8.76	12.73
	Disgusted*	10.13	.83	<.001	8.48	11.77
	Happy*	6.78	.90	<.001	5.00	8.56
	Sad*	16.56	1.27	<.001	14.04	19.08
	Surprised*	15.21	.77	<.001	13.69	16.74
	Ashamed	22.05	1.29	<.001	19.50	24.60
	Disappointed	34.13	1.18	<.001	31.79	36.47
	Frustrated	23.52	1.10	<.001	21.33	25.70
	Hurt	18.52	1.42	<.001	15.72	21.33
	Jealous	24.74	1.34	<.001	22.08	27.40
	Worried	23.37	1.19	<.001	21.01	25.73

Note. 'MC' is multiple-choice, 'FR' is free-report, * is a basic emotion, ** is based on estimated marginal means

Appendix F

Emotion Recognition Percent Agreement Interaction Effects Not Involving Group

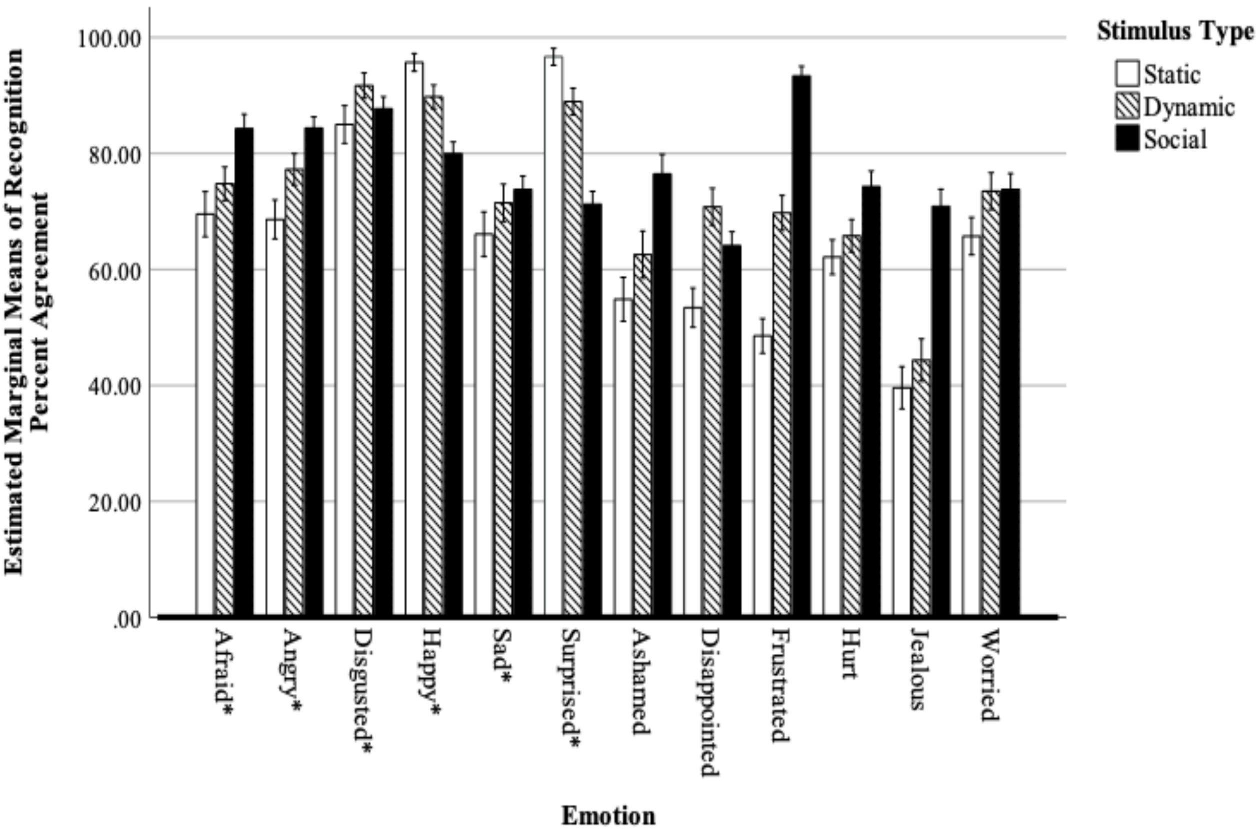
Stimulus Type × Emotion

As Figure F1 shows, the percent agreement for the static stimulus type was generally higher for basic emotions than complex emotions. Similarly, the percent agreement for dynamic stimulus type was generally higher for basic emotions compared to complex emotions; however, this trend did not appear to be as strong as that for the social stimulus type. Figure F1 also shows that the percent agreement for social stimulus type did not follow the same trend observed for other stimulus types. For three of the complex emotions (ashamed, frustrated, jealous), percent agreement for the social stimulus type appears markedly higher than the other stimulus types, whereas the superiority of the social stimulus type is only evident for two basic emotions (afraid, angry). Moreover, an opposite pattern is evident for two basic emotions (happy, surprised).

Inspection of the 95% CIs for the mean difference between percent agreement for each emotion (see Table F1) when making static-social stimulus and dynamic-social stimulus comparisons suggests patterns that are generally consistent with the above observations.

Figure F1

Two-Way Interaction of Stimulus Type and Emotion on Recognition Percent Agreement



Note. Error bars represent 95% confidence interval, * is a basic emotion

Table F1

Pairwise Comparison Outcomes Examining the Stimulus Type × Emotion Interaction on Recognition Percent Agreement

Stimulus Type Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Static-Dynamic	Afraid*	-5.62	1.97	.015	-10.41	-.84
	Angry*	-8.62	1.84	<.001	-13.10	-4.15
	Disgusted*	-6.59	1.58	<.001	-10.43	-2.74
	Happy*	5.62	1.17	<.001	2.79	8.46
	Sad*	-5.48	2.02	.023	-10.37	-.58
	Surprised*	7.71	1.19	<.001	4.84	10.59
	Ashamed	-7.50	2.19	.002	-12.81	-2.20
	Disappointed	-17.38	1.87	<.001	-21.92	-12.83
	Frustrated	-21.47	1.79	<.001	-25.81	-17.13
	Hurt	-3.74	1.76	.107	-8.01	.53
	Jealous	-4.51	1.85	.049	-9.00	-.01
	Worried	-7.94	1.86	<.001	-12.46	-3.43
Static-Social	Afraid*	-15.22	2.00	<.001	-20.08	-10.36
	Angry*	-15.90	1.84	<.001	-20.36	-11.43
	Disgusted*	-2.85	1.72	.301	-7.03	1.33
	Happy*	15.43	1.23	<.001	12.46	18.40
	Sad*	-7.87	2.10	.001	-12.97	-2.77
	Surprised*	25.19	1.12	<.001	22.46	27.92
	Ashamed	-21.39	2.22	<.001	-26.78	-16.00
	Disappointed	-11.00	1.89	<.001	-15.59	-6.42
	Frustrated	-45.20	1.70	<.001	-49.33	-41.06
	Hurt	-12.51	1.85	<.001	-16.99	-8.03
	Jealous	-31.01	1.93	<.001	-35.68	-26.33
	Worried	-8.49	1.94	<.001	-13.19	-3.79
Dynamic-Social	Afraid*	-9.59	1.57	<.001	-13.40	-5.79
	Angry*	-7.28	1.54	<.001	-11.02	-3.53
	Disgusted*	3.74	1.20	.007	.82	6.65
	Happy*	9.81	1.28	<.001	6.71	12.91

Table F1 Continued

Stimulus Type Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
	Sad*	-2.39	1.71	.494	-6.55	1.76
	Surprised*	17.48	1.41	<.001	14.05	20.91
	Ashamed	-13.88	1.99	<.001	-18.72	-9.05
	Disappointed	6.38	1.58	<.001	2.55	10.20
	Frustrated	-23.73	1.43	<.001	-27.20	-20.25
	Hurt	-8.78	1.80	<.001	-13.13	-4.42
	Jealous	-26.50	1.86	<.001	-31.01	-21.99
	Worried	-.55	1.85	1.00	-5.05	3.95

Note. * is a basic emotion, ** is based on estimated marginal means

Appendix G

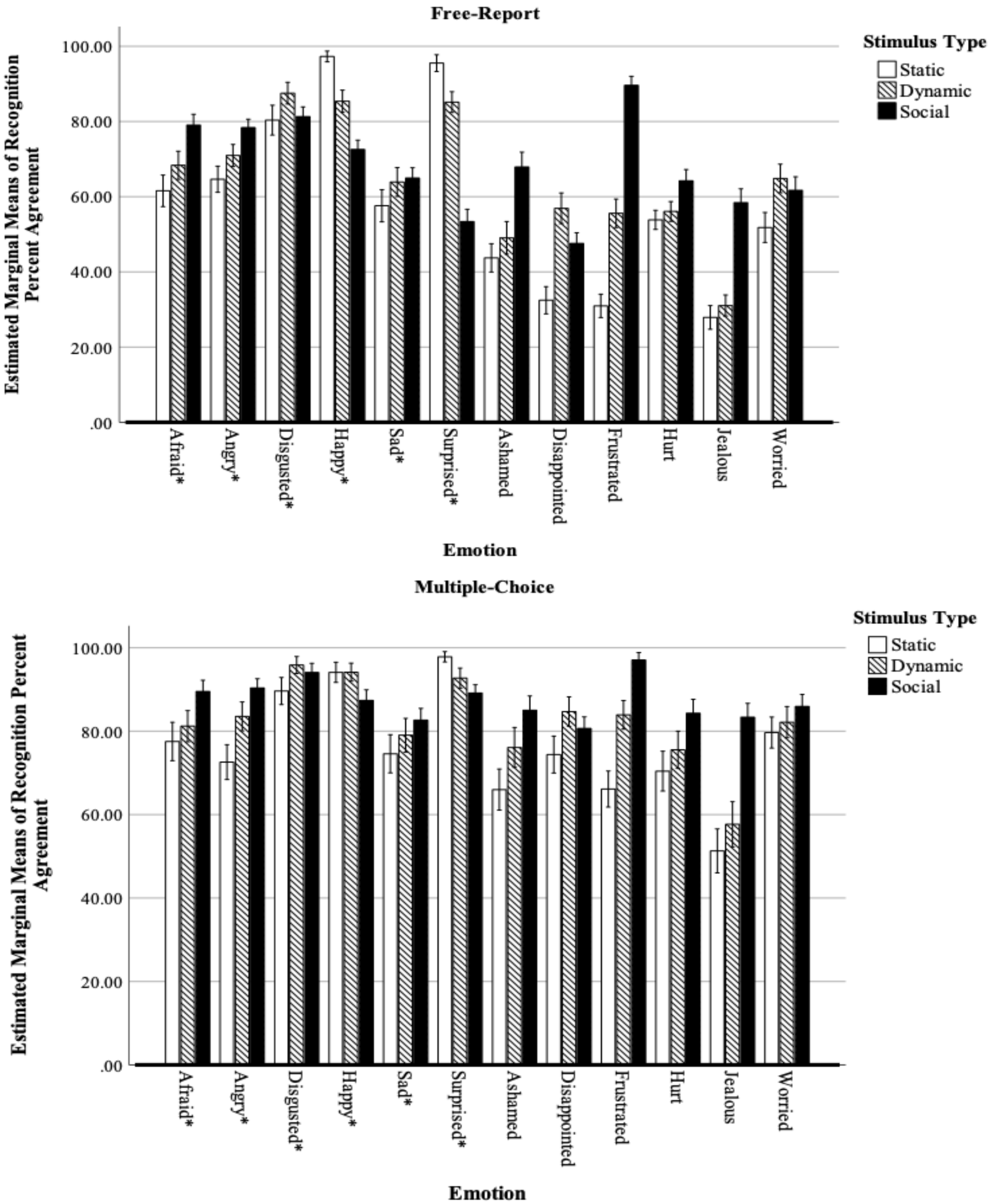
Emotion Recognition Percent Agreement Interaction Effects Not Involving Group

Response Format × Stimulus Type × Emotion

The significant three-way interaction between response format, stimulus type and emotion is extremely difficult to interpret, especially given that the emotion variable has 12 levels. Visual inspection of Figure G1 suggests one possible, but tentatively advanced, account of the interaction. The patterns in Figure G1 for the three stimulus types show higher percent agreement for the social task, followed by the dynamic and then static for both response formats, for 6 of the 12 emotions (afraid, angry, ashamed, frustrated, hurt, jealous). Happy, surprised, disappointed and worried appear to not follow this trend across both response formats.

Figure G1

Three-way Interaction of Emotion, Stimulus Type and Response Format on Recognition Percent Agreement



Note. Error bars represent 95% confidence interval, * is a basic emotion

Appendix H

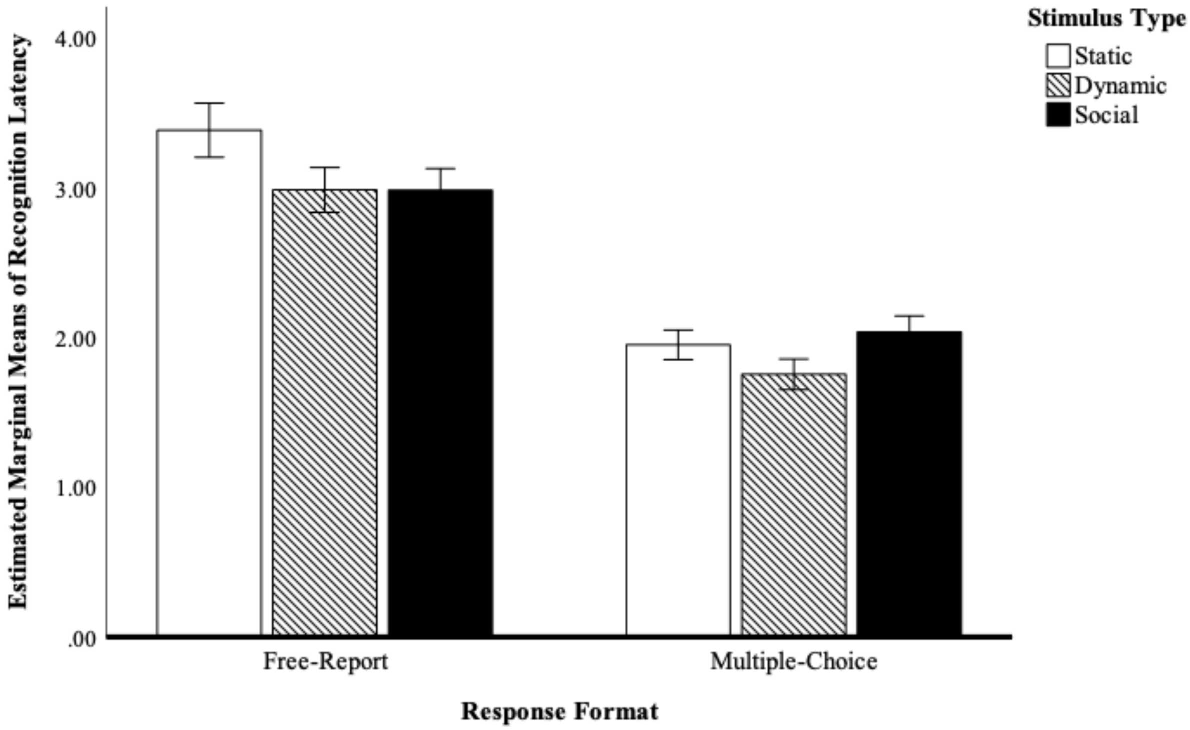
Emotion Recognition Latency Interaction Effects Not Involving Group

Response Format × Stimulus Type

As shown in Figure H1, and confirmed by inspection of the 95% CIs for the mean difference between mean latency for the three stimulus types shown in Table H1, there were longer mean latencies for free-report than multiple-choice across all stimulus types. The trends seen in Figure H1 varied across response formats, with multiple-choice showing shorter latencies for dynamic compared with static and social whereas the free-report response type shows longer latencies for static stimuli than for dynamic and social stimuli.

Figure H1

Two- Way Interaction of Response Format and Stimulus Type on Recognition Latency



Note. Error bars represent 95% confidence interval

Table H1

Pairwise Comparison Outcomes Examining the Stimulus Type × Response Format Interaction on Recognition Latency

Response Format	Stimulus Type	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
FR	Static-Dynamic	.40	.09	<.001	.17	.63
	Static-Social	.40	.08	<.001	.20	.60
	Dynamic-Social	.00	.07	1.00	-.18	.18
MC	Static-Dynamic	.20	.05	<.001	.07	.32
	Static-Social	-.09	.06	.370	-.22	.05
	Dynamic-Social	-.28	.05	<.001	-.41	-.16

Note. 'FR' is free report, 'MC' is multiple-choice, **based on estimated marginal means

Appendix I

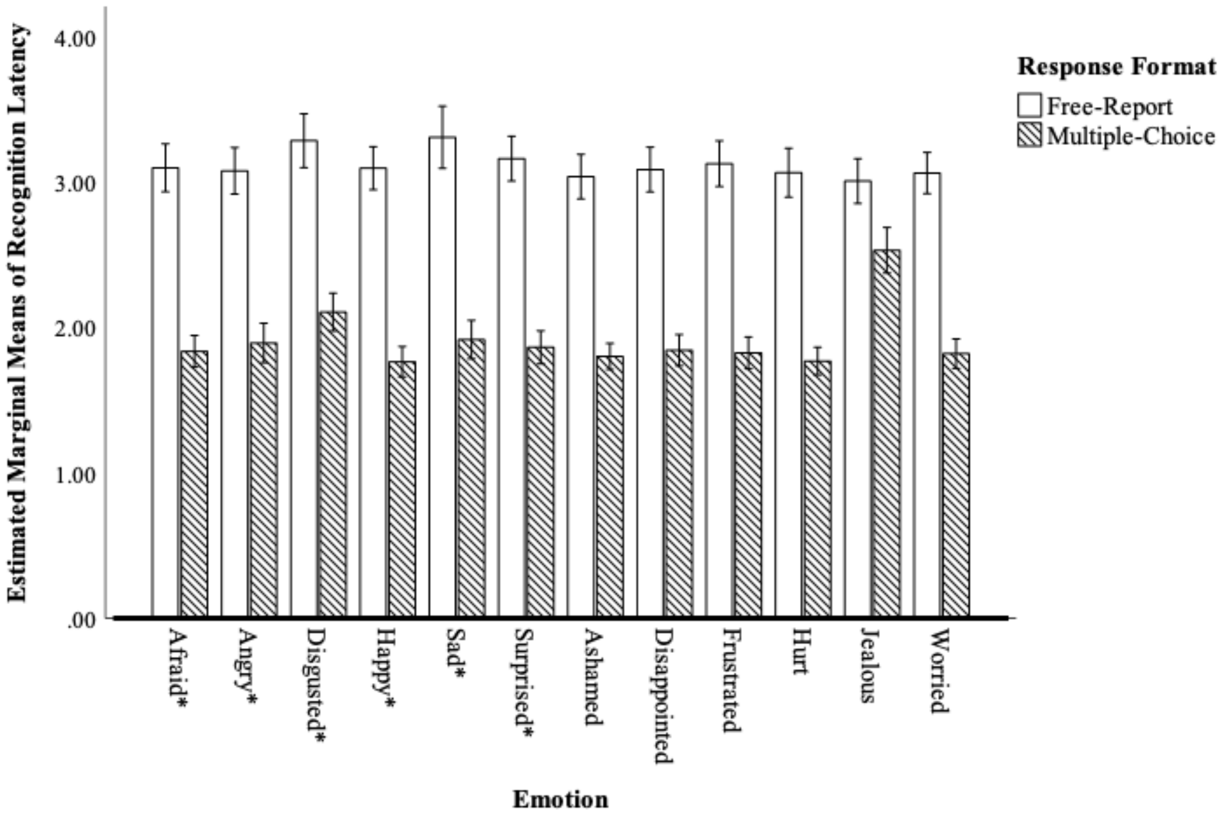
Emotion Recognition Latency Interaction Effects Not Involving Group

Response Format × Emotion

Figure I1 shows markedly longer mean latency for free-report than multiple-choice across all emotions (indicated by the main effect for response format), except for the emotion jealous for which the difference was less striking. Confirming this observation, inspection of the 95% CIs for the mean difference between mean latency for each emotion shown in Table I1 reveals that the mean difference between response formats for jealous was somewhat smaller than for any other emotion.

Figure I1

Two-Way Interaction of Response Format and Emotion on Recognition Latency



Note. Error bars represent 95% confidence interval

Table I1

Pairwise Comparison Outcomes Examining the Response Format × Emotion Interaction on Recognition Latency

Response Format Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
FR-MC	Afraid*	1.26	.07	<.001	1.13	1.40
	Angry*	1.19	.06	<.001	1.06	1.31
	Disgusted*	1.18	.10	<.001	.99	1.37
	Happy*	1.33	.06	<.001	1.22	1.45
	Sad*	1.39	.10	<.001	1.19	1.59
	Surprised*	1.30	.07	<.001	1.15	1.45
	Ashamed	1.24	.07	<.001	1.10	1.37
	Disappointed	1.24	.05	<.001	1.14	1.35
	Frustrated	1.30	.07	<.001	1.17	1.44
	Hurt	1.30	.08	<.001	1.15	1.45
	Jealous	.477	.084	<.001	.31	.64
	Worried	1.24	.064	<.001	1.12	1.37

Note. 'FR' is free report, 'MC' is multiple-choice, * is a basic emotion, ** is based on estimated marginal means

Appendix J

Emotion Recognition Latency Interaction Effects Not Involving Group

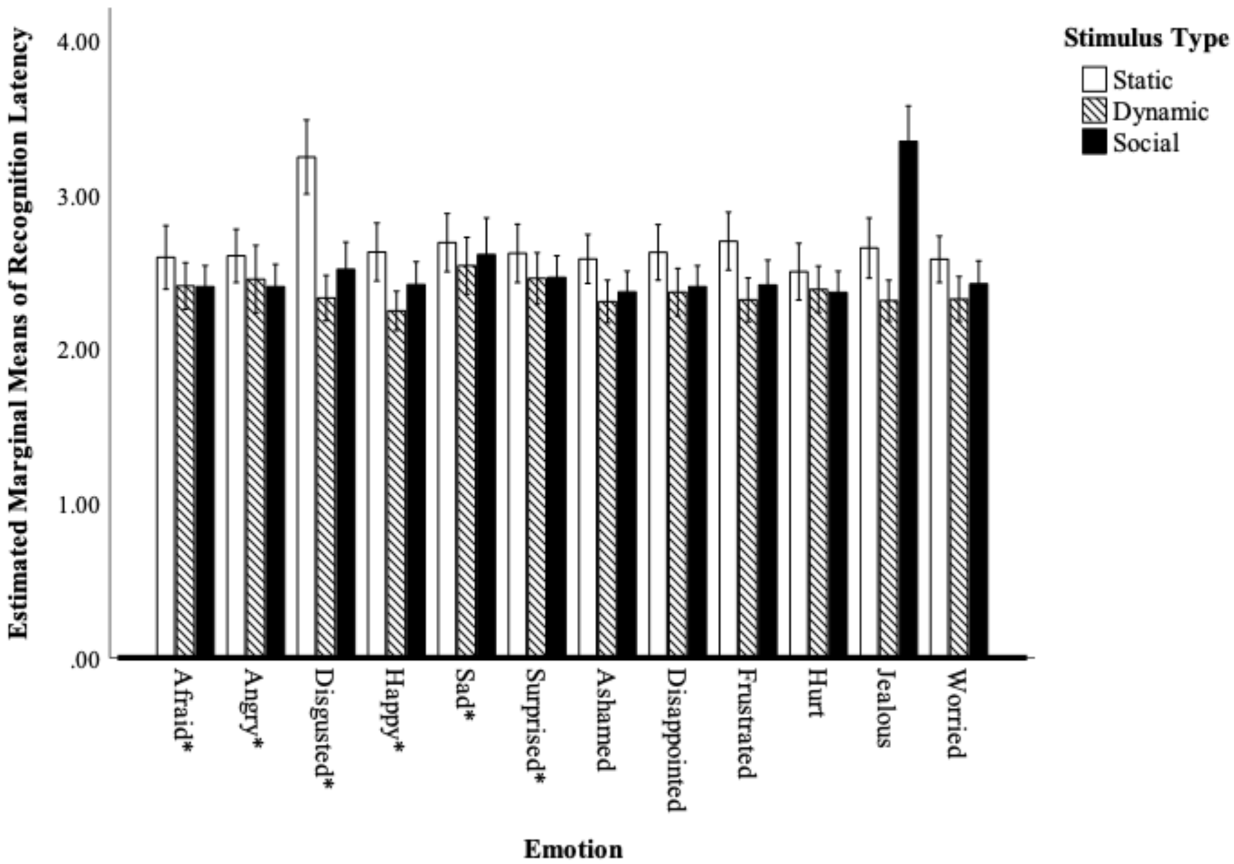
Stimulus Type × Emotion

A significant two-way relationship was demonstrated between stimulus type and Emotion. Although Figure J1 shows that means for the static stimulus type were generally higher across all emotions compared to the dynamic and social stimulus types, the error bars indicate considerable overlap for many of the emotions.

Confirming these observations, inspection of the 95% CIs for the mean difference between mean latency for each emotion as found in Table J1 revealed that for the static-dynamic and static-social comparisons there were 6 and 7 statistically significant comparisons, respectively (static-dynamic: disgusted, happy, ashamed, frustrated, jealous, worried; static-social: angry, disgusted, happy, ashamed, disappointed, frustrated, jealous). For the dynamic-social comparison there were only two statistically significant differences, those being for of the emotions happy and jealous. Note also that, for static stimuli, latency for disgusted was markedly longer than when presented in dynamic or social stimuli. In contrast, for social stimuli, latency for jealous was markedly longer than for the other two stimulus types.

Figure J1

Two-Way Interaction of Stimulus Type and Emotion on Recognition Latency



Note. Error bars represent 95% confidence interval

Table J1

Pairwise Comparison Outcomes Examining the Stimulus Type × Emotion Interaction on Recognition Latency

Stimulus Type Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Static-Dynamic	Afraid*	.19	.11	.108	-.04	.41
	Angry*	.15	.12	.205	-.08	.39
	Disgusted*	.91	.13	<.001	.67	1.16
	Happy*	.38	.09	<.001	.20	.56
	Sad*	.15	.09	.095	-.03	.33
	Surprised*	.16	.13	.199	-.09	.41
	Ashamed	.28	.10	.005	.09	.47
	Disappointed	.26	.10	.012	.06	.46
	Frustrated	.38	.11	<.001	.17	.59
	Hurt	.11	.11	.284	-.10	.33
	Jealous	.34	.10	<.001	.15	.53
	Worried	.26	.10	.010	.06	.45
Static-Social	Afraid*	19	.10	.056	-.01	.39
	Angry*	.20	.09	.023	.03	.37
	Disgusted*	.73	.15	<.001	.44	1.02
	Happy*	.21	.10	.038	.01	.41
	Sad*	.08	.14	.589	-.21	.36
	Surprised*	.16	.10	.118	-.04	.35
	Ashamed	.21	.08	.010	.05	.38
	Disappointed	.22	.08	.008	.06	.39
	Frustrated	.28	.11	.008	.08	.49
	Hurt	.14	.10	.173	-.06	.33
	Jealous	-.69	.14	<.001	-.97	-.41
	Worried	.16	.09	.071	-.01	.33
Dynamic-Social	Afraid*	.01	.09	.951	-.16	.17
	Angry*	.05	.11	.662	-.17	.27
	Disgusted*	-.19	.10	.057	-.38	.01
	Happy*	-.17	.08	.040	-.33	-.01

Table J1 Continued

Stimulus Type Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
	Sad*	-.07	.13	.587	-.34	.19
	Surprised*	-.01	.09	.942	-.18	.17
	Ashamed	-.06	.08	.419	-.22	.09
	Disappointed	-.04	.08	.650	-.19	.12
	Frustrated	-.10	.09	.289	-.28	.08
	Hurt	.02	.08	.794	-.13	.18
	Jealous	1.03	.12	<.001	.80	1.27
	Worried	-.10	.09	.266	-.27	.08

Note. * is a basic emotion, ** is based on estimated marginal means

Appendix K

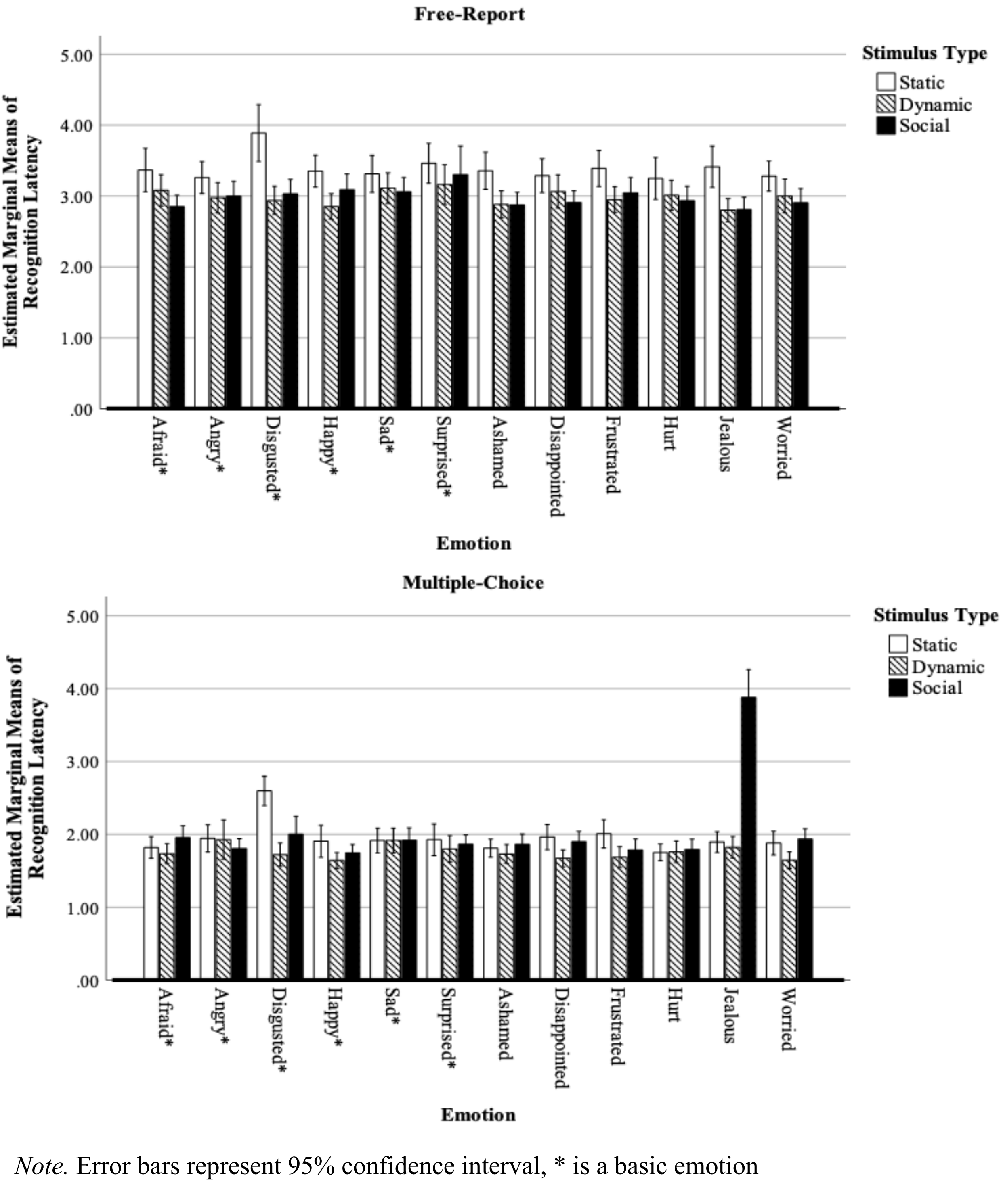
Emotion Recognition Latency Interaction Effects Not Involving Group

Response Format × Stimulus Type × Emotion

The significant three-way interaction between response format, stimulus type and emotion is extremely difficult to interpret, particularly given that the emotion variable has 12 levels. Visual inspection of Figure K1 suggests some possible trends. The patterns shown in Figure K1 show that the static stimuli have longer latencies compared to the dynamic and social stimulus types for all emotions in the free-report response format but only possibly three emotions for the multiple-choice response format. But additionally, under the multiple-choice format there was an unusually long latency for the complex emotion, jealous, in the social stimulus condition.

Figure K1

Three-Way Interaction of Response Format, Stimulus Type and Emotion on Recognition Latency



Note. Error bars represent 95% confidence interval, * is a basic emotion

Appendix L

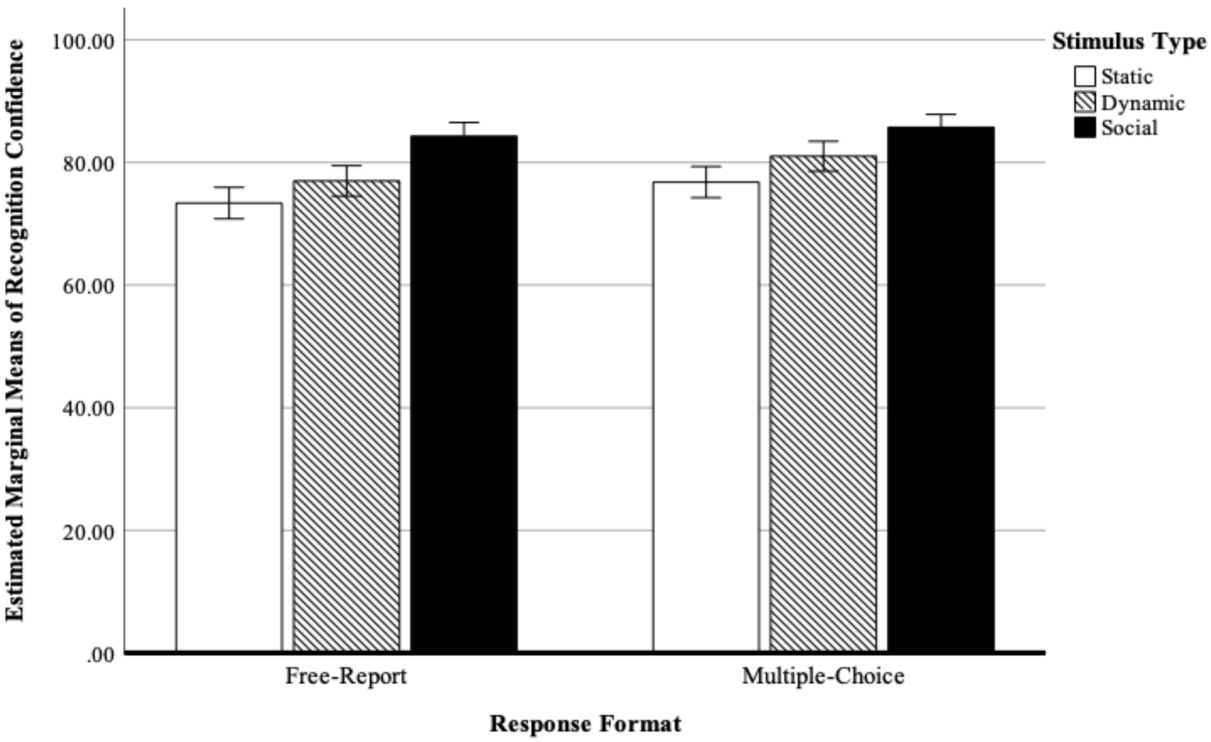
Emotion Recognition Confidence Ratings Interaction Effects Not Involving Group

Response Format × Stimulus Type

There was a significant two-way interaction between response format and stimulus type. Figure L1 suggests higher confidence for social than both dynamic and static stimuli across both response formats. However, that pattern appears more pronounced under the free-report than the multiple-choice format, a conclusion supported by examination of mean differences shown in Table L1.

Figure L1

Two-Way Interaction of Response Format and Stimulus Type on Recognition Confidence Ratings



Note. Error bars represent 95% confidence interval

Table L1

Pairwise Comparison Outcomes Examining the Response Format × Stimulus Type Interaction on Recognition Confidence Ratings

Response Format	Stimulus Type	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
FR	Static-Dynamic	-3.72	.78	<.001	-5.62	-1.82
	Static-Social	-11.14	1.01	<.001	-13.59	-8.70
	Dynamic-Social	-7.42	.92	<.001	-9.64	-5.20
MC	Static-Dynamic	-4.09	.74	<.001	-5.89	-2.28
	Static-Social	-10.03	.94	<.001	-12.33	-7.74
	Dynamic-Social	-5.95	.85	<.001	-8.01	-3.88

Note. ‘FR’ is free-report, ‘MC’ is multiple-choice, **based on estimated marginal means

Appendix M

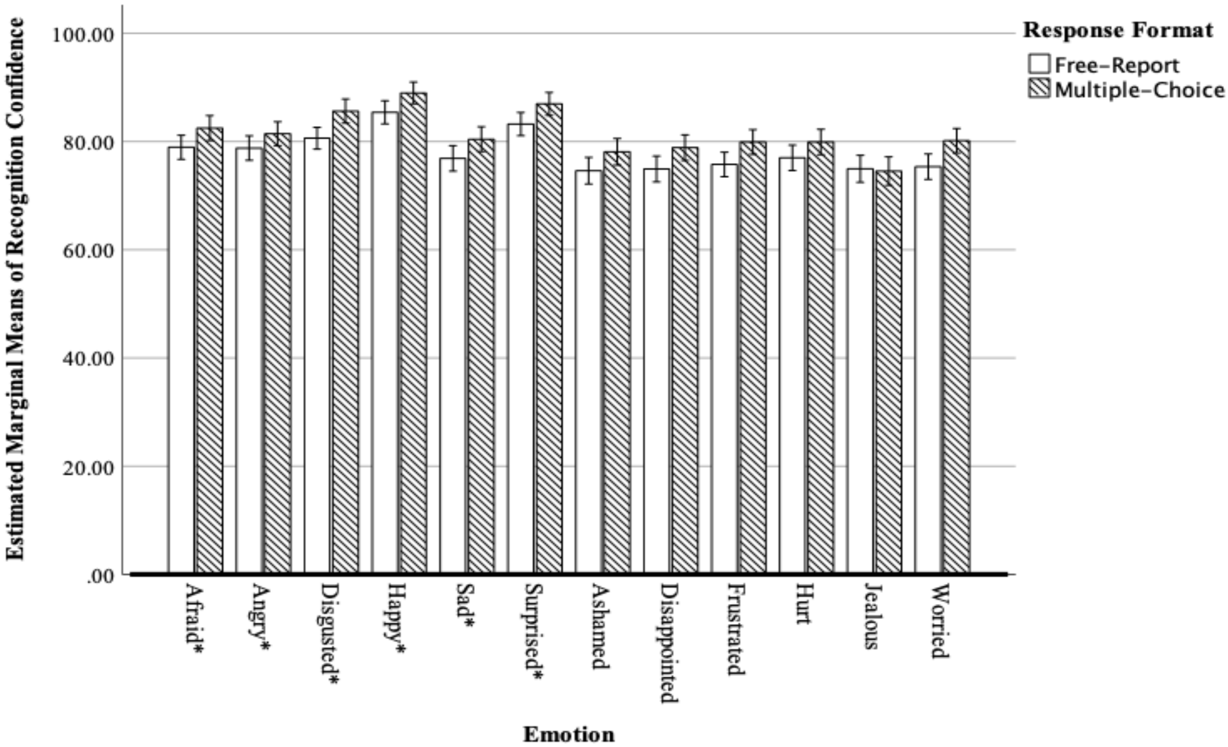
Emotion Recognition Confidence Ratings Interaction Effects Not Involving Group

Response Format × Emotion

There was a significant two-way interaction between response format and emotion. Figure M1 suggests consistently higher mean confidence for multiple-choice than free-report across all emotions, with the exception of the complex emotion jealous. Confirming this observation, inspection of the 95% CIs for the mean difference between emotions in each response format in Table M1 reveals that the mean difference between response formats for jealous was much smaller than for any other emotion.

Figure M1

Two-Way Interaction of Response Format and Emotion on Recognition Confidence Ratings



Note. Error bars represent 95% confidence interval, * is a basic emotion

Table M1

Pairwise Comparison Outcomes Examining the Response Format × Emotion Interaction on Recognition Confidence Ratings

Response Format Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
MC-FR	Afraid*	3.55	.52	<.001	2.52	4.57
	Angry*	2.65	.50	<.001	1.66	3.65
	Disgusted*	4.95	.47	<.001	4.02	5.88
	Happy*	3.58	.40	<.001	2.78	4.38
	Sad*	3.55	.55	<.001	2.46	4.63
	Surprised*	3.74	.46	<.001	2.83	4.66
	Ashamed	3.50	.69	<.001	2.13	4.87
	Disappointed	3.95	.61	<.001	2.73	5.16
	Frustrated	4.16	.59	<.001	2.99	5.34
	Hurt	2.92	.58	<.001	1.77	4.06
	Jealous	-.43	.71	.546	-1.83	.97
	Worried	4.78	.73	<.001	3.34	6.22

Note. ‘FR’ is free-report, ‘MC’ is multiple-choice, * is a basic emotion, **is based on estimated marginal means

Appendix N

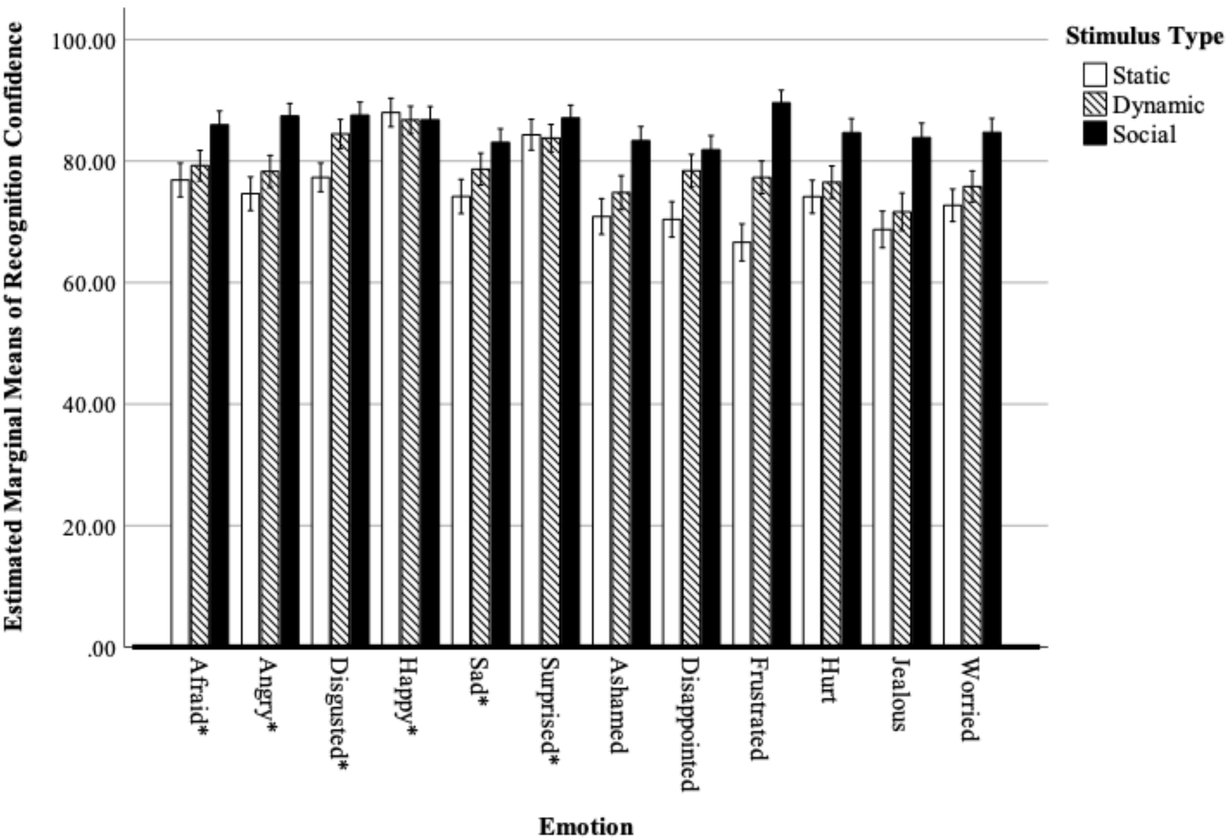
Emotion Recognition Confidence Ratings Interaction Effects Not Involving Group

Stimulus Type × Emotion

There was a significant two-way interaction between stimulus type and emotion as depicted in Figure N1. Given the overlap in confidence across the different stimulus types for many emotions it is difficult to interpret this interaction. However, from visual inspection of Figure N1 and examination of the comparison outcomes in Table N1, it seems that while confidence for the social stimulus type is generally higher for most emotions, this is clearly not the case for the emotion happy and perhaps for surprised.

Figure N1

Two-Way Interaction of Stimulus Type and Emotion on Recognition Confidence Ratings



Note. Error bars represent 95% confidence interval, * is a basic emotion

Table N1

Pairwise Comparison Examinations the Stimulus Type × Emotion Interaction on Recognition Confidence Ratings

Stimulus Type Comparison	Emotion	Mean Difference**	Standard Error	*p*	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Static-Dynamic	Afraid*	-2.37	1.09	.092	-5.00	.26
	Angry*	-3.67	1.03	.001	-6.15	-1.18
	Disgusted*	-7.06	.90	<.001	-9.23	-4.89
	Happy*	1.23	.86	.465	-.86	3.32
	Sad*	-4.50	.98	<.001	-6.87	-2.13
	Surprised*	.55	.82	1.000	-1.44	2.54
	Ashamed	-3.94	1.09	.001	-6.57	-1.30
	Disappointed	-8.01	1.06	<.001	-10.58	-5.45
	Frustrated	-10.70	1.16	<.001	-13.51	-7.89
	Hurt	-2.40	.93	.032	-4.64	-.15
	Jealous	-2.90	1.18	.046	-5.76	-.04
	Worried	-3.09	.92	.003	-5.33	-.85
Static-Social	Afraid*	-9.09	1.12	<.001	-11.80	-6.38
	Angry*	-12.79	1.17	<.001	-15.64	-9.94
	Disgusted*	-10.14	.91	<.001	-12.35	-7.93
	Happy*	1.22	.88	.514	-.93	3.36
	Sad*	-8.90	1.15	<.001	-11.70	-6.11
	Surprised*	-2.77	.96	.014	-5.09	-.45
	Ashamed	-12.49	1.25	<.001	-15.51	-9.47
	Disappointed	-11.47	1.29	<.001	-14.61	-8.34
	Frustrated	-22.98	1.50	<.001	-26.62	-19.34
	Hurt	-10.53	1.09	<.001	-13.18	-7.89
	Jealous	-15.11	1.33	<.001	-18.35	-11.88
	Worried	-12.01	1.18	<.001	-14.87	-9.15
Dynamic-Social	Afraid*	-6.72	1.08	<.001	-9.32	-4.11
	Angry*	-9.13	1.05	<.001	-11.66	-6.59
	Disgusted*	-3.07	.92	.003	-5.31	-.83
	Happy*	.02	.89	1.000	-2.15	2.18
	Sad*	-4.40	1.04	<.001	-6.93	-1.88
	Surprised*	-3.32	.90	<.001	-5.50	-1.14
	Ashamed	-8.55	1.09	<.001	-11.19	-5.92
	Disappointed	-3.46	1.09	.005	-6.09	-.83
	Frustrated	-12.28	1.15	<.001	-15.08	-9.48
	Hurt	-8.14	1.10	<.001	-10.81	-5.47
	Jealous	-12.22	1.24	<.001	-15.22	-9.22
	Worried	-8.92	1.10	<.001	-11.60	-6.24

Note. * is a basic emotion, ** is based on estimated marginal means

Appendix O

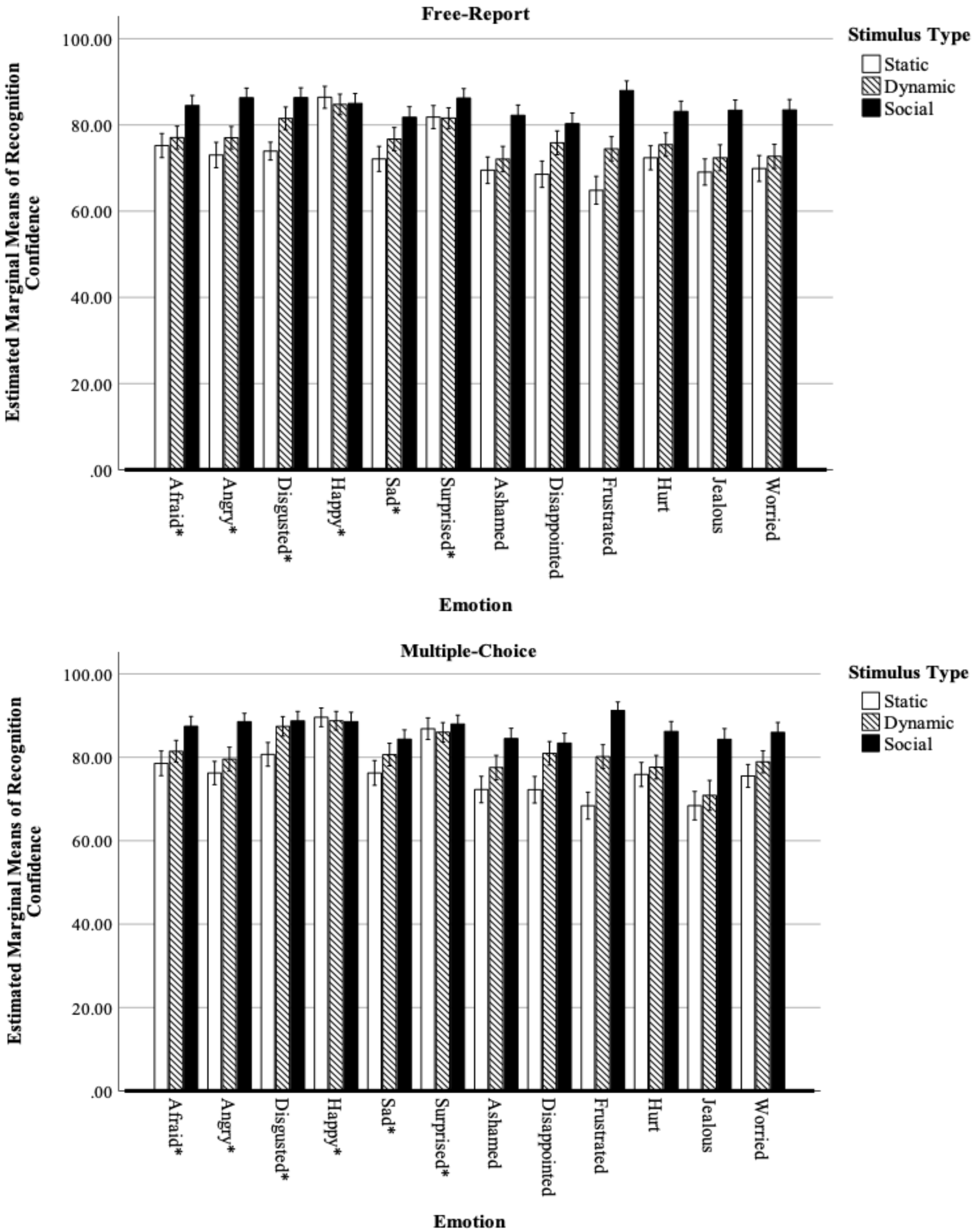
Emotion Recognition Confidence Ratings Interaction Effects Not Involving Group

Response Format × Stimulus Type × Emotion

As depicted in Figure O1, there was a significant three-way interaction between response format, stimulus type and emotion, which is particularly difficult to interpret given that the emotion variable has 12 levels. One possible interpretation may be that the social stimulus condition produced higher confidence for potentially all emotions, except for happy, under the free-report format but only for 10 of the 12 emotions (happy and surprised do not follow the trend) under the multiple-choice response format.

Figure O1

Three-Way Interaction of Response Format, Stimulus Type and Emotion on Recognition Confidence Ratings



Note. Error bars represent 95% confidence interval, * is a basic emotion