University of Reading

Implicit and explicit emotion processing in autism spectrum disorder: An investigation of various nonverbal communicative domains

Thesis submitted for the degree of Doctor of Philosophy
School of Psychology and Clinical Language Sciences

Florence Yik Nam Leung
September 2021

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Florence Yik Nam Leung

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr Fang Liu, for her continuous support, patience, and encouragement at every stage of this journey. Her guidance and immense knowledge have helped me in all the time of research and writing of this thesis. She has provided numerous opportunities for me to develop and has always believed in me before I believed in myself. I am forever grateful for her kindness and generosity, and I could not have imagined having a better supervisor for my PhD study.

I would also like to give my warmest thanks to my second supervisor, Prof Vesna Stojanovik, for her thoughtfulness, continuous optimism, and contributions of time to offer me valuable feedback towards improving my work. She has not only been unwavering supportive of my efforts and struggles in work but has always expressed genuine concern for my well-being. Without her insightful guidance, this PhD would not have been achievable.

My deep appreciation also goes to Prof Jacqueline Sin, who has been instrumental in guiding me and providing me insightful discussions throughout the systematic review process. I must also thank Dr Caitlin Dawson and Dr Chen Zhao who have contributed tremendously to this project with their efforts in screening and data extraction. In particular, I am deeply grateful to Dr Joey Jia Hoong Ong, not only for his contribution to this project, but also for his invaluable advice and feedback on my research. He has always been ready to help me with my questions and concerns and has taught me more than I could ever give him credit for here.

I would like to send my sincere thanks to all the participants who were so generous with their time and took part in my research studies. Thank you for sharing your stories with me and I have thoroughly enjoyed getting to know every one of you. This research would not have been possible without your incredible hard work.

I gratefully acknowledge the funding received towards my PhD from the European Research Council, which has provided me with invaluable opportunities and support for

learning, attending conferences, and conducting this research. I would also like to thank the School of Psychology and Clinical Language Sciences who has provided me with the excellent resources and facilities for collecting data, as well as the opportunities for development beyond research. Special thanks to the Centre for Autism for providing a warm and friendly platform for sharing and discussing research findings and ideas where I have gained insights into the range of related work in the autism field and received valuable feedback on my work. I must also thank them for their tremendous help with participant recruitment.

My special thanks and appreciations go to my fellow labmates in the CAASD lab and my PhD co-workers from office 215 for their constant emotional and moral support, as well as all the fun we have had and all the snacks we have shared in the last four years. I would like to especially thank Alice Wang, Anamarija Veić, and Nanthaka Fuseekul who have always been just a text a way and supported me all the way through this journey.

I owe my deepest gratitude to my Grandparents, Mum, Dad, Sister, Bobby, Mimi, and all my friends from home for their continuous and unparalleled love, help, and support. I am forever indebted to my parents for giving me the opportunities and experiences that have made me who I am. They have always selflessly encouraged me to explore new directions in life and seek my own destiny. This journey would not have been possible if not for them, and I dedicate this milestone to them.

And finally, I owe thanks to a very special person, my partner, Hong, for his love, care, and understanding throughout this long journey. He has not only been by my side living every minute of this journey but has also been a truly great supporter and listener who has endlessly listened to me rant and talk things out and is yet able to constantly put my worries to rest. Thank you for being my rock – without you, I would not have had the courage to embark on this journey in the first place.

Abstract

Emotions can be communicated through multiple domains (e.g., human faces, face-like objects, speech prosody, and song) and processed at implicit (e.g., priming) or explicit (e.g., recognition) levels. Much research has been done on emotion processing in autism spectrum disorder (ASD), yet findings have been highly variable. It remains unresolved whether emotion processing ability at the two levels is generalised across domains or specific to certain domain(s) in ASD. While a number of correlates have been proposed to be associated with emotion processing (e.g., cognitive processing style, pitch perception, and alexithymic trait), their corresponding role in ASD has been infrequently explored. This thesis conducted a systematic review and meta-analysis on emotion recognition in ASD and examined whether explicit emotion recognition and implicit emotion priming differed between autistic neurotypical (NT) individuals across domains from a developmental perspective, while testing the contribution of several related correlates to these processes. Regarding the role of domain and processing level, the ASD group showed intact emotion priming and recognition accuracy but impaired recognition speed and efficiency across domains. These results suggest a generalised emotion ability across domains (i.e., no specific impairments for any particular domain), while indicating a dissociation between implicit and explicit emotion processing in ASD (i.e., impaired recognition speed and efficiency at the explicit level but spared priming at the implicit level). With regards to developmental changes, the developmental trajectory of implicit and explicit emotion processing appears largely comparable between the ASD and NT groups across domains, with age-related improvements seen for recognition and an age-related decline observed for priming. In terms of related correlates, the underlying processes of explicit emotion recognition appear to differ between autistic and NT individuals, with respect to the contribution of cognitive processing style and pitch perception. Additionally, the impaired recognition speed and efficiency in ASD could not be attributed to co-occurring alexithymia.

These findings have shed light on the behavioural profile of emotion processing ability in ASD and the extent to which it differs from that in typical development, which has both theoretical and practical implications for emotion processing in ASD and typical development.

Table of Contents

List of Tables

List of Figures

Chapter 1: General introduction

Social communication difficulties are a hallmark symptom of autism spectrum disorder (ASD; American Psychiatric Association, 2013; World Health Organization, 2018). One manifestation of this impairment is the difficulty for autistic individuals to interpret nonverbal communicative functions, such as emotions (Trevisan & Birmingham, 2016). Various degrees of this impairment can add emotional stressors and risk factors associated with social rejection, bullying, and isolation, often seen among autistic individuals (Chin et al., 2019; T. L. Liu et al., 2019). These negative social circumstances may, in the long run, contribute to poorer social networks and increased risk for mental health problems, such as anxiety and depression (Mazurek, 2014; Tobin et al., 2014). Despite the extensive research efforts devoted to understanding emotion processing in the field, the scientific literature remains equivocal about the nature of how, and the extent to which, emotion processing in autistic individuals differs from that in neurotypical (NT) individuals. Further complicating the picture is the fact that emotions can be communicated via different domains across visual and auditory modalities, such as faces, speech prosody, and music (Ekman & Friesen, 1976; Juslin & Laukka, 2003; Livingstone & Russo, 2018). Moreover, the processing of emotions could be viewed along a continuum from a more implicit/unconscious level to a more explicit/conscious level (Lane, 2008). Thus, the aim of the present thesis is to provide a better understanding of emotion processing in autistic individuals compared to NT individuals across communicative domains and levels of processing.

This introductory chapter is organised as follows: Section 1.1 presents a brief overview of ASD; Section 1.2 elaborates on existing research concerning emotion processing across communicative domains and levels of processing in ASD and typical development, together with literature on the development of related skills in the two populations; Section 1.3 outlines potential correlates of emotion processing and their associations with ASD, as well as discussing the relationship between alexithymia and ASD in emotion processing; and finally, Section 1.4 outlines the aims of the thesis and provides an overview of the subsequent chapters of this thesis.

1.1. Autism spectrum disorder

Autism spectrum disorder (ASD) is a complex, pervasive neurodevelopmental disorder, characterised by (a) profound difficulties with social communication and interaction and (b) restricted, repetitive patterns of behaviours, interests, and activities (American Psychiatric Association, 2013; World Health Organization, 2018). ASD was first described in the 1940s by Leo Kanner (Kanner, 1943) and Hans Asperger (Asperger, 1943). Leo Kanner specifically described ASD as an *"example of inborn autistic disturbances of affective contact"* and that individuals' *"relation to people is altogether different"*, clearly stating that socioemotional processing challenges are part of ASD. The lack of sophisticated understanding of nonverbal communicative functions has been reported as particularly central and persistent across developmental levels in ASD (Seltzer et al., 2004; Shattuck et al., 2007). Although the clinical diagnosis of ASD has varied in line with constantly evolving research, the core description has remained consistent since its conception. The term ASD was redefined in the DSM-5 and ICD-11 (American Psychiatric Association, 2013; World Health Organization, 2018). Changes to criteria included combining previously held subtypes of ASD – autistic disorder, Asperger syndrome, pervasive developmental disorder-not otherwise specified (PDD-NOS) – into a single diagnosis of ASD. This change was driven by research illustrating the poor clarity of diagnostic criteria limiting reliability in assigning subcategory diagnoses, as well as constraints on treatment eligibility and coverage based on subtypes (see Rosen et al., 2021 for a review). The shift to consolidation is in favour of a dimensional approach, which classifies individuals on the autism spectrum based on the severity in the symptom domains (American Psychiatric Association, 2013; Georgiades et al., 2013; Ingram et al., 2008). Additionally, social and language deficits were collapsed into a single measure according to the latest changes to the diagnostic criteria, with a greater focus on social communicative behaviour, while remaining separate from the restricted and repetitive interests, behaviours, and activities domain. This change was driven by analytic findings of a single social-communication factor and the substantial overlap between social and communicative behaviours (see Rosen et al., 2021 for a review).

The prevalence rate of ASD is estimated to be approximately 1-2% of the general population globally (Alshaban et al., 2019; Centers for Disease Control and Prevention, 2020; Chiarotti & Venerosi, 2020; Kim et al., 2011; Sun et al., 2019), including the UK (Baron-Cohen et al., 2009; Brugha et al., 2016). Over the past decades, there has been a sharp rise in the prevalence of ASD worldwide (Chiarotti & Venerosi, 2020; Hong et al., 2020). This is presumably due to broadening of the diagnosis, heightened awareness in society, and better diagnostic tools, while the possibility of an increase in incidence remains speculative (Fombonne, 2005, 2020; Gernsbacher et al., 2005; Nevison, 2014; G. Russell et al., 2021; Wazana et al., 2007; K. Williams et al., 2014). The prevalence of ASD is higher in males than in females, with a ratio typically reported as affecting three to four times as many males as females (Fombonne, 2009; Loomes et al., 2017). This topic has prompted much discussion in the field, with a consistent view of the uneven male-to-female ratio being related to the different clinical presentation of ASD in females that may increase the risk of females with ASD being overlooked or misdiagnosed (Frazier et al., 2014; Loomes et al., 2017).

ASD is currently conceptualised as a spectrum disorder, meaning that there is a wide degree of heterogeneity in symptom severity as well as their impact on individuals' everyday lives within this diagnosis (Happé et al., 2006). Although the full aetiology and pathogenesis of this disorder have not yet been elucidated, the general consensus in the literature suggests genetic, epigenetic, and environmental factors are the primary determinants of this condition (Barak & Feng, 2016; Hallmayer et al., 2011; Lord et al., 2018; Ng et al., 2017; Tick et al., 2016). It is believed that these factors lead to structural and functional changes in the brain, which in turn cause differences in the way individuals process information, their cognitive functions, and eventually their behaviour (Hallmayer et al., 2011; Just et al., 2012; Maximo et al., 2014; Ng et al., 2017). There are currently no reliable diagnostic biomarkers for diagnosing ASD (Shen et al., 2020); ASD continues to be diagnosed by clinicians on the basis of behaviour according to standardised classification systems – DSM-5 and ICD-11 (American Psychiatric Association, 2013; World Health Organization, 2018), where the Autism Diagnostic Observation Schedule-Second Edition (ADOS-2; Lord et al., 2012) and the Autism Diagnostic Interview-Revised (ADI-R; Rutter, Couteur, et al., 2003) are generally considered the "gold standard" evaluative measures for diagnosing ASD.

1.1.1. Current directions in ASD research

In the field of autism science, ASD has long been conceptualised in terms of biologically derived functional deficits under the conventional medical paradigm, also known as the medical model of disability (Pellicano & den Houting, 2021). Within such paradigm, treatment typically targets at bringing an individual's abilities into accordance with the accepted norm, through remediating or eliminating the individual's impairment in order to enhance functioning (Marks, 1997). One key focus of ASD research has, thus, been centred on identifying neurodevelopmental mechanisms at the genetic, neurobiological, and cognitive levels, that might explain the behavioural manifestations of ASD. Despite its significance in advancing our understanding of ASD (Happé & Frith, 2020), the conventional medical approach has not gone unchallenged – given its (a) overfocus on deficits, (b) emphasis on the individual as opposed to social and environmental factors that might contribute to shaping autistic lives, and (c) lack of attention to autistic individuals' perspectives (Gernsbacher et al., 2006; Pellicano & den Houting, 2021; Robertson, 2010). These challenges have led to recent calls from both researchers and the autistic community to reconsider the way in which autism science should be approached (Pellicano & den Houting, 2021).

Over the last decade, the neurodiversity paradigm has gained much traction in many fields and disciplines as an important alternative to the conventional medical paradigm that puts right the critiques mentioned above (den Houting, 2019; Nicolaidis, 2012; Robertson, 2010). This paradigm portrays autism as one form of variation within a diversity of minds (Walker & Raymaker, 2021); that is, the way in which the structure and function of the human brain and mind develop can fall within the range considered as 'typical' neurodevelopment, as well as outside of this range considered to 'diverge' from the norm (Ecker et al., 2015). Two key assumptions have been outlined within the neurodiversity stance: (a) typical neurodevelopment is neither superior nor inferior to divergent neurodevelopment, and thus rejecting the view that divergence from the norm is a flaw requiring fixing, and (b) all individuals deserve to be treated with dignity and respect, independently of how they diverge from a putative norm, and should be valued for who they are and as they are (Pellicano & den Houting, 2021). Moreover, the neurodiversity paradigm is said to focus on strengths, and that any 'deficits' are best understood as a result of an environment that does not effectively accommodate those characteristics, rather than an individual's unique characteristics as suggested by the conventional medical approach (Graby, 2015; Graf et al., 2017). Fundamentally, the neurodiversity paradigm promotes autism acceptance, urging others to embrace ASD as an inherent and integral part of an autistic individual's identity and experience of the world (Pellicano & den Houting, 2021).

Aligned with the neurodiversity movement in viewing strengths, differences, and weaknesses associated with ASD as central to identity (Kapp et al., 2013; Robertson, 2010), the preference for using identity-first (e.g., “autistic person”) rather than person-first (e.g., “person with autism”) language to describe ASD has also been noted among the autism community (Kapp et al., 2013; Kenny et al., 2016; Orsini & Smith, 2010; Shakes & Cashin, 2019). The question of which terms to endorse has, in fact, been subject to much debate. Researchers, clinicians, and students have traditionally been guided to adopt person-first language as it is considered standard in writing across government documents, scientific journals, and various organisations’ publications (Crocker & Smith, 2019). Person-first language was originally designed as a response to dehumanisation and violence towards autistic and disabled people during earlier decades, by placing significance on the individual rather than their disability (Dunn & Andrews, 2015). However, emerging evidence has not found this approach to be efficacious, that may, in turn, has counteracting effects such as increased stigma (Gernsbacher, 2017). Importantly, two empirical studies specifically exploring language preferences (Kapp et al., 2013; Kenny et al., 2016), though limited by the representativeness of their samples, showed that identity-first terms are the most preferred language for the autistic community. By contrast, the specific person-first formulation of “person with autism” or “person with autism spectrum disorder/condition” have been regarded as the least preferred and most offensive language (Bury et al., 2020; Kapp et al., 2013; Kenny et al., 2016). The adoption of identity-first language has gained momentum within autistic associations and throughout literature, with the American Psychological Association’s website also acknowledging the use of identity-first language (American Psychological Association, 2018; https://www.apa.org/pi/disability/resources/choosing-words). Henceforth, identity-first terms “autistic” and “neurotypical” will be used throughout this thesis.

1.2. Existing literature on emotion processing in ASD and NT

Humans regularly exchange complex information to support the functioning of individual lives and wider society; an integral part of these exchanges is the communication of emotions (Carton et al., 1999; Cosmides & Tooby, 2000; Darwin, 1872; Ekman, 1992b; Keltner & Haidt, 2001; Shariff & Tracy, 2011). Emotions can be encoded using various nonverbal cues, where the decoding of these cues allows observers to recognise emotions expressed by others that helps the individual coordinate social interactions rapidly and appropriately (Ekman, 1993; Hwang & Matsumoto, 2019; Keltner & Haidt, 1999; K. R. Scherer, 1986; Shariff & Tracy, 2011). Among other facets of emotion science such as production of emotional expressions and induction of emotions, this thesis will focus on the perception of emotions – which refers to when an individual perceives or recognises emotions without necessarily feeling the emotions themselves.

As described earlier, a fundamental challenge for autistic individuals is social communication, a core part of which is the lack of sophisticated understanding of nonverbal communicative functions, such as the perception of emotional cues (American Psychiatric Association, 2013; Trevisan & Birmingham, 2016; World Health Organization, 2018). A large body of research on the emotion processing ability in ASD has, nonetheless, produced highly mixed results. Variations in experimental designs used to evaluate emotion processing have undoubtedly led to difficulties in integrating research findings, in order to create a cohesive picture of the emotion processing ability in autistic individuals. In the subsequent subsections, I provide an overview of the relevant literature on the different aspects of emotion processing in ASD and typical development, paying special focus on the different levels of processing, multiple domains of emotional communication, and the developmental course of this ability.

1.2.1. The role of processing levels

Emotion perception can happen at implicit or explicit levels (Celeghin et al., 2020; Clausi et al., 2017; Habel et al., 2007; Lane, 2008). Characterised by its fast, automatic, and stimulus-driven process, implicit processing of emotions does not require conscious awareness (C. D. Frith & Frith, 2008). In contrast, explicit processing of emotions refers to a slow, controlled, and attentionally-demanding process that requires declarative evaluation and involves higher cognitive resources (Birnboim, 2003; Lane, 2000; W. Schneider & Chein, 2003; W. Schneider & Shiffrin, 1977). It is noteworthy that there could be a dissociation between the two levels, such that impairment at the explicit level does not necessarily imply impairment at the implicit level (Mattavelli et al., 2021; Roux et al., 2010; Wagenbreth et al., 2016; Wieser et al., 2006).

1.2.1.1. Implicit emotion processing

Humans are constantly confronted with an immense amount of sensory information; yet, the attentional capacity to process this information is limited (Klingberg, 2000; Marois & Ivanoff, 2005). Consequently, only certain information is singled out as being appropriate for processing based on its valence, urgency, and significance to the organism according to appraisal theories (Desimone & Duncan, 1995; see also Scherer, 2009 for a summary). Within this context, emotional content is thought to be particularly salient and has higher processing priority in perception, due to both bottom-up (e.g., enhanced perceptual distinctiveness and biological preparedness) and top-down factors (e.g., past experience and prior knowledge) (Brosch et al., 2010; De Martino et al., 2009; Öhman et al., 2012; Yiend, 2010).

The appraisal process can and often proceeds automatically, which may operate on both conscious and unconscious inputs (Moors, 2010). For instance, task-irrelevant emotional information can be implicitly and automatically processed in faces when making an age estimation (i.e., where faces represent a conscious input; Passarotti et al., 2009), or when faces are presented under subliminal conditions (i.e., where faces represent an unconscious input; Kamio et al., 2006). Implicit emotion processing can be assessed using neurophysiological approaches (e.g., to measure neural sensitivity towards subliminally presented/unattended emotional expressions; H. Liu et al., 2010; Suslow et al., 2013; L. Zhou et al., 2019) and eye-tracking techniques (e.g., to examine automatic visual scanning patterns of emotional expressions; Van der Donck et al., 2021). Behavioural methods have also been used to measure measure implicit emotion processing: for example, it has also been shown that emotional meaning can be implicitly activated to influence subsequent behaviours such as emotional judgment through affective priming. This has been found across a number of different contexts, such as from verbal to nonverbal within modality (e.g., printed words priming facial expressions; Carroll & Young, 2005) and across modalities (e.g., printed words priming music/prosody; Goerlich et al., 2012), from nonverbal to verbal within modality (e.g., facial expressions priming printed words; Carroll & Young, 2005) and across modalities (e.g., music/prosody priming printed words; Goerlich et al., 2012; Steinbeis & Koelsch, 2011), as well as from nonverbal to nonverbal within modality (e.g., facial expressions priming facial expressions; Vanmarcke & Wagemans, 2017) and across modalities (e.g., music/prosody priming facial expressions; Pell, 2005; L. Zhou et al., 2019). These findings highlight the facilitatory role of emotional content in perception that allows it to be implicitly processed, where this implicitly processed information has the potential to prime conscious emotion recognition in typical development.

Implicit emotion processing has been proposed to be more impaired than explicit emotion processing in ASD (U. Frith, 2004). This is perhaps due to the fact that any difficulties resulting from impaired implicit learning of social information could be concealed on a behavioural (explicit) level through compensatory mechanisms (U. Frith, 2004; Rutherford &

McIntosh, 2007). Studies using various neural methods have consistently shown an impairment in implicit emotion processing in autistic individuals. For example, brain imaging studies have consistently reported reduced activation in brain regions towards task-irrelevant emotional information from facial and/or body expressions in autistic individuals relative to NT controls; these brain regions include the fusiform gyrus, medial prefrontal cortex (MPFC), superior temporal gyrus (STG), the amygdala, and the cerebellum (Ciaramidaro et al., 2018; Critchley et al., 2000; Kana et al., 2016). Furthermore, studies measuring event-related brain potentials (ERPs) have reported diminished mismatch negativity (MMN) response to emotionally spoken syllables during passive listening tasks in autistic individuals, indicating impaired automatic discrimination of emotional voices (Fan & Cheng, 2014; Lindström et al., 2018). Importantly, previous studies showed that despite of reduced neural sensitivity to emotional expressions, the ASD group recognised emotions equally accurately as the NT group (Van der Donck et al., 2019, 2020). Using eye-tracking techniques, studies have frequently shown atypical automatic scanning patterns of emotional faces in autistic individuals, such as during passive viewing of these stimuli (Nuske et al., 2014; Pelphrey et al., 2002). For example, whereas NT individuals tend to vary their scanning patterns in relation to the emotional content of expressions (e.g., with more fixations on the eyes for negative expressions and the mouth for positive expressions; Eisenbarth & Alpers, 2011; Wegrzyn et al., 2017), autistic individuals appear to show less of this differentiation (Åsberg Johnels et al., 2017). Again, despite of less exploratory visual strategy to process facial expressions (as demonstrated by smaller saccadic amplitudes), the ASD group recognised emotions from facial expressions just as accurately and quickly as the NT group (Van der Donck et al., 2021). These findings, together, give rise to the importance of studying implicit emotion processing in ASD, given that differences at the implicit level between autistic and NT individuals are not often reflected at the explicit level.

Implicit emotion processing in ASD, however, has rarely been studied using behavioural methods. One study employing the affective priming paradigm found that whereas liking ratings of Japanese ideographs were primed by subliminally presented emotional faces (e.g., happy faces priming higher liking ratings of the ideographs) in NT individuals but not in individuals with high-functioning pervasive developmental disorders (HFPDD) (Kamio et al., 2006). These results could be interpreted as that the failure to implicitly activate the emotional meaning of the unconsciously presented expressions may have resulted in the absence of priming effects observed in the HFPDD group. Conversely, in another study, valence judgment of emotional faces was implicitly primed by emotional face primes containing high or low spatial frequency information (e.g., positive facial expressions priming faster valence judgment of positive faces) in both autistic and NT individuals to similar extents (Vanmarcke & Wagemans, 2017). These results, diverging from those by Kamio et al. (2006), indicate that autistic individuals do not appear to show impairments in the implicit emotion processing of unconsciously presented expressions, and have therefore exhibited priming effects similar to those in NT individuals. Taken together, whereas previous studies implementing neural methods have provided consistent evidence for an impairment in the implicit appraisal of emotions and implicit discrimination between emotion expressions on the neural level in ASD, the influence of implicit emotion processing on the behavioural level in ASD relative to typical development remains largely unclear based on the scarce yet disparate findings from studies implementing behavioural methods in the existing literature.

1.2.1.2. Explicit emotion processing

In contrast to implicit emotion processing, explicit emotion processing indexes the conscious and strategic processing of emotions (Birnboim, 2003; Lane, 2000; W. Schneider & Chein, 2003; W. Schneider & Shiffrin, 1977). In the laboratory, this can be measured through the use of explicit instructions that direct participants to the emotional content of the stimuli presented in verbal and nonverbal tasks, and results show that NT participants excel in both. Nonverbal emotion recognition tasks, such as those that require making same/different judgments (e.g., Alonso-Recio et al., 2014; DeKosky et al., 1980; Greimel et al., 2014; Vannetzel et al., 2011) or matching expressions (e.g., Herba et al., 2006; Philip et al., 2010; Tanaka et al., 2012), tap into the perceptual stage of processing, where emotions are discriminated based on perceptual properties alone under conscious awareness (Adolphs, 2002; Palermo et al., 2013). By contrast, verbal emotion recognition tasks, such as those that require generating or assigning (from pre-generated options) emotional labels to emotional displays (e.g., Bombari et al., 2013; Boucher et al., 2000; Castelli, 2005; Elfenbein & Ambady, 2002; Mohn et al., 2011), are more cognitively demanding as they place greater reliance on emotion vocabulary (Palermo et al., 2013). These findings provide evidence that emotions can be explicitly categorised both verbally and nonverbally by NT individuals.

Explicit emotion processing has been investigated in both verbal and nonverbal contexts in autistic individuals, with the majority of studies implementing a verbal task. Findings have, however, been highly inconsistent across the different types of tasks. With respect to nonverbal tasks, including discrimination or matching between emotional expressions, both intact and impaired accuracy have been reported in autistic individuals (impaired: Philip et al., 2010; Tanaka et al., 2012; Vannetzel et al., 2011; intact: Lindström et al., 2018; Sasson et al., 2016). In terms of speed, studies have consistently reported slower discrimination between emotional expressions in autistic individuals compared to NT individuals (Greimel et al., 2014; Lindström et al., 2018; Sasson et al., 2016), whereas others have reported similar detection speed in autistic and NT individuals when the task required detection of the odd emotion expressions among other distractors (Isomura, Ogawa, et al., 2014; Kujala et al., 2005).

With regards to verbal tasks encompassing emotion labelling, discrimination, or matching with the use of pre-generated emotional labels, both intact and impaired accuracy have been reported in autistic individuals (impaired: Ciaramidaro et al., 2018; Griffiths et al., 2019; Loth et al., 2018; Oerlemans et al., 2014; Philip et al., 2010; intact: Baker et al., 2010; Ketelaars et al., 2016; Lindström et al., 2018; Rhodes et al., 2018). Similarly, in terms of speed, autistic individuals have been reported to show both slower but also comparable performance relative to NT individuals on verbal labelling and discrimination (slower: Berggren et al., 2016; Eack et al., 2015; Ketelaars et al., 2016; Kliemann et al., 2010; Sawyer et al., 2012; comparable: Akechi et al., 2010; Fink et al., 2014). Taken together, it appears that autistic individuals also have difficulties with explicit emotion processing. However, as can be seen, the picture remains highly unclear from the contradictory findings, which may have been confounded by the use of compensatory skills such as explicit cognitive or language-mediated processes to recognise emotions (Behrmann et al., 2006; J. B. Grossman et al., 2000; see U. Frith, 2004 and Harms et al., 2010 for reviews).

1.2.1.3. Summary

Although implicit and explicit processing of emotions appear to be dissociated processes, both impaired and preserved performance have been observed across the two levels in ASD. As noted earlier, while a number of studies have examined the implicit appraisal of emotions and implicit discrimination between emotional expressions in ASD, research into how emotional meaning is implicitly activated to prime subsequent behavioural judgment has been rarely explored. In addition, implicit emotion priming and explicit emotion recognition have seldom been studied within the same study in ASD – though one study by Kamio et al. (2006) provided insights into these two areas by showing impaired implicit priming of facial emotional cues on subsequent liking judgment but spared explicit recognition of these facial expressions. Thus, more research in these two areas (implicit and explicit emotion processing) is needed.

1.2.2. The role of nonverbal communicative domain

Given the significance of nonverbal emotional communication preceding the evolution of verbal abilities (Darwin, 1872; Masson & McCarthy, 1996; Parr et al., 2005), it is not surprising that emotions can be communicated through various nonverbal channels (Etcoff & Magee, 1992; Juslin & Laukka, 2003; Mohn et al., 2011; J. A. Russell et al., 2003; K. R. Scherer, 1986). Emotion recognition in different domains has been proposed to be underpinned by a unified multimodal processing ability (Borod et al., 2000) – that is, the capacity to process emotions in multiple domains stems from a more generalised latent emotion processing ability. In ASD, some studies suggest emotion perception impairments may be generalised (Lerner et al., 2013; Philip et al., 2010). In contrast, there is evidence suggesting emotion processing ability may be specific to domain(s) (Brosnan et al., 2015; Rosset et al., 2008), especially when drawing together findings from separate studies investigating different domains discretely (e.g., Griffiths et al., 2019; Quintin et al., 2011; Schelinski & von Kriegstein, 2019). The domains of particular interest for the present thesis are human faces and nonhuman faces (represented by artificial facial cues such as cartoon and animated faces) from the visual modality, and speech prosody and music from the auditory modality. In the following subsections, I provide an overview of the components that contribute to the discrimination between emotional expressions, as well as the existing findings on emotion processing in ASD, for each domain.

1.2.2.1. Human faces

The human face is thought to be special to the typical visual system and allows an individual to communicate a wide range of nuanced emotions (Ekman, 1993; Ekman & Friesen, 1976; McKone & Robbins, 2011; Parkinson, 2005). The earliest scientific exploration of the association between facial expressions and emotions dates back to Duchenne's pioneering work, which found that it was possible to elicit emotional expressions through electrical stimulation of facial muscles (Duchenne & Cuthbertson, 1862). Subsequently, Darwin (1872) hypothesised that all humans show emotions through similar expressions that could be traced across cultures and species. Building on this, Ekman (1992a) identified six basic emotions – anger, fear, happiness, sadness, disgust, and surprise – that are represented using a distinctive set of facial movements universally across cultures (though there have been controversies regarding the exact number of universal emotions; see Jack et al., 2016 and J. A. Russell, 1994)

Facial movement plays a critical role in facilitating facial emotion recognition (Bassili, 1978, 1979). The face movement pattern of each emotional expression could be described according to the Facial Action Coding System (FACS; Ekman & Friesen, 1976). For example, an angry expression comprises furrowed brows, wide eyes, tightened and pressed-together lips, whereas a fearful expression comprises raised and pulled-together brows, raised upper eyelids, tensed lower eyelids, and parted and stretched lips (Ekman & Friesen, 1976; Keltner, Dacher, Cordaro, 2015). Although the processing of featural information (i.e., concerning individual face parts) from these emotional expressions is sufficient to produce accurate recognition (Bombari et al., 2013; Leppänen & Hietanen, 2007), the processing of configural information (i.e., concerning the relations between face parts) has been found to play a more prominent role in typical emotion recognition (Bombari et al., 2013; Calder et al., 2000; Derntl et al., 2009; McKelvie, 1995; Prkachin, 2003). This has been demonstrated in the strategic and controlled visual scan paths exhibited by NT individuals, where a triangle subtending the eyes, nose, and mouth was generally traced across emotional faces (Pelphrey et al., 2002). Moreover, the extent to which featural and configural information contribute to emotion recognition may further depend on the individual emotion (Bombari et al., 2013). Specifically, anger requires both featural and configural information, fear and sadness rely on featural and configural information respectively, and happiness can be recognised by either featural or configural information (Bombari et al., 2013). Research shows that human faces do not seem to be so special in perception for autistic individuals (Kikuchi et al., 2009; Rosset et al., 2010; see also Simmons et al., 2009 for a detailed discussion). Although the literature has not been consistent regarding configural processing of faces in ASD, studies have generally shown qualitative differences in the way autistic individuals process faces (e.g., Dalton et al., 2005; Joseph & Tanaka, 2003; Pelphrey et al., 2002; see also Simmons et al., 2009 for a review). According to the facial emotion recognition model proposed by Adolphs (2002), the analysis of the emotion conveyed by facial expressions is preceded by visuo-perceptual processing of faces. In this sense, impairments in facial perception might hinder subsequent facial emotion recognition performance.

A wealth of research has sustained debate on the specific aspect of face perception in ASD, namely the processing of emotions from facial expressions (Harms et al., 2010; Lozier et al., 2014; Simmons et al., 2009; Uljarevic & Hamilton, 2013). As already mentioned (Section 1.2.1.1), studies examining the implicit processing of facial expressions have reported impairments in ASD (Ciaramidaro et al., 2018; Critchley et al., 2000; Kamio et al., 2006; Kana et al., 2016; Riby et al., 2012). By contrast, studies examining the explicit processing of facial expressions have yielded mixed results. Some studies have reported clear impairments across emotions in ASD for both accuracy (e.g., Eack et al., 2015; Griffiths et al., 2019; Pelphrey et al., 2002) and speed (e.g., Greimel et al., 2014; Sawyer et al., 2012); other studies observed emotion-specific impairments, particularly with negative emotions (e.g., Ashwin et al., 2006; Boraston et al., 2007; Tell et al., 2014; Yeung et al., 2020). Conversely, comparable performance between autistic and NT individuals has also been observed for both accuracy (e.g., Jones et al., 2011; Tracy et al., 2011) and speed (e.g., Fink et al., 2014; J. B. Grossman et al., 2000). Notwithstanding the mixed findings, previous meta-analyses have consistently demonstrated impairments in facial emotion recognition in ASD based on the great amount of evidence in the literature, providing more clarification on this topic within the human face domain (Lozier et al., 2014; Uljarevic & Hamilton, 2013).

1.2.2.2. Nonhuman faces

In contrast to human faces which provide naturalistic cues that imply social processing for the interpretation of communicative intent and emotional response, nonhuman faces such as cartoon, caricature, schematic, and pareidolic faces provide artificial emotional cues through face-like features (Donath, 2001). Although human and nonhuman faces have been proposed to possess differing social relevance (Rosset et al., 2010; Santos et al., 2009), there is evidence suggesting that both stimulus types access a common face recognition system and elicit similar patterns of neural activation (Tong et al., 2000). Moreover, previous studies have illustrated that NT individuals employ similar strategies for processing human and nonhuman faces, as both types of faces elicit the inversion effect (Guillon et al., 2016; Rosset et al., 2008) – an effect which disrupts the processing of facial features as configural information is preserved (Tanaka & Farah, 1993; Tanaka & Sengco, 1997). This suggests that both types of faces are processed configurally. Nonhuman facial stimuli have been shown to reliably convey the six basic emotions (Brosnan et al., 2015; Desmet, 2003), although NT individuals generally show better accuracy with human faces than nonhuman faces on emotion recognition tasks (Brosnan et al., 2015; Rosset et al., 2008).

It has been thought that the use of nonhuman faces as such allows a degree of social demand to be removed, and thus reduces stress and interference caused by potential social interaction when viewing human faces (see Riby & Hancock, 2009; Rosset et al., 2008). In essence, although studying emotion processing of human faces aids understanding of emotions induced by others in an interpersonal context, the use of nonhuman faces as experimental stimuli offers an avenue to distinguish emotion recognition deficits (if any) from impairments in social information or human face processing in ASD. Prior research has established a unique link between children with ASD and restricted interest in cartoons (Anthony et al., 2013; Grelotti et al., 2005; Kuo et al., 2014; Spiker et al., 2012). However, similar to the human face domain, literature on the processing of cartoon faces in ASD has also been mixed. Specifically, some studies have shown that children with ASD spent similar amount of time looking at cartoon faces and processed these faces in a configural manner just as their NT counterparts (Rosset et al., 2008; Van Der Geest, Kemner, Camfferman, et al., 2002; Van Der Geest, Kemner, Verbaten, et al., 2002). It was speculated that greater interest towards cartoons in ASD might have led to greater acquired expertise with cartoon faces, hence enabling the typical configural processing strategy to be employed (Rosset et al., 2008). By contrast, other studies have observed the opposite, where children with ASD spent less time looking at cartoon faces than their NT counterparts and processed schematic faces in a feature-based manner (Isomura et al., 2014; Riby & Hancock, 2009). Although the processing strategy for cartoon faces in ASD remains unclear, results from a handful of studies have yielded relatively consistent findings with respect to the preserved ability to accurately recognise emotions from cartoon faces in autistic individuals (Brosnan et al., 2015; Isomura, Ogawa, et al., 2014; Miyahara et al., 2007; Rosset et al., 2008), as well as showing comparable speed to NT individuals (Miyahara et al., 2007).

1.2.2.3. Speech prosody

The human voice is capable of making a wide variety of sounds and plays a significant role in social communication (Blasi et al., 2011). Whilst recognising facial expressions requires seeing the other person at relatively close proximity, voices can be heard from further away and also from all directions without having to turn the head. Thus, communicating emotions through the voice may be especially useful for situations where immediate action is required (Matsumoto et al., 2012). There are different types of information that can be obtained through the voice, including both verbal content and speech prosody (Kostic & Chadee, 2014). Focusing on the latter, speech prosody refers to the suprasegmental attributes of spoken language, such as intonation, tone, stress, and rhythm, which can be dissociated from verbal content (Banse & Scherer, 1996). The varying of prosodic features of speech such as intensity (i.e., loudness), voice quality (i.e., roughness), tempo (i.e., rate), and particularly fundamental frequency (i.e., pitch) is induced by autonomic changes and muscle activation patterns corresponding to the emotional states of the speaker (Belin et al., 2004). Speakers can, thus, portray additional meaning about their transient emotional states, including the six basic emotions (Banse & Scherer, 1996; Elfenbein & Ambady, 2002; Frick, 1985; Grandjean et al., 2006; Juslin & Laukka, 2003; Pell & Kotz, 2011; K. R. Scherer, 1986). For instance, a happy prosody is generally characterised by medium to high intensity, fast speech rate, and high pitch level with much variability and a rising contour, whereas a sad prosody is generally characterised by low intensity, slow speech rate, and low pitch level with little variability and a falling contour (Juslin & Laukka, 2003). Additionally, speakers can modulate their communicative intent through these features such that emotions are conveyed in a way that contradicts the verbal content (Cosmides, 1983; K. R. Scherer et al., 1984).

Decoding of emotions in speech prosody encompasses three independent stages: (i) extracting acoustic features, (ii) detecting meaningful connections between these features as an utterance unfolds, and (iii) conceptual processing of acoustic patterns in relation to emotion-specific knowledge held in long-term memory (Pell & Kotz, 2011; Schirmer & Kotz, 2006). Given that prosodic expressions are inherently dynamic and are dictated by their temporal structure, the integration of individual acoustic parameters to form different acoustic patterns is important for the differentiation between emotional expressions (Banse & Scherer, 1996; Pell, 2001; Sobin & Alpert, 1999). In other words, listeners have to track both absolute and relative changes in acoustic features (i.e., mean as well as variation) as speech unfolds to form discrete impressions about the speaker’s emotions (see Juslin & Laukka, 2003 for a detailed overview). The ability to spontaneously extract and integrate nonverbal acoustic features that accompany the verbal content is fundamental for successful emotion recognition.

Speech prosody has been shown to have greater perceptual saliency over verbal content in communicating emotions in typical development (Lin & Ding, 2019; Schwartz & Pell, 2012). Research has indicated that when NT individuals hear emotional prosody, the underlying emotional meaning is implicitly activated even when these meanings are not within attentional focus (Paulmann & Kotz, 2008; Pell & Skorup, 2008). By contrast, autistic individuals tend to rely on verbal content (Lindner & Rosén, 2006; Stewart et al., 2013) or contextual cues (Le Sourn-Bissaoui et al., 2013) to infer the speakers’ emotions rather than using the emotional prosody. Further to this, the investigations of implicit emotion processing of prosody have consistently revealed impairments in ASD (Fan & Cheng, 2014; Kujala et al., 2005; Lindström et al., 2018). By contrast, findings on explicit emotion processing based on prosody in ASD have been inconsistent. Whereas some studies demonstrated impairments across emotions (Doi et al., 2013; Schelinski & von Kriegstein, 2019; L. J. Taylor et al., 2015) or a specific emotion (e.g., happiness; Hubbard et al., 2017; J. E. Wang & Tsao, 2015), other studies reported no difference in accuracy between the ASD and NT groups (Baker et al., 2010; Boucher et al., 2000; Heikkinen et al., 2010; Ketelaars et al., 2016; O’Connor, 2007). It, therefore, remains unclear whether and to what extent autistic individuals are able to extract basic features of emotional speech prosody. Nonetheless, studies investigating the speed of emotion recognition have almost exclusively reported slower performance by autistic individuals compared to NT individuals on both discrimination (Kujala et al., 2005; Lindström et al., 2018) and labelling tasks (Ketelaars et al., 2016; Oerlemans et al., 2014; Waddington et al., 2018).

1.2.2.4. Music

The communication of emotion to promote social cohesion and well-being is generally regarded as the primary purpose of music (Gabrielsson & Juslin, 2003; Juslin, 2000, 2001; Juslin & Laukka, 2003; Snowdon et al., 2015; Trainor, 2010). Emotions can be communicated through different forms of musicality, including both instrumental music (e.g., piano, violin, cello, and clarinet) (Juslin & Laukka, 2003; Mohn et al., 2011; Paquette et al., 2013; Resnicow et al., 2004) and vocal music (i.e., singing) (Livingstone & Russo, 2018; K. R. Scherer et al., 2015; B. Zhang, Provost, Swedberg, & Essi, 2015). Emotional expressions in music have been hypothesised to be partly based on a code for vocal emotional expressions that served important functions throughout evolution (Juslin, 1997a, 2001). Consistent with this view, the way different emotions are conveyed in singing and instrumental music have been shown to have a high degree of similarity to that in speech prosody (Juslin & Laukka, 2003; Juslin & Sloboda, 2013; Nordström & Laukka, 2019; Scherer et al., 2015). Studies employing different analytic methods have noted some consistent relationships between musical structure (i.e., configurations of musical features) and emotional expression (Juslin, 1997; Juslin & Laukka, 2003; Schubert, 2004). The manipulation of composer-related features (e.g., mode, pitch, and rhythm) and performer-related features (e.g., tempo and sound-level) of a musical piece can, therefore, influence the perception of different emotions (Juslin & Laukka, 2003; Juslin & Sloboda, 2013).

A particular musical feature alone may not be a reliable indicator of any emotion, as it can be used in other emotional expressions in a similar manner (Juslin & Sloboda, 2013). For instance, a fearful piece is generally characterised by minor mode, low sound level, fast tempo, and high pitch with a wide range and an ascending contour (Juslin & Laukka, 2003; Juslin & Sloboda, 2013). Similarly, a sad piece is also characterised by minor mode, low sound level, slow tempo, and low pitch, but the pitch range is generally of a narrower range with a descending contour (Juslin & Laukka, 2003; Juslin & Sloboda, 2013). Moreover, different features appear to be important for different emotions, namely timbre for anger, mode for sadness, with tempo being more powerful than others overall (Juslin & Lindström, 2010). It has been outlined that successful emotional decoding in music entails the identification and discrimination of tones, perception of melody, rhythm and harmony, comprehension and analysis of compositional structure and content (i.e., emotion), whereby different musical features are combined in an additive manner (Eerola et al., 2013; Juslin & Lindström, 2010; Umemoto, 1990). The recognition of emotions from music has been shown to occur with high agreement within and between listeners (Bigand et al., 2005; Vieillard et al., 2008), cross-culturally (Balkwill & Thompson, 2016; Fritz, 2009; Laukka et al., 2013), and even when the excerpts are short in duration (Peretz et al., 1998).

Musical processing has been demonstrated to be relatively preserved, if not enhanced, in autistic individuals compared to NT individuals (Molnar-Szakacs & Heaton, 2012; Ouimet et al., 2012). In addition, extensive research has consistently indicated that autistic individuals display interest in music (Allen et al., 2009a, 2009b; Bhatara, Quintin, et al., 2013; Brownell, 2002), which was first described in the case studies by Kanner (1943). Studies investigating the processing of emotions in music in ASD have adopted both the dimensional approach and the discrete emotion approach. Under the dimensional approach, investigations mainly focus on ratings of emotional valence (e.g., from negative to positive with a neutral midpoint) and emotional arousal (e.g., intensity) (Bradley & Lang, 2000; J. A. Russell, 2003). By contrast, investigations under the discrete emotion approach assess the categorisation of basic emotions as those mentioned in Section 1.2.1.2 for the different domains. Despite the varying approaches in assessing emotion processing of music in ASD, studies on instrumental music have consistently demonstrated no difference between autistic and NT individuals in both the ratings of emotional valence and categorisation of emotions (Caria et al., 2011; Gebauer, Skewes, Westphael, et al., 2014; Heaton et al., 1999; Järvinen et al., 2016), or specifically when verbal ability is statistically controlled for (Quintin et al., 2011). However, a study by Kopec, Hillier, and Frye (2014) found that autistic individuals rated songs with lyrics lower in negative emotions than NT individuals. Together, it appears that autistic individuals are able to accurately recognise emotions just as their NT counterparts with instrumental music, but perhaps not for songs. It is, however, unclear whether the discrepant results for songs are reflective of impairments with this specific form of musicality (i.e., singing), the combination of different musicalities (i.e., instrumental with singing), or the semantic content of the lyrics.

1.2.2.5. Summary

Autistic individuals appear to show impairment when recognising emotions in human faces and perhaps also in songs but not in nonhuman faces and instrumental music, with mixed findings presented for the recognition of speech prosody. Importantly, while substantial amount of research has been conducted on emotion processing in the human face and speech prosody domains in ASD, research in the nonhuman and music (particularly song) domains remains scarce. Given the relatively limited, yet mixed, findings, there is a need to first establish whether impairments in ASD are found for each domain. Moreover, no previous research has examined the emotion processing ability across all four domains (human faces, nonhuman faces, speech prosody, and music/song) in ASD within the same study. Such investigation would aid understanding of whether the emotion processing ability in ASD generalises across domains or if it is specific only to certain domains, while eliminating any confounding effects of variations in participant and experimental variables arisen from comparing findings from separate studies.

1.2.3. The role of development

Neuroconstructivism has been a particularly important theoretical framework for understanding neurodevelopmental disorders (Ansari & Karmiloff-Smith, 2002; H. D'Souza & Karmiloff-Smith, 2017; Karmiloff-Smith, 1998; Sirois et al., 2008). Under the neuroconstructivist view, the way in which the brain develops and constructs cognition is a self-organisation process resulting from interactions between multiple subsystems, such that intrinsic factors (e.g., physiological, psychological, neural) and extrinsic factors (e.g., information cues, social context) constrain one another and sculpt this developmental process. In line with this account, initial impairment in one cognitive component within the highly interactive brain is likely to have cascading effects on other parts of the developing system. In other words, a basic-level deficit in the cognitive system will constrain the emergence of several higher-level cognitive functions, because these functions emerge from complex interactions in the brain (H. D'Souza & Karmiloff-Smith, 2017). As a result of cascading effects and multilevel interactions, children with neurodevelopmental disorders are likely to develop atypical neural and cognitive trajectories with numerous widespread impairments (Bishop, 1997; Karmiloff-Smith, 1998), rather than a set of impaired and intact modules as proposed by the neuropsychological account (Butterworth, 2010; Leslie & Thaiss, 1992; see H. D'Souza & Karmiloff-Smith, 2017 for a discussion). Although constraints in one developmental domain may spill over to influence changes across other domains, one domain may also facilitate another by opening up new opportunities for growth (e.g., compensatory mechanisms).

Given the progressive changes emerging from complex interactions in the brain, understanding emotion processing in ASD requires not only describing differences between autistic and NT individuals, but also tracking and understanding how differences emerge over the developmental course through longitudinal and cross-sectional designs (M. S. C. Thomas et al., 2009). For example, the cross-sectional developmental trajectories approach has been employed to contrast functions that link task performance separately with the clinical versus nonclinical groups, which enables developmental change to be compared between these groups (M. S. C. Thomas et al., 2009). In the case of emotion processing in ASD, development of this ability has been compared between autistic and NT groups using different approaches, including exploring relationships between age and task performance separately for each diagnostic group (Gepner et al., 2001; Greimel et al., 2014), comparing performance between diagnostic groups by age group (Kuusikko et al., 2009), and examining the interaction between diagnostic group and age group on performance (Rump et al., 2009) – findings will be further discussed below.

In typical development, the ability to decode emotional expressions emerges at an early age and improves from childhood to adulthood for both accuracy and speed at both implicit and explicit levels of processing (De Sonneville et al., 2002; Herba & Phillips, 2004; Mathersul et al., 2009; Pons et al., 2004; L. M. Williams et al., 2009). Age-related improvements are observed across communicative domains, including human faces (Durand et al., 2007; L. A. Thomas et al., 2007; Widen & Russell, 2008), nonhuman faces (Santos et al., 2009), speech prosody (Chronaki et al., 2015, 2018), and music (Vidas et al., 2018). Failure to acquire early fundamental emotion processing skills could hamper a child’s social development, as the child may lose out on opportunities to learn about and fine-tune their skills for the understanding and interpretation of others’ emotional states with accuracy and speed (Izard et al., 2001).

Different emotions have been shown to follow different developmental courses across different domains in typical development. For human faces, happiness and sadness appear to be earlier-emerging emotion categories, followed by anger and fear (Durand et al., 2007; L. A.

Thomas et al., 2007; Widen & Russell, 2008). For nonhuman faces, the development of specific emotions has not been well-investigated, though one study found no differences in the recognition accuracy across anger, happiness, and sadness among NT children and adolescents aged 4-15 years (Rosset et al., 2008). Although no age effects could be drawn from this study, it is possible that recognition may develop in parallel for these three emotions in the nonhuman face domain (Rosset et al., 2008); this remains an open question to be further ascertained. For speech prosody, anger recognition was found to reach adult-level accuracy the earliest among other basic emotions (Chronaki et al., 2015, 2018). For music, recognition accuracy appears to reach adult-level performance at similar rates for all emotions (Vidas et al., 2018). In general, children achieve adult-level performance for prototypical expressions around 10-11 years of age across domains (Bruce et al., 2000; Chronaki et al., 2015; Mondloch et al., 2003; Tonks et al., 2007; Van Lancker et al., 1989; Vidas et al., 2018), with considerable improvement in the recognition of more subtle (i.e., less intense) expressions beyond this age (Chronaki et al., 2015; Rump et al., 2009).

The development of implicit emotion processing in ASD has not been well-explored, nevertheless, a recent study using magnetoencephalography (MEG) has shown maturational course of functional connectivity for implicit processing of emotional faces is altered in ASD compared to typical development (Safar et al., 2021). In terms of explicit emotion processing, emotion recognition has been found to improve less overtime in ASD compared to typical development for both human faces and speech prosody (Gepner et al., 2001; Greimel et al., 2014; Kuusikko et al., 2009; Uono et al., 2011; Van Lancker et al., 1989). Specifically, both children with ASD and NT children appear to perform comparably and show similar improvement (Rump et al., 2009; Van Lancker et al., 1989). However, children with ASD seem to show a lack of improvement in proficiency beyond that acquired by late childhood, while NT children continue to develop skills relevant for the decoding of more subtle expressions

(Greimel et al., 2014; Rump et al., 2009; Van Lancker et al., 1989). In regard to the nonhuman face and music domains, the developmental trajectory has not been previously examined in ASD. Some evidence, nevertheless, arises from separate studies involving different age groups, which showed parallel development in the ASD and NT groups: comparable performance between autistic and NT individuals has been reported in children, adolescents, and adults for music (Gebauer, Skewes, Westphael, et al., 2014; Heaton et al., 1999; Quintin et al., 2011), and in children and adolescents for nonhuman faces (Brosnan et al., 2015; Miyahara et al., 2007; Rosset et al., 2008).

Taken together, implicit and explicit emotion processing appear to undergo different developmental courses in ASD compared to typical development. These findings suggest that any differences may be particularly prominent among adults, as demonstrated in Rump et al. (2009) where autistic individuals do not develop further skills as do NT individuals beyond late childhood. Moreover, while Safar et al. (2021) provided important evidence for the altered development of implicit emotion processing in ASD relative to typical development on the neural level, the development of how emotional meaning is implicitly activated on the behavioural level in ASD relative to typical development is yet to be explored. Moreover, based on the sparse data from separate studies on different ages, it appears that the developmental trajectory of emotion recognition may differ between ASD and typical development as a function of communicative domain, while the role of domain in development at the implicit level remains unknown. Further research is, therefore, required to provide insights into these areas.

1.3. Plausible correlates of emotion processing and their associations with ASD

The perception of social information is thought to encompass both domain-general and domain-specific processes. While historically evidence supporting the two different processes had long been contrasted as a dichotomy, more recent perspectives proposed that both processes seemingly contribute to social perception in an integrative, complementary manner (Capozzi & Ristic, 2020; Michael & D'Ausilio, 2015). In support of the domain-general account, a series of studies using the dot perspective task have converged to question the unique role of domain-specific social cognitive processes (e.g., mentalising) in social perception. In the dot perspective task (Samson et al., 2010), participants viewed an image of a room with red discs displayed on the walls, with a human avatar facing one of the walls that display either all of the discs (such that the avatar saw the same number of discs as the participant) or only some of the discs (such that the avatar saw a different number of discs than the participant). The task required participants to count the number of discs they could see while ignoring the avatar. It was found that participants' performance was hampered when the number of discs they saw mismatched that was seen by the avatar, suggesting the avatar's visual perspective was automatically processed and interfered with participants' own perspective (Samson et al., 2010). In a follow-up study (Qureshi et al., 2010), participants performed an executive task concurrently with the dot perspective task. It was found that additional cognitive load brought by the executive task had increased the interference from the avatar's perspective, such that participants continued to perform an irrelevant calculation of the avatar's perspective while simultaneously performing the executive task. These results provided further evidence that perspective calculation occurs in a relatively automatic manner and operates independently of executive function (Qureshi et al., 2010). Although these results are seemingly suggestive of a dedicated, automatic domain-specific process for mentalising in social orientation, there is evidence that the interference effects were found to be equally achieved when the avatar (i.e., a social cue) was replaced with an arrow (i.e., a nonsocial cue) using the same paradigm (Santiesteban et al., 2014). Likewise, previous work by Galfano et al. (2012) also showed that automatic orienting of visuospatial attention was equally mediated by eye gaze (i.e., a social cue) and arrow (i.e., a nonsocial cue). The behavioural equivalence in responses to social and nonsocial cues appeared indicative of the operation of domain-general attentional mechanisms responding to stimulus directionality, rather than any dedicated mechanisms for tracking social information (e.g., mentalising). Nonetheless, it has been suggested that although the equivalence in behavioural response implicates domain-general functions underlying responses to both social and nonsocial cues, it does not necessarily rule out the contribution of domain-specific processes specific to social information in social orientation (Capozzi & Ristic, 2020; Michael & D'Ausilio, 2015).

On the contrary, there is ample evidence to motivate the speculation that domain-specific processes are involved in attentional shifts in response to social stimulus directionality. For example, a qualitative dissociation between the type of attentional orienting mechanisms engaged by eye gaze (i.e., a pure location-based cueing effect) and arrow (i.e., a pure object-based cueing effect) was revealed (Marotta et al., 2012; see also Marotta et al., 2018). This finding supported the notion that social cues appear to function differently from nonsocial cues, even though they may also engage a common attention process. In a separate line of research, the computation of mental states was shown to be involved in social orienting. For instance, using a naturalistic paradigm, it was found that interference of others' visual perspective was modulated by their visual access to the stimuli; that is, participants only adopted the others' visuospatial perspective when their vision was not obstructed by wearing opaque goggles (Freundlieb et al., 2017). Similarly, previous research showed that the magnitude of gaze following is modulated by the others' visual access (Teufel et al., 2010) or intentional stance of the others' gaze behaviour (Wiese et al., 2012), suggesting that eye gaze cues have stronger effects on attention when it signals social information (e.g., mental state or intention). Importantly, while these findings show that domain-specific processes (e.g., mentalising) contribute to social orienting, they do not disregard the involvement of domain-general processes; rather, they show that perceived social relevance of available cues influence attentional responses.

As such, the available evidence converges to the interplay between domain-general (e.g., attentional mechanisms) and domain-specific (e.g., mentalising) processes in social perception (see Capozzi & Ristic, 2020 and Michael & D'Ausilio, 2015 for a detailed discussion). That is, while both social and nonsocial directional cues inform domain-general attentional processes, ascription of social relevance appears to complement or enhance the magnitude of the responses to these cues. The postulation of which functionally specific social cognitive processes can be subserved at least in part by domain-general processes has important implications for understanding emotion processing in ASD. With regards to this, it is important to consider whether any emotion processing difficulties stem from impaired domain-general processes and/or whether autistic individuals draw on the same domain-general mechanisms as NT individuals during emotion processing. So far, the role of attention as a domain-general mechanism underlying social perception has been emphasised. In addition to attention (Wong et al., 2005), other domain-general processes such as perceptual strategies are also thought to play an important underlying role in emotion processing, which will be the focus of the present work.

Although autistic individuals may show intact emotion recognition under some circumstances, they appear to employ strategies differing from those used by NT individuals when processing emotional expressions (Hobson et al., 1988; Pelphrey et al., 2002; see Harms et al., 2010 for a review), and perhaps also with emotional nonhuman faces (Isomura, Ogawa, et al., 2014). Specifically, autistic individuals may have a preference for local, detail-based information processing (Baron-Cohen, 2008), which may underlie the atypical use of featural processing for faces in ASD (Gross, 2005; Pelphrey et al., 2002). As such, the relevant literature on cognitive processing style in ASD and how it may relate to differences in processing emotions within the visual modality will be introduced. Less is known about the strategies employed by autistic individuals for processing emotions in the auditory modality, and importantly, whether these strategies differ from those used by NT individuals. Given that the emotion processing in the auditory modality (i.e., speech prosody and music) is one of the key parts of this thesis, I will also review the existing findings on the potential correlates of the proficiency in extracting emotional cues through acoustic features in ASD and typical development, such as musical expertise and pitch perception abilities in this section. Finally, this thesis also takes into consideration the current debate on whether emotion recognition impairments in ASD may be better explained by co-occurring alexithymia (Bird & Cook, 2013). Thus, a brief overview of alexithymia and how it may relate to emotion recognition impairments in ASD will be given. Investigation into these potential correlates across modalities will provide a fuller picture of the underlying nature of emotion processing in autistic individuals.

1.3.1. Factors relating to visual emotion processing

1.3.1.1. Cognitive processing style

As introduced in Section 1.2.2.1, facial emotion recognition depends on both individual facial features and their configurations, with configural information playing a seemingly more prominent role than featural information in typical emotion recognition (Bombari et al., 2013; Calder et al., 2000; Derntl et al., 2009; Leppänen & Hietanen, 2007; McKelvie, 1995; Prkachin, 2003). Here, configural processing refers to the perception of the face as a whole and has been shown to rely on global facial features (Goffaux et al., 2005; Goffaux & Rossion, 2006). In typical development, recognition of emotions in human faces and schematic faces has been found to be poorer when stimuli are presented in an inverted orientation, due to the disrupted processing of facial features as a global unit (Derntl et al., 2009; Fallshore & Bartholow, 2003).

Given that successful processing and interpretation of facial expressions require a more global approach to information processing, it has been proposed that facial emotion recognition impairments in autistic individuals may be associated with their peculiar local information processing style (Baron-Cohen, 2008; Behrmann et al., 2006).

Studies have provided evidence for superior processing of local details in autistic individuals, such as faster identification of embedded figures, lower susceptibility to visual illusions, and enhanced performance on the block design task compared to NT individuals (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1993; see Simmons et al., 2009 for a review). This phenomenon has been primarily addressed by two theoretical accounts: the Weak Central Coherence account (WCC; Frith, 1989; Frith & Happé, 1994; Happé & Frith, 2006) and the Enhanced Perceptual Functioning account (EPF; Mottron et al., 2006). The WCC account, while not explicitly attempting to explain emotion recognition differences, posits that the local processing style in ASD results from a deficit in the higher-order processes that are responsible for the integration of separate components into a unified whole (U. Frith, 1989; U. Frith & Happé, 1994). In contrast, the EPF account proposes that the local processing style in ASD reflects an overdevelopment of low-level perceptual operations and that local processing is mandatory in ASD while global processing is more optional (Mottron et al., 2006). The two accounts are broadly similar in accounting for locally oriented processing but differ in their emphasis, with the WCC account emphasising that local processing results from a global deficit (U. Frith, 1989; U. Frith & Happé, 1994; Happé & Frith, 2006) and the EPF account emphasising that local processing is the default setting in perception in ASD without a global deficit (Mottron et al., 2006). Findings from research into the processing styles of autistic individuals have been inconsistent (Behrmann et al., 2006; Simmons et al., 2009; Van der Hallen et al., 2015). Nonetheless, previous meta-analytic work suggests that despite the limited evidence for an impairment, global processing appears to take longer and require more effort in autistic individuals compared to NT individuals (Van der Hallen et al., 2015). It has been suggested that local processing might be better conceptualised as an initial strategy employed for a specific task or in a specific context by autistic individuals (D. D'Souza et al., 2016; Van der Hallen et al., 2015).

Among the scarce research examining the relationship between facial emotion recognition and cognitive processing style, Gross (2005) found that in comparison to NT children, children with ASD had greater difficulty recognising facial emotional expressions, coupled with lower engagement in global processing on the global-local task employed. These findings demonstrate a relationship between cognitive processing style and emotion recognition ability in ASD. Conversely, findings from Oerlemans et al. (2013) suggested that processing style might not explain the significantly poorer performance in facial and prosodic emotion recognition in children with ASD, given that the cognitive processing style assessed on the feature identification task did not differ between children with ASD and NT children. Although a weak correlation between local processing style and facial emotion recognition was observed in the ASD proband, this correlation was not found in their sibling group nor NT group (Oerlemans et al., 2013). Furthermore, local processing style appeared not to be familial, as indicated by a non-significant sibling correlation, by which the authors concluded that processing style and social cognition might be relatively independent constructs within ASD (Oerlemans et al., 2013). Given the sparse and disparate literature, the relationship between cognitive processing style and social cognition remains unclear. Specifically, the question of whether group differences (if any) in the recognition of facial expressions can be explained by the different processing styles shown by autistic and NT individuals will be addressed in this thesis.

1.3.2. Factors relating to auditory emotion processing

1.3.2.1. Musical expertise

The effects of musical training on auditory emotion recognition have been repeatedly documented in typical development. Within the musical domain, research has indicated that the length of musical training is positively correlated with accuracy of emotion recognition from instrumental musical excerpts (Castro & Lima, 2014; Lima & Castro, 2011a; Livingstone et al., 2010). Moreover, musical expertise has also been shown to be associated with cross-domain benefits to speech prosody. For example, musicians showed better accuracy than non-musicians when recognising emotions from speech prosody for both single and multiagent stimuli (Correia et al., 2020; Farmer, Jicol, & Petrini, 2020; Lima & Castro, 2011b; Thompson, Schellenberg, & Husain, 2004; see Martins, Pinheiro, & Lima, 2021 for a review). This facilitatory effect has been demonstrated across age groups (Lima & Castro, 2011b; W. F. Thompson et al., 2004). Specifically, not only adult musicians performed better than adult non-musicians at recognising emotions from speech prosody, 6-year-old children who received keyboard lessons on a weekly basis also recognised emotions from speech prosody with greater accuracy compared to children of the same age who received no training in the previous year (W. F. Thompson et al., 2004). The effects of musical training have been shown to be limited to the auditory modality only, such that facilitatory effects do not seem to extend to facial and audio-visual emotion recognition (Correia et al., 2020; Farmer et al., 2020; see Martins, Pinheiro, & Lima, 2021 for a review). Together, these findings suggest that learning music improves emotion recognition in music, which could be transferred to speech prosody.

In addition to the effects of musical training, Correia and colleagues (2020) also examined whether naturally good musical abilities (i.e., in the absence of training) were related to enhancements in emotion recognition. There was a positive association between music perception abilities and emotion recognition in vocalisations (e.g., laughter, crying) and speech prosody, which was the case even when musical training was held constant (Correia et al., 2020). Importantly, musically-untrained participants with good musical abilities showed similar emotion recognition to that of musically-trained participants, who were found to exhibit better emotion recognition skills for vocalisations and speech prosody (Correia et al., 2020). These findings suggest that untrained participants with good musical perception abilities recognised vocal emotions just as well as highly trained musicians (Correia et al., 2020). Thus, musical perception abilities, aside from musical training, may also be an important correlate of auditory emotion processing to be considered.

While no studies to date have examined the effects of musical perception abilities on auditory emotion recognition in autistic individuals, the effects of musical training have been investigated in one study (Quintin et al., 2011). In this study, it was found that musical training was not a significant predictor of emotion recognition from instrumental musical excerpts, both across and within the ASD and NT groups (Quintin et al., 2011). However, the impact of musical training was investigated by dividing the sample into musician versus non-musician groups based on the criteria of having played at least one musical instrument and received at least two years of musical training (Quintin et al., 2011). Although information about the mean number of years of musical training in the musician and non-musician groups was not available, the requirement of two years of musical training for the musician group is relatively low in comparison with other studies investigating the effects of musical training (e.g., ≥ 13 years; Marques et al., 2007; Parbery-Clark et al., 2009; Schön et al., 2004; W. F. Thompson et al., 2004). Altogether, it would be worthwhile to further examine the effects of musical training, as well as musical perception abilities, on emotion recognition in music. Given the cross-domain benefits to emotional prosody seen in NT individuals, it would also be of interest to explore if this is also the case in autistic individuals. Finally, whether the effects of musical training and perception abilities are evident at both implicit and explicit levels across the speech prosody and music domains in the two groups also remains to be explored.

1.3.2.2. Low-level pitch perception

As mentioned earlier, cross-domain benefits of musical expertise appear to be limited to the auditory modality, which likely occur due to enhancements in the low-level perception of acoustic features (Kannyo & DeLong, 2011; Mankel et al., 2020). Here, I focus on the processing of pitch information, as there is ample evidence demonstrating the importance of pitch as a cue to emotional meaning in speech (Banse & Scherer, 1996; Juslin & Laukka, 2003; Pell & Kotz, 2011) and in both instrumental and vocal music (Hakanpää et al., 2019b, 2019a; Juslin & Laukka, 2003; Quinto et al., 2014; Schellenberg et al., 2000). Thus, the ability to detect changes in pitch is thought to be crucial for the categorisation of emotional types.

In typical development, pitch processing ability is an important mechanism supporting emotional prosody recognition (Globerson et al., 2013, 2015). Specifically, psychoacoustic thresholds obtained from pitch tasks that require differentiating and naming pitches (e.g., high vs. low, glide vs. non-glide, ascending vs. descending), but not pitch tasks that require same-different judgments, were found to be significant correlates of explicit emotional prosody recognition (Globerson et al., 2013, 2015). The underlying role of pitch processing ability in implicit processing of emotional prosody, however, has not been explored.

The relationship between pitch perception ability and emotion recognition in music, however, has not been well-established, as reflected in the lack of association between the two variables in previous studies (Lévêque et al., 2018; L. Zhou et al., 2019). Notably, in Lévêque et al. (2018), the threshold data were obtained from simple same-different pitch discrimination (rather than pitch tasks that require differentiating and naming pitches, c.f., Globerson et al., 2013, 2015). It is possible that the simple same-different pitch task may not have an intrinsic relationship with auditory emotion processing, resulting in a lack of correlation observed, as noted in the speech domain (Globerson et al., 2013, 2015). Although Zhou et al. (2019) correlated thresholds obtained from pitch tasks that required differentiating and naming pitches and performance on emotion recognition in music, this was done separately for the amusic and control groups. The small sample size in each group (N ≤ 16) may have resulted in insufficient power to detect this correlation. Together, the question of whether pitch perception (indexed by pitch discrimination and naming) serves as an underlying mechanism of emotion processing of music remains open and needs to be further elucidated.

A number of studies have demonstrated intact or even enhanced pitch processing abilities in autistic individuals (Bonnel et al., 2003; Chowdhury et al., 2017; Globerson et al., 2015; Heaton et al., 2001; Jones et al., 2009; see also Haesen et al., 2011a; Kellerman et al., 2005; Ouimet et al., 2012 for reviews), although impaired pitch discrimination has also been reported (Bhatara, Babikian, et al., 2013; Jiang et al., 2015; Kargas et al., 2015; Schelinski & von Kriegstein, 2019; Sota et al., 2018). The few studies investigating the relationship between pitch perception abilities and prosodic emotion recognition have provided mixed results. For instance, Globerson et al. (2015) showed that high-low pitch discrimination and naming abilities for non-vocal stimuli were strongly associated with prosodic emotion recognition in both ASD and NT groups, with a more pronounced association found in the ASD group. Based on these findings, the two groups seemed to have employed similar processing strategies for extracting pitch information in emotional prosody, despite poorer emotion recognition observed in the ASD group (Globerson et al., 2015). By contrast, Schelinski & von Kriegstein (2019) found that whereas high-low pitch discrimination and naming abilities for vocal stimuli were associated with prosodic emotion recognition in the NT group, no such association was found in the ASD group. It was postulated that vocal pitch information was perhaps not available for prosodic emotion recognition in autistic individuals like it was in NT individuals

(Schelinski & von Kriegstein, 2019). Again, the underlying role of pitch processing ability in implicit processing of emotional prosody is yet to be examined. Furthermore, no studies have previously investigated the association between pitch processing ability and emotion processing of music across both implicit and explicit levels in ASD. Taken together, further investigation into these relationships will contribute to the better understanding of whether and how pitch processing underlies implicit and explicit emotion processing of speech prosody and music in both ASD and typical development.

1.3.3. The case for alexithymia in ASD research

Alexithymia is a subclinical cognitive-affective disturbance characterised by impairments in the experience, regulation, and communication of emotions and is best conceptualised as a dimensional personality trait (Connelly & Denney, 2007; Nemiah & Sifneos, 1970; J. D. A. Parker et al., 2008; Suslow & Donges, 2017; G. J. Taylor & Bagby, 2000). The alexithymia construct was first introduced to characterise emotional deficits of psychosomatic patients, who were typically presented with an incapacity to verbalise their emotions (Ruesch, 1948; Sifneos, 1973). This difficulty is thought to be related to patients' unawareness of feelings and emotions, as well as confusions about emotions and bodily sensations (Poquérusse et al., 2018). There is currently no formal diagnosis for alexithymia, although several self-report measures can help to identify its signs, such as the 20-Item Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994) and the Bermond-Vorst Alexithymia Questionnaire (BVAQ; Vorst & Bermond, 2001).

The prevalence rate of alexithymia is approximately 10% of the general population, which has remained stable over time since the first demographic assessment conducted in Finland (Hiirola et al., 2017; Salminen et al., 1999). The prevalence of alexithymia is higher in males than in females, with a ratio reported as affecting twice as many males as females (Salminen et al., 1999). Alexithymia has been found to be present in a range of neuropsychiatric disorders, including post-traumatic stress disorder (Frewen, Dozois, et al., 2008; Frewen, Lanius, et al., 2008), depression and anxiety disorders (Marchesi et al., 2000), eating disorders (Westwood et al., 2017), and schizophrenia (van ’t Wout et al., 2007). In particular, it has been suggested that 50% of autistic individuals have alexithymia (Berthoz & Hill, 2005; Milosavljevic et al., 2016).

The most recent contributions to the debate of whether autistic individuals have atypical emotion recognition suggested that alexithymia, but not ASD per se, may account for individuals’ emotion recognition difficulties (Bird & Cook, 2013; Cook et al., 2013; Kinnaird et al., 2019). Specifically, Bird & Cook (2013) argued that the heterogeneity in the emotional competence within the ASD population can partly be attributed to the high co-occurring rate of alexithymia in autistic individuals, rather than being a symptom of ASD per se. This is because alexithymia has also been found to be associated with emotion perception, behaviourally and neurologically, in the general population (Gündel et al., 2004; Jessimer & Markham, 1997; Prkachin et al., 2009). The contribution of alexithymia to the better understanding of emotion processing in ASD has gained tentative traction in recent years (see Sivathasan et al., 2020 for a recent review), with studies predominantly assessing its contribution to the processing of facial expressions (Brewer et al., 2016; Keating et al., 2021; Stephenson et al., 2019), but less so with prosodic (Heaton et al., 2012) and musical expressions (Allen et al., 2013). Given the high incidence of alexithymia in ASD as well as its plausible role in explaining emotion recognition difficulties, it is important that the contribution of alexithymia is taken into consideration when assessing the multiple facets of emotion processing ability in autistic individuals.

1.4. Aims and rationale

The overarching aim of this thesis was to elucidate whether differences in implicit and explicit processing of emotions between autistic and NT individuals were generalised across communicative domains (i.e., domain-general) or specific to certain domain(s) (i.e., domain-specific), in order to enhance understanding of the strengths and weaknesses of emotion processing in autistic individuals. This will be assessed in individuals across the age span, including children, adolescents, and adults, to shed light on the developmental trajectory of emotion processing in ASD compared to typical development. A secondary aim of this thesis is to examine the strategies employed for emotion processing in autistic and NT individuals, as well as the contribution of alexithymia to emotion processing. The research questions are:

1. Does emotion processing ability generalise across communicative domains (human face, nonhuman face, speech prosody, and music/song) in ASD?
2. Does emotion processing ability generalise across implicit and explicit levels of processing in ASD?
3. Does the developmental trajectory of emotion processing differ between autistic and NT individuals?
4. Is there a relationship between emotion processing and several related factors (cognitive processing style, musical training, musical perception, pitch perception, alexithymic trait, and autistic trait) in ASD and typical development?

1.4.1. Structure of thesis

Given the highly variable findings regarding emotion recognition in ASD as a whole, the objective of Chapter 2 was to consolidate the expanding literature on this topic through a systematic review and meta-analysis. Contradictory findings are not uncommon in ASD research, potentially owing to inadequate statistical power and intrinsic heterogeneity in the diagnosis itself. Thus, combining findings from multiple small studies provides an opportunity to investigate this topic with increased power. Although previous meta-analytic works have provided strong evidence for the presence of facial emotion recognition impairments in ASD (Lozier et al., 2014; Uljarevic & Hamilton, 2013), there remains a lack of consensus about the general emotion recognition ability in ASD when considering evidence from the growing literature on other domains (as presented in Section 1.2.2). I investigated in this review whether emotion recognition of the six basic emotions is impaired both in terms of recognition accuracy and speed in ASD based on data not only for the human face domain but also the other less examined domains (i.e., nonhuman face, speech prosody, and music). In addition, I examined the effects of population characteristics (i.e., age, verbal IQ, nonverbal IQ, and full-scale IQ), experimental design parameters (i.e., stimulus domain, task demand), and study quality factors (i.e., verbal, nonverbal, and full-scale IQ matching) on group differences seen across different studies.

Chapter 3 was an experimental study examining explicit emotion processing via the widely used forced-choice paradigm, where participants selected an emotional label that best described the emotion conveyed by the stimulus from a set of alternatives. The purpose of this study was to investigate emotion recognition accuracy, speed, and efficiency for four basic emotions (anger, fear, happiness, and sadness) across visual and auditory stimuli (human faces, face-like objects, speech prosody, and song) between autistic and NT individuals ranging from children to adults. This study was the first of its kind to directly compare performance in the four communicative domains using the same participant samples that were matched on age, gender, verbal ability, and nonverbal ability under the same experimental paradigm. The findings of this study corresponded to the issues and recommendations raised in the review in Chapter 2, by contributing to the under-researched topics in the literature on emotion recognition in ASD across communicative domains and over the developmental course.

Chapter 4 investigated implicit emotion processing from a behavioural perspective through employing a cross-modal affective priming paradigm. Considering reports of facial expressions serving a more dominant communicative domain in multimodal processing of emotions (Collignon et al., 2008; Lin & Ding, 2019), a particular interest was to examine how implicitly processed emotions conveyed by auditory stimuli influence emotional judgment in visual stimuli. Implicit emotion priming of auditory cues (speech prosody and song) on emotional judgment in visual targets (human faces and face-like objects), to my knowledge, has not been directly compared between autistic and NT individuals across the age span. The findings of this study, therefore, provided insights into the impact of implicit activation of emotional meaning of different expressions on subsequent emotional behaviours and whether this impact was moderated by domain in the two populations.

In Chapter 5, several potential correlates of emotion processing were examined in ASD and typical development. First, the relationship between cognitive processing style (assessed on the Navon task: Navon, 1977) and emotion recognition from faces and objects (assessed in Chapter 3) was examined. This investigation aimed to explore whether any emotion recognition differences across visual stimuli between autistic and NT individuals were related to differences in their perceptual processing styles. Secondly, whether emotion processing of speech prosody and song (assessed in Chapters 3 and 4) were related to years of musical training, musical perception ability (assessed on the Montreal Battery of Evaluation of Amusia: Peretz et al., 2003; or the Montreal Battery of Evaluation of Musical Abilities: Peretz et al., 2013), and pitch perception abilities (assessed on a pitch direction discrimination task: F. Liu et al., 2010, 2012). This investigation aimed to determine whether proficiency in extracting auditory perceptual information, particularly pitch changes, was related to emotion processing of auditory stimuli in autistic and NT individuals. Finally, the contribution of alexithymic traits (assessed on the TAS-20: Bagby et al., 1994) and autistic traits (assessed on the Autism-

Spectrum Quotient: Baron-Cohen et al., 2001) to overall performance on emotion priming (assessed in Chapter 4) and recognition (assessed in Chapter 3) was examined. This investigation added to the current debate on whether co-occurring alexithymia could account for emotion recognition impairments in ASD (Bird & Cook, 2013), while extending beyond the human face domain that has predominantly been investigated in previous studies.

Chapter 6 brought the aforementioned chapters together through evaluating, discussing and interpreting the overall results in the context of previous research in ASD. The implications of these findings and their contributions to the field of emotion processing in ASD were then reviewed, followed by a discussion of the limitations of the research presented in this thesis and suggestions for future research.

Chapter 2: A systematic review and meta-analysis of emotion recognition across nonverbal communicative domains in ASD

This chapter examines the emotion recognition ability across domains and modalities in individuals with autism spectrum disorder (ASD) relative to neurotypical (NT) individuals, through a systematic review of the interdisciplinary literature, and through a meta-analysis. The moderating effects of age, full-scale, verbal, and nonverbal IQ, stimulus domain, and task demand on emotion recognition impairments in ASD were assessed. Any shortcomings and gaps in research arisen from the literature would inform the design of studies in the subsequent chapters.

2.1. Introduction

An expanding literature has investigated emotion recognition in autism spectrum disorder (ASD). Findings, however, have been highly variable. Although previous meta-analytic works have provided strong evidence for the presence of emotion recognition impairments in individuals with autism spectrum disorder (ASD), this was limited to the visual modality, specifically human faces and body gestures (Lozier et al., 2014; Uljarevic & Hamilton, 2013). There remains a lack of consensus about the general emotion recognition ability in ASD when considering evidence from the growing literature on other stimulus domains, such as those involving nonhuman faces (e.g., Brosnan et al., 2015; Miyahara et al., 2007; Rosset et al., 2008), speech prosody (e.g., Baker et al., 2010; Heikkinen et al., 2010; Schelinski & von Kriegstein, 2019), and music (e.g., Järvinen et al., 2016; Quintin et al., 2011). In addition, questions about whether impairments vary by age, IQ, stimulus domain, task demand, and emotion have also been raised previously (Harms et al., 2010; Nuske et al., 2013). Thus, this review sought to provide comprehensive answers to (i) whether the ability to recognise emotions differs between autistic and neurotypical (NT) individuals when data available in the different domains are incorporated, and (ii) whether factors such as age, IQ, stimulus domain, and task demand moderate any identified impairments, by systematically examining the interdisciplinary literature and by using meta-analysis.

2.1.1. Past findings of emotion processing in ASD

2.1.1.1. Stimulus domain

As introduced in Section 1.2.2, within the visual modality, difficulty in identifying and understanding emotions through facial expressions is widely recognised as the most common social-cognitive impairments in ASD (see Harms et al., 2010 for a discussion). Nevertheless, there has also been contradictory evidence suggesting that this ability is intact (Jones et al., 2011; Tracy et al., 2011). In contrast, the ability to recognise emotions from nonhuman facial stimuli such as cartoons, caricatures, and schematic faces have been consistently reported to be intact in ASD (Brosnan et al., 2015; Isomura, Ogawa, et al., 2014; Miyahara et al., 2007; Rosset et al., 2008). It has been postulated that the greater interest towards such stimuli may have led to greater acquired expertise with processing the faces of these stimuli, and hence resulting in the comparable performance in emotion recognition between autistic and NT individuals (Rosset et al., 2008). This notion is further supported by previous evidence demonstrating a unique link between autistic individuals and their restricted interest in cartoons (Anthony et al., 2013; Grelotti et al., 2005; Kuo et al., 2014; Spiker et al., 2012).

Within the auditory modality, studies investigating emotional prosody in ASD have yielded mixed findings, with some reporting clear impairments (Doi et al., 2013; Schelinski & von Kriegstein, 2019; L. J. Taylor et al., 2015) and others reporting no impairments (Baker et al., 2010; Heikkinen et al., 2010). These discrepancies may in part be explained by variations in the methodologies used across studies, such as presenting prosodic stimuli that elicit emotions (in)congruent to its verbal content or in neutral semantics (J. E. Wang & Tsao, 2015), implementing low-pass filtering methods that eliminate verbal content (R. B. Grossman et al., 2010), and presenting emotional prosody in the presence of contextual cues (Le Sourn-Bissaoui et al., 2013). Despite that emotions in music and speech prosody are expressed using similar cues (Coutinho & Dibben, 2013; Juslin & Laukka, 2003), the ability to recognise emotions from music seems unimpaired in ASD (Gebauer, Skewes, Westphael, et al., 2014; Heaton et al., 1999), or specifically when verbal IQ was statistically controlled for (Quintin et al., 2011). This is not surprising given that music has been widely documented as a domain of preserved abilities in autistic individuals (Heaton, 2009; Mottron et al., 2000; Quintin et al., 2013). Taken together, it appears that the emotion recognition ability of autistic individuals may be moderated by stimulus domain, which could be considered a potential contributing factor to the mixed results in the literature.

2.1.1.2. Specific emotions

Alongside studies showing general impairments across all emotions in ASD (e.g., Lindner & Rosén, 2006; Sawyer et al., 2012; Y. Song et al., 2020), a body of work suggests that emotion recognition impairments in ASD may be specific to certain emotions. In the face domain, this is supported by numerous studies reporting specific impairments in the recognition of negative emotions, particularly for fearful expressions (Ashwin et al., 2006; Pelphrey et al., 2002; Tell et al., 2014; Uono et al., 2011, 2013; S. Wallace et al., 2008) and sad expressions (Ashwin et al., 2006; Boraston et al., 2007). According to the amygdala theory of autism, the poor recognition of fear and other negative emotions has been postulated as a result of amygdala dysfunction in ASD (Ashwin et al., 2006; Baron-Cohen et al., 2000; Howard et al., 2000). However, in the speech prosody domain, emotion-specific impairments have been noted in the recognition of happiness, a positive emotion (Hubbard et al., 2017; J. E. Wang & Tsao, 2015). Moreover, no general or specific impairments have been reported for nonhuman faces

(Brosnan et al., 2015; Rosset et al., 2008) nor music (Heaton et al., 1999; Järvinen et al., 2016; Quintin et al., 2011). Thus, the claim of specific impairments for negative emotions does not seem to be supported by findings in the nonhuman face, speech prosody, and music domains.

The different emotion-specific impairments observed might be attributed to the varying difficulty levels of identifying different expressions across domains. Certain emotions are thought to be more robustly expressed in one domain compared to the other in typical development (Keltner et al., 2016). For instance, happiness was found to be the most accurately recognised emotion from human faces, followed by anger and sadness, with fear being the least accurately recognised (Elfenbein & Ambady, 2002). Conversely, across the speech prosody and music domains, anger and sadness were found to be better recognised than fear and happiness (Chronaki et al., 2018; Elfenbein & Ambady, 2002; Juslin & Laukka, 2003). With these in mind, if emotion-specific difficulties in ASD were present, it is plausible that they may differ across domains given their distinctive perceptual differences. The current review will examine this possibility by analysing each emotion type separately across and within domains, as well as for the following moderating factors.

2.1.1.3. Age

Research on emotion has predominantly focused on six basic emotions, including anger, fear, happiness, sadness, disgust, and surprise (Ekman & Friesen, 1976; Prinz, 2004). The ability to explicitly label these emotional expressions emerges during early childhood and develops through adolescence to adulthood (De Sonneville et al., 2002; Herba & Phillips, 2004). As noted in Section 1.2.3, several cross-sectional studies involving participants from different age groups provide some evidence that emotion recognition improves less over time in ASD than in typical development (Gepner et al., 2001; Kuusikko et al., 2009; Van Lancker et al., 1989). Specifically, children in both ASD and NT groups show similar improvement and perform comparably on emotion recognition (Rump et al., 2009; Van Lancker et al., 1989). However, adolescents with ASD do not seem to develop proficiency beyond those acquired by late childhood, while NT adolescents continue to refine emotion recognition skills (Greimel et al., 2014; Rump et al., 2009; Van Lancker et al., 1989). Consequently, adults with ASD do not reach the level of proficiency demonstrated by NT adults (Rump et al., 2009). The different developmental trajectories between ASD and NT groups may have, therefore, contributed to the inconsistent group differences observed in the literature. The evidence reviewed here came exclusively from the human face and speech prosody domains, where no studies have directly examined the developmental trajectory of emotion recognition in the nonhuman face and music domains in ASD. Separate studies representing performance by different age groups, nonetheless, provide some insights into this: studies investigating emotion recognition of nonhuman faces have indicated intact performance in both children and adolescents with ASD (Brosnan et al., 2015; Miyahara et al., 2007; Rosset et al., 2008), and the same for music across children, adolescents, and adults with ASD (Heaton et al., 1999; Järvinen et al., 2016; Quintin et al., 2011). These findings, by contrast, appear to suggest that age did not moderate emotion recognition ability of autistic individuals. Considering the contradictory observations discussed, it is necessary to evaluate whether the moderating effects of age could account for the heterogeneity in the literature.

2.1.1.4. IQ

Emotion recognition encompasses both verbal and nonverbal intellectual functioning: while nonverbal ability enables the perception and integration of nonverbal characteristics that differentiate emotion expressions, verbal ability allows the interpretation and assignment of an appropriate emotional label to their meaning (Davitz & Beldoch, 1964). Indeed, intellectual ability has not only been identified as a significant correlate of emotion recognition ability but has also shown significant predictive power in typical development (Khawar et al., 2013). Previous studies indicated that the correlation between IQ and emotion recognition also extends to ASD, with IQ being an important predictor of emotion recognition ability regardless of diagnosis (full-scale IQ: Jones et al., 2011; nonverbal IQ: Salomone et al., 2019). However, it has also been shown that the association between IQ and emotion recognition is significantly more prominent in the ASD group compared to the NT group (verbal IQ: Dyck et al., 2006a), or that this association is uniquely present in the ASD group only but not in the NT group (full-scale IQ: Tanaka et al., 2012; verbal IQ: Atkinson, 2009; Quintin et al., 2011; Wallace et al., 2008). These findings suggest that emotion recognition may involve higher-level analytical processes that autistic individuals could employ to aid performance, whereas NT individuals may opt for intuitive rather than analytical strategies (J. B. Grossman et al., 2000). In particular, verbal IQ may contribute to the intact performance of the ASD group in studies requiring explicit knowledge about emotional labels (see Trevisan & Birmingham, 2016 for a review). IQ may, therefore, constitute a compensatory mechanism for emotion recognition especially in autistic individuals who have higher cognitive functioning (Harms et al., 2010), which may explain individual differences in emotion recognition within ASD (see Nuske et al., 2013 for a review). In this sense, autistic individuals who have higher IQ may be relatively less impaired in emotion recognition when compared against their NT counterparts, i.e., higher IQ of the ASD group may be associated with smaller/no group differences.

Nonetheless, contradictory findings have also been reported in previous studies. Rommelse et al. (2015) examined three groups of ASD and NT participants with below average, average, or above average IQ, and found that indeed in absolute terms, autistic individuals who had below average IQ performed worse than those who had higher IQ on facial and prosodic emotion recognition tasks. However, in relative terms, a larger group difference in performance on these tasks was found between the ASD and NT groups with above average

IQ than the groups with lower IQ (Rommelse et al., 2015). These findings suggest that autistic individuals who have higher IQ may be more severely impaired in emotion recognition when compared against their NT counterparts, i.e., higher IQ of the ASD group may be associated with larger group differences. Furthermore, in other studies, IQ has been shown to be unrelated to the ability to recognise emotions in autistic individuals (full-scale IQ: Heaton et al., 2012; Parron et al., 2008), discounting the potential moderating role of IQ on group differences in emotion recognition. Given the disparate findings regarding the role of IQ in moderating performance by autistic individuals relative to their NT counterparts, the present study sought to examine whether the different measures of IQ (full-scale IQ, verbal IQ, and nonverbal IQ) could account for the heterogeneity among results across studies.

2.1.1.5. Task demands

The majority of emotion recognition studies in ASD implemented a verbal task, which involved identifying, labelling, discriminating, matching, or detecting different emotions using verbal cues. Such tasks are cognitively demanding as they place great reliance on emotion vocabulary necessary for labelling emotions following perception (Palermo et al., 2013). As briefly outlined in Section 1.2.1.2, findings from studies employing a verbal task have been inconsistent. Implementing a forced-choice task that involves participants selecting their response to the target emotion from pre-generated options, some studies reported significant group differences (e.g., Griffiths et al., 2019; Philip et al., 2010), while others did not (e.g., Baker et al., 2010; Rhodes et al., 2018). Studies using an open-ended labelling task, where participants spontaneously generate an emotional label that describes the target emotion, have similarly reported inconsistent findings (significant group difference: Boucher et al., 2000; Castelli, 2005; no group difference: Hobson et al., 1989). Verbal discrimination tasks, requiring participants to judge whether a stimulus displays the same or different emotion than the given label, have consistently reported group differences (e.g., Oerlemans et al., 2014; Waddington et al., 2018). Verbal matching tasks involving participants selecting a stimulus among a set that matches the given label have also reported significant group differences (e.g., Loth et al., 2018). However, verbal detection tasks which require participants to detect the stimulus displaying the target emotion among other distractors, have found no group differences (e.g., Shafritz et al., 2015).

In contrast to verbal tasks, nonverbal emotion recognition tasks tap only into the perceptual stage where emotions are discriminated based on perceptual properties alone (Adolphs, 2002; Palermo et al., 2013). Despite the seemingly reduced task demands, findings have also been inconsistent. Nonverbal discrimination tasks, requiring participants to indicate whether the stimuli within a pair display the same or different emotions, have reported group differences with human faces (e.g., Greimel et al., 2014; Sasson et al., 2016; Vannetzel et al., 2011), but not with speech prosody (e.g., Lindström et al., 2018). Nonverbal matching tasks, where participants are shown an emotional stimulus and then choose a stimulus from a set that displays the same expression, have found group differences (e.g., Philip et al., 2010; Tanaka et al., 2012). By contrast, nonverbal detection tasks, which involve participants detecting the odd emotion expression among other distractors, have found no group differences (Isomura, Ogawa, et al., 2014; Kujala et al., 2005). Upon these mixed results, the moderating role of task demands in accounting for the heterogeneity in the literature is yet to be elucidated.

2.1.2. Prior reviews of emotion processing in ASD

To date, three systematic reviews have investigated emotion recognition in ASD (Harms et al., 2010; Lozier et al., 2014; Uljarevic & Hamilton, 2013). The two meta-analyses, bringing together 48 papers on facial and body expressions (Uljarevic & Hamilton, 2013) and 43 papers on facial expressions (Lozier et al., 2014), revealed general emotion recognition impairments in ASD with varying severity across emotions. Age did not moderate emotion recognition in Uljarevic and Hamilton (2013) but moderated recognition of fear, sadness, and disgust in Lozier et al. (2014). The magnitude of impairments could not be accounted for by either full-scale IQ (Lozier et al., 2014; Uljarevic & Hamilton, 2013) or verbal IQ (Lozier et al., 2014). The effect of task on emotion recognition was examined in Uljarevic and Hamilton (2013) and was not found to be a significant moderator; that is, there appeared no overall differences in performance between emotion matching (i.e., requiring perceptual demands) and emotion labelling (i.e., requiring verbal and perceptual demands) tasks. To my knowledge, no reviews have yet evaluated emotion recognition in ASD across domains (human faces, nonhuman faces, speech prosody, and music) and modalities (visual and auditory). In addition, the effects of domain have not been systematically examined, especially when impairments in the human face and speech prosody domains do not seem to generalise to the nonhuman face nor to the music domain. Reviewing the growing body of literature, particularly with the increased attention to emotion recognition in the auditory domain during the last decade (Järvinen et al., 2016; Schelinski & von Kriegstein, 2019; L. J. Taylor et al., 2015; Waddington et al., 2018; J. E. Wang & Tsao, 2015), may therefore further our understanding of the general emotion recognition ability in ASD as a whole.

2.1.3. Aims and purpose

This systematic review and meta-analysis aimed to investigate whether the ability to accurately recognise basic emotions in the different domains across visual (human and nonhuman face) and auditory modalities (speech prosody and music) in autistic individuals differs from that in NT individuals (see Table 2.1 for a detailed summary of the similarities and differences between prior reviews and the present study). The possible influence of study quality factors on these observed group differences were also examined. One specific focus was on the influence of IQ matching, given that group differences in emotion recognition could be confounded by differences in IQ rather than diagnosis per se (see Harms et al., 2010 for a detailed discussion). To identify factors contributing to the mixed results, the present work investigated several potential moderators, including age, full-scale/verbal/nonverbal IQ, stimulus domain, and task demand. In addition to recognition accuracy, response times (where available) were also examined separately in the main meta-analyses to address the discrepancies in the literature that show either slower (Greimel et al., 2014; Ketelaars et al., 2016; Lindström et al., 2018; Sawyer et al., 2012) or comparable (Akechi et al., 2010; Fink et al., 2014; J. B. Grossman et al., 2000) speed of emotion recognition in ASD relative to typical development. Specific review questions include:

1. Is there evidence for a general emotion recognition impairment across emotions in ASD?
2. Are these results influenced by IQ matching procedures?
3. Is there evidence for emotion recognition impairments across ages, full-scale/verbal/nonverbal IQ, domains, or tasks in ASD?

Table 2. 1. Summary of similarities and differences between prior reviews and the current study.

	Harms et al. (2010)	Uljarevic & Hamilton (2013)	Lozier et al. (2014)	The present study
Data synthesis approach				
Systematic review	✓	✓	✓	✓
Meta-analysis		✓	✓	✓
Quality assessment				
Critical appraisal of studies				✓
Assessment of publication bias		✓	✓	✓
Domain(s) covered				
Human faces	✓	✓	✓	✓
Nonhuman faces	✓	✓	✓	✓
Body expressions		✓		
Speech prosody				✓
Music				✓
Moderator(s) covered				
Age	✓	✓	✓	✓
Full-scale IQ		✓	✓	✓
Verbal IQ			✓	✓
Nonverbal IQ				✓
Task selection	✓	✓		✓
Study quality factor(s) covered				
IQ-matching protocol	✓			✓
Performance index				
Accuracy	✓	✓	✓	✓
Response time	✓			✓

Note. A ✓ indicates a particular item was assessed or conducted.

2.2. Method

This study was carried out in accordance with the recommended procedures for conducting systematic reviews and meta-analyses (i.e., the Preferred Reporting Items for Systematic reviews and Meta-Analyses [PRISMA] guidelines ; Moher et al., 2009). The protocol for this systematic review was registered on PROSPERO at https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019091703.

2.2.1. Search strategy

To identify all eligible studies, a comprehensive search was conducted using Embase, Medline, PubMed, Web of Science, and Google Scholar, from the earliest record to August 2019. The search strategy was designed to combine ASD and emotion recognition relating to the domains of interest using three blocks of search terms: (a) synonyms and terms used to describe ASD; (b) emotion recognition; and (c) human face, nonhuman face, speech prosody, and music (see Table 2.2 for full details). The initial search was completed on 27th April 2018; additional studies were included until 1st August 2019 via an updated search.

Table 2. 2. Search terms and synonyms (truncated where possible) used in database search.

Term Block	PsycINFO, Medline, PubMed, Embase	Google Scholar
Participant	Autis* OR Asperger* OR ASD OR pervasive developmental disorder	Autism OR Asperger
Emotion	Emotion* OR affect OR social OR communicat* OR happy OR happiness OR angry OR anger OR sad OR sadness OR fear OR fearful OR frightened OR scared OR threat* OR disgust OR surprise*	Emotion* OR affect
Recognition	Recogn* OR percept* OR perceive* OR process* OR read* OR understand*	Recognition OR perception OR process OR understanding
Face	Face OR faces OR facial OR object OR objects OR visual	Face* OR facial OR object OR visual
Speech Prosody	Speech OR language OR linguistic* OR prosod* OR intonation OR inflection OR voice OR vocal* OR sound*	Speech OR language OR prosody OR intonation OR voice
Music	Music* OR melod* OR auditory OR acoustic* OR listen* OR pitch OR fundamental frequency OR frequency OR intensity	Music OR pitch OR auditory OR acoustic*

2.2.2. Study selection and eligibility criteria

The inclusion criteria of the review required studies that:

1. included an ASD group with mean IQ ≥ 70 and a sample size ≥ 2. The ASD diagnosis encompassed terms including ASD, autism, Asperger's syndrome, and pervasive developmental disorder-not otherwise specified (PDD-NOS), depending on the diagnostic tools/scales used in the studies.
2. included a NT control group for comparison.
3. implemented an emotion recognition paradigm.
4. used stimuli that conveyed emotions through human faces, nonhuman faces, speech prosody, or music.
5. included an objective measure of unimodal emotion recognition accuracy or response time.
6. provided summary statistics (i.e., means and standard deviations) of each group or inferential statistics (i.e., *t*, *F*, or *z* statistics) of group difference for at least one of the

six basic emotions (anger, fear, happiness, sadness, disgust, and surprise), or for an overall composite with all six emotions combined. In cases where data were not present in eligible studies (e.g., due to reporting styles), first and/or corresponding authors were contacted in an attempt to obtain sufficient information to calculate effect sizes. If this information was not obtained upon two request attempts, studies were excluded from this review and analysis.

7. published in English and in peer-reviewed journals. This criterion was set because (a) professional translators are needed for review teams to include non-English-language studies (Neimann Rasmussen & Montgomery, 2018), which requires extra costs and resources, given that translation software such as Google Translate produces uneven accuracy across different languages (Aiken, 2019); (b) there is no evidence of a language bias regarding effect estimates or conclusions in meta-analyses relying exclusively on English-language studies (Dobrescu et al., 2021; Morrison et al., 2012); (c) to limit the present review to a structured and replicable database search, as there is currently no "gold standard" for systematic search of grey literature (Adams et al., 2016; Paez, 2017). The inclusion of grey literature in systematic reviews has been a subject of debate. While its benefits include providing a more complete view of available evidence in a broader context, the inclusion of data from unpublished studies itself has also been suggested to introduce bias (e.g., due to their possibly lower methodological quality, more favourable results being provided more readily upon request; Sterne et al., 2011). The inclusion of grey literature has also been shown to represent only a small proportion of included studies and rarely impact results and conclusions of reviews (Hartling et al., 2017; Schmucker et al., 2017). Nevertheless, publication bias can pose a threat to the validity of any reviews and will, thus, be assessed in the present study (see Section 2.3.5).

All citations retrieved from the searches were imported into EndNote X8 (Clarivate Analytics, 2016). The resulting set was screened for eligibility against inclusion criteria using a three-stage process of reviewing the titles, followed by abstracts, and full-texts. Each stage involved two review authors (FYNL and CD/AV/JO) independently screening the retrieved items. To ensure rigour, a random 25% of the total sample of all titles, abstracts, and full texts were doubly and independently screened by two review authors. Any disagreement over the eligibility of particular studies was resolved through discussion by all review authors.

2.2.3. Data extraction

The data of full texts were extracted independently by four review authors (FYNL, CD, JO, and CZ). A random 10% of studies were doubly extracted by FYNL and CD to ensure accuracy of the data extracted. The predefined data extraction form included items of study setting, study characteristics, intellectual level of participants, diagnostic instruments, group matching procedures, type of task, stimulus type and domain, emotions assessed, and the study results. With regards to studies that included stimuli with more than one intensity level, the most prototypical and well-validated stimuli presented at the highest intensity were extracted. One study included stimuli with two different presentation durations (2000ms vs. 50ms; Otsuka, Uono, Yoshimura, Zhao, & Toichi, 2017), and only the data with the 2000ms presentation time were extracted. This was to ensure greater homogeneity of the data extracted across studies, since most studies used a presentation time greater than 50ms. The accuracy and response time results were extracted based on summary statistics (i.e., means and standard deviations) of each group or inferential statistics (i.e., *t, F, or z* statistics) of group difference in terms of performance on each individual emotion assessed (i.e., anger, fear, happiness, sadness, disgust, and/or surprise) and/or performance across all six emotions (i.e., the composite). Studies with multiple tasks, domains, and emotions were extracted as individual datasets (e.g., 1. verbal, face, angry; 2. verbal, face, fearful; 3. verbal, speech, angry; 4. verbal, speech, fearful) for separate analyses, as outlined in Section 2.2.5.

2.2.4. Quality assessment

The Critical Appraisal Skills Programme (CASP, 2017) tool was used to assess the methodological quality and the overall risk of bias of all included studies. The CASP tool is an accessible and commonly used critical appraisal tool in systematic review (e.g., Tosto et al., 2015; Westwood et al., 2016) and comprises seven checklists to guide the appraisal process in a consistent and systematic way for different types of evidence. The CASP checklist for case-control studies was specifically chosen given its suitability for evaluating the included studies of this review, which followed a case-control research design (i.e., studies must have included an ASD group and a NT control group as outlined in Section 2.2.2). A score was assigned for each criterion on the checklist using a three-point rating system developed by Duggleby et al. (2010). A score of 1 point denotes a high risk of bias and is given to papers that provide little to no justification for a particular issue; 2 points, a moderate risk of bias where the issue was addressed somewhat but not fully elaborated; 3 points, a low risk of bias with issues concerned being extensively justified and explained. Additional prompts were further adapted for criteria 3 (relating to diagnostic instruments used), 4 (relating to matching procedures), 7 (relating to the reporting of effect sizes), 8 (relating to the reporting of *p*-values), and 10 (relating to the generalisability in terms of sample demographics) by the review team to optimise objectivity in the quality assessment (see Appendix B). The quality assessment was performed by four review authors (FYNL, CD, JO, and CZ), with a random 10% of studies doubly assessed by FYNL and CD. Any discrepancies over were discussed and resolved among the team. The CASP scores were not used as a means of excluding papers but to provide indications of the quality across the included studies, as well as on the collective research evidence.

2.2.5. Analysis plan

The main meta-analyses were conducted on the 14 sets of effect sizes for accuracy and response time (seven each) separately, in order to examine group differences in emotion recognition for anger, fear, happiness, sadness, disgust, surprise, and the composite. All analyses were performed in R (RStudio Team, 2018). The standardised mean difference (SMD) was computed as the overall pooled effect for each dataset using the *compute.es* package (Deeks et al., 2001; Del Re, 2013). SMDs (given as Hedges' *g*) and their corresponding sampling variances for each dataset were calculated from summary statistics (Hedges, 1981). In cases where non-positive sampling variances (i.e., 0) were reported, datasets were omitted from the meta-analysis (Deeks et al., 2019; Vesterinen et al., 2014).

Random-effects models were run to analyse the SMDs from individual datasets and to compute the estimated pooled SMD together with its 95% confidence interval (CI) using the *metafor* package in R (Viechtbauer, 2010). The greater the magnitude of the pooled SMD, the larger the difference in effect between the ASD and NT groups. An SMD < 0.40 was interpreted as a small effect size, 0.40-0.70 as moderate, and > 0.70 as large (Schünemann et al., 2019). Statistical heterogeneity was quantified using the I^2 statistic (Higgins et al., 2003) and the 95% prediction interval (PI; or "plausible value interval") of the estimated pooled SMD (IntHout et al., 2016; Spineli & Pandis, 2020). Values of I^2 were classified as < 50% (low; possibly unimportant) and ≥ 50% (high; considerable concern). Since I^2 is not an absolute indicator of heterogeneity (Borenstein et al., 2017), prediction intervals were also included, which estimate heterogeneity by providing the expected ranges of true effects in future similar studies (IntHout et al., 2016; Spineli & Pandis, 2020).

As there were often multiple datasets per study, the count of studies was referred to as *k* and the count of datasets as *N*. For multiple datasets in each study, the following prioritisation criteria were employed in the main meta-analyses so that effect sizes would not be computed multiple times based on data from the same sample: (a) where the same group of participants were tested on multiple domains, human faces were prioritised, followed by speech prosody, nonhuman faces and music; (b) in cases where multiple tasks were reported, priority was given to verbal over nonverbal tasks, and labelling over discrimination, matching and detection tasks; and (c) where multiple NT groups were present, priority was given to the NT group matched with the ASD group on chronological age and full-scale IQ (if available), followed by verbal mental age (or verbal IQ; if available) and performance mental age (or nonverbal IQ; if available). These prioritisation criteria were chosen in order to reduce heterogeneity of the data analysed in the main meta-analyses, since the majority of the included studies matched groups on chronological age and full-scale IQ and examined facial emotion recognition using verbal labelling tasks.

To establish the significance of plausible moderating factors in emotion recognition, a series of planned subgroup analyses and meta-regressions was conducted as guided by previous literature. Planned subgroup analyses for each emotion were conducted based on domain (human face, nonhuman face, speech prosody, and music) and task demand (verbal and nonverbal). Following the computation of the pooled SMDs for each subgroup, effect sizes were compared to determine differences between subgroups using the Cochran's *Q* test (Borenstein & Higgins, 2013; Çoğaltay & Karadağ, 2015). To examine the moderating effects of age and full-scale IQ, verbal IQ, nonverbal IQ on group differences for each emotion, univariate meta-regressions were conducted to assess whether there were significant associations between each of these moderators and performance accuracy (S. G. Thompson & Higgins, 2002) when sufficient data were available (e.g., ≥10 studies for meta-regressions; Deeks et al., 2019).

We performed sensitivity analyses to test the robustness of the main meta-analysis results on accuracy, by including only studies meeting specific quality standards, i.e., those that had undertaken full-scale, verbal, and/or nonverbal IQ matching. Specifically, sensitivity analyses were first performed to assess the impact of IQ matching following the removal of datasets that did not undertake full-scale IQ matching, nonverbal IQ matching, and verbal IQ matching (on verbal tasks only, given that verbal IQ has been proposed to be an important contributor to performance on tasks requiring explicit verbal knowledge in ASD as discussed earlier). Results obtained from the sensitivity analyses were subsequently compared against the results of the main meta-analyses in order to assess the sustainability of the results after focusing on datasets with more controlled study designs. Moreover, through informal comparisons between the results of the main meta-analyses and sensitivity analyses, whether the pooled SMDs and heterogeneity were strengthened or weakened following exclusion of datasets aforementioned was evaluated. In addition, to test the impact of IQ matching on the robustness of the results of the full-scale IQ, verbal IQ, and nonverbal IQ moderator analyses, further sensitivity analyses were performed following the removal of datasets that did not undertake full-scale, verbal, and nonverbal IQ matching, respectively.

Publication bias (Easterbrook et al., 1991) was assessed with funnel plots and the Egger's test of funnel plot asymmetry (Egger et al., 1997; Page et al., 2019). Trim-and-fill analyses using the R_0 estimator were conducted to establish how each respective mean effect size would change if any identified bias were to be removed (Duval & Tweedie, 2000), in addition to test of the null hypothesis that the number of missing studies is zero (Duval, 2005)

2.3. Results

2.3.1. Study characteristics

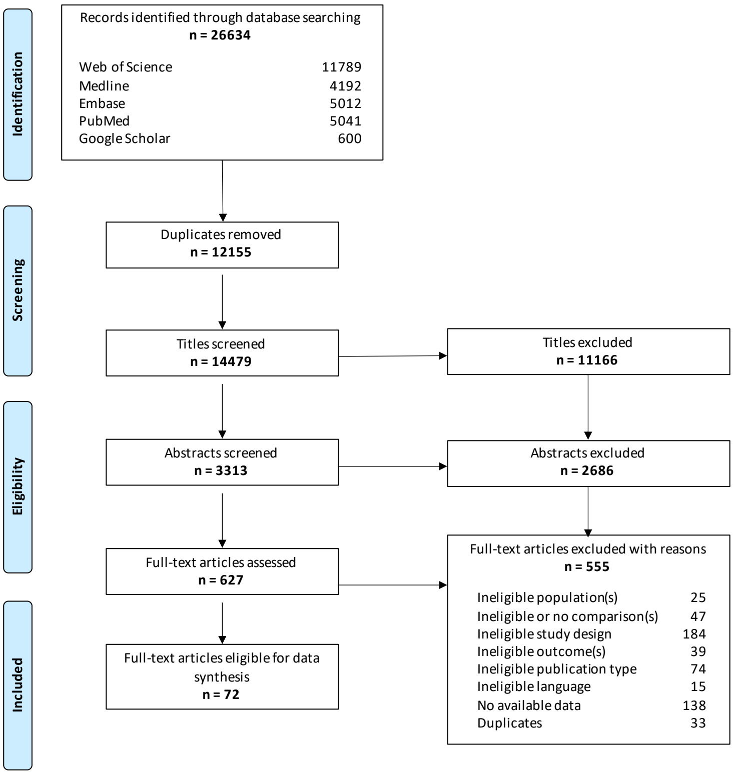


Figure 2. 1. Study selection process. PRISMA diagram of the combined initial and update literature search and screening process.

Figure 2.1 shows the PRISMA flowchart of the study selection process. A total of 72 papers with 1868 ASD and 2232 NT participants were included (see Table 2.3 for a summary of study/participant characteristics). Table 2.4 provides a detailed summary of the included studies. One paper (Fridenson-Hayo et al., 2016) reported results from three separate international sites that all met the inclusion criteria and were therefore considered to be three independent samples in the meta-analysis. This resulted in a total of 74 studies and 332 unique datasets: 73 studies contributed 259 unique accuracy datasets and 27 studies contributed 73 unique response time datasets for the main meta-analyses (see Appendix A for a full reference list of the included studies).

Table 2. 3. Characteristics of the 72 studies included in the meta-analysis separated by age groups.

	ASD				NT			
	N	**Mean (SD)**	**Min**	**Max**	**N**	**Mean (SD)**	**Min**	**Max**
Child								
Study sample size	42	26.98 (22.70)	10	114	42	34.59 (40.07)	12	220
Age	42	11.46 (2.85)	5.09	17.30	42	11.22 (2.94)	4.94	16.90
Male (%)	42	83.81 (10.53)	51.43	100.00	42	72.16 (16.45)	40.00	100.00
Verbal IQ	16	103.57 (7.21)	91.72	115.80	14	109.49 (6.25)	95.92	117.60
Performance IQ	12	103.78 (4.29)	97.20	112.50	11	107.59 (4.32)	100.56	113.17
Full-scale IQ	26	103.56 (5.09)	93.89	111.00	24	107.55 (4.53)	98.50	116.20
Adult								
Study sample size	30	22.70 (9.92)	5	46	30	23.67 (10.98)	5	53
Age	30	27.72 (6.62)	17.60	44.86	30	27.57 (5.84)	18.00	39.89
Male (%)	29	78.89 (19.65)	0.00	100.00	29	72.51 (22.92)	0.00	100.00
Verbal IQ	15	109.14 (5.01)	98.20	117.00	12	111.05 (4.06)	104.10	118.80
Performance IQ	15	104.80 (9.53)	86.50	128.47	12	111.16 (6.69)	98.00	126.40
Full-scale IQ	23	107.92 (5.28)	100.75	119.90	18	111.21 (4.31)	100.56	117.40

Note. Studies reporting IQ scaled scores (Akechi et al, 2009, 2010; Davidson, Hilvert, Misiunaite, Kerby, & Giordano, 2019; Fridenson-Hayo et al., 2016; Stewart, McAdam, Ota, Peppé, & Cleland, 2013; Taylor, Maybery, Grayndler, & Whitehouse, 2015; Tell, Davidson, & Camras, 2014) or raw scores (Griffiths et al., 2017; Wallace, Coleman, & Bailey, 2008) instead of standardised scores were not included in this table

Table 2. 4. Summary of included studies on emotion recognition in ASD. * indicates datasets included in the main analyses. - indicates information that was not available. X indicates unused datasets due to prioritisation and non-positive sampling variances.

Row	**Study**	**Number (males)**	**Diagnostic instrument**	**Mean age in years (SD/range)**	**IQ mean (SD/range)**	**Number (males)**	**Matching procedure**	**Mean age in years (SD range)**	**IQ mean (SD range)**	**Domain (Subdomain)**	**Type**	**Stimuli**	**Emotions**	**Mean accuracy ASD (SD)**	**Mean accuracy control (SD)**
1	Akechi et al. (2009)	14 (10)	ASQ-J	12.10 (2.00)	FSIQ: 98.90 (16.10) VIQ: 10.50 (3.10) NVIQ: 9.10 (3.00)	14 (10)	Age Gender FSIQ VIQ NVIQ	11.90 (1.90)	FSIQ: 101.30 (12.80) VIQ: 11.40 (2.40) NVIQ: 9.10 (2.90)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful	-	-
2*	Akechi et al (2010)	14 (10)	ASQ-J	13.70 (2.30)	FSIQ: 96.80 (16.80) VIQ: 9.40 (3.40) NVIQ: 9.60 (2.70)	14 (8)	Age Gender FSIQ VIQ NVIQ	12.30 (2.10)	FSIQ: 104.10 (8.10) VIQ: 11.10 (2.90) NVIQ: 10.20 (2.30)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful	0.88 (0.16) 0.81 (0.14)	0.95 (0.09) 0.85 (0.11)
3*	Baker (2010)	19 (13)	ADI-R; ADOS; ASDS	12.80 (1.47)	FSIQ: - (85-115)	19 (13)	Age Gender	12.20 (1.43)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Nonsense passages with prosodic patterns implying different emotions	Angry Happy Sad	0.70 (0.22) 0.78 (0.20) 0.73 (0.24)	0.67 (0.21) 0.84 (0.15) 0.75 (0.14)
4*	Bast et al. (2019)	23 (18)	ADI-R; ADOS	15.90 (2.80)	FSIQ: 100.00 (16.10) VIQ: 102.40 (16.60) NVIQ: 02.70 (18.40)	24 (19)	Age Gender FSIQ VIQ NVIQ	15.80 (2.40)	FSIQ: 108/60 (14.40) VIQ: 103.30 (13.90) NVIQ: 110.80 (13.80)	Human faces (dynamic)	Verbal (forced-choice labelling)	Videos from the Movie for the Assessment of Social Cognition (Dziobek et al., 2006)	Angry Fearful Happy	0.39 (0.49) 0.59 (0.50) 0.50 (0.51)	0.66 (0.48) 0.92 (0.28) 0.67 (0.48)

5*	Berggren et al. (2016)	35 (18)	ADOS; ICD-10	11.60 (1.80)	FSIQ: 103.80 (11.90)	32 (18)	Age Gender FSIQ	11.70 (1.80)	FSIQ: 102.90 (8.50)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Swedish version; Bölte et al. 2002)	Angry Fearful Happy Sad Disgust Surprise	3.80 (1.80) 3.70 (1.80) 8.20 (1.40) 5.40 (1.60) 4.20 (1.50) 4.70 (1.80)	4.60 (1.90) 4.90 (2.20) 9.80 (1.30) 5.30 (1.30) 4.60 (1.00) 5.40 (1.80)
6*	Boggs et al. (2010)	17 (15)	ADI-R; DSM-IV-TR	15.47 (2.38)	FSIQ: 105.94 (10.12) VIQ: 105.35 99.71)	17 (15)	Age Gender FSIQ VIQ	15.12 (2.15)	FSIQ: 108.41 (10.06) VIQ: 108.59 (9.64)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of four females expressing the different emotions	Angry Fearful Happy Sad Surprise	0.90 (0.18) 0.49 (0.23) 0.99 (0.06) 0.94 (0.14) 0.91 (0.15)	0.96 (0.13) 0.54 (0.24) 1.00 (0) 0.96 (0.1) 0.91 (0.15)
7*	Boucher et al. (2000)	19 (16)	DSM-IV	9.58 (1.00)	-	19 (10)	Age Gender FSIQ	6.33 (0.80)	-	Speech prosody (excerpt)	Verbal (free-choice labelling)	Audio tapes of a woman reciting the days of the week or months of the year in each of the emotions	Overall	13.50 (7.79)	13.05 (1.84)
8a*	Brennand et al. (2011)	15 (14)	-	14.50 (2.70)	VIQ: 92.50 (21.70)	15 (12)	Age	13.30 (1.67)	VIQ: 117.60 (13.60)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Psuedo sentences consisting of phonemes and phonotactics common in European languages that elicited different emotions by German actors	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad	55.00 (19.36) 40.80 (19.36) 30.80 (15.88) 47.50 (17.82)	58.30 (19.36) 53.30 (19.36) 45.00 (15.88) 50.00 (17.82)

8b	Brennand et al. (2011)	15 (14)	-	14.50 (2.70)	VIQ: 92.50 (21.70)	32 (28)	Not matched	-	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Psuedo sentences consisting of phonemes and phonotactics common in European languages that elicited different emotions by German actors	Angry Fearful Happy Sad	55.00 (19.36) 40.80 (19.36) 30.80 (15.88) 47.50 (17.82)	65.20 (19.80) 48.00 (19.23) 48.00 (15.84) 65.20 (19.80)
9a*	Brewer et al. (2016)	14 (13)	ADOS; AQ	44.86 (13.06)	-	13 (13)	Gender FSIQ	31.62 (9.66)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli posed by male control participants under the standard condition of the production task (i.e. free production)	Angry Fearful Happy Sad Disgust Surprise	0.31 (0.18) 0.24 (0.16) 0.9 (0.12) 0.49 (0.14) 0.47 (0.14) 0.41 (0.15)	0.57 (0.25) 0.28 (0.18) 0.88 (0.17) 0.44 (0.1) 0.53 (0.21) 0.44 (0.16)
9b	Brewer et al. (2016)	14 (13)	ADOS; AQ	44.86 (13.06)	-	13 (13)	Gender FSIQ	31.62 (9.66)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli posed by male control participants under the communicate condition of the production task (i.e. experimenter guessed the posers' expressed emotions)	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad ^{X} Disgust ^{X} Surprise	0.38 (0.24) 0.29 (0.13) 0.84 (0.13) 0.54 (0.21) 0.64 (0.14) 0.54 (0.14)	0.61 (0.22) 0.42 (0.25) 0.82 (0.2) 0.46 (0.18) 0.55 (0.15) 0.52 (0.20)

9c	Brewer et al. (2016)	14 (13)	ADOS; AQ	44.86 (13.06)	-	13 (13)	Gender FSIQ	31.62 (9.66)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli posed by male control participants under the mirror condition of the production task (i.e. production while watching own expressions in a camera)	^{X} Angry	0.34 (0.19)	0.64 (0.27)
													^{X} Fearful	0.34 (0.21)	0.45 (0.26)
													^{X} Happy	0.93 (0.11)	0.85 (0.18)
													^{X} Sad	0.53 (0.19)	0.53 (0.19)
													^{X} Disgust	0.41 (0.11)	0.50 (0.24)
													^{X} Surprise	0.62 (0.19)	0.55 (0.21)
10	Campbell et al. (2006)	13 (11)	ICD-10	13.16 (1.75)	VIQ: 96.07 (17.86)	13 (11)	Age Gender VIQ	13.32 (2.08)	VIQ: 95.92 (16.41)	Human faces (static)	Verbal (forced-choice labelling)	The Ekman-Friesen Test of Affect Recognition	Overall	71.39 (10.2)	80.69 (10.65)
11*	Ciaramidaro et al. (2018)	33 (31)	ADI-R; ADOS; ICD-10	18.76 (4.98)	NVIQ: 105.82 (13.75)	25 (21)	Gender VIQ	19.86 (3.45)	FSIQ: 109.00 (12.55) NVIQ: 109.00 (12.55)	Human faces (static)	Verbal (discrimination)	Photographs from the Karolinska Directed Emotional Faces	Angry	26.60 (3.18)	28.76 (2.33)
													Fearful	-	-
12*	Corden et al. (2008)	18 (13)	ADOS	32.90 (13.35)	FSIQ: 119.90 (11.10) VIQ: 116.30 (9.14) NVIQ: 117.10 (14.56)	17 (13)	Age Gender FSIQ VIQ NVIQ	31.90 (11.30)	FSIQ: 117.40 (8.26) VIQ: 115.10 (8.37) NVIQ: 115.90 (8.87)	Human faces (static)	Verbal (forced-choice labelling)	Halftone images of emotionally expressive faces	Angry	8.50 (1.34)	8.70 (1.10)
													Fearful	7.50 (1.41)	8.80 (1.33)
													Happy	9.90 (0.24)	10.00 (0)
													Sad	7.70 (1.19)	8.20 (1.25)
													Disgust	6.60 (2.57)	8.10 (2.26)
													Surprise	8.40 (2.20)	8.80 (1.29)
13*	Couture et al. (2010)	36 (29)	ADI-R	20.90 (5.70)	FSIQ: 101.20 (17.80)	41 (34)	Age Gender FSIQ	22.90 (5.60)	FSIQ: 109.40 (15.10)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of movie stills (Adolphs & Tranel, 2003)	Angry	0.52 (0.24)	0.67 (0.19)
													Fearful	0.62 (0.19)	0.62 (0.17)
													Happy	-	-
													Sad	0.69 (0.27)	0.80 (0.19)
													Disgust	-	-
													Surprise	-	-

14a*	Davidson et al. (2019)	23 (19)	-	11.08 (1.75)	FSIQ: 97.00 (16.30) VIQ: 46.00 (13.30) NVIQ: 49.00 (13.10)	23 (19)	Age Gender NVIQ	11.50 (2.08)	FSIQ: 107.50 (10.50) VIQ: 58.00 (8.00) NVIQ: 50.10 (6.20)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry	0.91 (0.28)	0.94 (0.25)
													Fearful	0.54 (0.49)	0.52 (0.49)
													Happy	0.63 (0.49)	0.76 (0.43)
													Sad	0.60 (0.37)	0.90 (0.38)
14b	Davidson et al. (2019)	23 (19)	-	11.08 (1.75)	FSIQ: 97.00 (16.30) VIQ: 46.00 (13.30) NVIQ: 49.00 (13.10)	23 (19)	Age Gender NVIQ	11.50 (2.08)	FSIQ: 107.50 (10.50) VIQ: 58.00 (8.00) NVIQ: 50.10 (6.20)	Non-human faces (static)	Verbal (forced-choice labelling)	Canine faces from the internet	Angry	0.80 (0.24)	0.83 (0.24)
													Fearful	0.68 (0.29)	0.70 (0.39)
													Happy	0.89 (0.29)	0.89 (0.21)
													Sad	0.82 (0.32)	0.78 (0.22)
15*	Doi et al. (2013)	20 (20)	AQ-J; DSM-IV	32.10 (7.30)	FSIQ: 104.20 (15.30) VIQ: 109.90 (15.50) NVIQ: 92.70 (19.90)	20 (20)	Age Gender FSIQ VIQ	33.50 (4.70)	FSIQ: 107.20 (13.90) VIQ: 104.10 (13.80) NVIQ: 109.90 (12.40)	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs at from the ATR DB99 and ATR-Promotions database	Angry	80.80 (17.90)	91.20 (7.20)
													Happy	87.90 (15.70)	91.30 (5.70)
													Sad	84.40 (19.60)	90.50 (7.50)
15b	Doi et al. (2013)	20 (20)	AQ-J; DSM-IV	32.10 (7.30)	FSIQ: 104.20 (15.30) VIQ: 109.90 (15.50) NVIQ: 92.70 (19.90)	20 (20)	Age Gender FSIQ VIQ	33.50 (4.70)	FSIQ: 107.20 (13.90) VIQ: 104.10 (13.80) NVIQ: 109.90 (12.40)	Speech prosody (utterance)	Verbal (forced-choice labelling)	Spoken words uttering a common family name in Japan in different emotions	Angry	90.80 (16.60)	88.80 (18.90)
													Happy	53.30 (30.90)	85.00 (17.00)
													Sad	87.50 (14.20)	92.50 (11.80)
16*	Eack et al. (2015)	45 (40)	ADOS	24.64 (5.72)	FSIQ: 112.60 (15.74)	30 (22)	Age Gender	26.40 (5.80)	FSIQ: 105.53 (7.01)	Human faces (static)	Verbal (forced-choice labelling)	Penn Emotion Recognition Test – 40 (Kohler et al. 2003)	Angry	4.90 (1.55)	5.18 (1.3)
													Fearful	6.59 (1.64)	7.22 (1.05)
													Happy	7.37 (0.68)	7.91 (0.31)
													Sad	6.04 (1.70)	6.77 (1.14)
17*	Evers et al. (2014)	22 (22)	ADI-R; DSM-IV-TR	7.85 (0.88)	FSIQ: 94.36 (11.93)	22(22)	Age Gender FSIQ	7.95 (0.68)	FSIQ: 98.50 (7.78)	Human Faces (static)	Verbal (forced-choice labelling)	California Facial Expressions Database	Angry	0.71 (0.22)	0.67 (0.26)
													Fearful	0.76 (0.21)	0.75 (0.21)
													Happy	0.92 (0.11)	0.98 (0.04)
													Sad	0.52 (0.23)	0.56 (0.26)
18*	Fink et al. (2014)	114 (76)	DSM-IV	10.65 (1.23)	VIQ: 103.58 (14.440	145 (94)	Age Gender	10.32 (1.32)	VIQ: 110.56 (15.78)	Human Faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Angry	3.43 (0.86)	3.52 (0.78)
													Fearful	2.79 (0.91)	2.94 (0.68)
													Happy	3.71 (0.78)	3.76 (0.52)
													Sad	2.40 (1.04)	2.69 (1.00)

19a*	Fridenson-Hayo et al. (2016)	20 (18)	ADOS-2; DSM-IV-TR; ICD-10	7.45 (1.31)	VIQ: 11.15 (4.26) NVIQ: 12.50 (2.96)	22 (19)	Age Gender VIQ NVIQ	7.50 (1.47)	VIQ: 11.82 (2.99) NVIQ: 11.55 (2.30)	Human faces (Dynamic)	Verbal (forced-choice labelling)	Video clips from Mindreading	Overall	0.70 (0.18)	0.86 (0.12)
19b*	Fridenson-Hayo et al. (2016)	16 (15)	ADOS-2; DSM-IV-TR; ICD-10	8.58 (1.03)	VIQ: 11.38 (3.56) NVIQ: 11.44 (2.48)	18 (13)	Age Gender VIQ NVIQ	7.80 (1.42)	VIQ: 12.22 (2.71) NVIQ: 9.72 (3.12)	Human faces (Dynamic)	Verbal (forced-choice labelling)	Video clips from Mindreading emotions obtained from the Mindreading database	Overall	0.69 (0.14)	0.84 (0.13)
19c*	Fridenson-Hayo et al. (2016)	19 (15)	ADOS-2; DSM-IV-TR; ICD-10	6.97 (1.03)	VIQ: 9.05 (1.90) NVIQ: 11.00 (2.79)	18 (15)	Age Gender VIQ NVIQ	7.36 (1.20)	VIQ: 10.11 (1.76) NVIQ: 11.83 (2.70)	Human faces (dynamic)	Verbal (forced-choice labelling)	Video clips from Mindreading	Overall	0.74 (0.13)	0.74 (0.12)
19d	Fridenson-Hayo et al. (2016)	20 (18)	ADOS-2; DSM-IV-TR; ICD-10	7.45 (1.31)	VIQ: 11.15 (4.26) NVIQ: 12.50 (2.96)	22 (19)	Age Gender VIQ NVIQ	7.50 (1.47)	VIQ: 11.82 (2.99) NVIQ: 11.55 (2.30)	Speech Prosody (sentence)	Verbal (forced-choice labelling)	Audio clips from The EU-Emotion Stimulus Set	Overall	0.68 (0.22)	0.73 (0.13)
19e	Fridenson-Hayo et al. (2016)	16 (15)	ADOS-2; DSM-IV-TR; ICD-10	8.58 (1.03)	VIQ: 11.38 (3.56) NVIQ: 11.44 (2.48)	18 (13)	Age Gender VIQ NVIQ	7.80 (1.42)	VIQ: 12.22 (2.71) NVIQ: 9.72 (3.12)	Speech Prosody (sentence)	Verbal (forced-choice labelling)	Audio clips from The EU-Emotion Stimulus Set	Overall	0.64 (0.14)	0.74 (0.16)
19f	Fridenson-Hayo et al. (2016)	19 (15)	ADOS-2; DSM-IV-TR; ICD-10;	6.97 (1.03)	VIQ: 9.05 (1.90) NVIQ: 11.00 (2.79)	18 (15)	Age Gender VIQ NVIQ	7.36 (1.21)	VIQ: 10.11 (1.76) NVIQ: 11.83 (2.70)	Speech Prosody (sentence)	Verbal (forced-choice labelling)	Audio clips from The EU-Emotion Stimulus Set	Overall	0.69 (0.13)	0.69 (0.13)

20a*	Griemel et al. (2014)	38 (38)	ADI-R; ADOS; DSM-IV; ICD-10	21.10 (9.50)	FSIQ: 107.70 (13.20)	37 (37)	Age Gender FSIQ	20.60 (7.00)	FSIQ: 113.00 (10.20)	Human faces (static)	Nonverbal (discrimination)	Identification of Facial Emotion task from the Amsterdam Neuropsychological Task battery (De Sonneville. 2001)	Angry Fearful Happy Sad	5.58 (4.50) 4.00 (3.04) 1.53 (1.96) 8.21 (5.52)	2.46 (2.27) 1.89 (2.11) 1.00 (1.08) 4.95 (3.53)
20b	Griemel et al. (2013)	38 (38)	ADI-R; ADOS; DSM-IV; ICD-10	21.10 (9.50)	FSIQ: 107.70 (13.20)	18(18)	Gender FSIQ	10.5 (1.3)	FSIQ: 111.7 (15.6)	Human faces (static)	Nonverbal (discrimination)	Identification of Facial Emotion task from the Amsterdam Neuropsychological Task battery (De Sonneville. 2001)	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad	5.58 (4.50) 4.00 (3.04) 1.53 (1.96) 8.21 (5.52)	6.28 (3.86) 4.39 (2.62) 1.11 (1.08) 9.72 (4.88)
21*	Griffiths et al. (2017)	66 (58)	-	11.24 (2.91)	VIQ: 127.14 (24.26) NVIQ: 37.08 (11.61)	70 (35)	Age VIQ NVIQ	11.24 (2.49)	VIQ: 135.17 (24.99) NVIQ: 40.16 (9.72)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful Happy Sad Disgust Surprise Overall	0.60 (0.29) 0.46 (0.31) 0.81 (0.24) 0.70 (0.27) 0.49 (0.32) 0.58 (0.28) 0.61 (0.31)	0.7 (0.25) 0.44 (0.32) 0.87 (0.18) 0.8 (0.23) 0.68 (0.26) 0.63 (0.26) 0.69 (0.29)
22*	Grossman et al. (2000)	13 (13)	ICD-10	11.80 (3.27)	FSIQ: 106.40 (18.42) VIQ: 115.80 (15.63) NVIQ: 97.20 (21.59)	13 (12)	Age FSIQ VIQ	11.50 (1.90)	FSIQ: 116.20 (12.75) VIQ: 115.00 (11.92) NVIQ: 111.80 (14.22)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful Happy Sad Surprise	0.75 (0.29) 0.48 (0.31) 0.96 (0.09) 0.81 (0.23) 0.75 (0.31)	0.89 (0.17) 0.50 (0.23) 1.00 (0) 0.87 (0.17) 0.89 (0.17)
23*	He et al. (2019)	21 (17)	DSM-5	5.09 (0.95)	-	21 (17)	Age Gender	4.94 (0.90)	-	Human faces (dynamic)	Verbal (forced-choice labelling)	Short film scenes from CASIA Chinese Natural Emotional Audio-Visual Database	Happy Sad	t = -2.52 t = -3.44	p = .021 p = .003

24*	Heaton et al. (2012)	20 (15)	AQ; DSM	33.70 (12.77)	FSIQ: 109.10 (18.43) VIQ: 106.40 (17.45) NVIQ: 109.50 (17.95)	20 (15)	Age Gender FSIQ VIQ NVIQ	33.60 (12.06)	FSIQ: 109.50 (15.11) VIQ: 109.00 (12.84) NVIQ: 105.15 (12.89)	Speech prosody (utterance)	Verbal (forced-choice labelling)	Vocal recordings of four actors expressing the different emotions verbally	Overall	60.42 (9.78)	85.16 (8.92)
25*	Heikkinen et al (2010)	12 (9)	ADI-R; ADOS-G; ICD-10	14.50 (-)	VIQ: 107.00 (-) NVIQ: 105.00 (-)	15 (8)	Age	14.30 (-)	-	Speech prosody (excerpt)	Verbal (forced-choice labelling)	Emotional speech data in Finnish from the MediaTeam, University of Oulu's emotional speech corpus database	Angry Happy Sad	13.08 (2.28) 16.58 (1.68) 15.17 (2.21)	13.67 (2.47) 17.07 (1.28) 16.40 (1.12)
26*	Hubbard et al. (2017)	22 (20)	ADOS; DSM-IV; DSM-5	25.91 (5.34)	FSIQ: 111.32 (11.20)	30 (10)	Not matched	22.53 (7.37)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Phrases portraying each of the emotions recorded by NT talkers during the production task	Angry Happy Sad	0.50 (0.16) 0.29 (0.15) 0.37 (0.20)	0.49 (0.18) 0.42 (0.14) 0.39 (0.20)
27a*	Isomura et al. (2014)	20 (15)	AQ-J; DSM-IV; ICD-10	9.02 (0.98)	FSIQ: 102.60 (16.00)	23 (12)	Age FSIQ	9.06 (1.21)	FSIQ: 105.50 (13.70)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures of 1 target and 2 distractors drawn in black against a white background	Angry Happy	97.50 (6.10) 95.90 (7.50)	99.30 (2.40) 98.00 (3.80)
27b	Isomura et al. (2014)	20 (15)	AQ-J; DSM-IV; ICD-10	9.02 (0.98)	FSIQ: 102.60 (16.00)	23 (12)	Age FSIQ	9.06 (1.21)	FSIQ: 105.50 (13.70)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures of 1 target and 5 distractors drawn in black against a white background	^{X} Angry ^{X} Happy	98.30 (4.40) 97.70 (4.00)	99.60 (1.70) 98.00 (6.70)

27c	Isomura et al. (2014)	20 (15)	AQ-J; DSM-IV; ICD-10	9.02 (0.98)	FSIQ: 102.60 (16.00)	23 (12)	Age FSIQ	9.06 (1.21)	FSIQ: 105.50 (13.70)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures of 1 target and 11 distractors drawn in black against a white background	^{X} Angry ^{X} Happy	99.60 (1.90) 99.20 (2.60)	99.60 (1.70) 98.20 (4.30)
28*	Isomura et al. (2014)	10 (8)	AQ-J; DSM-IV; ICD-10	10.47 (1.10)	FSIQ: 103.40 (13.85)	14 (11)	Age FSIQ	10.09 (1.30)	103.30 (9.28)	Non-human faces (static)	Nonverbal (detection)	Schematic pictures drawn in black against a white background	^{X} Angry ^{X} Happy	0.71 (0.05) 0.66 (0.05)	0.69 (0.04) 0.58 (0.05)
29*	Jaervinen et al. (2016)	17 (13)	ADI-R; ADOS	10.60 (-)	-	20 (8)	Age NVIQ	10.70 (-)	-	Music (segment)	Verbal (forced-choice labelling)	Novel musical pieces eliciting each of the emotions	Fearful Happy Sad	77.21 (27.68) 94.85 (9.94) 66.18 927.87)	91.88 (10.94) 99.38 (2.8) 85 (14.96)
30*	Ketelaars et al. (2016)	31 (0)	DSM-IV-TR	41.35 (11.22)	FSIQ: 105.80 (15.44)	28 (0)	Age Gender	39.89 (13.20)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Speech samples of semantically neutral sentences from the Amsterdam Neuropsychological Task battery	Angry Fearful Happy Sad	91.40 (9.97) 50.00 (23.27) 83.33 (12.91) 87.10 (15.64)	93.45 (9.97) 47.32 (22.8) 88.39 (10.72) 88.1 (15.62)
31*	Kim et al (2015)	19 (13)	ASSQ; SCQ; SRS	11.10 (2.50)	FSIQ: 110.60 (15.30) VIQ: 114.40 (16.30) NVIQ: 107.20 (15.80)	23 (16)	Age Gender FSIQ VIQ NVIQ	11.50 (2.30)	FSIQ: 115.20 (10.30) VIQ: 117.60 (10.30) NVIQ: 110.60 (13.00)	Human faces (dynamic)	Verbal (forced-choice labelling)	Avatar recordings eliciting each of the emotions	Angry Fearful Happy Sad Disgust Surprise	62.40 (10.9) 52.00 (15.00) 50.70 (7.10) 56.00 (9.30) 60.40 (13.50) 54.40 (6.10)	56.00 (18.80) 52.20 (23.40) 83.20 (21.50) 57.60 (18.00) 58.20 (22.50) 57.10 (24.10)
32*	Kliemann et al. (2012)	16 (16)	ADI-R; AQ; ASDI; DSM-IV	30.44 (6.34)	VIQ: 108.06 (7.38) NVIQ: 128.47 (10.82)	17 (17)	Age Gender FSIQ VIQ NVIQ	30.47 (6.23)	VIQ: 108.12 (14.76) NVIQ: 126.40 (8.94)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Fearful Happy	0.92 (0.07) 0.95 90.05)	0.97 (0.03) 0.98 (0.01)

33*	Kliemann et al. (2010)	17 (12)	ADI; ADOS; AQ	32.70 (8.20)	VIQ: 104.50 (15.60)	19 (14)	Age Gender VIQ	30.40 (5.90)	VIQ: 110.40 (12.90)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Fearful Happy	9039 (10.75) 94.37 (5.07)	96.44 (4.11) 96.58 (3.56)
34*	Król & Król (2019)	21 (19)	ADOS; ICD-10	16.27 (4.84)	FSIQ: 109.43 (17.67) VIQ: 110.14 (18.26) NVIQ: 107.24 (18.58)	23 (18)	Age Gender FSIQ VIQ NVIQ	16.31 (2.69)	FSIQ: 112.30 (10.59) VIQ: 10.00 (11.77) NVIQ: 113.17 (11.77)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from FACES database (Ebner et al., 2010)	Angry Fearful Happy Sad Disgust	0.86 (0.23) 0.88 (0.32) 0.95 (0.16) 0.71 (0.34) 0.83 (0.34)	0.81 (0.33) 0.91 (0.25) 1.00 (0.00) 0.82 (0.29) 1.00 (0.00)
35a*	Kujala et al. (2005)	8 (4)	DSM-IV; ICD-10	33.00 (22-43)	FSIQ: 114.00 (99-140)	8 (4)	Age Gender	32 (-)	-	Speech Prosody (utterance)	Verbal (forced-choice labelling)	Finnish word ('Saara') uttered by a female speaker with different emotional connotations	Sad	25.00 (38.00)	63.00 (42.00)
35b	Kujala et al. (2005)	8 (4)	DSM-IV; ICD-10	33.00 (22-43)	FSIQ: 114.00 (99-140)	8 (4)	Age Gender	32 (-)	-	Speech Prosody (utterance)	Nonverbal (detection)	Finnish word ('Saara') uttered by a female speaker with different emotional connotations	Sad	93.00 (14.00)	95.00 (8.00)
36*	Law Smith et al. (2010)	21 (21)	ADI-R; ADOS; DSM-IV	15.33 (2.20)	NVIQ: 100.67 (12.22)	16 (16)	Age Gender NVIQ	14.76 (2.08)	NVIQ: 100.56 (11.69)	Human faces (dynamic)	Verbal (forced-choice labelling)	Video clips of actors depicting each of the emotions	Angry Fearful Happy Sad Disgust Surprise Overall	- - - - F = -8.357 - F = -5.67	- - - - p = .007 - p = .023
37*	Li et al. (2017)	34 (30)	DSM-5; ADI-R	9.27 (2.23)	FSIQ: 109.94 (20.82)	39 (29)	Age Gender FSIQ	10.05 (3.20)	FSIQ: 113.03 (16.83)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Chinese facial affective picture system	Angry Fearful Happy Sad	0.47 (0.24) 0.34 (0.22) 0.81 (0.20) 0.50 (0.24)	0.66 (0.21) 0.47 (0.3) 0.9 (0.19) 0.77 (0.19)

38	Lindström et al. (2018)	15 (15)	ADI-R; DSM-IV; DSM-5; ICD-10	10.40 (-)	VIQ: 108.00 (14.72) NVIQ: 98.00 (12.89)	16 (16)	Age Gender VIQ	10.10 (-)	VIQ: 116.00 (15.30) NVIQ: 108.00 (12.90)	Speech prosody (utterance)	Nonverbal (discrimination)	Finnish word (‘Saara’) uttered by a female speaker with different emotional connotations	Sad	0.98 (0.04)	0.97 (0.06)
39*	Loth et al. (2018)	46 (34)	AQ	30.20 (9.40)	FSIQ: 116.00 (87-135) VIQ: 113.90 (85-160)	53 (33)	Age Gender FSIQ VIQ NVIQ	27.50 (7.80)	FSIQ: 115.50 (85-143) VIQ: 114.00 (74-146)	Human faces (static)	Verbal (matching)	Images sourced from films made in non-English speaking countries	Overall	0.74 (0.14)	0.89 (0.09)
40a*	O'Connor (2007)	18 (16)	DSM-IV; Gillberg & Gillberg (1989) criteria	26.90 (7.80)	-	18 (16)	Age Gender	25.20 (6.50)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs selected from the Mind Reading Emotions Library (Baron-Cohen et al. 2003)	Angry	0.94 (0.09)	0.97 (0.05)
													Happy	0.97 (0.06)	0.99 (0.03)
													Sad	0.87 (0.11)	0.91 (0.08)
40b	O'Connor (2007)	18 (16)	DSM-IV; Gillberg & Gillberg (1989) criteria	26.90 (7.80)	-	18 (16)	Age Gender	25.20 (6.50)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	Semantically neutral sentences (“I want to go to the other movies”) spoken in each of the emotions by six female and eight male actors	Angry	0.84 (0.15)	0.84 (0.11)
													Happy	0.77 (0.18)	0.81 (0.15)
													Sad	0.86 (0.13)	0.89 (0.09)
41a*	Oerlemans et al. (2014)	90 (73)	ADI-R; CRS-R; SCQ	10.60 (2.05)	FSIQ: 103.20 (13.70) VIQ: 101.55 (15.70) NVIQ: 105.20 (17.25)	139 (62)	-	9.20 (1.90)	FSIQ: 107.40 (11.70) VIQ: 108.80 (13.10) NVIQ: 106.40 (14.70)	Human faces (static)	Verbal (discrimination)	Emotional faces eliciting each of the emotions	Angry	82.50 (1.20)	85.70 (1.00)
													Fearful	82.60 (0.70)	84.80 (1.70)
													Happy	93.70 (0.70)	95.70 (0.60)
													Sad	79.20 (1.20)	81.60 (0.90)

41b	Oerlemans et al. (2014)	66 (54)	ADI-R; CRS-R; SCQ	11.55 (1.35)	FSIQ: 102.56 (13.30) VIQ: 1.1.20 (16.15) NVIQ: 104.25 (15.85)	72 (33)	-	10.70 (1.10)	FSIQ: 106.10 (11.30) VIQ: 108.20 (12.50) NVIQ: 104.10 (14.60)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Sentences with neutral content spoken in the tone of each emotion (Vingerhoets et al, 2003)	Angry	82.30 (1.80)	84.80 (1.70)
													Fearful	33.40 (2.20)	31.10 (2.00)
													Happy	77.40 (1.90)	77.40 (1.80)
													Sad	72.60 (2.70)	68.60 (2.50)
42a*	Otsuka et al. (2017)	21 (14)	AQ; DSM-IV	25.24 (5.75)	FSIQ: 112.00 (9.92) VIQ: 115.05 (11.16) NVIQ: 105.38 (12.46)	21 (14)	Age Gender FSIQ VIQ NVIQ	24.90 (6.32)	FSIQ: 113.57 (11.58) VIQ: 113.43 (12.35) NVIQ: 110.81 (12.38)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988) presented for 2000ms	Angry	70.24 (23.21)	75 (19.36)
													Fearful	55.95 (31.53)	71.43 (24.09)
													Happy	97.62 (7.52)	100 (0)
													Sad	72.62 (30.52)	80.95 (23.59)
													Disgust	50 (36.23)	45.24 (28.08)
													Surprise	97.62 (7.52)	100 (0)
													Overall	73.41 (12.11)	78.57 (8.57)
42b	Otsuka et al. (2017)	21 (14)	AQ; DSM-IV	25.24 (5.75)	FSIQ: 112.00 (9.92) VIQ: 115.05 (11.16) NVIQ: 105.38 (12.46)	21 (14)	Age Gender FSIQ VIQ NVIQ	24.90 (6.32)	FSIQ: 113.57 (11.58) VIQ: 113.43 (12.35) NVIQ: 110.81 (12.38)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988) presented for 50ms	^{X} Angry	32.14 (22.56)	41.67 (21.41)
													^{X} Fearful	19.05 (22.23)	19.05 (17.51)
													^{X} Happy	89.29 (28.03)	94.05 (10.91)
													^{X} Sad	55.95 (31.53)	64.29 (29.12)
													^{X} Disgust	10.71 (18.66)	23.81 (30.08)
													^{X} Surprise	88.1 (18.74)	94.05 (17.51)
													^{X} Overall	49.6 (13.94)	56.15 (8.4)
43*	Pelphrey et al. (2002)	5 (5)	ADI-R; ADOS; DSM-IV; ICD-10	25.20 (-)	FSIQ: 100.75 (7.69) VIQ: 117.00 (23.12) NVIQ: 86.50 (9.57)	5 (5)	Gender	28.20 (-)	-	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry	0.60 (0.29)	0.90 (0.14)
													Fearful	0.65 (0.22)	0.95 (0.11)
													Happy	0.95 (0.11)	1.00 (0.00)
													Sad	0.85 (0.22)	0.95 (0.09)
													Disgust	0.70 (0.21)	0.80 (0.33)
													Surprise	0.80 (0.27)	0.95 (0.32)
													Overall	0.76 (0.12)	0.93 (0.08)

44a*	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Angry Fearful Happy Sad Disgust Surprise Overall	61.30 (25.99) 58.26 (26.05) 94.78 (11.63) 65.65 (24.09) 65.65 (27.44) 79.57 (22.66) 70.74 (14.79)	90.43 (7.06) 80.87 (16.49) 100 (0.00) 82.61 (14.21) 84.75 (17.29) 90.87 (9.96) 88.3 (5.47)
44b	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988)	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad ^{X} Disgust	62.17 (34.14) 82.65 (23.12) 98.17 (4.82) 83.3 (26.08) 72.09 (29.39)	88.17 (13.51) 86.96 (26.25) 100 (0) 95.04 (13.34) 89.48 (18.85)
44c	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Human faces (static)	Nonverbal (matching)	Emotional faces from the Japanese and Caucasian Facial Expressions of Emotion (Matsumoto & Ekman, 1988)	Angry Fearful Happy Sad Disgust	85.65 (17.84) 85.78 (14.970 96.91 (7.42) 77.04 (23.16) 72.7 (21.5)	95.09 (8.15) 93.17 (10.51) 96.91 (14.8) 94.43 (11.2) 86.35 (20.4)
44d	Philip et al. (2010)	23 (16)	AQ; DSM-IV	32.50 (10.90)	FSIQ: 101.50 (18.50) VIQ: 98.20 (15.80) NVIQ: 104.40 (18.60)	23 (17)	Age Gender	32.40 (11.10)	FSIQ: 111.20 (8.50) VIQ: 106.80 (8.80) NVIQ: 113.40 (10.40)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Strings of numbers spoken in an emotional tone from Calder Vocal Emotion (Calder et al., 2004)	Angry Fearful Happy Sad Disgust	63.48 (20.14) 59.57 (25.85) 57.83 (20.66) 75.22 (15.34) 48.26 930.55)	83.48 (14.96) 77.83 (11.66) 76.96 (14.6) 80 (14.14) 76.52 (17.48)
45	Quintin et al. (2011)	26 (20)	DSM-IV	13.58 (1.92)	FSIQ: 97.00 (15.00) VIQ: 94.00 (19.00) NVIQ: 101.00 (13.00)	26 (12)	Age VIQ	13.50 (2.17)	FSIQ: 108.00 (12.00) VIQ: 107.00 (13.00) NVIQ: 107.00 (15.00)	Music (excerpt)	Verbal (forced-choice labelling)	Music clips eliciting each of the emotions	Fearful Happy Sad	4.38 (1.06) 3.77 91.21) 3.65 (1.41)	4.69 (0.55) 3.85 (0.97) 4.23 (0.82)

46*	Rhodes et al. (2018)	19 (17)	ADOS-2; DSM-IV	12.25 (1.92)	FSIQ: 107.30 (12.40) VIQ: 101.90 (11.40) NVIQ: 112.50 (15.30)	19 (14)	Age FSIQ VIQ NVIQ	12.25 (1.83)	FSIQ: 107.80 (5.00) VIQ: 103.10 (6.30) NVIQ: 110.40 (5.40)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad	0.92 (0.17) 0.83 (0.15) 0.96 (0.09) 0.80 (0.24)	0.95 (0.10) 0.86 (0.15) 0.99 (0.06) 0.78 (0.22)
47a*	Rigby et al. (2018)	16 (11)	-	27.80 (7.80)	FSIQ: 106.30 (10.80) VIQ: 107.80 (13.90) NVIQ: 103.90 (15.70)	16 (11)	Age Gender FSIQ VIQ NVIQ	27.30 (7.50)	FSIQ: 113.40 (11.40) VIQ: 110.70 (9.50) NVIQ: 113.30 (11.40)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of female Caucasian actors from Pilz et al. (2006)	Angry Surprise	93.30 (7.10) 96.10 (6.60)	97.70 (3.00) 96.30 (4.70)
47b	Rigby et al. (2018)	16 (11)	-	27.80 (7.80)	FSIQ: 106.30 (10.80) VIQ: 107.80 (13.90) NVIQ: 103.90 (15.70)	16 (11)	Age Gender FSIQ VIQ NVIQ	27.30 (7.50)	FSIQ: 113.40 (11.40) VIQ: 110.70 (9.50) NVIQ: 113.30 (11.40)	Human faces (dynamic)	Verbal (forced-choice labelling)	Photographs of female Caucasian actors from Pilz et al. (2006)	^{X} Angry ^{X} Surprise	97.50 (2.70) 95.50 (5.10)	95.50 (4.80) 98.00 (2.80)
48*	Sasson et al. (2016a)	21 (18)	ADOS	23.43 (4.36)	FSIQ: 101.48 (16.97)	39 (23)	Gender FSIQ	35.87 (9.33)	FSIQ: 100.56 (14.87)	Human faces (static)	Verbal (forced-choice labelling)	Photographs of an individual's face from Kohler et al., 2003	^{X} Angry Fearful ^{X} Happy Sad	1.00 (0) 0.71 (0.46) 1.00 (0) 0.92 (0.28)	1.00 (0.00) 0.79 (0.41) 1.00 (0.00) 0.97 (0.26)
49*	Sasson et al. (2016b)	21 (19)	ADOS; AQ	23.81 (4.58)	FSIQ: 11.56 (12.09)	28 (25)	Age Gender FSIQ	23.75 (6.60)	FSIQ: 116.71 (10.38)	Human faces (static)	Nonverbal (discrimination)	Faces eliciting each of the emotions	Angry Happy	84.66 (13.82) 91.01 (11.58)	86.31 (10.85) 96.63 (5.31)
50*	Sawyer et al. (2012)	30 (20)	DSM	21.60 (9.80)	FSIQ: 108.10 (17.90) VIQ: 109.70 (19.10) NVIQ: 104.30 (18.20)	24 (7)	Age FSIQ VIQ NVIQ	24.00 (9.20)	FSIQ: 114.10 (13.00) VIQ: 113.40 (12.80) NVIQ: 111.40 (12.80)	Human faces (static)	Verbal (forced-choice labelling)	Ekman facial affect set	Overall	69.97 (11.09)	86.77 (11.04)
51a*	Schaller & Rauh (2017)	23 (23)	ADI-R; ADOS	15.72 (1.25)	NVIQ: 105.65 (11.47)	22 (22)	Age Gender NVIQ	15.85 (0.97)	NVIQ: 103.77 (11.09)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Bölte et al. 2006)	Angry Fearful Overall	F = -6.48 F = -6.11 0.81 (0.07)	p = .015 p = .018 0.86 (0.05)

51b	Schaller & Rauh (2017)	23 (23)	ADI-R; ADOS	15.72 (1.25)	NVIQ: 105.65 (11.47)	22 (22)	Age Gender NVIQ	15.85 (0.97)	NVIQ: 103.77 (11.09)	Human faces (dynamic)	Verbal (forced-choice labelling)	Facially Expressed Emotion Labelling (FEEL; Kessler et al., 2002)	^{X} Overall	0.83 (0.09)	0.79 (0.11)
52*	Schelinski & von Kriegstein (2019)	16 (13)	ADI-R; ADOS; AQ; ICD-10	33.75 (10.12)	FSIQ: 110.31 (13.79) VIQ: 110.75 (12.35) NVIQ: 107.38 (17.55)	16 (13)	Age Gender FSIQ VIQ NVIQ	33.69 (9.58)	FSIQ: 111.50 (10.97) VIQ: 108.75 (12.59) NVIQ: 112.69 (9.59)	Speech prosody (utterance)	Verbal (forced-choice labelling)	Two-syllabic semantically neutral German nouns spoken in each one of the emotions by one female and one male actor	Angry Fearful Happy Sad Disgust	84.98 (13.15) 68.37 927.82) 64.10 (30.93) 51.45 (32.05) 50.18 (22.77)	92.48 (6.58) 89.85 (6.71) 80.54 (9.93) 81.64 (10.76) 65.97 (21.78)
53*	Shafritz et al. (2015)	15 (12)	ADI; ADOS	18.10 (-)	FSIQ: 101.50 (18.60) VIQ: 105.70 (18.80) NVIQ: 103.50 (17.40)	15 (12)	Age Gender VIQ NVIQ	18.40 (12-23)	FSIQ: 115.20 (9.30) VIQ: 118.80 (14.90) NVIQ: 108.00 (8.10)	Human faces (static)	Verbal (detection)	Ekman facial affect set	Fearful Happy	92.40 (7.68) 95.00 (6.45)	93.21 (8.05) 96.14 (3.42)
54a*	Shanok et al. (2019)	12 (9)	GARS-2	5.75 (0.97)	-	16 (10)	-	5.50 (1.41)	-	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad	77.77 (16.42) 86.11 (17.17) 94.44 (12.98) 69.44 (22.49)	91.67 (14.99) 93.75 (13.44) 100.00 (0.00) 95.83 (11.39)
54b	Shanok et al. (2019)	12 (9)	GARS-2	5.75 (0.97)	-	16 (10)	-	5.50 (1.41)	-	Human faces (static)	Verbal (forced-choice labelling)	Mothers replicating photographs from the NimStim Face Stimulus Set	^{X} Angry ^{X} Fearful ^{X} Happy ^{X} Sad	97.22 (14.91) 83.33 (22.48) 100.00 (0.00) 86.11 (17.17)	91.67 (14.99) 93.75 (13.44) 100.00 (0.00) 93.75 (13.44)
55a*	Sinzig et al. (2008)	19 (17)	ASD: ADI-R; ADOS DSM-IV	13.60 (3.40)	FSIQ: 111.00 (19.10)	29 (22)	Age FSIQ	12.80 (2.90)	FSIQ: 109.00 (12.90)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Bölte et al. 2006)	Angry Fearful Happy Sad Disgust Surprise Overall	5.40 (1.60) 3.20 (1.30) 7.60 (1.60) 5.80 (1.70) 4.40 (1.70) 4.90 (1.30) 0.76 (0.09)	6.1 (1.10) 2.9 (1.30) 7.7 (1.10) 5.9 (1.60) 4.1 (1.50) 5.2 (0.90) 0.78 (0.07)

55b	Sinzig et al. (2008)	21 (20)	ASD+ADHD: ADI-R; ADOS; DCL-TES; DSM-IV	11.60 (3.70)	FSIQ: 102.00 (13.10)	29 (22)	Age FSIQ	12.80 (2.90)	FSIQ: 109.00 (12.90)	Human faces (static)	Verbal (forced-choice labelling)	Black and white photographs from the Frankfurt Test for Facial Affect Recognition (Bölte et al. 2006)	^{X} Angry	5.10 (2.00)	6.1 (1.10)
													^{X} Fearful	3.50 (2.80)	2.9 (1.30)
													^{X} Happy	6.70 (2.20)	7.7 (1.10)
													^{X} Sad	4.90 (2.30)	5.9 (1.60)
													^{X} Disgust	4.00 (1.80)	4.1 (1.50)
													^{X} Surprise	3.60 (1.80)	5.2 (0.90)
													^{X} Overall	0.72 (0.10)	0.78 (0.07)
56*	Stephenson et al. (2019)	30 (-)	ADOS-2; AQ; DSM-5	24.52 (6.04)	FSIQ: 112.36 (10.63)	46 (-)	FSIQ	20.93 (2.03)	111.95 (8.21)	Human faces (dynamic)	Verbal (forced-choice labelling)	Dynamic faces from the Amsterdam Dynamic Facial Expression Set; van der Schalk et al., 2011)	Angry	0.93 (0.26)	0.95 (0.22)
													Fearful	0.99 (0.09)	0.99 (0.10)
													Happy	1.00 (0.05)	1.00 (0.05)
57*	Stewart et al. (2013)	11 (7)	DSM-IV	27.20 (7.50)	VIQ: 14.90 (6.20)	14 (8)	Age Gender VIQ	26.40 (5.60)	VIQ: 18.10 (4.00)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from JAFFE database (Lyons et al 1999)	Overall	0.55 (0.08)	0.68 (0.07)
58a*	Tanaka et al. (2012)	66 (56)	ADI-R; ADOS-G; DSM-IV	11.90 (4.00)	FSIQ: 106.80 (20.90)	68 (43)	Age FSIQ	11.90 (3.10)	FSIQ: 106.80 (7.80)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry	82.40 (21.20)	89.30 (13.40)
													Fearful	67.50 (22.00)	69.20 (20.80)
													Happy	98.10 (6.70)	97.50 (5.80)
													Sad	91.70 (13.50)	94.70 (9.70)
													Disgust	87.50 (18.00)	92.80 (12.20)
													Surprise	86.60 (18.60)	87.80 (14.60)
													Overall	F = -2.86	p = .09
58b	Tanaka et al. (2012)	67 (57)	ADI-R; ADOS-G; DSM-IV	12.00 (4.00)	FSIQ: 106.80 (20.90)	66 (42)	Age FSIQ	11.90 (3.10)	FSIQ: 106.80 (7.80)	Human faces (static)	Nonverbal (matching)	Photographs from the NimStim Face Stimulus Set	Angry	84.30 (15.30)	94.20 (9.90)
													Fearful	74.10 (22.20)	86.90 (16.90)
													Happy	96.80 (9.70)	99.50 (2.90)
													Sad	65.90 (22.40)	82.30 (16.80)
													Disgust	66.70 (20.10)	75.80 (16.60)

59a*	Taylor et al. (2015)	17 (12)	ADOS-G; DSM-IV	9.67 (2.25)	VIQ: 91.72 (15.33) NVIQ: 10.41 (2.94)	54 (26)	Age NVIQ	8.94 (1.92)	-	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs posed by both children and adults from movie files featured on Mind Reading DVD (Baron-Cohen, 2002)	Overall	62.50 (11.97)	70.75 (10.99)
59b	Taylor et al. (2015)	17 (12)	ADOS-G; DSM-IV	9.67 (2.25)	VIQ: 91.72 (15.33) NVIQ: 10.41 (2.94)	54 (26)	Age NVIQ	8.94 (1.92)	-	Speech prosody (sentence)	Verbal (forced-choice labelling)	The sentence "Oh I'm going out of the room now but I'll be back later" spoken each of the emotions with Australian accents	Overall	64.22 (13.26)	71.06 (12.71)
60*	Tell et al. (2014)	22 (17)	ADI-R; ADOS-G; DSM-IV	10.31 (-)	FSIQ: 102.68 (6.64) VIQ: 9.32 (1.69)	22 (17)	Age Gender	9.80 (-)	VIQ: 10.64 (1.67)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry Fearful Happy Sad	63.30 (4.90) 49.30 (7.10) 80.20 (4.80) 45.00 (5.90)	65.80 (4.70) 73.20 (6.40) 79.60 (4.70) 54.50 (6.90)
61*	Tottenham et al. (2014)	33 (30)	ADOS; AQ	15.00 (6.00)	FSIQ: 111.00 (-)	53 (35)	FSIQ VIQ	16.00 (8.00)	FSIQ: 103.00 (-)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the NimStim Face Stimulus Set	Angry	0.93 (0.15)	0.90 (0.24)
62*	Uono et al. (2011)	28 (23)	DSM-IV-TR	17.60 (5.20)	FSIQ: 103.30 (13.40) VIQ: 105.20 (14.70) NVIQ: 100.10 (13.30)	28 (24)	Age Gender	18.00 (4.00)	-	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	Angry Fearful Happy Sad Disgust Surprise Overall	60.70 (21.70) 32.10 (26.46) 98.70 (4.23) 79.90 (16.93) 39.30 (24.34) 95.30 (10.05) 67.70 (8.47)	53.80 (22.75) 51.80 (19.58) 98.20 (4.76) 76.30 (19.58) 46.40 (22.75) 94.20 (9.00) 70.10 (7.94)

63a*	Uono et al. (2013)	18 (15)	AS: CARS; DSM-IV	18.60 (6.50)	FSIQ: 106.00 (11.90) VIQ 108.30 (11.90) NVIQ: 101.90 (13.90)	18 (14)	Age Gender	18.80 (3.60)	-	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	Angry	59.00 (24.18)	54.2 (22.06)
													Fearful	40.30 (30.12	54.9 (16.12)
													Happy	97.90 (4.67)	97.9 (4.67)
													Sad	84.70 (14.42)	76.4 (18.24)
													Disgust	48.60 (25.46)	52.1 (22.06)
													Surprise	97.20 (5.52)	95.8 97.64)
													Overall	71.30 (8.06)	71.9 (6.79)
63b	Uono et al. (2013)	18 (12)	PDD-NOS: CARS; DSM-IV	19.80 (4.70)	FSIQ: 101.00 (14.10) VIQ: 103.10 (17.70) NVIQ: 98.20 (11.30)	18 (14)	Age Gender	18.80 (3.60)	-	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	^{X} Angry	62.50 (16.97)	54.2 (22.06)
													^{X} Fearful	31.30 (22.06)	54.9 (16.12)
													^{X} Happy	99.30 (2.97)	97.9 (4.67)
													^{X} Sad	75.70 (16.12)	76.4 (18.24)
													^{X} Disgust	32.60 (22.91)	52.1 (22.06)
													^{X} Surprise	93.80 (11.46)	95.8 97.64)
													^{X} Overall	65.90 (7.64)	71.9 (6.79)
64*	Vannetzel et al. (2011)	10 (9)	ADI-R; CARS; DSM-IV	9.60 (1.70)	FSIQ: above 70	35 (30)	Age	8.40 (1.80)	-	Human faces (static)	Nonverbal (discrimination)	Ekman facial affect set	Angry	33.30 (20.54)	80.00 (26.55)
													Happy	70.00 (15.48)	85.70 (23.01)
													Sad	53.30 (19.28)	84.80 (20.06)
65a*	Waddington et al. (2018)	89 (69)	ADI-R	12.32 (2.48)	FSIQ: 101.51 (14.67)	220 (110)	Gender	13.11 (2.35)	FSIQ: 105.5 (12.42)	Human faces (static)	Verbal (discrimination)	Identification of Facial Emotion task from the Amsterdam Neuropsychological Task battery (DeSonneville, 1999)	Angry	85.24 (10.77)	88.24 (9.48)
													Fearful	84.52 (12.67)	89.85 (10.20)
													Happy	94.58 (6.12)	96.00 (4.58)
													Sad	79.98 (12.98)	84.40 (10.19)

65b	Waddington et al. (2018)	89 (69)	ADI-R	12.32 (2.48)	FSIQ: 101.51 (14.67)	220 (110)	Gender	13.11 (2.35)	FSIQ: 105.5 (12.42)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Affective Prosodytask from the Amsterdam Neuropsychological Task battery (DeSonneville, 1999)	Angry Fearful Happy Sad	84.93 (16.76) 36.47 (21.84) 79.00 (17.02) 72.73 (22.68)	87.20 (14.61) 40.36 (20.73) 81.06 (15.69) 72.77 (21.24)
66a*	Wallace et al. (2008)	26 (23)	ADI-R; ICD-10	32.00 (9.00)	VIQ: 148.00 (13.00) NVIQ: 101.00 (18.00)	26 (23)	Age Gender FSIQ VIQ NVIQ	31.00 (9.00)	VIQ: 153.00 (9.00) NVIQ: 98.00 (12.00)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion	Angry Fearful Happy Sad Disgust Surprise	0.62 (0.17) 0.62 (0.24) 0.93 (0.13) 0.68 (0.22) 0.72 (0.23) 0.68 (0.22)	0.73 (0.22) 0.85 (0.17) 0.99 (0.70) 0.86 (0.11) 0.88 (0.12) 0.84 (0.22)
66b	Wallace et al. (2008)	26 (23)	ADI-R; ICD-10	32.00 (9.00)	FSIQ: 122.00 (10.00) VIQ: 118.00 (14.00) NVIQ: 122.00 (7.00)	26 (23)	Age Gender FSIQ VIQ NVIQ	31.00 (9.00)	FSIQ: 117.00 (13.00) VIQ: 115.00 (14.00) NVIQ: 116.00 (13.00)	Human faces (static)	Verbal (forced-choice labelling)	Emotional faces from the Ekman facial affect set and the Japanese and Caucasian Facial Expressions of Emotion presented in a piecemeal fashion starting from the eyes or the mouth	^{X} Angry ^{X} Fearful ^{X} Disgust ^{X} Surprise	0.70 (0.22) 0.76 (0.22) 0.72 (0.30) 0.65 (0.27)	0.82 (0.27) 0.92 (0.11) 0.91 (0.13) 0.75 (0.25)
67a*	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Neutral sentences spoken in each one of the emotions	Angry Happy Sad	83.72 (14.07) 88.01 (14.54) 96.29 (10.13)	85.72 (14.45) 95.55 (6.25) 99.15 (2.35)

67b	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Emotional sentences spoken in matching emotional prosody	^{X} Angry ^{X} Happy ^{X} Sad	84.95 (15.59) 74.47 (16.42) 97.71 (4.79)	86.94 (11.11) 81.33 (19.24) 98.57 (2.57)
67c	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Neutral words spoken in each one of the emotions	^{X} Angry ^{X} Happy ^{X} Sad	84.90 (17.83) 53.78 (30.05) 94.23 (12.46)	79.12 (19.6) 63.12 (27.18) 94.23 (3.68)
67d	Wang & Tsao (2015)	25 (25)	DSM-IV-TR	8.15 (1.17)	FSIQ: 107.12 (11.14) VIQ: 107.08 (10.57)	25 (25)	Age Gender FSIQ VIQ	8.20 (1.04)	FSIQ: 112.96 (9.91) VIQ: 113.68 (8.47)	Speech prosody (sentence)	Verbal (forced-choice labelling)	Emotional words spoken in matching emotional prosody	^{X} Angry ^{X} Happy ^{X} Sad	73.00 (17.54) 73.01 (15.44) 94.23 (7.38)	69.45 (18.62) 83.00 (10.5) 97.11 (3.35)
68*	Wang & Adolphs (2017)	18 (18)	ADI-R; ADOS-2; DSM-5; ICD-10	30.80 (7.40)	FSIQ: 105.00 (13.30)	15 (11)	Age FSIQ VIQ NVIQ	35.10 (11.40)	FSIQ: 107.00 (8.69)	Human faces (static)	Verbal (forced-choice labelling)	Facial photographs obtained from the STOIC database (Roy et al., 2007)	Fearful Happy	0.97 (0.03) 0.96 (0.05)	0.99 (0.03) 0.98 (0.04)
69*	Wingenbach et al. (2017)	12 (9)	AQ; SCQ	17.30 (0.75)	-	12 (9)	Age Gender	16.90 (0.29)	-	Human faces (dynamic)	Verbal (forced-choice labelling)	Facial emotional video from The Amsterdam Dynamic Facial Expression Set (van der Schalk et al., 2011)	Angry Fearful Happy Sad Disgust Surprise	0.74 (0.08) 0.47 (0.08) 0.96 (0.02) 0.78 (0.04) 0.77 (0.07) 0.94 (0.03)	0.89 (0.08) 0.78 (0.05) 0.95 (0.05) 0.88 (0.02) 0.84 (0.05) 0.93 (0.03)
70*	Wong et al. (2012)	19 (16)	ADI-R	11.28 (1.48)	FSIQ: 118.21 (14.93)	21 (15)	Age FSIQ	10.24 (1.81)	FSIQ: 113.43 (11.21)	Human faces (static)	Verbal (forced-choice labelling)	Facial stimuli from the Standardized Penn Emotion Recognition Set (Gur et al., 2002)	Angry Fearful ^{X} Happy Sad Disgust	2.89 (0.66) 3.05 (0.91) 4.00 (0) 3.16 (0.96) 1.16 (0.96)	2.71 (0.96) 2.9 (1.14) 4.00 (0) 3.19 (0.98) 1.52 (1.25)

71*	Wright et al. (2008)	35 (33)	ADI-R; ADOS; AQ; ICD-10	11.31 (2.17)	FSIQ: 104.63 (17.99) VIQ: 105.66 (21.01) NVIQ: 103.03 (16.09)	35 (33)	Age Gender FSIQ VIQ NVIQ	11.57 (1.94)	FSIQ: 103.86 (16.26) VIQ: 105.74 (16.31) NVIQ: 100.94 (16.39)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Ekman's Facial Expressions of Emotion: Stimuli and Tests (FEEST)	Angry	5.57 (2.05)	6.57 (1.80)
													Fearful	4.77 (2.38)	4.74 (2.97)
													Happy	9.51 (1.17)	9.91 (0.28)
													Sad	6.91 (2.42)	6.43 (2.71)
													Disgust	5.40 (2.98)	5.03 (2.74)
													Surprise	8.09 (2.17)	8.77 (2.13)
													Overall	40.26 (9.56)	41.51 (7.99)
72*	Yeung et al. (2014)	18 (15)	ADI-R; DSM-IV	9.61 (3.13)	FSIQ: 101.33 (10.85)	18 (11)	Age Gender FSIQ	10.72 (3.61)	FSIQ: 107.06 (9.35)	Human faces (static)	Verbal (forced-choice labelling)	Photographs from the Karolinska Directed Emotional Faces	Angry	0.29 (0.16)	0.40 (0.17)
													Fearful	0.12 (0.11)	0.19 (0.14)
													Happy	0.95 (0.28)	1.12 (0.17)
													Sad	0.44 (0.19)	0.57 (0.17)
													Disgust	0.10 (0.11)	0.23 (0.19)
													Surprise	0.53 (0.21)	0.67 (0.14)
													Overall	0.41 (0.14)	0.53 (0.12)

Abbreviations: ADI-R, Autism Diagnostic Instrument Revised; ADOS, Autism Diagnostic Observation Schedule; ADI-R-III, AADOS-G, Autism Diagnostic Observation Schedule—Generic; ADOS-2, Autism Diagnostic Observation Schedule, Second Edition; ASDI, Autism Spectrum Diagnostic Interview; ASDS, Asperger Syndrome Diagnostic Scale; ASQ-J, Autism Screening Questionnaire, Japanese Version; AQ, Autism Spectrum Quotient Questionnaire; AQ-J, Autism Spectrum Quotient, Japanese Version; ASSQ, Autism Spectrum Screening Questionnaire; CARS, Childhood Autism Rating Scale; DCL-TES, Diagnostic Checklist for Pervasive Developmental Disorders (*Diagnostik Chekliste fr Tiefgreifende Entwicklungsstörungen*); CRS-R, Conners' Rating Scales-Revised; DSM, The Diagnostic and Statistical Manual of Mental Disorders; DSM-IV, The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; DSM-IV-TR, The Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision; DSM-5, The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition; GARS-2, Gilliam Autism Rating Scale, Second Edition; ICD-10, International Statistical Classification of Diseases, Tenth Revision; SCQ, Social Communication Questionnaire; SRS, Social Responsiveness Scale.

2.3.2. Quality assessment

The results of the critical appraisal process are summarised in Figure 2.2. The overall quality of the included studies was high, with total scores ranging from 28 to 35 out of the maximum score of 36 on the CASP case control study quality assessment (CASP, 2017). The strengths of the literature included the formulation of a clearly focused research question and the implementation of the appropriate methods to address the research aims. Selection bias was determined by the recruitment criteria of participants with ASD and NT controls. The use of standard diagnostic instruments during the selection of participants with ASD was of particular importance to reflect the representativeness of the clinical group. Among the 11 papers with moderate to high risk of bias, eight did not specify how the clinical group was diagnosed and three had very small samples. The lack of consistent matching procedures across studies might have compromised the generalisability of the findings. Thirty-nine papers did not match the ASD and NT groups on all three of the essential background measures, namely age, gender, and IQ, while two papers did not match groups on any of these measures. Measurement bias was observed in eight papers which could be characterised by unvalidated emotional stimuli, unstandardized experimental methods implemented across participants or conditions, and inappropriate measurement methods. Inappropriate reporting styles lowered the quality of a substantial proportion of the included studies. These related to the reporting of effect sizes which were missing in 31 papers as well as the reporting of imprecise *p*-values of inferential statistical tests in 23 papers. The reliability of results was also of concern for nine papers. Among those, five papers concluded their findings that were not justified by the inferential statistical tests conducted. Two papers did not conduct follow-up post-hoc tests to determine group differences within an interaction. One paper included contradictory information, e.g., groups described as matched by verbal IQ yet with a statistically significant group difference in verbal IQ. The remaining paper drew results based on a limited number of trials. The generalisability of the study findings to clinical populations was driven by the coverage of the range of participant demographics in terms of age and gender. While the sample in two papers covered both genders from different age groups, 61 papers covered both genders from only one age group (e.g., both female and male children) and six papers covered only one gender from the same age group (e.g., male adults only).

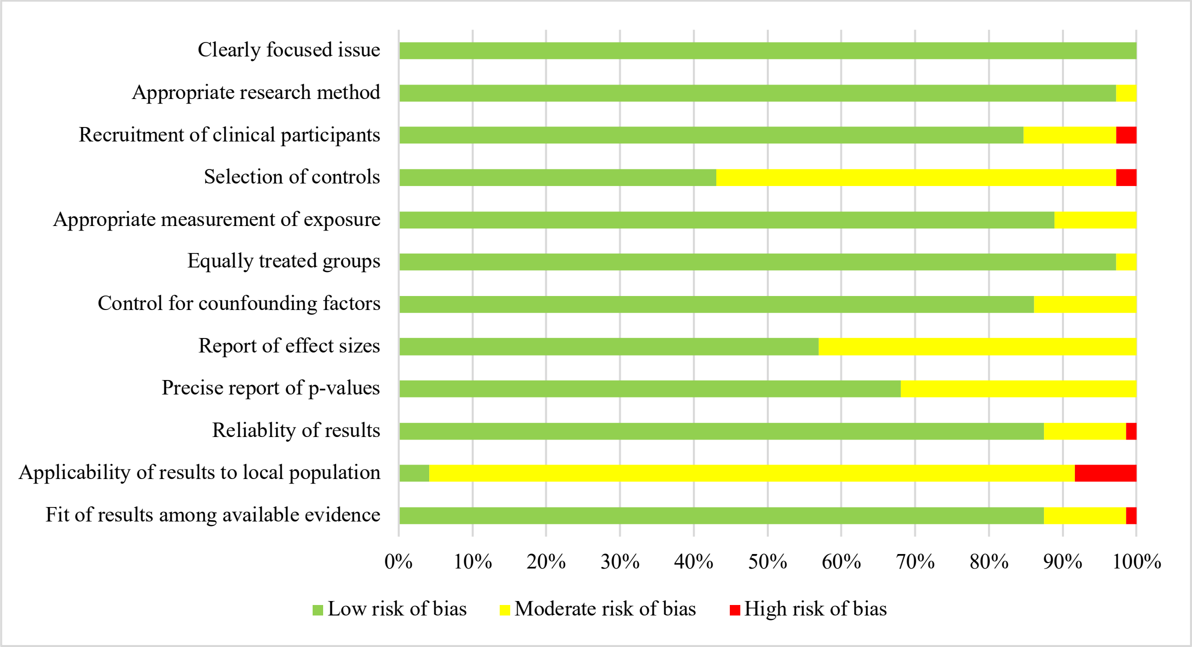


Figure 2. 2. Risk of bias graph. Reviewers' judgements about each risk of bias item presented as percentages across 72 included studies using the CASP (2017) checklist.

2.3.3. Main meta-analyses

2.3.3.1. Group differences in emotion recognition accuracy

Table 2. 5. Results summary of main meta-analyses on all accuracy datasets.

Emotion	***N***	**SMD**	**95% CI**	**95% PI**	***p***	**I^2**
Anger	**52**	**-0.42*****	**[-0.59, -0.24]**	**[-1.54, 0.70]**	**< 0.001**	**81.40%**
Fear	**46**	**-0.47*****	**[-0.66, -0.28]**	**[-1.60, 0.66]**	**< 0.001**	**82.07%**
Happiness	**52**	**-0.45*****	**[-0.61, -0.28]**	**[-1.51, 0.61]**	**< 0.001**	**79.03%**
Sadness	**48**	**-0.47*****	**[-0.65, -0.28]**	**[-1.57, 0.64]**	**< 0.001**	**80.87%**
Disgust	**20**	**-0.41*****	**[-0.58, -0.24]**	**[-0.89, 0.07]**	**< 0.001**	**38.89%**
Surprise	**19**	**-0.23****	**[-0.36, -0.10]**	**[-0.36, -0.10]**	**0.001**	**0.00%**
Composite	**22**	**-0.77*****	**[-1.03, -0.50]**	**[-1.87, 0.33]**	**< 0.001**	**77.56%**

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font.

Results from the main meta-analyses on accuracy were summarised in Table 2.5. The pooled SMDs for group differences in recognition accuracy were significant across all six basic emotions: anger ($N = 52$, SMD = -0.42, 95% CI [-0.59, -0.24], 95% PI [-1.54, 0.70], $p < 0.001$, $I^2 = 81.40\%$), fear ($N = 46$, SMD = -0.47, 95% CI [-0.66, -0.28], 95% PI [-1.60, 0.66], $p < 0.001$, $I^2 = 82.07\%$), happiness ($N = 52$, SMD = -0.45, 95% CI [-0.61, -0.28], 95% PI [-1.51, 0.61], $p < 0.001$, $I^2 = 79.03\%$), sadness ($N = 48$, SMD = -0.47, 95% CI [-0.65, -0.28], 95% PI [-1.57, 0.64], $p < 0.001$, $I^2 = 80.87\%$), disgust ($N = 20$, SMD = -0.41, 95% CI [-0.58, -0.24], 95% PI [-0.89, 0.07], $p < 0.001$, $I^2 = 38.89\%$), and surprise ($N = 19$, SMD = -0.23, 95% CI [-0.36, -0.10], 95% PI [-0.36, -0.10], $p = 0.001$, $I^2 = 0.00\%$). Significant group differences in recognition accuracy were also observed for the six-emotion composite score ($N = 22$, SMD = -0.77, 95% CI [-1.03, -0.50], 95% PI [-1.87, 0.33], $p < 0.001$, $I^2 = 77.56\%$), as shown in Figure 2.3. Depending on the emotion type, the pooled SMDs represented small (i.e., surprise), moderate (i.e., anger, fear, happiness, sadness, disgust), and large effects (i.e., composite), as can be seen in Figure 2.4. While heterogeneity was not observed for surprise, it was low for disgust but considerably high for anger, fear, happiness, sadness, and the composite. These results indicate that overall, the ASD group showed lower accuracy than the NT group in emotion recognition across all emotion types, and the size of these group differences varied by emotion type. Importantly, there was substantial heterogeneity across studies for most emotion types. Exploration for potential contributors to the observed high heterogeneity via moderator analyses is reported in Section 2.3.4.

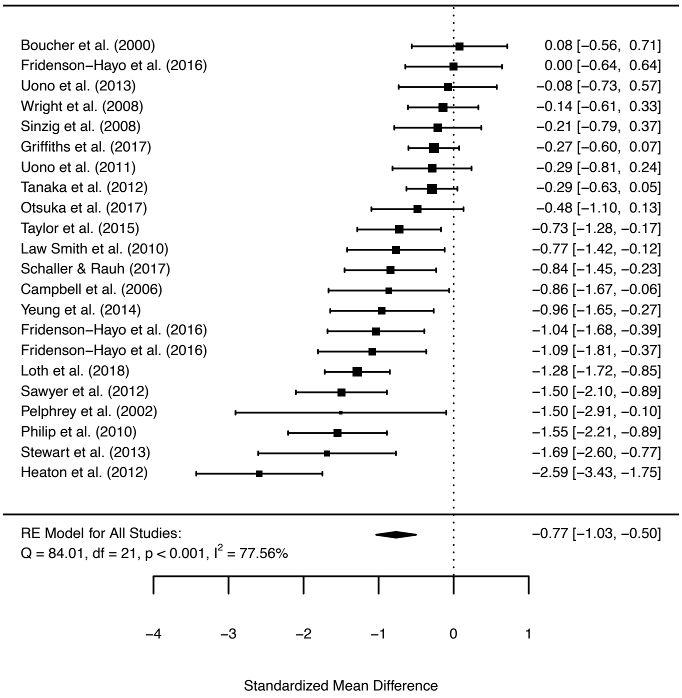


Figure 2. 3. Forest plot for six-emotion composite accuracy results. Twenty-two datasets contributed to this. The pooled SMDs indicated that the ASD groups performed worse than the NT groups, representing a large effect of -0.77 (95% CI [-1.03, -0.50]). The I^2 shows that there is significant heterogeneity between studies ($I^2 = 77.56\%$).

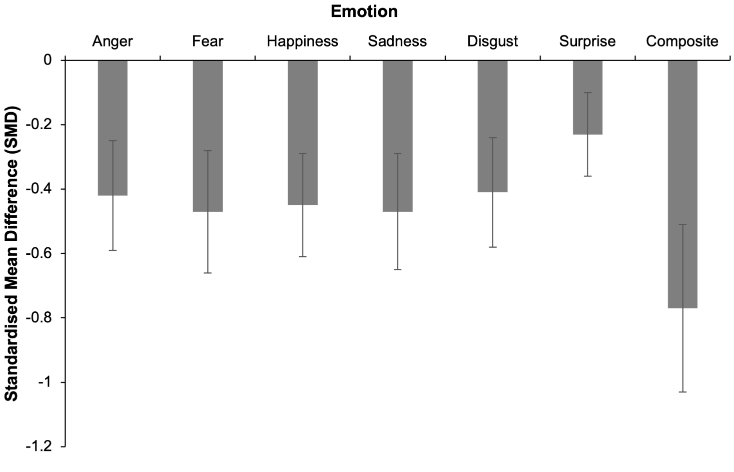


Figure 2. 4. Bar graph with means and 95% confidence interval error bars for the standardised mean difference (SMD) for each of the seven emotion categories.

The influence of IQ matching on meta-analysis results

To examine the impact of IQ matching on the robustness of the results in the main meta-analyses on accuracy, sensitivity analyses were performed on datasets that had implemented full-scale IQ matching (on all tasks), nonverbal IQ matching (on all tasks), and verbal IQ matching (on verbal tasks only) for each emotion category (see Table 2.6 for full results).

Table 2. 6. Results summary of sensitivity analyses on accuracy datasets with full-scale IQ matching and nonverbal IQ matching across both verbal and nonverbal tasks and verbal IQ matching on verbal tasks.

Emotion	*N*	SMD	95% CI	95% PI	*p*	I^2
Full-Scale IQ-Matched/ All Tasks						
Anger	**30**	**-0.28*****	**[-0.44, -0.12]**	**[-0.93, 0.37]**	**< 0.001**	**55.43%**
Fear	**26**	**-0.25****	**[-0.42, -0.08]**	**[-0.91, 0.40]**	**0.003**	**55.73%**
Happiness	**27**	**-0.38*****	**[-0.57, -0.19]**	**[-1.18, 0.41]**	**< 0.001**	**64.04%**
Sadness	**23**	**-0.27****	**[-0.46, -0.07]**	**[-1.00, 0.47]**	**0.007**	**61.87%**
Disgust	**13**	**-0.31****	**[-0.52, 0.11]**	**[-0.80, 0.17]**	**0.002**	**38.42%**
Surprise	**13**	**-0.28*****	**[-0.44, -0.13]**	**[-0.44, -0.13]**	**< 0.001**	**0.00%**
Composite	**9**	**-0.79****	**[-1.31, -0.26]**	**[-2.34, 0.77]**	**0.003**	**88.24%**
Nonverbal IQ-matched/ All Tasks						
Anger	**13**	**-0.35*****	**[-0.53, -0.17]**	**[-0.67, -0.03]**	**< 0.001**	**17.73%**
Fear	**17**	**-0.41*****	**[-0.62, -0.21]**	**[-1.02, 0.19]**	**< 0.001**	**48.22%**
Happiness	**15**	**-0.47*****	**[-0.64, -0.30]**	**[-0.81, -0.13]**	**< 0.001**	**20.21%**
Sadness	**11**	**-0.44****	**[-0.70, -0.18]**	**[-1.11, 0.24]**	**0.001**	**55.21%**
Disgust	**9**	**-0.44****	**[-0.73, -0.14]**	**[-1.17, 0.29]**	**0.004**	**59.22%**
Surprise	**7**	**-0.28****	**[-0.47, -0.08]**	**[-0.47, -0.08]**	**0.005**	**0.00%**
Composite	**12**	**-0.86*****	**[-1.23, -0.49]**	**[-2.05, 0.33]**	**< 0.001**	**80.70%**
Verbal IQ-matched/ Verbal Tasks						
Anger	**17**	**-0.35*****	**[-0.51, -0.18]**	**[-0.75, 0.05]**	**< 0.001**	**28.58%**
Fear	**18**	**-0.39*****	**[-0.58, -0.20]**	**[-0.93, 0.15]**	**< 0.001**	**41.49%**
Happiness	**18**	**-0.42*****	**[-0.57, -0.28]**	**[-0.65, -0.20]**	**< 0.001**	**7.85%**
Sadness	**13**	**-0.35****	**[-0.55, -0.15]**	**[-0.82, 0.12]**	**0.001**	**35.71%**
Disgust	**8**	**-0.39***	**[-0.69, -0.08]**	**[-1.11, 0.34]**	**0.014**	**59.30%**
Surprise	**9**	**-0.27****	**[-0.45, -0.09]**	**[-0.45, -0.09]**	**0.004**	**0.00%**
Composite	**11**	**-0.95*****	**[-1.39, -0.51]**	**[-2.33, 0.43]**	**< 0.001**	**83.74%**

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font. As no FSIQ-matched datasets on nonverbal tasks were available for disgust, surprise, and the composite, further analyses on the role of FSIQ beyond verbal tasks (e.g., FSIQ-matching on nonverbal tasks) were precluded.

In the sensitivity analyses including datasets that employed full-scale IQ matching, the pooled SMDs remained significant for all emotions: anger ($N = 30$, SMD = -0.28, 95% CI [-0.44, -0.12], 95% PI [-0.93, 0.37], $p < 0.001$, $I^2 = 55.43\%$), fear ($N = 26$, SMD = -0.25, 95% CI [-0.42, 0.08], 95% PI [-0.91, 0.40], $p = 0.003$, $I^2 = 55.73\%$), happiness ($N = 27$, SMD = -0.38, 95% CI [-0.57, -0.19], 95% PI [-1.18, 0.41], $p < 0.001$, $I^2 = 64.04\%$), sadness ($N = 23$, SMD = -0.27, 95% CI [-0.46, -0.07], 95% PI [-1.00, 0.47], $p = 0.007$, $I^2 = 61.87\%$), disgust ($N = 13$, SMD = -0.31, 95% CI [-0.52, 0.11], 95% PI [-0.80, 0.17], $p = 0.002$, $I^2 = 38.42\%$), surprise ($N = 13$, SMD = -0.28, 95% CI [-0.44, -0.13], 95% PI [-0.44, -0.13], $p < 0.001$, $I^2 = 0.00\%$), and the composite ($N = 9$, SMD = -0.79, 95% CI [-1.31, -0.26], 95% PI [-2.34, 0.77], *p* = 0.003, I^2 = 88.24%). In comparison with the main meta-analysis results, the size of the pooled SMDs decreased from moderate to small for anger, fear, happiness, sadness, and disgust after removing datasets without implementing full-scale IQ matching. The small effect for surprise and the large effect for the composite remained unchanged. Heterogeneity, although reduced, remained considerably high for anger, fear, happiness, and sadness, while the low heterogeneity for disgust and nil heterogeneity for surprise remained unchanged. Heterogeneity for the composite remained considerably high. These results indicate that, in comparison to the main meta-analysis results based on the full datasets, the sizes of the group differences were weakened when only including studies that matched groups on full-scale IQ. In addition, the reduced, yet high, heterogeneity observed here suggests that full-scale IQ matching may explain some of the variability across studies, but only to a certain extent.

In the sensitivity analyses including datasets that employed nonverbal IQ matching, the pooled SMDs remained significant for all emotions: anger (*N* = 13, SMD = -0.35, 95% CI [-0.53, -0.17], 95% PI [-0.67, -0.03], *p* < 0.001, I^2 = 17.73%), fear (*N* = 17, SMD = -0.41, 95% CI [-0.62, -0.21], 95% PI [-1.02, 0.19], *p* < 0.001, I^2 = 48.22%), happiness (*N* = 15, SMD = -0.47, 95% CI [-0.64, -0.30], 95% PI [-0.81, -0.13], *p* < 0.001, I^2 = 20.21%), sadness (*N* = 11, SMD = -0.44, 95% CI [-0.70, -0.18], 95% PI [-1.11, 0.24], *p* = 0.001, I^2 = 55.21%), disgust (*N* = 9, SMD = -0.44, 95% CI [-0.73, -0.14], 95% PI [-1.17, 0.29], *p* = 0.004, I^2 = 59.22%), surprise (*N* = 7, SMD = -0.28, 95% CI [-0.47, -0.08], 95% PI [-0.47, -0.08], *p* = 0.005, I^2 = 0.00%), and the composite (*N* = 12, SMD = -0.86, 95% CI [-1.23, -0.49], 95% PI [-2.05, 0.33], *p* < 0.001, I^2 = 80.70%). In comparison with the main meta-analysis results, the size of the pooled SMDs decreased from moderate to small for anger upon removing datasets without implementing nonverbal IQ matching. The small effect for surprise, moderate effects for fear, happiness, sadness, and disgust, and the large effect for the composite remained unchanged. Heterogeneity substantially reduced from high to low for anger, fear, and happiness but increased from low to high for disgust. The nil heterogeneity for surprise and the high heterogeneity for sadness as well as the composite remained unchanged. These results indicate that nonverbal IQ matching may explain some of the variability across studies, but only for certain emotions.

Given that verbal ability is required to label emotions in verbal tasks, sensitivity analyses examining the influence of verbal IQ matching were conducted on datasets of verbal tasks. The analyses showed that the pooled SMDs from VIQ-matched datasets using verbal tasks remained significant for all emotions: anger ($N = 17$, SMD = -0.35, 95% CI [-0.51, -0.18], 95% PI [-0.75, 0.05], $p < 0.001$, $I^2 = 28.58\%$), fear ($N = 18$, SMD = -0.39, 95% CI [-0.58, -0.20], 95% PI [-0.93, 0.15], $p < 0.001$, $I^2 = 41.49\%$), happiness ($N = 18$, SMD = -0.42, 95% CI [-0.57, -0.28], 95% PI [-0.65, -0.20], $p < 0.001$, $I^2 = 7.85\%$), sadness ($N = 13$, SMD = -0.35, 95% CI [-0.55, -0.15], 95% PI [-0.82, 0.12], $p = 0.001$, $I^2 = 35.71\%$), disgust ($N = 8$, SMD = -0.39, 95% CI [-0.69, -0.08], 95% PI [-1.11, 0.34], $p = 0.014$, $I^2 = 59.30\%$), surprise ($N = 9$, SMD = -0.27, 95% CI [-0.45, -0.091], 95% PI [-0.45, 0.09], $p = 0.004$, $I^2 = 0.00\%$), and the composite ($N = 11$, SMD = -0.95, 95% CI [-1.39, -0.51], 95% PI [-2.33, 0.43], $p < 0.001$, $I^2 = 83.74\%$). In comparison with the main meta-analysis results, the sizes of the pooled SMDs for anger, fear, sadness, and disgust decreased from moderate to small upon removing datasets without implementing verbal IQ matching. The effect for the composite was particularly strengthened and remained large, while the small effect for surprise and moderate effect for happiness remained unchanged. Heterogeneity substantially reduced from high to low for anger, fear, happiness, and sadness but increased from low to high for disgust. The nil heterogeneity for surprise and high heterogeneity for the composite remained unchanged. These results indicate that verbal IQ matching for verbal tasks may explain the variability across studies for most emotions.

Overall, these three sets of sensitivity analyses confirmed the robustness of the results of the main meta-analyses on accuracy: the pooled SMDs for group differences in recognition accuracy remained significant across all six basic emotions and the composite after the removal of datasets without IQ matching. The magnitude of group differences in emotion recognition was, nevertheless, weakened for most emotions when groups were matched on full-scale IQ (anger, fear, happiness, sadness, and disgust) and verbal IQ (anger, fear, sadness, and disgust), but remained relatively unchanged with nonverbal IQ matching (except for anger). Notably, group differences were particularly strengthened for the composite upon verbal IQ matching. Moreover, whether groups were matched on IQ appears to be important sources of variability across studies for anger, fear, happiness, and sadness, with verbal IQ matching showing a greater contribution above and beyond that by nonverbal and full-scale IQ matching.

2.3.3.2. Group differences in emotion recognition response time

Table 2. 7. Results summary of main meta-analyses on response times datasets.

Emotion	*N*	SMD	95% CI	95% PI	*p*	I^2
Anger	**18**	**0.45***	**[0.02, 0.88]**	**[-1.34, 2.24]**	**0.041**	**93.08%**
Fear	**17**	**0.57***	**[0.07, 1.07]**	**[-1.48, 2.62]**	**0.026**	**94.46%**
Happiness	18	0.33	[-0.26, 0.91]	[-2.17, 2.77]	0.271	96.09%
Sadness	**12**	**0.70***	**[0.08, 1.32]**	**[-1.37, 2.88]**	**0.027**	**95.57%**
Disgust	2	0.14	[-0.89, 1.17]	[-1.56, 1.83]	0.794	85.10%
Surprise	3	0.37	[-0.33, 1.08]	[-0.88, 1.62]	0.300	71.17%
Composite	**3**	**0.45***	**[0.09, 0.81]**	**[0.07, 0.85]**	**0.014**	**0.00%**

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font.

The pooled SMDs for group differences in recognition response time were significant and represented moderate effects for anger (*N* = 18, SMD = 0.45, 95% CI [0.02, 0.88], 95% PI [-1.34, 2.24], $p = 0.041$, $I^2 = 93.08\%$), fear (*N* = 17, SMD = 0.57, 95% CI [0.07, 1.07], 95% PI [-1.48, 2.62], $p = 0.026$, $I^2 = 94.46\%$), sadness (*N* = 12, SMD = 0.70, 95% CI [0.08, 1.32], 95% PI [-1.37, 2.88], $p = 0.027$, $I^2 = 95.57\%$), and the composite (*N* = 3, SMD = 0.45, 95% CI [0.09, 0.81], 95% PI [0.07, 0.85], $p = 0.014$, $I^2 = 0.00\%$). Heterogeneity was not observed for the composite but was substantially high for anger, fear, and sadness. The group differences in recognition response time, however, did not reach significance for happiness, disgust, and surprise (see Table 2.7 for full results). These results indicate that the ASD group was generally slower than the NT group at recognising anger, fear, sadness, and the composite emotions, with substantial amount of heterogeneity across studies. The two groups, nevertheless, showed comparable response time when recognising emotions of happiness, disgust, and surprise.

2.3.4. Moderator analyses

The following moderator analyses were only performed on recognition accuracy, but not on response time (due to limited data available).

2.3.4.1. Age

Table 2. 8. Results summary of meta-regressions on accuracy datasets with age (mean age of the ASD groups) as a moderator across the seven emotion categories, respectively.

Emotion	***N***	**β (SE)**	**Q_M**	***p***	**R^2**	**I^2**
Anger	52	0.00 (0.01)	0.13	0.723	0.00%	81.40%
Fear	46	0.00 (0.01)	0.13	0.716	0.00%	82.07%
Happiness	52	0.01 (0.01)	0.51	0.473	0.00%	79.03%
Sadness	48	0.01 (0.01)	1.10	0.294	0.48%	80.87%
Disgust	20	-0.01 (0.01)	1.31	0.252	7.04%	38.89%
Surprise	19	-0.01 (0.01)	0.70	0.401	0.00%	0.00%
Composite	**22**	**-0.05 (0.01)**	**16.22*****	**< 0.001**	**57.43%**	**77.56%**

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font.

Meta-regressions revealed that age (i.e., mean age of the ASD groups) was a significant predictor of the magnitude of the pooled SMD for the composite ($N = 22$, Q_M (1) = 16.22, $p <$ 0.001), accounting for 57.43% of the variability in the difference in SMDs among studies. As seen in Figure 2.5, age was negatively associated with the SMDs for the composite (β = -0.05, SE = 0.01): the older the participants with ASD, the greater difference in their recognition accuracy of the composite emotions relative to that in NT participants. Age was, however, not a significant predictor for the recognition of other emotions (see Table 2.8 for full results). These results suggest that the group difference increases with age for the composite measure, but not for the individual emotions.

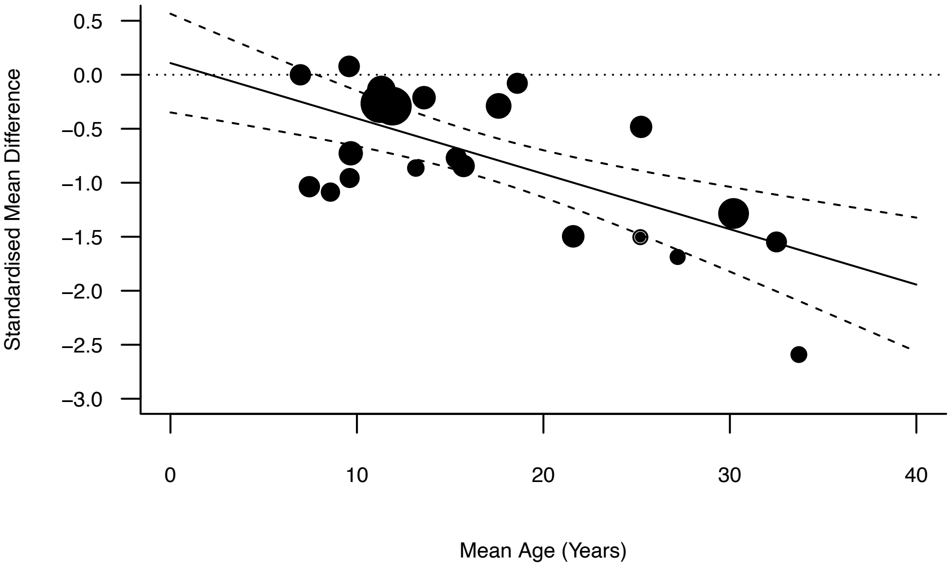


Figure 2. 5. Meta-regression scatterplot showing the standardised mean difference of the individual studies plotted against the mean age of the respective ASD group for the composite. Each point represents a study. The radius of the points is drawn proportional to the inverse of the standard errors (i.e., the precision of the effect estimates with larger/more precise studies shown as larger points). The predicted average standardised mean difference as a function of age is shown as the fitted regression line with corresponding 95% confidence interval bounds.

2.3.4.2. IQ

Separate meta-regressions across all the available datasets revealed that full-scale IQ, verbal IQ, and nonverbal IQ (i.e., mean standard scores of the ASD groups) were not significant moderators of the magnitude of the pooled SMDs for any of the seven emotion types (see Table 2.9 for full results). These results indicate that group differences in emotion recognition accuracy were not moderated by full-scale IQ, verbal IQ, or nonverbal IQ.

Table 2. 9. Results summary of meta-regressions on accuracy datasets with full-scale IQ, verbal IQ, and nonverbal IQ (mean standard scores of the ASD groups) as a moderator across the seven emotion categories, respectively.

Emotion	N	β (SE)	Q_M	p	R^2	I^2
Full-Scale IQ						
Anger	40	0.03 (0.02)	2.40	0.121	4.10%	80.88%
Fear	35	0.02 (0.02)	0.87	0.352	0.00%	80.95%
Happiness	38	-0.01 (0.02)	0.19	0.664	0.00%	83.46%
Sadness	35	0.01 (0.02)	0.47	0.494	0.00%	79.32%
Disgust	15	0.01 (0.02)	0.29	0.593	0.00%	29.11%
Surprise	15	0.01 (0.02)	0.44	0.507	0.00%	0.00%
Composite	12	-0.01 (0.05)	0.02	0.881	0.00%	85.93%
Verbal IQ						
Anger	20	0.05 (0.03)	2.51	0.114	11.04%	85.31%
Fear	20	0.01 (0.02)	0.20	0.656	0.00%	67.84%
Happiness	23	0.01 (0.02)	0.47	0.492	0.00%	84.16%
Sadness	20	0.02 (0.03)	0.68	0.408	0.00%	81.52%
Disgust	9	0.02 (0.03)	0.64	0.424	0.00%	45.67%
Surprise	11	0.00 (0.02)	0.00	0.947	0.00%	0.00%
Composite	11	-0.01 (0.03)	0.03	0.874	0.00%	83.48%
Nonverbal IQ						
Anger	20	0.02 (0.03)	0.45	0.503	0.00%	85.17%
Fear	19	0.00 (0.02)	0.01	0.924	0.00%	65.87%
Happiness	20	-0.02 (0.02)	0.51	0.475	0.00%	84.33%
Sadness	18	-0.01 (0.03)	0.09	0.759	0.00%	82.26%
Disgust	11	0.00 (0.02)	0.00	0.961	0.00%	54.10%
Surprise	11	0.01 (0.02)	0.12	0.731	0.00%	12.16%
Composite	10	-0.03 (0.05)	0.40	0.525	0.00%	82.98%

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font.

The influence of IQ matching on meta-regression results

To examine the impact of IQ matching on the robustness of the meta-regression results regarding IQ as shown above, sensitivity analyses were carried out on datasets that had implemented full-scale, verbal, or nonverbal IQ matching for each emotion category (see Table 2.10 for full results).

Table 2. 10. Results summary of sensitivity analyses on accuracy with full-scale IQ, verbal IQ, and nonverbal IQ as moderators across the seven emotion categories, after removing datasets without undertaking full-scale IQ matching, verbal IQ matching, and nonverbal IQ matching, respectively.

Emotion	N	β (SE)	Q_M	p	R^2	I^2
Full-Scale IQ on Full-Scale IQ-Matched						
Anger	28	0.02 (0.02)	1.54	0.214	4.91%	53.37%
Fear	23	0.01 (0.02)	0.18	0.672	0.00%	52.46%
Happiness	24	0.00 (0.02)	0.03	0.859	0.00%	68.85%
Sadness	21	0.00 (0.02)	0.07	0.791	0.00%	60.30%
Disgust	11	0.00 (0.02)	0.00	0.967	0.00%	40.05%

Surprise	11	0.01 (0.02)	0.41	0.521	0.00%	0.00%
Composite	8	-0.03 (0.07)	0.18	0.670	0.00%	89.36%
Verbal IQ on Verbal IQ-Matched						
Anger	12	0.02 (0.02)	1.23	0.267	21.38%	20.32%
Fear	14	0.00 (0.02)	0.04	0.840	0.00%	28.47%
Happiness	**15**	**-0.03 (0.01)**	**5.68***	**0.017**	**100.00%**	**0.00%**
Sadness	12	-0.01 (0.02)	0.13	0.719	0.00%	25.62%
Disgust	6	-0.01 (0.05)	0.03	0.858	0.00%	58.56%
Surprise	7	-0.01 (0.03)	0.27	0.606	0.00%	0.005%
Composite	6	0.00 (0.06)	0.00	0.993	0.00%	89.62%
Nonverbal IQ on Nonverbal IQ-Matched						
Anger	10	0.04 (0.03)	1.90	0.053	26.64%	31.90%
Fear	12	-0.01 (0.02)	0.25	0.168	0.00%	49.86%
Happiness	11	-0.01 (0.02)	0.70	0.618	0.00%	44.94%
Sadness	8	0.02 (0.04)	0.23	0.402	0.00%	65.17%
Disgust	8	0.00 (0.04)	0.01	0.631	0.00%	64.96%
Surprise	6	0.02 (0.03)	0.88	0.924	0.00%	0.00%
Composite	6	-0.21 (0.11)	3.73	0.349	40.47%	79.24%

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font.

In the sensitivity analyses including datasets that employed full-scale IQ matching and nonverbal IQ matching, full-scale IQ and nonverbal IQ remained as a nonsignificant predictor of the magnitude of the pooled SMDs for all emotions.

In the sensitivity analyses including datasets that employed verbal IQ matching, verbal IQ was a significant predictor of the magnitude of the pooled SMDs for happiness ($N = 15$, Q_M (1) = 5.68, $p = 0.017$), but not for other emotions. The negative association between verbal IQ and the SMDs for happiness (β = -0.03, SE = 0.01) indicated that the higher the verbal IQ of the participants with ASD, the lower the SMDs (i.e., poorer recognition accuracy of happiness in the ASD group relative to the NT group) (see Figure 2.6 for a scatter plot).

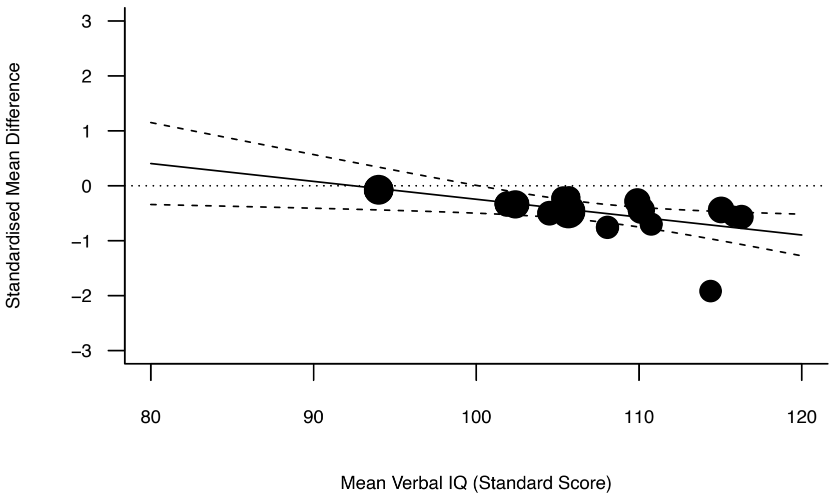


Figure 2. 6. Scatter plot showing the SMD of the individual studies plotted against the mean verbal IQ score of the respective ASD group for happiness in datasets that implemented verbal IQ matching between groups. Each point represents a study. The radius of the points is drawn proportional to the inverse of the standard errors (i.e., the precision of the effect estimates with larger/more precise studies shown as larger points). The predicted average standardised mean difference as a function of verbal IQ is shown as the fitted regression line with corresponding 95% confidence interval bounds.

2.3.4.3. Stimulus domain

Subgroup analyses were performed on human face, nonhuman face, speech prosody, and music domains separately for each emotion type (see Table 2.11 for full results).

Table 2. 11. Results summary for within-subgroup analyses and between-subgroup comparisons across domains.

Emotion	**Moderator**	***N***	**SMD**	**95% CI**	**95% PI**	**Q**	***p***	**I²**
Anger						4.66	0.098	
	Human face	**43**	**-0.47*****	**[-0.68, -0.27]**	**[1.68, 0.73]**		**< 0.001**	**84.26%**
	Nonhuman face	3	-0.11	[-0.49, 0.28]	[-0.53, 0.31]		0.591	6.09%
	Speech prosody	**12**	**-0.34***	**[-0.64, -0.03]**	**[-1.26, 0.59]**		**0.030**	**74.48%**
Fear						0.01	0.994	
	Human face	**41**	**-0.47*****	**[-0.68, -0.26]**	**[-1.69, 0.74]**		**< 0.001**	**84.26%**
	Speech prosody	6	-0.22	[-0.86, 0.43]	[-1.83, 1.39]		0.508	91.29%
	Music	**2**	**-0.50***	**[-0.92, -0.08]**	**[-0.92, -0.08]**		**0.021**	**0.00%**
Happiness						1.47	0.481	

	Human face	**41**	**-0.47*****	**[-0.66, -0.27]**	**[-1.57, 0.64]**		**< 0.001**	**81.36%**
	Nonhuman face	3	0.34	[-0.74, 1.43]	[-1.71, 2.40]		0.535	86.85%
	Speech prosody	**12**	**-0.52*****	**[-0.76, -0.28]**	**[-1.16, 0.13]**		**< 0.001**	**57.90%**
	Music	2	-0.32	[-0.86, 0.23]	[-1.04, 0.41]		0.252	38.73%
Sadness						1.91	0.384	
	Human face	**37**	**-0.48*****	**[-0.71, -0.26]**	**[-1.72, 0.75]**		**< 0.001**	**85.19%**
	Speech prosody	14	-0.16	[-0.50, 0.19]	[-1.34, 1.03]		0.372	81.58%
	Music	**2**	**-0.63****	**[-1.06, -0.21]**	**[-1.06, -0.21]**		**0.004**	**0.00%**
Disgust						0.61	0.435	
	Human face	**19**	**-0.40*****	**[-0.57, -0.23]**	**[-0.89, 0.10]**		**< 0.001**	**40.73%**
	Speech prosody	**2**	**-0.93*****	**[-1.40, -0.46]**	**[-1.40, -0.46]**		**< 0.001**	**0.00%**
Surprise								
	Human face	**19**	**-0.23****	**[-0.36, -0.10]**	**[-0.36, -0.10]**		**0.001**	**0.00%**
Composite						0.15	0.696	
	Human face	**20**	**-0.72*****	**[-0.95, -0.49]**	**[-1.57, 0.13]**		**< 0.001**	**67.76%**
	Speech prosody	6	-0.63	[-1.38, 0.23]	[-2.49, 1.23]		0.097	87.33%

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font. Not all domains had available datasets for each of the emotion categories. Domain subgroup comparisons were precluded for surprise, due to the limited datasets available for domains other than human faces

For studies on emotion recognition in human faces, subgroup analyses revealed that the ASD groups were significantly worse than NT controls across all emotions: anger (*N* = 43, SMD = -0.47, 95% CI [-0.68, -0.27], 95% PI [1.68, 0.73], $p < 0.001$, $I^2 = 84.26\%$), fear (*N* = 41, SMD = -0.47, 95% CI [-0.68, -0.26], 95% PI [-1.69, 0.74], $p < 0.001$, $I^2 = 84.26\%$), happiness (*N* = 41, SMD = -0.47, 95% CI [-0.66, -0.27], 95% PI [-1.57, 0.64], $p < 0.001$, $I^2 = 81.36\%$), sadness (*N* = 37, SMD = -0.48, 95% CI [-0.71, -0.26], 95% PI [-1.72, 0.75], $p < 0.001$, $I^2 = 85.19\%$), disgust (*N* = 19, SMD = -0.40, 95% CI [-0.57, -0.23], 95% PI [-0.89, 0.10], $p < 0.001$, $I^2 = 40.73\%$), surprise (*N* = 19, SMD = -0.23, 95% CI [-0.36, -0.10], 95% PI [-0.36, -0.10], $p = 0.001$, $I^2 = 0.00\%$), and the composite (*N* = 20, SMD = -0.72, 95% CI [-0.95, -0.49], 95% PI [-1.57, 0.13], $p < 0.001$, $I^2 = 67.76\%$). The pooled SMDs represented small (i.e., surprise), moderate (i.e., anger, fear, happiness, sadness, disgust), and large effects (i.e., composite). The results for disgust and surprise were associated with low heterogeneity, whereas the results for the other emotions were associated with high heterogeneity. These results suggest that, for studies investigating human faces alone, group differences were statistically significant across all emotions: (a) the ASD group showed lower recognition accuracy compared to the NT group across all seven emotion types; (b) the size of these group differences varied by emotion type; and (c) substantial heterogeneity across studies remained for most emotions.

In contrast to the above results on human faces, no significant group differences were found in the nonhuman face subgroup analyses for emotions with sufficient datasets, including anger ($N = 3$, SMD = -0.11, 95% CI [-0.49, 0.28], 95% PI [-0.53, 0.31], $p = 0.591$, $I^2 = 6.09\%$) and happiness ($N = 3$, SMD = 0.34, 95% CI [-0.74, 1.43], 95% PI [-1.71, 2.40], $p = 0.535$, $I^2 = 86.85\%$). These results indicate that group differences were not evident in studies investigating nonhuman faces alone, where the ASD and NT groups showed comparable accuracy in the recognition of anger and happiness, specifically.

Within the auditory modality, mixed results were obtained from the speech prosody subgroup analyses. The ASD groups were significantly worse than NT controls at recognition of anger ($N = 12$, SMD = -0.34, 95% CI [-0.64, -0.03], 95% PI [-1.26, 0.59], $p = 0.030$, $I^2 = 74.48\%$), happiness ($N = 12$, SMD = -0.52, 95% CI [-0.76, -0.28], 95% PI [-1.16, 0.13], $p < 0.001$, $I^2 = 57.90\%$), and disgust ($N = 2$, SMD = -0.93, 95% CI [-1.40, -0.46], 95% PI [-1.40, -0.46], $p < 0.001$, $I^2 = 0.00\%$) in speech prosody. These significant pooled SMDs represented small (i.e., anger), moderate (i.e., happiness), and large effects (i.e., disgust). While no heterogeneity was observed for disgust, results for anger and happiness were associated with high heterogeneity. The pooled SMDs for fear, sadness, and the composite, however, did not reach significance, while no datasets on surprise were available precluding analysis. These results indicate that, for studies investigating speech prosody alone, lower accuracy by the ASD group compared to the NT group was consistently found in the recognition of anger, happiness, and disgust, with the size of these group differences and amount of heterogeneity varying by emotion type. By contrast, the ASD and NT groups showed comparable accuracy in the recognition of fear, sadness, and the composite.

Similar to the speech domain, mixed results were obtained from the music subgroup analyses. The ASD groups were significantly worse than NT controls at recognising fear ($N = 2$, SMD = -0.50, 95% CI [-0.92, -0.08], 95% PI [-0.92, -0.08], $p = 0.021$, $I^2 = 0.00\%$) and sadness ($N = 2$, SMD = -0.63, 95% CI [-1.06, -0.21], 95% PI [-1.06, -0.21], $p = 0.004$, $I^2 = 0.00\%$) in music. The pooled SMD for happiness was, however, not significant, while there were no available datasets for anger, disgust, surprise, nor the composite. These results indicate that, for studies investigating music alone, lower accuracy by the ASD group compared to the NT group was consistently found in the recognition of fear and sadness, both with a moderate effect size and no heterogeneity across studies. By contrast, the ASD and NT groups showed comparable accuracy in the recognition of happiness in music.

The test for subgroup differences revealed that the pooled SMDs did not differ significantly across human faces, speech prosody, nonhuman faces, and music for any of the emotions (see Table 2.11 for full results). It should, however, be noted that subgroup comparison was not available for surprise due to a lack of datasets on domains other than human faces. Subgroup comparison incorporating all four domains was only feasible for happiness, with significant group differences observed for human faces and speech prosody, but not for nonhuman faces or music (see Figure 2.7 for forest plots).

.

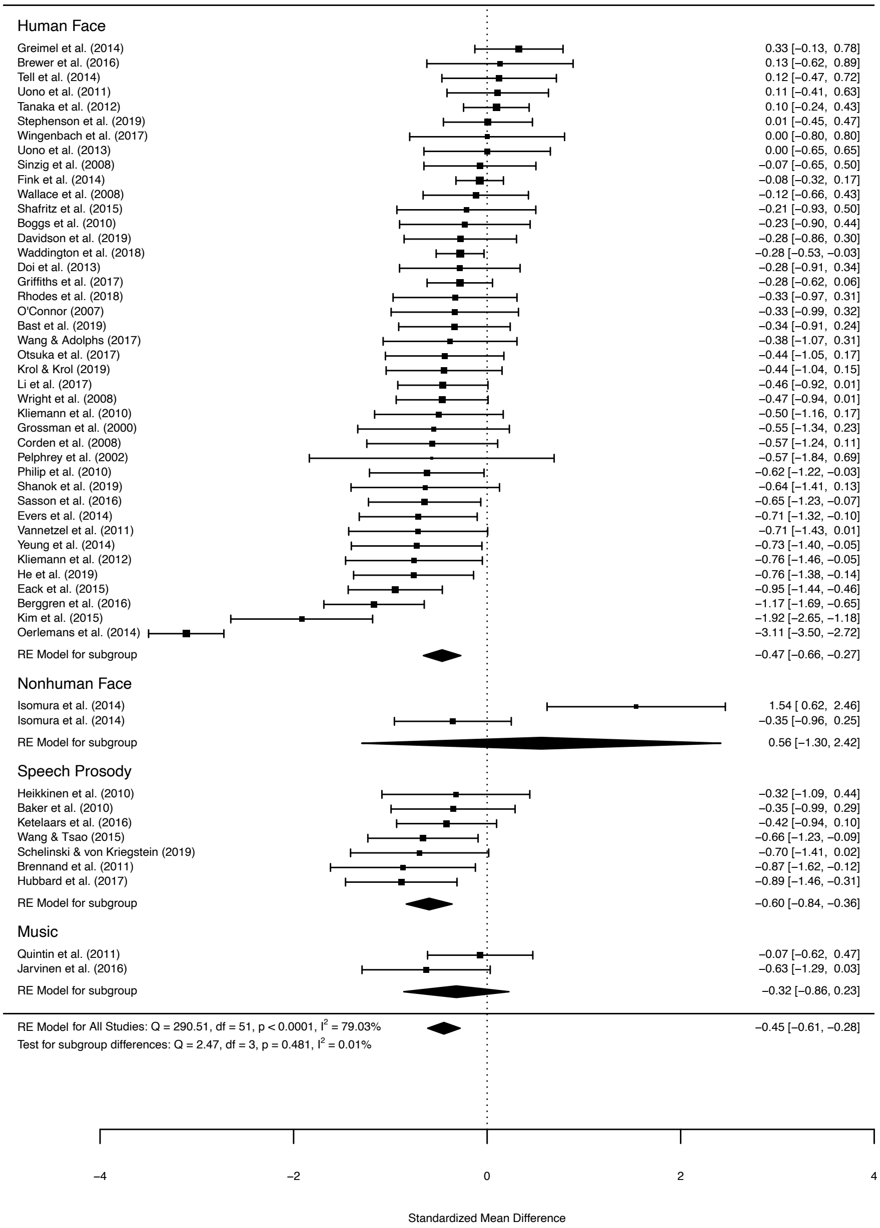


Figure 2. 7. Forest plots for happiness results by domain subgroup. This was the only emotion category that assessed all of the domains of interest. Forty-one datasets contributed to the human face subgroup, seven datasets to the speech prosody subgroup, two datasets to the nonhuman face subgroup, and two datasets to the music subgroup. The pooled SMDs was statistically significant for human faces and speech prosody with moderate effects but not for music and nonhuman faces. The test for subgroup differences however showed that the pooled SMDs did not differ significantly across subgroups.

2.3.4.4. Task demand

Subgroup analyses were done on verbal and nonverbal tasks separately for each emotion type (see Table 2.12 for full results).

Table 2. 12. Results summary for within-subgroup analyses and between-subgroup comparisons across task demands.

Emotion	Moderator	*N*	SMD	95% CI	95% PI	Q	*p*	I²
Anger						0.32	0.573	
	Verbal	**47**	**-0.44*****	**[-0.62, -0.27]**	**[-1.49, 0.61]**		**< 0.001**	**79.75%**
	Nonverbal	7	-0.34	[-0.97, 0.28]	[-2.00, 1.32]		0.286	88.51%
Fear						**25.38*****	**< 0.001**	
	Verbal	**45**	**-0.50*****	**[-0.68, -0.32]**	**[-1.54, 0.55]**		**< 0.001**	**79.56%**
	Nonverbal	3	-0.14	[-1.06, 0.78]	[-1.91, 1.64]		0.770	91.61%
Happiness						1.49	0.222	
	Verbal	**47**	**-0.49*****	**[-0.66, -0.32]**	**[-1.50, 0.52]**		**< 0.001**	**77.85%**
	Nonverbal	7	-0.08	[-0.57, 0.42]	[-1.34, 1.18]		0.760	81.89%
Sadness						0.17	0.685	
	Verbal	**45**	**-0.48*****	**[-0.66, -0.31]**	**[-1.52, 0.55]**		**< 0.001**	**78.88%**
	Nonverbal	6	-0.43	[-1.10, 0.24]	[-2.07, 1.21]		0.260	87.06%
Disgust								
	Verbal	**20**	**-0.41*****	**[-0.58, -0.24]**	**[-0.89, 0.07]**		**< 0.001**	**38.89%**
	Nonverbal	**2**	**-0.53****	**[-0.83, -0.23]**	**[-0.83, -0.23]**		**0.001**	**0.00%**
Surprise								
	Verbal	**19**	**-0.23****	**[-0.36, -0.10]**	**[-0.36, -0.10]**		**0.001**	**0.00%**
Composite								
	Verbal	**22**	**-0.77*****	**[-1.03, -0.50]**	**[-1.87, 0.33]**		**< 0.001**	**77.56%**

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant results are highlighted using bold font. Not all task demands had available datasets for each of the emotion categories. Task demand subgroup comparisons were precluded for disgust, surprise, and the composite, due to the limited datasets available for nonverbal tasks and prioritisation of multiple datasets from the same sample.

For verbal tasks, the pooled SMDs for group differences in recognition accuracy were significant across all emotions: anger ($N = 47$, SMD = -0.44, 95% CI [-0.62, -0.27], 95% PI [-1.49, 0.61], $p < 0.001$, $I^2 = 79.75\%$), fear ($N = 45$, SMD = -0.50, 95% CI [-0.68, -0.32], 95% PI [-1.54, 0.55], $p < 0.001$, $I^2 = 79.56\%$), happiness ($N = 47$, SMD = -0.49, 95% CI [-0.66, -0.32], 95% PI [-1.50, 0.52], $p < 0.001$, $I^2 = 77.85\%$), sadness ($N = 45$, SMD = -0.48, 95% CI [-0.66, -0.31], 95% PI [-1.52, 0.55], $p < 0.001$, $I^2 = 78.88\%$), disgust ($N = 20$, SMD = -0.41, 95% CI [-0.58, -0.24], 95% PI [-0.89, 0.07], $p < 0.001$, $I^2 = 38.89\%$), surprise ($N = 19$, SMD = -0.23, 95% CI [-0.36, -0.10], 95% PI [-0.36, -0.10], $p = 0.001$, $I^2 = 0.00\%$), and the composite ($N = 22$, SMD = -0.77, 95% CI [-1.03, -0.50], 95% PI [-1.87, 0.33], $p < 0.001$, $I^2 = 77.56\%$). These results were associated with small (i.e., surprise), moderate (i.e., anger, fear, happiness, sadness, disgust), and large effects (i.e., composite). While heterogeneity was not observed for surprise, it was low for disgust and high for anger, fear, happiness, sadness, and the composite. Thus, focusing on studies employing a verbal task, subgroup analyses suggested significant group differences across all emotion types: (a) the ASD group showed lower accuracy than the NT group; (b) the sizes of these group differences varied by emotion type; and (c) substantial heterogeneity remained for most emotions.

For nonverbal tasks, the pooled SMDs for group differences in recognition accuracy were only significant for disgust, with a moderate effect and no heterogeneity ($N = 2$, SMD = -0.53, 95% CI [-0.83, -0.23], 95% PI [-0.83, -0.23], $p = 0.001$, $I^2 = 0.00\%$). No significant group differences were found for anger, fear, happiness, or sadness, while no datasets were available for surprise or the composite precluding analysis. These results indicate that, for nonverbal tasks, the ASD group performed worse than the NT group only for recognition of disgust, but not for recognition of anger, fear, happiness, or sadness.

The tests for subgroup differences revealed that the pooled SMDs differed between verbal and nonverbal tasks only for fear ($Q(1) = 25.38$, $p < 0.001$), but not for anger, happiness, or sadness (see Table 2.12 for full results). Due to the limited datasets available for nonverbal tasks relating to surprise and the composite and prioritisation of multiple datasets from the same sample for disgust, subgroup comparisons were precluded for these emotions.

2.3.5. Publication bias

The Egger's test identified significant potential bias in the studies contributing to the pooled SMDs for fear ($z = -4.10$, $p < 0.0001$) and the composite ($z = -2.60$, $p = 0.009$), but not for anger ($z = -1.26$, $p = 0.207$), happiness ($z = -0.47$, $p = 0.635$), sadness ($z = -1.67$, $p = 0.095$), disgust ($z = -0.75$, $p = 0.453$), or surprise ($z = -0.47$, $p = 0.637$). For fear, an estimated number of one study was missing on the right side of the funnel plot. A trim-and-fill procedure corrected the observed pooled SMD in the meta-analysis to -0.44 (95% CI [-0.65, -0.22], 95%

PI [-1.77, 0.90], *p* < 0.001), which remained significant albeit with non-significant results in the test of null hypothesis that the number of missing studies is zero (*p* = 0.250). For the composite, the test of null hypothesis that the number of missing studies is zero was significant (*p* = 0.016). An estimated number of five studies were missing on the right side of the funnel plot (Figure 2.8). The trim-and-fill imputed mean effect for the composite remained significant with the observed pooled SMD corrected to -0.53 (95% CI [-0.84, -0.22], 95% PI [-2.02, 0.96], *p* = 0.001).

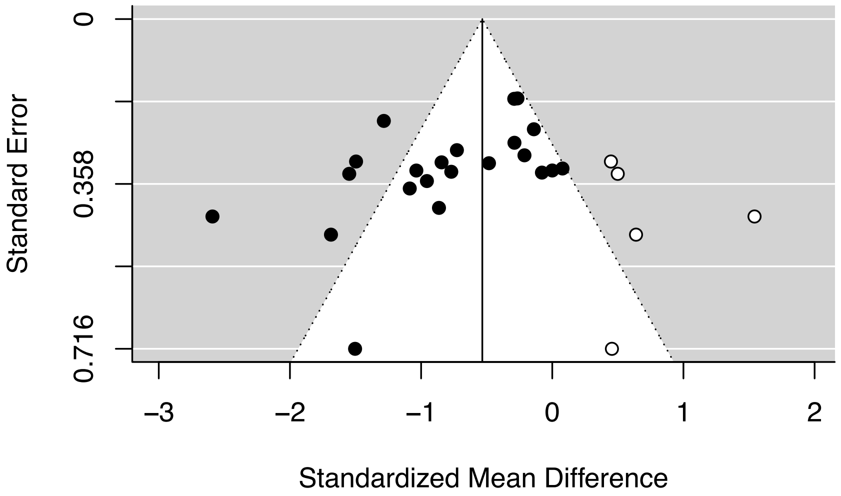


Figure 2. 8. Funnel plot of studies included in the meta-analysis for the composite. Black dots indicate observed studies and white dots indicate imputed studies correcting for funnel plot asymmetry.

2.4. Discussion

The current systematic review and meta-analysis evaluated the recognition of six basic emotions and their composite in ASD relative to typical development, across domains of human and nonhuman faces, speech prosody, and music, while identifying a number of potential moderating factors (age, IQ, domain, and task demand) that might have contributed to the mixed findings in the literature.

Combining non-overlapping datasets across the four domains, the main meta-analyses suggested emotion recognition impairments across all emotions in the ASD group, who also showed longer response times than NT controls for anger, fear, sadness, and the composite. Sensitivity analyses confirmed the robustness of the observed impairments in emotion recognition accuracy, as significant group differences remained for all emotions after removing datasets without implementing IQ matching. Nevertheless, the magnitude of these group differences was weakened for a subset of emotions across datasets with full-scale IQ matching (i.e., anger, fear, happiness, sadness, and disgust) and verbal IQ matching (i.e., anger, fear, sadness, and disgust), but less so for nonverbal IQ matching (i.e., anger). This indicates that while group differences in emotion recognition accuracy are not due to an absence of IQ matching, the magnitude of these differences could have been inflated due to the lack of IQ matching.

Moderator analyses indicated that age predicted the magnitude of group differences for the composite (but not for the individual emotions): the older the participants, the more pronounced the group differences. Domain was another significant moderator, as autistic individuals showed impaired recognition accuracy compared to controls across all emotions with human faces, but only for particular emotions with speech prosody (i.e., anger, happiness, and disgust) and music (i.e., fear and sadness), while no impairment was observed with nonhuman faces. Task demands also modulated group differences, with verbal tasks revealing group differences across all emotions and nonverbal tasks suggesting impairment of the ASD group for the recognition of disgust only. Finally, moderator analyses suggested no significant moderating effect of full-scale IQ, verbal IQ, and nonverbal IQ on group differences for any of the emotions. The nil effect of these IQ measures remained for datasets that had undertaken IQ matching, though there was one exception where verbal IQ significantly predicted the magnitude of group differences for happiness when focusing on verbal-IQ-matched datasets.

Specifically, the higher the verbal IQ of the ASD group, the more prominent the group differences when recognising happiness. These findings will be further discussed in the subsections below.

2.4.1. Age-related factors

The current finding of more pronounced impairments in adults for the composite has been described in Harms et al. (2010). Indeed, cross-sectional literature shows a lack of improvement in emotion recognition skills in ASD beyond late childhood, whereas maturation of skills continues through adulthood in typical development (Rump et al., 2009; Uono et al., 2011). In the present work, the age effect was only seen in the composite measure, which reflects greater task complexity as sophisticated categorical skills are required to distinguish all six basic emotions. The focus on the composite score also explains the discrepancy in the age effect between the current and prior reviews. In prior reviews, the overall measure comprised any numbers or combinations of emotions examined in the individual studies (Lozier et al., 2014; Uljarevic & Hamilton, 2013), which may reflect a reduced task complexity as fewer emotions were involved. As a result, no age effects were observed on the overall measures in prior reviews (Lozier et al., 2014; Uljarevic & Hamilton, 2013). Nevertheless, the age effects were found for individual emotions including fear, sadness, and disgust in Lozier et al. (2014), where the data came from studies that examined all six emotions. Thus, it is plausible that studies involving more emotions are more sensitive in revealing the age effects on differences between the ASD and NT groups.

In addition, research on the development of emotion recognition skills suggests that children generally achieve adult-level performance with prototypical expressions around 10-11 years of age across domains (Bruce et al., 2000; Chronaki et al., 2015; Mondloch et al., 2003; Tonks et al., 2007; Van Lancker et al., 1989; Vidas et al., 2018). However, the recognition of more subtle expressions requires considerable improvement in proficiency beyond this age (Chronaki et al., 2015; Rump et al., 2009). Given children with ASD have stronger preference for non-social over social stimuli (Gale et al., 2019) and musical over speech stimuli (Blackstock, 1978), as well as their lack of social motivation (Chevallier et al., 2012), it might be the case that they are not exposed to enough opportunities to develop and harness sophisticated emotional understanding like NT children do throughout adolescence and adulthood. Future longitudinal studies are needed to track the emotional development in autistic individuals, so effective interventions can be employed to increase emotional skills in ASD (Vogan et al., 2018).

2.4.2. IQ-related factors

Sensitivity analyses conducted on datasets undertaken full-scale/verbal/nonverbal IQ matching confirmed the robustness of the main meta-analysis results by showing that the group differences for all emotions remained significant. However, weakened magnitudes of group differences were observed for the majority of emotions among datasets with full-scale IQ matching (i.e., anger, fear, happiness, sadness, and disgust) and verbal IQ matching (i.e., anger, fear, sadness, and disgust), but less so for datasets with nonverbal IQ matching (i.e., anger). This indicates that matching groups on IQ could lower the magnitude of the observed group differences as well as reduce heterogeneity across studies. Thus, it is important for future studies to incorporate IQ matching procedures in the design, in order to provide more precise estimates of group effects on emotion recognition performance that are not influenced by differences in IQ between the ASD and NT groups.

In the moderator analyses regarding IQ, the meta-regressions on all available datasets corroborate previous meta-analytic works in failing to detect effects of full-scale IQ and verbal IQ on group differences for all emotions (Lozier et al., 2014; Uljarevic & Hamilton, 2013). Additionally, the current results suggested no significant effects of nonverbal IQ on group differences across studies. Focusing on datasets that matched groups on full-scale/

verbal/nonverbal IQ, sensitivity analyses still suggested non-significant effects of IQ on group differences across all emotions with only one exception – the higher verbal IQ of participants with ASD, the larger group differences in the recognition of happiness. These findings indicate that the magnitude of group differences in emotion recognition is largely unrelated to the intellectual level of autistic individuals. It should, however, be noted that this lack of IQ effect does not imply an absent relationship between IQ and emotion recognition in *absolute* terms; in other words, it does not mean that autistic individuals who have lower IQ perform similarly to autistic individuals who have higher IQ. Rather, these results indicate that the difference in performance between autistic individuals and their NT counterparts is largely unaffected by IQ differences in *relative* terms.

The effects of IQ on emotion recognition have predominantly been studied in *absolute* terms but rarely in *relative* terms in the autism literature. Among the limited studies investigating the *relative* effects of IQ on emotion recognition, Rommelse et al. (2015) found that group differences in performance on social cognition tasks (encompassing face recognition, facial and prosodic emotion recognition) were larger for autistic individuals who had higher full-scale IQ in comparison to autistic individuals who had lower full-scale IQ. The present finding of a lack of IQ effects across studies, thus, stands in contrast to the findings of Rommelse et al. (2015) – except for the observation of more prominent impairments for the recognition of happiness in autistic individuals who have higher verbal IQ when the groups are matched on this measure. It is particularly curious as to why such effects were only observed for happiness but not for other emotions across studies. Notably, in Rommelse et al. (2015), the ASD and NT groups were carefully matched on full-scale IQ, verbal IQ, nonverbal IQ, as well as verbal-nonverbal IQ discrepancy. The lack of IQ effects in previous reviews (Lozier et al., 2014; Uljarevic & Hamilton, 2013) and the present work may be explained by the potential confounds of verbal-nonverbal IQ discrepancy between ASD and NT groups, since IQ matching (mostly) did not influence the moderating role of IQ in the present findings. Given the substantial variability in cognitive profiles in ASD (Nowell et al., 2015; Tager-Flusberg & Joseph, 2003), it is plausible that different cognitive profiles influence performance in different ways such that individuals with discrepantly higher verbal IQ may make use of verbal abilities to succeed on labelling tasks, while performance may be hindered in individuals with discrepantly lower verbal IQ. For example, within the same pair of ASD and NT participants, despite having the same verbal IQ score, if the participant with ASD had a verbal > nonverbal IQ profile and the NT participant had a nonverbal > verbal IQ profile, the participant with ASD may be able to use their verbal ability to compensate for their emotion recognition difficulties and perform comparably to their NT counterparts. By contrast, if the participant with ASD had a nonverbal > verbal IQ profile while the NT participant had a verbal > nonverbal IQ profile, the participant with ASD may not be as readily able to make use of their verbal ability as a compensatory strategy and thus perform worse than their NT counterparts. The separate effects of full-scale IQ, verbal IQ, and nonverbal IQ on group differences may not fully account for the potential effects of verbal-nonverbal IQ discrepancy. As research investigating the effects of IQ on group differences in *relative* terms is limited, more research is warranted. Future research should also consider the intertwining relationship between verbal-nonverbal IQ discrepancy matching and IQ characteristics on emotion recognition in autistic individuals *relative* to NT individuals. This area of research has important implications for clinical practice if autistic individuals who have higher IQ indeed have more severe impairments in *relative* terms, as their emotion recognition capacities may risk being overestimated by their social environment due to their high cognitive abilities, which can contribute to the development of behavioural problems (Howlin, 1998).

2.4.3. Domain- and emotion-related factors

Examining data from four different domains, different conclusions were reached regarding whether emotion recognition impairments in ASD are specific to certain emotions. In the human face domain, impairments were observed across all emotions, similar to findings of previous meta-analyses (Lozier et al., 2014; Uljarevic & Hamilton, 2013). For nonhuman faces, however, no impairments were observed for any emotions. Whereas recognition of anger, happiness, and disgust was impaired for speech prosody, fear and sadness were the most difficult to recognise for autistic individuals when it came to music. Thus, in the visual modality, autistic individuals seem to have a general ability (in cases of nonhuman faces) or impairment (in cases of human faces) to recognise different emotions. In the auditory modality (in cases of speech and music), however, impairments were specific to certain emotions.

The different results for human faces versus nonhuman faces may be related to the specific perceptual processing strategies used to process these stimuli in autistic and NT individuals. As described in Section 1.2.2.1, configuration information has been found to play a particularly prominent role in typical emotion recognition for both human and nonhuman faces (Bombari et al., 2013; Calder et al., 2000; Derntl et al., 2009; McKelvie, 1995; Prkachin, 2003; Rosset et al., 2008). In support of this, an eye-tracking study demonstrated that the visual scan paths exhibited by NT individuals were strategic and controlled, which traced across the core features of emotional faces (Pelphrey et al., 2002). By contrast, the scan paths of autistic individuals were undirected and disorganised, where attention was often drawn to unimportant features (e.g., an ear) when processing emotional faces (Pelphrey et al., 2002). These qualitative differences may be attributed to the atypical featural processing for human faces (Deruelle et al., 2004; Hernandez et al., 2009) and/or avoidance of the eyes in autistic individuals (Frazier et al., 2017; Tanaka & Sung, 2016), which may in turn hamper emotion recognition from human faces. For nonhuman faces, both the use of typical configural processing with human cartoon faces (Rosset et al., 2008) and atypical featural processing with schematic faces (Isomura, Ogawa, et al., 2014) have previously been noted in ASD. Interestingly, regardless of the different processing strategies employed, autistic individuals performed comparably to their NT counterparts in those studies, specifically on both emotion recognition and detection tasks (Isomura et al., 2014; Rosset et al., 2008).

Although these findings together suggest impairments in facial emotion recognition are likely due to the atypical processing strategies employed by autistic individuals, the relationship between processing strategy and emotion processing of nonhuman faces is less clear. Critically, two questions are yet to be scrutinised in future research. First, given the disparity in the use of different processing strategies for nonhuman faces, how emotional expressions are decoded from nonhuman facial stimuli and whether they differ as a function of different types of nonhuman faces in autistic individuals need to be further examined, particularly through eye tracking techniques for more precise comparisons with that for human faces. Secondly, since intact emotion processing of nonhuman faces was observed regardless of the processing strategy employed by autistic individuals, it raises the question as to whether the advantage of configural processing is less critical for nonhuman faces (which does not share the same special status in perception as with human faces; Akdeniz, 2020; Rosset et al., 2010), and thus the use of either processing strategy may result in comparable performance. All in all, the literature on emotion processing of nonhuman faces is yet to be expanded, while further investigation should be directed to uncovering the processing strategies across the two visual domains and how they relate to emotion processing specifically.

The current findings outline the preserved skills to decode some but not all basic emotions from prosody and music in ASD. Compared to the visual modality, less is known about how auditory emotions are processed in ASD that may underlie the emotion-specific impairments observed. Interestingly, research has shown that autistic individuals perform just as well as NT individuals when emotions are solely cued by prosody (Brennand et al., 2011; Le Sourn-Bissaoui et al., 2013), though they tended to rely on alternative cues when presented alongside prosody for emotion recognition such as contextual cues and verbal content (Le Sourn-Bissaoui et al., 2013; Lindner & Rosén, 2006; Stewart et al., 2013). It is, therefore, plausible that autistic individuals are able to extract basic sensory features of emotional prosody, such as pitch – an important acoustic feature for inferring emotions in speech prosody (Juslin & Laukka, 2003; Pell & Kotz, 2011). As introduced in Section 1.3.2.2, pitch processing appears to be preserved, if not enhanced, in autistic individuals (Bonnel et al., 2003; Chowdhury et al., 2017; Globerson et al., 2015; but also see Bhatara, Babikian, et al., 2013; Kargas et al., 2015). While pitch processing has been shown to be related to prosodic emotion recognition in NT individuals (Globerson et al., 2013, 2015; Schelinski & von Kriegstein, 2019), this relationship is not clear for autistic individuals. Previous studies have reported both a significant relationship between pitch processing ability and prosodic emotion recognition (Globerson et al., 2015) and an absent relationship between the two variables in ASD (Schelinski & von Kriegstein, 2019). Thus, whether or not autistic individuals employ similar strategies to NT individuals such as pitch processing during prosodic emotion recognition remains to be elucidated. Additionally, where research has consistently reported anger and sadness being easier emotions to be recognised than fear and happiness in speech prosody (Chronaki et al., 2018; Elfenbein & Ambady, 2002; Juslin & Laukka, 2003), the specific impairments for anger, happiness, and disgust cannot seem to be fully explained by the difficulty levels of identifying these expressions. Thus, it raises another important question regarding how strategies employed for prosodic emotion recognition may contribute to the intact recognition of some but not all emotions in ASD. The way in which musical emotions are processed in autistic individuals relative to NT individuals has not been explored. Given the intricate overlap between speech prosody and music in communicating emotional signals

(Arbib, 2013; Patel, 2010), whether processing strategies are shared across these auditory domains are even less understood in ASD, and therefore more research is warranted in this area (Molnar-Szakacs & Heaton, 2012).

2.4.4. Experimental factors

Contrary to Uljarevic and Hamilton's (2013) findings of a lack of differences between labelling and matching paradigms, the present results revealed consistent impairments for verbal tasks but not for nonverbal tasks when detection and discrimination paradigms were also included. In addition, the moderating effect of task demand was particularly evident for the recognition of fear, with group differences being significantly more pronounced for verbal tasks compared to nonverbal tasks. Although verbal and nonverbal tasks are thought to involve the same core emotion recognition systems (Herba & Phillips, 2004; Phan et al., 2002), the two tasks also have substantial differences in their demands. Specifically, nonverbal tasks could be completed by discriminating perceptual characteristics between emotional expressions without necessarily understanding the emotional meaning of these expressions (Adolphs, 2002; Palermo et al., 2013), which may therefore lack sensitivity to detect group differences in the emotion understanding of different expressions. By contrast, verbal tasks not only require decoding emotion expressions based on perceptual properties as do nonverbal tasks, but also require access to emotion vocabulary for assigning emotional labels to the expressions (Palermo et al., 2013). The present findings, therefore, highlight that the particular difficulties autistic individuals have may be due to the linguistic demands for labelling of emotion expressions. It is, however, noteworthy that the inconsistent findings for nonverbal tasks may be due to a lack of sufficient data. More research is needed to establish whether and how different tasks moderate group differences.

2.4.5. Limitations and future research directions

Insufficient reporting of study data (*k* = 138) has resulted in limitation of the present analyses. In an era of open science, it is recommended that authors make their data accessible, in order to optimise data usage in the field. A considerable number of papers did not (consistently) report exact *p*-values, undermining accurate interpretation of results in relation to study hypotheses. To infer the importance of results, it is also recommended that researchers shift towards a meta-analytic thinking orientation and report effect sizes and confidence intervals along with statistical test results (Henson, 2006). The reporting of effect sizes (including nonsignificant results) allows explicit comparisons to be made between studies and enables all relevant data to be included in future syntheses, thus reducing potential influence of publication bias.

The generalisability of the present study sample characteristics may be compromised due to a lack of coverage of ages within individual studies. Only 8% of the papers studied both children and adults with ASD. When investigating the developmental trajectory of emotion recognition in ASD, it is important to control for heterogeneity brought by different tasks from different studies. More cross-sectional and longitudinal studies would be favoured to examine the effects of age on emotion recognition in ASD. Furthermore, with the growing number of interventions available for enhancing social communication skills in ASD (see Berggren et al., 2018 and Kouo & Egel, 2016 for reviews), it is likely that these studies have included participants who had undergone training programs prior to the experiments. It is speculated that any long-term training effects brought into experimental settings may have obscured the true effects of emotion recognition ability in ASD observed across studies in a subtle and inconsistent way. It may, therefore, be worthwhile for new studies to take note of the interventions that participants with ASD may have undertaken, such as in Wright et al. (2008).

The results of the main meta-analyses were based on evidence disproportionately distributed across domains (i.e., with the majority on human faces) and tasks (i.e., with the majority employing a verbal labelling task). The lack of data for the different categories further led to the decision on prioritising datasets based on domains and tasks that were most commonly used across studies, in an attempt to reduce heterogeneity. Subgroup comparisons, nonetheless, did not find significant differences between domain subgroups across emotions nor between task subgroups for happiness, sadness, and anger. This suggests that the impact of such prioritisation on the overall results is likely to be limited. Future meta-analytic work with more data in these under-researched areas will provide clearer indications of results and further insights into the effects of these moderators with increased statistical power.

A number of studies have explored emotion recognition ability at varying intensities in ASD. As intensity level was based on either ratings obtained from validation or different morphed continua used (e.g., between neutral and an emotional expression, Law Smith et al., 2010; Otsuka et al., 2017; or between emotions within a pair such as fear vs. happy, S. Wang & Adolphs, 2017), it was not possible to reliably compare results at low (25%) or intermediate (50%) intensity levels across studies. Thus, only findings at the highest intensity level were included in the current meta-analyses. One may, therefore, expect stronger effects to be revealed with data collated from lower intensity stimuli. However, despite taking a potentially less sensitive measure in detecting subtle group differences (cf. Mazefsky & Oswald, 2007; Rump et al., 2009), general emotion recognition impairments across emotions were still observed in the ASD groups.

Finally, there has been an ongoing debate on whether co-occurring alexithymia, but not ASD per se, may be responsible for emotion recognition difficulties documented in the ASD population (Bird & Cook, 2013; Cook et al., 2013; Kinnaird et al., 2019; see also Sivathasan et al., 2020 for a recent review). Alexithymia, characterised by a lack of fluency in identifying and describing one's own emotions and feelings (Bird & Cook, 2013), has been reported to be highly comorbid with ASD, affecting approximately 50% of the ASD population (Berthoz & Hill, 2005; Milosavljevic et al., 2016). Several studies have provided supporting evidence for the alexithymia hypothesis, by noting the significant relevance of alexithymia to emotion recognition difficulties with facial and prosodic expressions in ASD (Cook et al., 2013; Heaton et al., 2012; Ketelaars et al., 2016). Conversely, other studies have reported a lack of contribution of alexithymia to emotion recognition difficulties in ASD (Kliemann et al., 2013; Stephenson et al., 2019). Although there have been increasing efforts to consider alexithymia as a potential candidate in accounting for emotion recognition performance in autistic individuals over the past decade (Cook et al., 2013; Heaton et al., 2012; Keating et al., 2021; Ketelaars et al., 2016; Kliemann et al., 2013; Milosavljevic et al., 2016), available data for the different emotion types examined in the current review remain scarce. Adhering to the recommendations that meta-regressions should generally not be performed on data from less than 10 studies (Deeks et al., 2019), it was not possible to examine the moderating effect of alexithymia on group differences in emotion recognition in the present study. To enable a meta-analysis of the findings, future studies would need to include measures of alexithymia when investigating emotion recognition in ASD.

2.4.6. Chapter summary

In summary, this quantitative synthesis of the current literature found that autistic individuals demonstrate general emotion recognition impairments across six basic emotions and their composite, who also showed longer response times than NT individuals for anger, fear, sadness, and the composite. The general impairments in recognition accuracy were shown to be robust and were not driven by differences in IQ matching and stimulus presentation time restriction – though the severity of these impairments was less pronounced for a subset of emotions which full-scale IQ matching (i.e., anger, fear, happiness, sadness, and disgust) and verbal IQ matching (i.e., anger, fear, sadness, and disgust) had been undertaken. The results suggest that sample characteristics and experimental designs interact to give rise to the heterogeneity seen in the literature. By investigating all these factors simultaneously in a large dataset, with rigorous inclusion criteria and robust analysis procedures, the moderating effects of age, domain, task demand on emotion recognition in ASD relative to typical development were found. The group effect was more pronounced in adults with increased task demands. Although insufficient data prevent reliable conclusions to be drawn on the effect of domain, impairments were consistently found for human faces but not for nonhuman faces, with impairments for speech prosody and music being specific to certain emotions. Task demands moderated emotion recognition, with autistic individuals performing worse on tasks that required verbal knowledge about emotions. Full-scale/verbal/nonverbal IQ, by contrast, were not important moderators of group effects. Further work is needed to extend the literature particularly on emotion recognition of prosodic, musical, and nonhuman facial emotions in ASD, in order to draw an unbiased comparison across domains. Future research should also focus on the processing strategies autistic individuals employ for the different types of stimuli during both bottom-up and top-down processes. Given the positive consequences of learned strategies in fulfilling life experiences, such investigations will provide insights into the optimal contexts for autistic individuals to accomplish successful social interactions in daily life.

Chapter 3: A behavioural study of emotion recognition across multiple nonverbal communicative domains in ASD and NT

This chapter examines the domain generality and specificity of explicit emotion processing in autism spectrum disorder (ASD) – that is, whether emotion recognition ability generalises across communicative domains or it is specific to certain domain(s) in autistic individuals. The present study design takes into consideration the issues and recommendations discussed in Chapter 2 by examining (a) the recognition of emotions in domains other than human faces, and (b) the developmental trajectory of emotion recognition cross domains. In addition, the ASD and neurotypical (NT) control groups were individually matched on age, gender, verbal, and nonverbal abilities, to control for confounding effects and study quality limitations as outlined in Chapter 2.

3.1. Introduction

Emotions can be communicated via different channels through different visual and auditory cues, such as facial expressions (Ekman & Friesen, 1976; Etcoff & Magee, 1992; J. A. Russell et al., 2003), speech prosody (Juslin & Laukka, 2003; Koolagudi & Rao, 2012a, 2012b), instrumental music (Argstatter, 2016; Mohn et al., 2011), and singing (Atmaja & Akagi, 2020; Livingstone & Russo, 2018; K. R. Scherer et al., 2015; B. Zhang, Essl, et al., 2015). It is, therefore, paramount to examine the ability to process emotions from different sources when considering the general emotion recognition ability in autism spectrum disorder (ASD).

Collating data from the existing literature, Chapter 2 showed that emotion recognition may be moderated by communicative domain in ASD. The findings suggested general impairments in emotion recognition accuracy for human faces (i.e., for all six basic emotions and their composite), but only for a subset of emotions for speech prosody (i.e., anger, happiness, and disgust) and music (i.e., fear and sadness) in ASD. Conversely, no impairments were noted for nonhuman faces such as cartoon, caricature, and schematic faces in ASD. While it is tempting to suggest that emotion recognition impairments in ASD may be specific to some domains, this conclusion is not only based on data from separate studies of different samples and designs but is also limited by the insufficient evidence in some domains. Crucially, the distribution of research data across the different domains was highly disproportionate (see Section 2.3.4.3); there was insufficient data for domains other than human faces, particularly for nonhuman faces (with only three datasets available for anger and happiness, respectively) and music (with only two datasets available for fear, happiness, and sadness, respectively). The present study was, therefore, set out to address these issues by investigating the general emotion recognition ability in autistic individuals from a multi-domain perspective within the same study, while taking into account the potential effects of specific emotions.

The findings from Chapter 2 also highlighted the moderating role of age in the severity of emotion recognition impairments in ASD, a specific case for the composite of all six basic emotions (anger, fear, happiness, sadness, disgust, and surprise) but not for the individual emotions. Namely, differences in emotion recognition accuracy between autistic individuals and their NT counterparts were more pronounced in adults than in children. However, due to limited data available, it was not possible to investigate the moderating effects of age by domain. Further to this, the significant age effects observed for the six-emotion composite were only representative for the human face and speech prosody domains (combined), as no composite data were available for the nonhuman face and music domains to contribute to this analysis. As discussed in Section 1.2.3, the developmental trajectory of emotion recognition has been directly compared between autistic and NT individuals but only in the human face and speech prosody domains. These studies have consistently noted that emotion recognition improves less overtime in ASD compared to typical development (Gepner et al., 2001; Greimel et al., 2014; Kuusikko et al., 2009; Rump et al., 2009; Uono et al., 2011; Van Lancker et al., 1989), which may result in particularly prominent group differences among adults. It remains largely unknown whether the developmental trajectory of emotion recognition differs between autistic and NT individuals for nonhuman faces and music, though some evidence from separate studies involving different age groups provided some insights into this. In the nonhuman face domain, children and adolescents with ASD have shown comparable emotion recognition performance to their NT counterparts (Brosnan et al., 2015; Miyahara et al., 2007; Rosset et al., 2008). Similarly, in the music domain, children, adolescents, and adults with ASD performed similarly to their NT counterparts (Gebauer, Skewes, Westphael, et al., 2014; Heaton et al., 1999; Quintin et al., 2011). These findings appear to suggest intact ability to recognise emotions from nonhuman faces and music in ASD regardless of age. Taken together, it is plausible that the moderating effects of age on emotion recognition ability in ASD may differ as a function of domain, where the sparse research in understanding emotion recognition ability in ASD from a developmental perspective for the different domains underscores the need for an investigation of this topic.

Previous studies have not directly compared emotion recognition in the four domains of interest, namely human faces, nonhuman faces, speech prosody, and music, using the same participant sample within the same paradigm. The present study, thus, aimed to examine whether autistic individuals show emotion recognition impairments across these domains, and whether any impairments are modulated by age and specific emotion or the interaction of both. Such comparisons would give rise to the understanding of the domain specificity (or otherwise) of emotion recognition impairments in ASD, while controlling for the moderating effects of participant and experiment factors. The specific research questions are:

1. Are there differences in emotion recognition between the ASD and NT groups?

2. Are there differences in the developmental trajectory of emotion recognition between the ASD and NT groups?
3. Are differences, if any, in emotion recognition modulated by stimulus domain (human face vs. nonhuman face vs. speech prosody vs. song) and emotion (angry vs. scared vs. happy vs. sad)?

Based on previous research, it was hypothesised that there will be no group difference in emotion recognition for nonhuman faces and music, but the ASD group will be less proficient than the NT group in recognising emotions from human faces and speech prosody. Group differences in these domains may be modulated by specific emotion as a function of domain. There is evidence from previous reports suggesting specific impairments for negative emotions, particularly fear, in the human face domain (Pelphrey et al., 2002; S. Wallace et al., 2008) but for happiness in the speech prosody domain (Hubbard et al., 2017; J. E. Wang & Tsao, 2015). With fear and happiness being the more difficult emotions to recognise from human faces and speech prosody respectively in typical development (Chronaki et al., 2018; Elfenbein & Ambady, 2002; Pell, Paulmann, et al., 2009), it may be the case that autistic individuals show particular impairments for expressions decoded at higher difficulty levels respective to domains. Moreover, group differences for human faces and speech prosody will be particularly noticeable in adults due to discrepancy in the development of more fine-grained emotion recognition during adolescence between groups (Rump et al., 2009).

3.2. Methods

3.2.1. Participants

The sample consisted of 76 native English speakers: 14 children with ASD and 14 NT children aged 7-11 years, 11 adolescents with ASD and 11 NT adolescents aged 12-15 years, and 13 adults with ASD and 13 NT adults aged 16-56 years – note that the present study followed the classification of age cohorts in the Autism Spectrum Quotient (AQ; (Auyeung et al., 2008; Baron-Cohen et al., 2001, 2006). An optimal sample size was estimated through an a priori power analysis using G*Power 3.1 (Faul et al., 2007). It was determined that with a medium effect size (*f* = 0.25; Cohen, 1988), an alpha level of 0.05, and power of 0.80, a total sample size of 342 was required to detect the Diagnostic group (ASD vs. NT) × Age group (child vs. adolescent vs. adult) × Condition (face vs. object vs. prosody vs. song) × Emotion (angry vs. scared vs. happy vs. sad) interaction in the current study. Yet, due to challenges with participant recruitment, this study was only able to test 76 participants as aforementioned. Participants were recruited via word-of-mouth, local charities and organisations associated with supporting people with autism, campus advertisements, and social media. Written informed consent was obtained from all adult participants and caregivers of child and adolescent participants prior to entering the study. Assent was also obtained from child and adolescent participants. Children received a toy and caregivers and adults were provided with financial compensation for their participation. The study received ethical approval from the University of Reading Research Ethics Committee.

All participants reported normal hearing and normal or corrected-to-normal vision. All participants passed a hearing screening at octave frequencies between 500 and 4000 Hz at 25 dB HL except one male NT adolescent, who had reported normal hearing but did not pass at 4000 Hz in the left ear. It was ensured that the auditory stimuli presented to this participant was at a comfortable and audible volume. The emotion recognition performance of this participant was 91% in the song condition and 78% in the prosody condition, which confirmed that his performance was not affected by the reduced hearing ability in the left ear. Thus, the data of this participant remained in all analyses. All participants in the ASD groups had previously received an official clinical diagnosis of an ASD from clinicians. None of the NT participants had been diagnosed with ASD, as confirmed by their AQ scores (Auyeung et al., 2008; Baron-Cohen et al., 2001, 2006) and none reported a family history of ASD. NT participants in each age group were individually matched to the ASD group on chronological age and gender. The two groups were also statistically matched on their receptive vocabulary measured on the Receptive One-Word Picture Vocabulary Test – Fourth Edition (ROWPVT-4; Martin & Brownell, 2011), and nonverbal reasoning ability measured on the Raven's Standard Progressive Matrices (RSPM; Raven, 1983) on a group level. Table 3.1 summarises the participants' demographic characteristics.

Table 3. 1. Means and standard deviations for ASD and NT groups on demographic variables and background measures.

	ASD		NT			
	M	SD	M	SD	*t*	*p*
Child						
N	14		14			
Gender (M:F)	13:1		13:1			
Age (years)	9.47	1.27	9.39	1.28	0.17	0.863
ROWPVT-4 (raw)	129.29	15.30	124.79	16.41	0.75	0.460
ROWPVT-4 (standard score)	122.57	11.84	118.29	14.62	0.85	0.402
RSPM (raw)	42.50	7.25	43.71	6.06	-0.48	0.635
RSPM (percentile)	72.50	25.40	80.36	22.31	-0.87	0.393
AQ	**103.29**	**14.8**	**40.85**	**14.3**	**11.15**	**< 0.00**
EQ	**18.86**	**6.18**	**42.69**	**5.79**	**-10.35**	**< 0.001**
SQ	31.43	10.92	24.54	6.16	2.04	0.055
Adolescent						
N	11		11			
Gender (M:F)	8:3		8:3			
Age (years)	13.87	1.26	13.79	1.12	0.16	0.871
ROWPVT-4 (raw)	149.64	19.89	161.64	16.47	-1.54	0.140
ROWPVT-4 (standard score)	119.00	20.35	131.55	18.81	-1.50	0.149
RSPM (raw)	47.27	7.02	49.00	6.21	-0.61	0.548
RSPM (percentile)	56.82	30.02	64.09	29.22	-0.58	0.571
AQ	**38.36**	**3.38**	**16.82**	**7.69**	**8.50**	**< 0.001**
EQ	**12.55**	**5.65**	**40.27**	**11.25**	**-7.30**	**< 0.001**
SQ	45.55	13.72	35.45	11.82	1.85	0.080
Adult						
N	13		13			
Gender (M:F)	6:7		6:7			
Age (years)	35.80	14.39	35.95	14.17	-0.03	0.979
ROWPVT-4 (raw)	171.08	12.27	172.46	9.51	-0.32	0.751
ROWPVT-4 (standard score)	108.46	11.30	111.23	14.54	-0.54	0.593
RSPM (raw)	53.62	3.80	53.15	3.8	0.31	0.760
RSPM (percentile)	55.38	25.78	47.69	31.60	0.68	0.503
AQ	**41.46**	**4.29**	**14.31**	**6.29**	**12.86**	**< 0.001**
EQ	**20.15**	**5.63**	**50.38**	**13.82**	**-7.31**	**< 0.001**
SQ	**85.00**	**31.87**	**48.77**	**16.90**	**3.62**	**0.002**
TAS-20	**69.85**	**8.76**	**45.54**	**16.51**	**4.57**	**< 0.001**

Note. ROWPVT-4 = Receptive One-Word Picture Vocabulary Test – Fourth Edition (Martin & Brownell, 2011); RSPM = Raven's Standard Progressive Matrices (Raven 1983); AQ = Autism-Spectrum Quotient (Auyeung et al., 2008; Baron-Cohen et al., 2001, 2006); EQ = Empathy Quotient (Auyeung, Allison, Wheelwright, & Baron-Cohen, 2012; Auyeung et la., 2009; Baron-Cohen & Wheelwright, 2004); SQ = Systemizing Quotient (Auyeung et al., 2009; 2012; Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003); TAS-20 = The 20-Item Toronto Alexithymia Scale (Bagby, Parker, & Taylor, 1994); MBEMA = Montreal Battery of Evaluation of Musical Ability; MBEA = Montreal Battery of Evaluation of Amusia; significant differences between ASD and NT groups for the individual measures are highlighted using bold font.

3.2.2. Background measures

Before the experiment, each participant or caregiver on behalf of the child participants completed the AQ (Auyeung et al., 2008; Baron-Cohen et al., 2006, 2001), Empathy Quotient (EQ; Auyeung, Allison, Wheelwright, & Baron-Cohen, 2012; Auyeung et al., 2009; Baron-Cohen & Wheelwright, 2004), and Systemising Quotient (SQ; Auyeung et al., 2012, 2009; Simon Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003) respective to the age of the participants as a measure of their autistic traits, levels of empathy, and systemising cognitive styles, respectively. Additionally, adult participants completed the 20-item Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994) as a measure of their alexithymic traits. A battery of background tasks was administered in addition to the main emotion experiments in a random order, either during one single session or across multiple sessions. The background tasks measured receptive vocabulary (ROWPVT-4; Martin & Brownell, 2011), nonverbal reasoning ability (RSPM; Raven, 1983), cognitive processing style (Navon; Navon, 1977), musical (MBEA or MBEMA; Peretz et al., 2003, 2013), and pitch processing (full task details will be described in Chapter 5).

3.2.3. Design

This study adopted a mixed design, with between-subjects factors Diagnostic group (ASD vs. NT) and Age group (Child vs. Adolescent vs. Adult) and within-subjects factors Condition (Human face vs. Face-like object vs. Speech prosody vs. Song) and Emotion (Angry vs. Scared vs. Happy vs. Sad) as independent variables. Several steps were taken to closely match stimuli within the same modality in the present study. Face-like objects were selected as a representative of the nonhuman face domain. Previous studies investigating emotion recognition of nonhuman faces in ASD have mostly employed schematic and cartoons faces (Brosnan et al., 2015; Isomura, Ito, et al., 2014; Miyahara et al., 2007; Rosset et al., 2008). These stimuli are line drawings which may not represent the same level of featural complexity and variability as human faces, and thus may result in enhanced emotion recognition (Kendall et al., 2016). Accordingly, face-like objects may offer an alternative to address these issues for more valid comparisons with human faces, given their higher complexity presented in realistic images, as well as their appearance in a wide variety of scenes. Importantly, face-like objects can provide emotional cues through face-like features (Donath, 2001; L. F. Zhou & Meng, 2020). Likewise, the use of sung stimuli as a representative of the music domain allows for a close match in acoustic similarity to spoken stimuli upon controlling for possible confounding effects of timbre that may be brought in by instrumental music (Hailstone et al., 2009).

Although six basic emotions – anger, fear, happiness, sadness, disgust, and surprise – have been proposed to be universally recognised and associate with distinctive expressions across domains (Ekman & Friesen, 1976; Humphreys et al., 2007; Mohn et al., 2011; K. R. Scherer, 2003), disgust and surprise are not commonly expressed in music and have been shown to be recognised poorly in listener studies (Juslin et al., 2008; Juslin & Laukka, 2003; Kallinen, 2005). Therefore, four of the six basic emotions – anger, fear, happiness, and sadness – were selected for inclusion, providing their representativeness of expressions conveyed across all the domains of interest in the present study.

The dependent variable was performance on the simple forced-choice emotion labelling task, indexed by accuracy, RT, and efficiency, respectively. Importantly, this type of task was selected given its sensitivity in detecting differences in emotion recognition between the ASD and NT groups. As demonstrated in Chapter 2, group differences were consistently found on verbal tasks across studies for all emotions. Indeed, various types of verbal tasks have been employed to assess visual emotion recognition in previous studies (as introduced in Section 2.1.1.5), some of which may, however, be particularly complex to implement for auditory stimuli. For example, in detection tasks, the presentation of an auditory target among multiple distractors would create a demanding auditory scenario within a multi-talker environment

(Bronkhorst, 2015), which may have an impact on emotion recognition. The trial-by-trial presentation of a single stimulus within the forced-choice emotion labelling format, thus, allows for appropriate comparisons to be made on performance across modalities. Importantly, this type of task has been frequently employed in previous studies to assess emotion recognition of different types of stimuli across modalities (e.g., Livingstone & Russo, 2018) and is, thus, deemed appropriate to address the research questions of the present study.

The majority of studies take performance accuracy as the only measure to reflect emotion recognition ability in ASD. As fast decoding of emotions is required in order to adapt one's behaviour to the social situation, delayed emotion recognition might be problematic, even if accuracy is not affected. The few studies that examined response time (RT) have revealed both typical (Akechi et al., 2010; Fink et al., 2014; Waddington et al., 2018) and slower recognition speed in ASD (Greimel et al., 2014; Ketelaars et al., 2016; Sawyer et al., 2012), which was specifically found for anger, fear, and sadness in Chapter 2. Furthermore, studies including both accuracy and speed measures have reported impaired accuracy with slower RT mostly for faces (Berggren et al., 2016; Eack et al., 2015; Greimel et al., 2014; Kliemann et al., 2010), intact accuracy with slower RT mostly for speech prosody (Ketelaars et al., 2016; Lindström et al., 2018; Waddington et al., 2018), and intact accuracy and RT for both instrumental music (Quintin et al., 2011) and cartoon faces (Miyahara et al., 2007). It is plausible that the particular observation of intact accuracy with slower RT for speech prosody demonstrates potential speed-accuracy trade off (e.g., slower RT to facilitate accuracy; Chittka et al., 2009) that is not found across domains. Given that accuracy and speed are equally important in emotion recognition tasks, it is possible that participants (or groups) may adopt different response strategies (e.g., slower response to facilitate accuracy vs. faster response resulting in reduced accuracy). The use of a single measure of speed-accuracy composite score (SACS; Charbonneau et al., 2013; Collignon et al., 2010), could therefore discard speed-accuracy trade-off effects on performance. In other words, response biases favouring either accuracy or response time would result in lower scores than responses that are made accurately and quickly. This SACS measure, in addition to the separate measures of accuracy and RT, could inform whether any group differences are due to response biases rather than an impairment per se. The present study, thus, tested for differences in terms of accuracy, RT, and SACS.

3.2.4. Stimuli

3.2.4.1. Auditory stimuli

The auditory emotion set consisted of spoken and sung stimuli extracted from the audiovisual files from the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS; Livingstone & Russo, 2018). The audiovisual files depicted 12 male and 12 female actors saying or singing two sentences in North American English accent: "*dogs are sitting by the door*" and "*kids are talking by the door*". Stimuli that presented one of the four basic emotions (angry, scared, happy, and sad) expressed at normal and strong intensity levels were selected for this study. The spoken and sung sentences were segmented to only include the last word, "door". Initial durations of these stimuli varied but were subsequently equalised to 500ms and the intensity of all stimuli was normalised to 75 dB using Praat software (Boersma & Weenink, 2004).

An acoustic analysis was conducted on the speech and song stimuli with Praat. As duration and intensity were controlled for, particular attention was paid to the mean and excursion size of the fundamental frequency (f0), which is the main parameter signifying emotions in prosody (Bänziger & Scherer, 2005; Lieberman & Michaels, 1962; Quam & Swingley, 2012; Scherer, 1986) – see Table 3.2 for a summary of the acoustic characteristics for each auditory stimulus type by actor gender and emotion.

A three-way ANOVA with Actor gender (male vs. female), Stimulus type (prosody vs. song), Emotion (angry vs. scared vs. happy vs. sad), and all possible interactions as predictors, was conducted on mean f0 and excursion size, respectively.

Concerning the mean f0 (in Hz), there was a significant main effect of Actor gender ($F(1, 16) = 56.48, p < 0.001$), as overall female actors (M = 334.05, SD = 67.45) had a higher mean f0 than male actors (M = 180.83, SD = 31.26).

Concerning the excursion size (in semitones) of f0, there was a significant main effect of Stimulus type ($F(1, 16) = 123.69, p < 0.001$), as well as a significant interaction of Stimulus type × Emotion ($F(3, 16) = 4.18$, $p = 0.023$). Post-hoc analyses revealed no significant differences between emotions among each sound type. This interaction was, however, driven by the larger excursion size observed for prosody compared to song for the angry (prosody: M = 9.85, SD = 2.51; song: M = 2.64, SD = 1.31; $t(3) = 6.05, p = 0.009$), happy (prosody: M = 12.02, SD = 3.45; song: M = 2.41, SD = 0.33; $t(3) = 5.92$, $p = 0.010$), and sad emotions (prosody: M = 9.08, SD = 0.33; song: M = 2.89, SD = 1.47; $t(3) = 7.93, p = 0.004$), but not for the scared emotion.

Table 3. 2. Descriptive statistics of the mean and excursion size of the fundamental frequency for each auditory stimulus type by actor gender and emotion.

		Mean (Hz)				**Excursion size (semitone)**			
		Prosody		**Song**		**Prosody**		**Song**	
		M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Female	Angry	361.03	154.15	352.73	5.49	11.54	1.78	2.29	1.49
	Scared	368.55	21.61	366.77	9.38	7.20	3.74	3.47	0.67
	Happy	258.19	56.80	349.41	5.18	10.63	2.50	2.26	0.16
	Sad	297.26	136.78	348.68	1.15	9.14	0.24	3.84	1.70
Male	Angry	206.48	13.45	178.94	1.14	8.16	2.07	3.00	1.57
	Scared	222.21	12.19	176.98	4.34	6.35	1.83	2.73	0.58
	Happy	180.81	-	171.83	-	16.18	-	2.87	-
	Sad	131.15	49.43	173.74	3.40	9.03	0.50	1.94	0.06

3.2.4.2. Visual stimuli

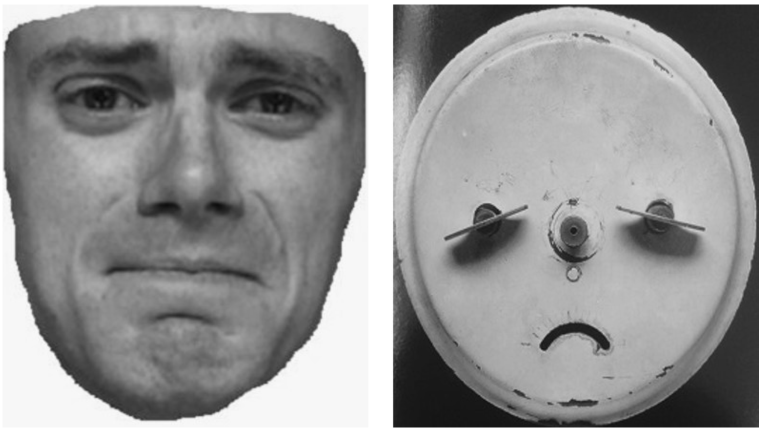


Figure 3. 1. Example of human face (left) and face-like object (right) stimuli presenting a sad emotion.

The visual emotion set comprised images of human faces and face-like objects, with each portraying one of the four basic emotions (angry, scared, happy, and sad) (Figure 3.1). The human face stimuli were also obtained from the audiovisual files from RAVDESS (Livingstone & Russo, 2018). The faces were extracted by taking screenshots of the actors when s/he was not talking but expressing the corresponding emotion that matched the spoken and sung emotions. Overly expressive faces were excluded to avoid ceiling effects. The faces were cropped using Microsoft Paint to only include expressive features of the face while removing other characteristics such as hair and ears. The face-like object stimuli were selected from the book *Faces* that comprised photographs of everyday objects that resemble human faces (Robert & Robert, 2000) and from the internet. All images of human faces and face-like objects were converted into grey scale using *ImageJ*, an image processing program (Abràmoff et al., 2004). They were then sized to 260 × 300 pixels using a MATLAB script, and the contrast and luminance of these images were normalised using the SHINE toolbox in MATLAB (Willenbockel et al., 2010).

3.2.3.3. Stimulus validation

Table 3. 3. Descriptive statistics of recognition rate and intensity rating for each stimulus type by emotion on the validation task.

		Recognition rate		Intensity rating	
		M	**SD**	**M**	**SD**
Face	Angry	0.85	0.12	4.30	0.73
	Scared	0.85	0.13	4.60	0.65
	Happy	0.95	0.05	4.41	0.60
	Sad	0.81	0.11	4.28	0.74
Object	Angry	0.86	0.12	5.03	0.62
	Scared	0.84	0.13	4.78	0.79
	Happy	0.95	0.06	4.67	0.71
	Sad	0.82	0.10	4.52	0.56
Prosody	Angry	0.91	0.15	5.19	0.91
	Scared	0.72	0.28	5.06	0.70
	Happy	0.74	0.24	3.86	1.04
	Sad	0.90	0.15	5.67	0.75
Song	Angry	0.88	0.17	4.83	0.67
	Scared	0.71	0.25	4.51	0.83
	Happy	0.72	0.33	4.36	0.84
	Sad	0.90	0.15	4.59	0.97

All stimuli were validated by 20 independent judges (3 males, 17 females; M = 25.27 years, SD = 7.34), who were asked to choose the appropriate label for each stimulus from the four emotional labels (angry, scared, happy, and sad). Judges also rated the intensity of the emotion expressed in each stimulus using a 7-point Likert scale, ranging from 1 (not intense at all) to 7 (very intense). The final set consisted of a total of 64 facial and 64 face-like object, 16 spoken, and 16 sung stimuli. These stimuli received an overall average recognition rate of 0.86 (SD = 0.35), demonstrating reliability of the emotional content of the selected stimuli. Emotion expressed at intermediate intensities with an overall average rating of 4.61 (SD = 1.38) were selected to avoid ceiling effects – see Table 3.3 for a full summary of the validation results for each condition by emotion.

A two-way ANOVA with Stimulus type (face vs. face-like object vs. prosody vs. song), Emotion (angry vs. scared vs. happy vs. sad), and their interaction as predictors was conducted on recognition rate. The analysis revealed significant main effects of Stimulus type (F(3, 304) = 2.74, p = 0.044), Emotion (F(3, 304) = 4.18, p = 0.006), as well as a significant interaction of Stimulus type × Emotion ($F(9, 304) = 4.41$, $p < 0.001$). Post-hoc analyses revealed that for face and object stimuli, the recognition rate for the happy emotion was the highest among other emotions. For prosody stimuli, the recognition rates for the angry and sad emotions were higher than those for the scared and happy emotions. For song stimuli, the recognition rate was higher for the angry and sad emotions than for the scared emotion – see Table 3.4 for full results summary.

A separate two-way ANOVA with Stimulus type (face vs. face-like object vs. prosody vs. song), Emotion (angry vs. scared vs. happy vs. sad), and their interaction as predictors was conducted on intensity rating. The analysis revealed significant main effects of Stimulus type ($F(3, 304) = 7.51$, $p < 0.001$), Emotion ($F(3, 304) = 7.17$, $p < 0.001$), as well as a significant interaction of Stimulus type × Emotion ($F(9, 304) = 5.48$, $p < 0.001$). Post-hoc analysis showed no difference in the intensity rating across emotions for face and song stimuli. For object stimuli, the intensity rating was higher for the angry emotion than for the happy and sad emotions. For prosody stimuli, the intensity rating for the happy emotion was lowest among other emotions – see Table 3.4 for full results summary.

Table 3. 4. Results of post-hoc analyses with Benjamini-Hochberg corrections for the interaction between Stimulus type and Emotion on validation recognition rate and intensity rating, respectively.

			Face		**Object**		**Prosody**		**Song**	
		df	*t*	*p*	*t*	*p*	*t*	*p*	*t*	*p*
Recognition rate										
Angry	Scared	19	0.13	0.896	0.31	0.762	**2.88**	**0.029**	**2.67**	**0.046**
Angry	Happy	19	**-2.87**	**0.020**	**-2.78**	**0.024**	**2.67**	**0.031**	1.98	0.093
Angry	Sad	19	1.11	0.422	1.25	0.339	0.33	0.888	-0.44	0.799
Scared	Happy	19	**-3.27**	**0.012**	**-4.01**	**0.002**	-0.14	0.888	-0.15	0.881
Scared	Sad	19	0.79	0.528	0.71	0.584	**-2.40**	**0.040**	**-3.00**	**0.044**
Happy	Sad	19	**4.95**	**< 0.001**	**5.30**	**< 0.001**	**-3.11**	**0.029**	-2.27	0.070
Intensity rating										
Angry	Scared	19	-1.91	0.213	1.81	0.129	0.59	0.561	2.37	0.088
Angry	Happy	19	-0.60	0.664	**3.53**	**0.007**	**5.51**	**< 0.001**	2.35	0.088
Angry	Sad	19	0.14	0.887	**6.44**	**< 0.001**	-1.83	0.099	1.28	0.432
Scared	Happy	19	1.33	0.400	0.76	0.457	**5.19**	**< 0.001**	0.89	0.458
Scared	Sad	19	1.95	0.213	1.85	0.129	**-2.99**	**0.011**	-0.43	0.670
Happy	Sad	19	0.69	0.664	1.55	0.164	**-6.69**	**< 0.001**	-1.03	0.458

Note. Significant effects and interactions are highlighted using bold font.

3.2.5. Apparatus

The experiment was run using E-prime 2.0 (E. Schneider & Zuccoloto, 2007). Auditory stimuli were presented through Sennheiser HD280 pro headphones and a Roland RUBIX22 USB Audio Interface at a comfortable volume. Responses were made on a Cedrus RB-740 response pad with coloured key cap lenses indicating each of the corresponding emotion categories: red for angry, green for scared, yellow for happy, and blue for sad. The position of the corresponding keys was reversed between participants but was held constant throughout the experiment for each participant (i.e., on both the recognition and priming tasks which will be described in Chapter 4). The experiment was conducted in a sound-proof booth.

3.2.6. Procedure

The simple forced-choice emotion labelling task assessed participants' emotion recognition across four stimulus types in separate blocks, with a total of 320 trials: 32 in the song, 32 in the prosody, 128 in the face, and 128 in the face-like object condition. Two practice trials preceded the start of each condition. Trials were pseudo-randomised with each stimulus presenting twice over separate blocks within each condition. Two versions of pseudo-randomisation were adopted and counterbalanced between participants. On each trial, an auditory stimulus (prosody or song) or a visual stimulus (face or face-like object) was presented. Participants were instructed to decide as quickly and accurately as possible which of the four emotional labels (angry, scared, happy, or sad) best described the emotion presented in the stimulus – note that the label "scared" was used instead of "fearful" as suggested by child participants in the pilot study to be a more commonly used term. Responses were made by pressing the corresponding key of the chosen emotion on the response pad. No response time limits were stipulated and no feedback was given regarding the accuracy of judgment on each trial. The presentation of the stimuli was terminated as soon as a response was made. The order of the condition presented to participants was counterbalanced across participants.

3.2.7. Statistical analyses

All analyses were performed in R (RStudio Team, 2018). Three performance measures of interest were computed: mean accuracy, mean response time (RT), and mean speed-accuracy composite score (SACS). As all participants selected the correct emotional label at above the chance level of 0.25 overall and for each condition respectively, all accuracy data were retained for analyses. Mean accuracy was calculated for each participant by condition and emotion. RT, measured from stimulus onset, was based on correct responses only. RTs less than 150ms or more than 2.5 SD of the mean of each participant for each condition were excluded. The exclusion of incorrect responses (4905 observations; 20% of total observations) and RT outliers (665 observations; 3% of correct-response observations), resulted in a dataset of 18750 observations across all participants (77% of the original number of observations). The mean RTs for each participant by condition and emotion were subsequently calculated. With SACS, to take into account both accuracy and RT measures, the mean accuracy and RTs were normalised (M = 0, SD = 1) by condition and the normalised RTs were subsequently subtracted from the normalised mean accuracy scores [Z(Accuracy) – Z(RT) = SACS]. A high composite score reflects efficient performance (i.e., high accuracy coupled with short RTs), and a low composite score reflects poor performance (i.e., low accuracy coupled with long RTs).

Three separate linear mixed effects models were constructed using the *lme4* package (Bates et al., 2015) to analyse the three dependent measures: arcsine-transformed mean accuracy, log-transformed mean RT, and mean SACS. Each model included Diagnostic group (ASD vs. NT), Age group (child vs. adolescent vs. adult), Condition (face vs. face-like object vs. prosody vs. song), Emotion (angry vs. scared vs. happy vs. sad), and all possible interactions as fixed effects. The maximal random effects structure was initially specified with by-subject intercept and by-subject slopes for condition and emotion for all three models. However, steps were taken to address convergence issues for the mean accuracy and mean SACS models following recommendations by Brown (2020). Where necessary, this included removing the correlation between the by-subject intercept and by-subject slopes, increasing the number of iterations and/or turning off derivation calculations. Model specifications are listed in the corresponding results summary tables.

For all the linear mixed effects analyses, the statistical significance of the fixed effects were obtained using the *anova*() function from the *lmerTest* package (Kuznetsova et al., 2017). For effect sizes, partial eta-squared (η^2_p) was computed for each fixed effect using the *effectsize* package (Ben-Shachar et al., 2020). A $\eta^2_p \geq 0.01$ was interpreted as a small effect size, a $\eta^2_p \geq 0.09$ as a medium effect size, and a $\eta^2_p \geq 0.25$ as a large effect size (Cohen et al., 2013). Significant effects and interactions emerging from the mixed effects models, were followed up through post-hoc pairwise comparisons using the *pairwise_t_test* function from the *rstatix* package (Kassambara, 2020). Correction of the post-hoc tests for multiple comparisons was performed with Benjamini-Hochberg (false discovery rate) procedure (Benjamini & Hochberg, 1995).

To explore whether there were any systematic group differences between the ASD and NT groups in error patterns, a 4 × 4 confusion matrix was calculated between the target emotion and response provided for each condition. This matrix was calculated separately for each diagnostic group. A chi-squared test (χ^2) was used to compare differences in the off-diagonal elements (i.e., representing a mismatch between the target emotion and response provided) between the two groups.

3.3. Results

3.3.1. Accuracy

Table 3. 5. Linear mixed effects model results for diagnostic group, age group, condition, emotion, and their interactions on mean arcsine-transformed accuracy.

Fixed effects	df	*F*	*p*	η^2_p

Diagnostic group	1	69.99	0.48	0.489	0.01
Age group	**2**	**69.99**	**6.84**	**0.002**	**0.16**
Condition	**3**	**114.22**	**21.54**	**< 0.001**	**0.36**
Emotion	**3**	**79.18**	**95.92**	**< 0.001**	**0.78**
Diagnostic group × age group	2	69.99	0.13	0.876	0.00
Diagnostic group × condition	3	114.22	1.37	0.255	0.03
Age group x condition	6	114.22	1.83	0.099	0.09
Diagnostic group × emotion	3	79.18	0.30	0.822	0.01
Age group × emotion	**6**	**79.18**	**3.56**	**0.004**	**0.21**
Condition × emotion	**9**	**769.06**	**25.05**	**< 0.001**	**0.23**
Diagnostic group × age group × condition	6	114.22	0.22	0.970	0.01
Diagnostic group × age group × emotion	6	79.18	1.49	0.192	0.10
Diagnostic group × condition × emotion	**9**	**769.06**	**2.36**	**0.012**	**0.03**
Age group × condition × emotion	**18**	**769.06**	**1.85**	**0.017**	**0.04**
Diagnostic group × age group × condition × emotion	18	769.06	0.46	0.973	0.01

Note. R model equation: lmer (asin (sqrt (Accuracy)) ~ Diagnostic Group * Age Group * Condition * Emotion + (1 | Subject) + (0 + Condition + Emotion | Subject), control = lmerControl(optCtr = list(maxfun = 1e9), calc.derivs = FALSE)); significant effects and interactions are highlighted using bold font.

Table 3.5 displays the full results summary of the linear mixed effects analysis on arcsine-transformed mean accuracy data. The analysis showed significant main effects of Age group (F(2, 69.99) = 6.84, p = 0.002, η^2_{p} = 0.16), Condition (F(3, 114.22) = 21.54, p < 0.001, η^2_{p} = 0.36), and Emotion (F(3, 79.18) = 95.92, p < 0.001, η^2_{p} = 0.78); the interactions of Age group × Emotion (F(6, 79.18) = 3.56, p = 0.004, η^2_{p} = 0.21), Condition × Emotion (F(9, 769.06) = 25.05, p < 0.001, η^2_{p} = 0.23), Diagnostic group × Condition × Emotion (F(9, 769.06) = 2.36, p = 0.012, η^2_{p} = 0.03), and Age × Condition × Emotion (F(18, 769.06) = 1.85, p = 0.017, η^2_{p} = 0.04) were also significant. No other factors or interactions were significant. I will unpack both the three-way interactions below.

3.3.1.1. Diagnostic group × Condition × Emotion

This Diagnostic group × Condition × Emotion interaction was examined using linear mixed effects models for each condition, with mean arcsine-transformed accuracy as the dependent measure. Each model included Diagnostic group, Emotion, and their two-way interaction as fixed effects, and with by-subject intercept as random effects. The results of the models are summarised in Table 3.6.

Table 3. 6. Linear mixed effects model results for diagnostic group, emotion, and their interactions on mean arcsine-transformed accuracy for each condition separately.

	Fixed effects	df		*F*	*p*	η^2_p
Face						
	Diagnostic group	1	74	0.86	0.356	0.01
	Emotion	**3**	**222**	**63.31**	**< 0.001**	**0.46**
	Diagnostic group × emotion	3	222	0.65	0.584	0.00
Object						
	Diagnostic group	1	74	0.34	0.559	0.00
	Emotion	**3**	**222**	**44.16**	**< 0.001**	**0.37**
	Diagnostic group × emotion	**3**	**222**	**2.97**	**0.033**	**0.04**
Prosody						
	Diagnostic group	1	74	0.175	0.677	0.00
	Emotion	**3**	**222**	**40.82**	**< 0.001**	**0.36**
	Diagnostic group × emotion	3	222	2.29	0.079	0.03
Song						
	Diagnostic group	1	74	1.53	0.219	0.02
	Emotion	**3**	**222**	**31.21**	**< 0.001**	**0.30**
	Diagnostic group × emotion	3	222	0.72	0.542	0.00

Note. R model equation: lmer (asin (sqrt (Accuracy)) ~ Diagnostic Group * Emotion + (1 | Subject)); significant effects and interactions are highlighted using bold font.

Faces

When presented with faces, there was a significant effect of Emotion ($F(3, 222) = 63.31$, $p < 0.001$, $\eta^2_p = 0.46$). Across the ASD and NT groups, the happy emotion (M = 0.93, SD = 0.11) was the most accurately recognised emotion from faces, followed by angry (M = 0.77, SD = 0.19), with scared (M = 0.71, SD = 0.22) and sad emotions (M = 0.72, SD = 0.18) being the least accurately recognised (happy vs. angry: $t(75) = 10.31$, $p < 0.001$; happy vs. scared: $t(75) = 12.16$, $p < 0.001$; happy vs. sad: $t(75) = 13.39$, $p < 0.001$; angry vs. scared: $t(75) = 2.69$, $p = 0.013$; angry vs. sad: $t(75) = 2.62$, $p = 0.013$).

Objects

When presented with objects, there was a significant effect of Emotion ($F(3, 222) = 44.16$, $p < 0.001$, $\eta^2_p = 0.37$), as well as an interaction of Diagnostic group × Emotion ($F(3, 222) = 2.97$, $p = 0.033$, $\eta^2_p = 0.04$). While the ASD and NT groups did not differ in accuracy across emotions, the two groups showed different performance among emotions.

In the ASD group, the happy emotion (M = 0.94, SD = 0.13) was the most accurately recognised emotion from objects, followed by scared (M = 0.87, SD = 0.20), with angry (M =

0.77 SD = 0.18) and sad (M = 0.79, SD = 0.17) emotions being the least accurately recognised (happy vs. scared: $t(37) = 2.61$, $p = 0.020$; happy vs. angry: $t(37) = 6.98$, $p < 0.001$; happy vs. sad: $t(37) = 6.46$, $p < 0.001$; scared vs. angry: $t(27) = 3.22$, $p = 0.005$; scared vs. sad: $t(37) = 2.52$, $p = 0.020$).

In the NT group, the happy emotion (M = 0.93, SD = 0.11) was also the most accurately recognised emotion from objects, but was followed by scared (M = 0.85, SD = 0.13) and sad (M = 0.85, SD = 0.13), with the angry emotion (M = 0.72, SD = 0.14) being the least accurately recognised (happy vs. scared: $t(37) = 6.96$, $p < 0.001$; happy vs. sad: $t(37) = 4.72$, $p < 0.001$; happy vs. angry: $t(37) = 12.84$, $p < 0.001$; scared vs. angry: $t(37) = 7.03$, $p < 0.001$; sad vs. angry: $t(37) = 5.86$, $p < 0.001$) – see Figure 3.2.

Thus, whereas the ASD group recognised anger and sadness equally poorly, the NT group recognised sadness better than anger with anger being the least accurately recognised.

Prosody

When presented with prosody, there was a significant effect of Emotion ($F(3, 222) = 40.82$, $p < 0.001$, $\eta^2_p = 0.36$). Across the ASD and NT groups, the happy emotion (M = 0.91, SD = 0.15) was the most accurately recognised emotion from prosody, followed by angry (M = 0.80, SD = 0.20) and sad (M = 0.80, SD = 0.19), with the scared emotion (M = 0.61, SD = 0.24) being the least accurately recognised (happy vs. angry: $t(75) = 5.06$, $p < 0.001$; happy vs. sad: $t(75) = 5.58$, $p < 0.001$; happy vs. scared: $t(75) = 11.04$, $p < 0.001$; angry vs. scared: $t(75) = 5.42$, $p < 0.001$; sad vs. scared: $t(75) = 5.59$, $p < 0.001$).

Song

When presented with song, there was a significant effect of Emotion ($F(3, 222) = 31.21$, $p < 0.001$, $\eta^2_p = 0.30$). Across the ASD and NT groups, the scared emotion (M = 0.52, SD = 0.22) was the least accurately recognised emotion from song, which differed significantly from angry (M = 0.79, SD = 0.21), happy (M = 0.77, SD = 0.21), and sad (M = 0.76, SD = 0.24) emotions (scared vs. angry: $t(75) = -8.86, p < 0.001$; scared vs. happy: $t(75) = -8.03, p < 0.001$, scared vs. sad: $t(75) = -7.64, p < 0.001$).

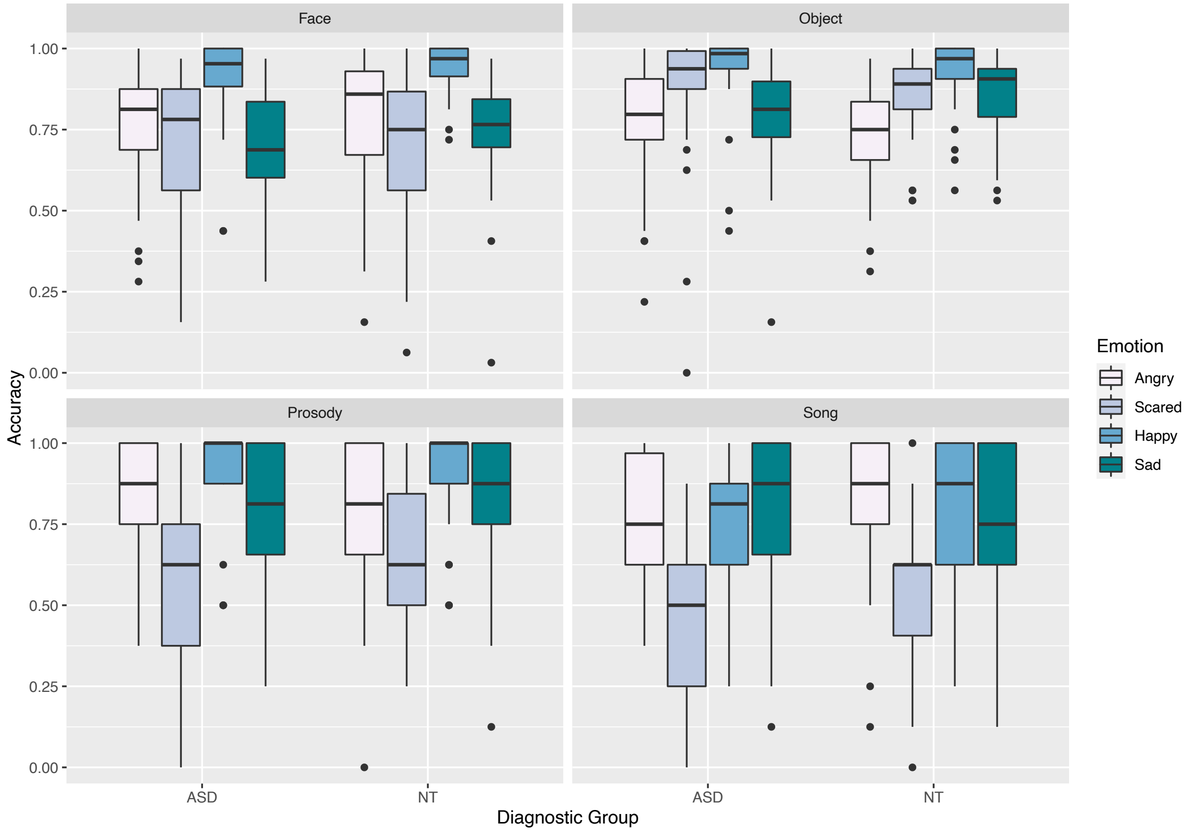


Figure 3. 2. Boxplots of mean accuracy for the ASD and NT groups in the emotion recognition task across the four conditions (face, object, prosody, and song) as a function of the four different emotions (angry, scared, happy, and sad).

Summary

The ASD and NT groups generally performed comparably across conditions and emotions but there were some idiosyncratic differences for objects but not for faces, prosody, and song across emotions. Across groups, happiness was the most accurately recognised emotion for all conditions, while sadness was the least accurately recognised emotion from faces, anger from objects, and fear from prosody and song.

Misclassification patterns between diagnostic groups by condition

Table 3. 7. Confusion matrix for faces, face-like objects, prosody, and song by diagnostic group. Rows show each of the four target emotions and the columns show the proportion of responses (in percentage) that were given for the correct answer (in bold) and the three alternative response options.

		ASD				NT			
		Angry	Scared	Happy	Sad	Angry	Scared	Happy	Sad
Face									
	Angry	**76.56**	7.73	4.28	11.43	**77.88**	6.74	4.11	11.27
	Scared	9.46*	**70.81**	4.61	15.13	5.84	**70.64**	5.35	18.17
	Happy	2.63*	2.96	**91.04**	3.37	0.82	1.73	**95.07**	2.38
	Sad	9.70*	12.66	8.72	**68.91**	6.50	12.01	7.15	**74.34**
Object									
	Angry	**76.15**	8.96	2.96	11.92*	**72.20**	10.03	1.89	15.87
	Scared	6.99	**87.42**	3.21	2.38***	4.69	**85.36**	3.62	6.33
	Happy	0.58	2.96	**93.59**	2.88	0.74	2.06	**93.01**	4.19
	Sad	5.43	10.94	4.44**	**79.19**	4.61	7.98	1.97	**85.44**
Prosody									
	Angry	**74.67**	11.84	7.57	5.92	**82.57**	8.55	6.25	2.63
	Scared	24.01	**49.67**	7.24	19.08	21.05	**54.28**	3.29	21.38
	Happy	3.29	14.14	**75.66**	6.91	2.63	9.21	**78.95**	9.21
	Sad	3.29	11.84	8.55	**76.32**	3.95	15.46	4.93	**75.66**
Song									
	Angry	**81.58**	7.89	7.57	2.96	**78.29**	8.88	10.20	2.63
	Scared	13.49	**55.59**	10.20	20.72	8.88	**66.78**	6.58	17.76
	Happy	2.96	1.97	**92.11**	2.96	4.61	2.63	**89.80**	2.96
	Sad	4.61	11.51	3.95	**79.93**	5.26	11.18	3.95	**79.61**

Note. Asterisks (*) indicate significant differences in the frequencies of response provided between the ASD and NT groups with * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

To examine whether the ASD group were making systematic errors (i.e., confusing two emotions) that was not seen in the NT groups, the patterns of emotion misclassification were explored between the two groups. Table 3.7 presents the accuracy rates of responses to each emotion across conditions by diagnostic group. Results showed that for faces, the ASD group identified scared, happy, and sad faces as angry significantly more frequently compared to the NT group (scared-angry: $\chi^2(1) = 10.41$, $p = 0.010$; happy-angry: $\chi^2(1) = 11.52$, $p = 0.010$; sad-angry: and $\chi^2(1) = 7.72$, $p = 0.029$). For objects, the ASD group identified sad objects as happy more frequently compared to the NT group ($\chi^2(1) = 11.54$, $p = 0.006$), which agrees with the Diagnostic group × Emotion interaction observed above – that is, the sad emotion was least accurately recognised emotion in the ASD group, which was not seen in the NT group. Conversely, the ASD group identified angry and scared objects less frequently compared to the NT group (angry-sad: $\chi^2(1) = -6.82$, $p = 0.048$; scared-sad: $\chi^2(1) = -21.74$, $p < 0.00$). The misclassification pattern did not differ between the ASD and NT groups across emotions for prosody and song.

3.3.1.2. Age group × Condition × Emotion

The Age group × Condition × Emotion interaction was examined using linear mixed effects models for each condition, with mean arcsine-transformed accuracy as the dependent measure. Each model included Age group, Emotion, and their two-way interaction as fixed effects, and with by-subject intercept as random effects. The results of the models are summarised in Table 3.8.

Table 3. 8. Linear mixed effects model results for age group, emotion, and their interactions on mean arcsine-transformed accuracy for each condition separately.

	Fixed effects	**df**		***F***	***p***	**η^2_p**
Face						
	Age group	**2**	**73**	**9.89**	**< 0.001**	**0.21**
	Emotion	**3**	**219**	**67.67**	**< 0.001**	**0.48**
	Age group × emotion	**6**	**219**	**3.90**	**0.001**	**0.10**
Object						
	Age group	2	73	0.96	0.388	0.03
	Emotion	**3**	**219**	**45.23**	**< 0.001**	**0.38**
	Age group × emotion	**6**	**219**	**2.59**	**0.019**	**0.07**
Prosody						
	Age group	**2**	**73**	**5.67**	**0.005**	**0.13**
	Emotion	**3**	**219**	**39.12**	**< 0.001**	**0.35**
	Age group × emotion	6	219	0.82	0.557	0.02
Song						
	Age group	**2**	**73**	**5.89**	**0.004**	**0.14**
	Emotion	**3**	**219**	**33.33**	**< 0.001**	**0.31**
	Age group × emotion	**6**	**219**	**2.97**	**0.008**	**0.08**

Note. R model equation: lmer (asin (sqrt (Accuracy)) ~ Age Group * Emotion + (1 | Subject)); significant effects and interactions are highlighted using bold font.

Faces

When presented with faces, the effects of Age group ($F(2, 73) = 9.89, p < 0.001, \eta^2_p = 0.21$) and Emotion ($F(3, 219) = 67.67, p < 0.001, \eta^2_p = 0.48$), as well as the interaction of Age group × Emotion ($F(6, 219) = 3.90, p = 0.001, \eta^2_p = 0.10$) were all significant. It was found that the recognition of the happy and sad emotions from faces reached adult-level accuracy the earliest by late childhood, as no age-related differences were observed for these emotions.

These were followed by the recognition of the angry emotion: children (M = 0.72, SD = 0.19) showed lower accuracy for angry faces than adults (M = 0.83, SD = 0.16) ($t(51.91) = -2.48$, $p = 0.049$), while adolescents showed no difference compared to either group. By contrast, the recognition of the scared emotion from faces followed a more protracted developmental trajectory, as children (M = 0.56, SD = 0.23) showed lower accuracy for scared faces than adolescents (M = 0.73, SD = 0.17), who in turn showed lower accuracy than adults (M = 0.85, SD = 0.12) (child vs. adolescent: $t(47.54) = -3.05$, $p = 0.006$; child vs. adult: $t(48.99) = -6.23$, $p < 0.001$; adolescent vs. adult: $t(41.01) = -2.86$, $p = 0.007$).

Examining the performance across emotions for faces, the happy emotion was the most accurately recognised emotion in all three age groups: children (happy, M = 0.90, SD = 0.13; angry, M = 0.72, SD = 0.19; sad, M = 0.69, SD = 0.20; scared, M = 0.56, SD = 0.23; happy vs. angry: $t(27) = 6.79$, $p < 0.001$; happy vs. sad: $t(27) = 6.82$, $p < 0.001$; happy vs. scared: $t(27) = 10.77$, $p < 0.001$), adolescents (happy, M = 0.93, SD = 0.13; angry, M = 0.77, SD = 0.20; scared, M = 0.73, SD = 0.17; sad, M = 0.67, SD = 0.18; happy vs. angry: $t(21) = 6.25$, $p < 0.001$; happy vs. scared: $t(21) = 6.22$, $p < 0.001$; happy vs. sad: $t(21) = 9.85$, $p < 0.001$), and adults (happy, M = 0.97, SD = 0.04; angry, M = 0.72, SD = 0.19; scared, M = 0.56, SD = 0.23; sad, M = 0.69, SD = 0.20; happy vs. angry: $t(25) = 4.94$, $p < 0.001$; happy vs. scared: $t(25) = 5.43$, $p < 0.001$; happy vs. sad: $t(25) = 7.37$, $p < 0.001$).

The three age groups, however, showed different accuracy patterns among the other emotions. In children, the angry and sad emotions were recognised equally accurately from faces, with the scared emotion being the least accurately recognised emotion (angry vs. scared: $t(27) = 4.85$, $p < 0.001$; sad vs. scared: $t(27) = 3.00$, $p = 0.007$).

In adolescents, the angry emotion was recognised more accurately than the sad emotion from faces ($t(21) = 2.37$, $p = 0.041$), with the recognition of the scared emotion showing no difference to that of either emotion.

In adults, the angry, scared, and sad emotions were recognised equally accurately from faces – see Figure 3.3.

These accuracy patterns among emotions further reinforced the more protracted improvement with fear in development, whereas fear was the least accurately recognised emotion from faces in children, the recognition of fear reached similar levels to that of anger and sadness over time. Happiness remained the most accurately recognised emotion across ages.

Objects

When presented with objects, there was a significant effect of Emotion ($F(3, 219) = 45.23$, $p < 0.001$, $\eta^2_p = 0.38$), as well as an interaction of Age group × Emotion ($F(6, 219) = 2.59$, $p = 0.019$, $\eta^2_p = 0.07$). It was found that the recognition of angry, scared, and happy emotions from objects reached adult-level accuracy by late childhood, as no age-related differences were observed for these emotions. These were followed by the recognition of the sad emotion from objects: children (M = 0.79, SD = 0.13) showed lower accuracy for sad objects than adults (M = 0.89, SD = 0.09) ($t(51.56) = 3.07$, $p = 0.010$), while adolescents showed no difference compared to either group.

Examining the performance across emotions for objects, the happy emotion was the most accurately recognised emotion in all three age groups: children (happy, M = 0.93, SD = 0.10; angry, M = 0.76, SD = 0.16; scared, M = 0.81, SD = 0.21; sad, M = 0.79, SD = 0.13; happy vs. angry: $t(27) = 7.82$, $p < 0.001$; happy vs. scared: $t(27) = 3.98$, $p < 0.001$; happy vs. sad: $t(27) = 6.71$, $p < 0.001$), adolescents (happy, M = 0.94, SD = 0.12, scared, M = 0.90, SD = 0.15; angry, M = 0.74, SD = 0.17; sad, M = 0.79, SD = 0.20; happy vs. scared: $t(21) = 2.40$, $p = 0.031$; happy vs. angry: $t(21) = 8.64$, $p < 0.001$; happy vs. sad: $t(21) = 4.88$, $p < 0.001$), and adults (happy, M = 0.93, SD = 0.13; scared, M = 0.88, SD = 0.12; sad, M = 0.89, SD = 0.09;

angry, M = 0.74, SD = 0.17; happy vs. scared: $t(25) = 2.68$, $p = 0.019$; happy vs. sad: $t(25) = 2.55$, $p = 0.021$; happy vs. angry: $t(25) = 6.30$, $p < 0.001$).

The three age groups, however, showed different accuracy patterns among the other emotions. In children, the recognition accuracy did not differ across the angry, scared, and sad emotions for objects.

In adolescents, the scared emotion was recognised more accurately than the angry and sad emotions, which did not differ from each other (scared vs. angry: $t(21) = 5.95$, $p < 0.001$; scared vs. sad: $t(21) = 3.78$, $p = 0.002$).

In adults, the scared and sad emotions were recognised equally accurately, with the angry emotion being the least accurately recognised (scared vs, angry: $t(25) = 4.26$, $p < 0.001$; sad vs. angry: $t(25) = 3.61$, $p = 0.003$) – see Figure 3.3.

These accuracy patterns among emotions outlined that while the recognition of fear, followed by sadness, exceeded that of anger with increasing age, the recognition of anger from objects remained poor across ages. Happiness was the most accurately recognised emotion across ages.

Prosody

When presented with prosody, there was a significant effect of Age group ($F(2, 73) = 5.67$, $p = 0.005$, $\eta^2_p = 0.13$), with children (M = 0.73, SD = 0.26) and adolescents (M = 0.77, SD = 0.21) generally showing lower accuracy than adults (M = 0.84, SD = 0.17) across emotions (Child vs. Adolescent: $t(197.79) = -1.19$, $p = 0.237$; Child vs. Adult: $t(206.12) = -3.50$, $p = 0.002$; Adolescent vs. Adult: $t(180.31) = -2.36$, $p = 0.029$).

The effect of Emotion was also significant ($F(3, 219) = 39.12$, $p < 0.001$, $\eta^2_p = 0.35$). Across age groups, the happy emotion (M = 0.91, SD = 0.15) was the most accurately recognised emotion from prosody, followed by angry (M = 0.80, SD = 0.20) and sad (M =

0.80, SD = 0.19), with the scared emotion (M = 0.61, SD = 0.24) being the least accurately recognised (happy vs. angry: *t*(146.18) = 4.35, *p* < 0.001; happy vs. sad: *t*(148.03) = 4.46, *p* < 0.001; happy vs. scared: *t*(147.46) = 10.12, *p* < 0.001; angry vs. scared: *t*(149.85) = 5.26, *p* < 0.001; sad vs. scared: *t*(149.96) = 5.40, *p* < 0.001).

Song

When presented with song, there were significant effects of Age group (*F*(2, 73) = 5.89, *p* = 0.004, η^2_p = 0.14) and Emotion group (*F*(3, 219) = 33.33, *p* < 0.001, η^2_p = 0.31) as well as a significant interaction of Age group × Emotion group (*F*(6, 219) = 2.97, *p* = 0.008, η^2_p = 0.08). It was found that the recognition of angry, happy, and sad emotions from song reached adult-level performance by late childhood, as no age-related differences were found for these emotions. These were followed by the recognition of the scared emotion: while no difference was observed between children (M = 0.40, SD = 0.19) and adolescents (M = 0.46, SD = 0.21), both groups showed lower accuracy for scared song than adults (M = 0.70, SD = 0.13) (child vs. adult: *t*(51.79) = -6.36, *p* < 0.001; adolescent vs. adult: *t*(34.35) = -4.18, *p* < 0.001).

Examining the performance across emotions for song, children and adolescents showed similar patterns that differed from those exhibited by adults. In both children and adolescents, the scared emotion was the least accurately recognised emotion (children: scared, M = 0.40, SD = 0.19; angry, M = 0.74, SD = 0.23; happy, M = 0.75, SD = 0.18; sad, M = 0.71, SD = 0.23; scared vs. angry: *t*(27) = -6.24, *p* < 0.001; scared vs. happy: *t*(27) = -6.78, *p* < 0.001; scared vs. sad: *t*(27) = -5.21, *p* < 0.001; adolescents: scared, M = 0.46, SD = 0.21; angry, M = 0.75, SD = 0.21; happy, M = 0.83, SD = 0.23; sad, M = 0.76, SD = 0.27; scared vs. angry: *t*(21) = -4.46, *p* < 0.001; scared vs. happy: *t*(21) = 8.53, *p* < 0.001; scared vs. sad: *t*(21) = -5.28, *p* < 0.001). The recognition accuracy did not differ across the angry, happy, and sad emotions in both children and adolescents.

By contrast, in adults, the scared emotion (M = 0.70, SD = 0.13) was less accurately recognised from song than the angry (M = 0.87, SD = 0.15) and sad emotions (M = 0.81, SD = 0.22) (scared vs. angry: $t(25) = -4.65$, $p < 0.001$; scared vs. sad: $t(25) = -2.94$, $p = 0.021$), where the recognition of the happy emotion did not differ from any of these emotions – see Figure 3.3.

These accuracy patterns among emotions reinforced the more protracted development of fear recognition for song: whereas fear was the least accurately recognised emotion from song at a younger age, the recognition of this emotion gradually reached similar levels as that of happiness.

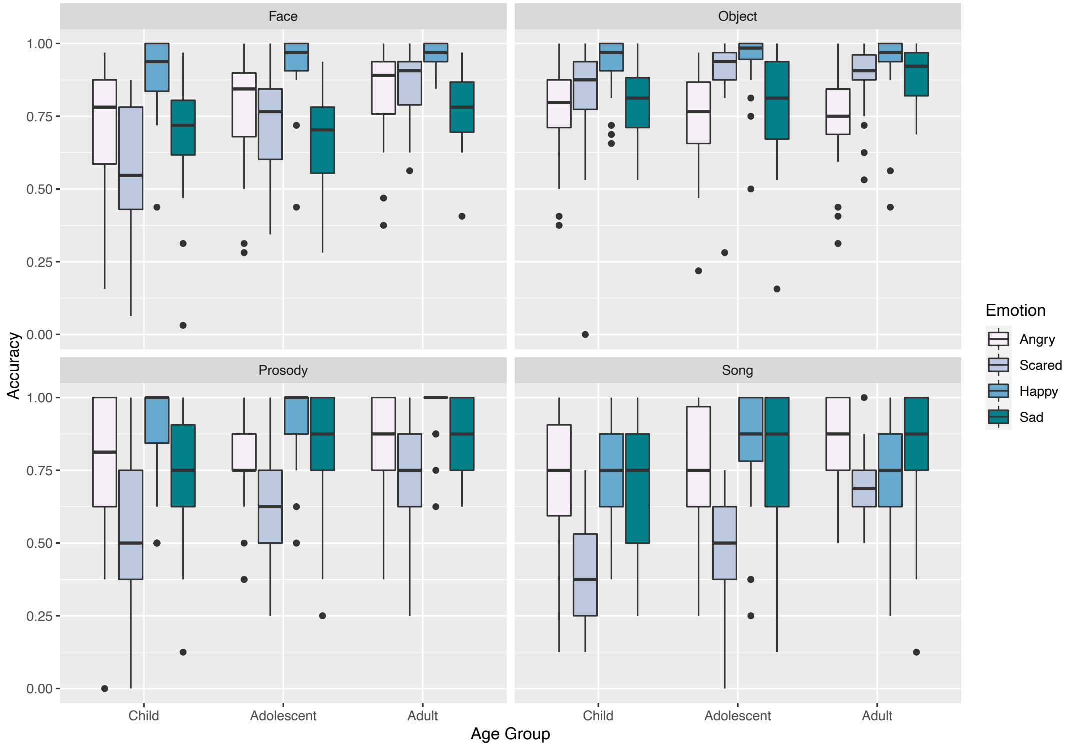


Figure 3. 3. Boxplots of mean accuracy for children, adolescents, and adults in the emotion recognition task across the four conditions (face, object, prosody, and song) as a function of the four different emotions (angry, scared, happy, and sad).

Summary

Adult-level recognition accuracy appeared to be well-achieved by late childhood across some emotions for faces (i.e., happiness and sadness), objects (i.e., anger, fear, and happiness), and song (i.e., anger, happiness, and sadness). By contrast, the recognition accuracy continued to improve for the other emotions for faces (i.e., anger and fear), objects (i.e., sadness), and song (i.e., fear) and across emotions for prosody beyond late childhood. Across groups, happiness was the most accurately recognised emotion for all conditions, while sadness was the least accurately recognised emotion from faces, anger from objects, and fear from prosody and song.

3.3.2. Response time

Table 3. 9. Linear mixed effects model results for diagnostic group, age group, condition, emotion, and their interactions on mean log-transformed RT.

Fixed effects	df		*F*	*p*	η^2_p
Diagnostic group	**1**	**69.99**	**10.71**	**0.002**	**0.13**
Age group	**2**	**69.99**	**5.30**	**0.007**	**0.13**
Condition	**3**	**69.93**	**17.57**	**< 0.001**	**0.43**
Emotion	**3**	**69.82**	**40.56**	**< 0.001**	**0.64**
Diagnostic group × age group	2	69.99	1.66	0.197	0.05
Diagnostic group × condition	3	69.93	0.84	0.475	0.03
Age group × condition	**6**	**69.93**	**4.56**	**< 0.001**	**0.28**
Diagnostic group x emotion	3	69.82	0.78	0.508	0.03
Age group × emotion	6	69.83	0.50	0.808	0.04
Condition × emotion	**9**	**627.28**	**4.76**	**< 0.001**	**0.06**
Diagnostic group × age group × condition	6	69.93	0.90	0.500	0.07
Diagnostic group × age group × emotion	6	69.83	0.58	0.742	0.05
Diagnostic group × condition × emotion	9	627.28	0.64	0.764	0.00
Age group × condition x emotion	**18**	**627.25**	**1.68**	**0.039**	**0.05**
Diagnostic group × age group × condition × emotion	18	627.25	1.07	0.377	0.03

Note. R model equation: lmer (log(Response time) ~ Diagnostic Group * Age Group * Condition * Emotion + (1 + Condition + Emotion | Subject)); significant effects and interactions are highlighted using bold font.

Table 3.9 displays the full model results summary of the linear mixed effects analysis on log-transformed mean RT data. The analysis showed a significant main effect of Diagnostic group ($F(1, 69.99) = 10.71$, $p = 0.002$, $\eta^2_p = 0.13$). In general, the RT was slower for the ASD group (M = 1717.69, SD = 531.03) than for the NT group (M = 1373.56, SD = 360.26) (Figure 3.4). The main effects of Age group ($F(2, 69.99) = 5.30$, $p = 0.007$, $\eta^2_p = 0.13$), Condition ($F(3, 69.93) = 17.57$, $p < 0.001$, $\eta^2_p = 0.43$), and Emotion ($F(3, 69.82) = 40.56$, $p < 0.001$, $\eta^2_p = 0.64$), as well as the interactions of Age group × Condition ($F(6, 69.93) = 4.56$, $p < 0.001$, $\eta^2_p = 0.28$) and Condition × Emotion ($F(9, 627.28) = 4.76$, $p < 0.001$, $\eta^2_p = 0.06$) were also significant, all of which were further qualified by a three-way interaction of Age group × Condition × Emotion ($F(18, 627.25) = 1.68$, $p = 0.039$, $\eta^2_p = 0.05$). No other factors and interactions were significant. In the following subsection, I will unpack the three-way interaction.

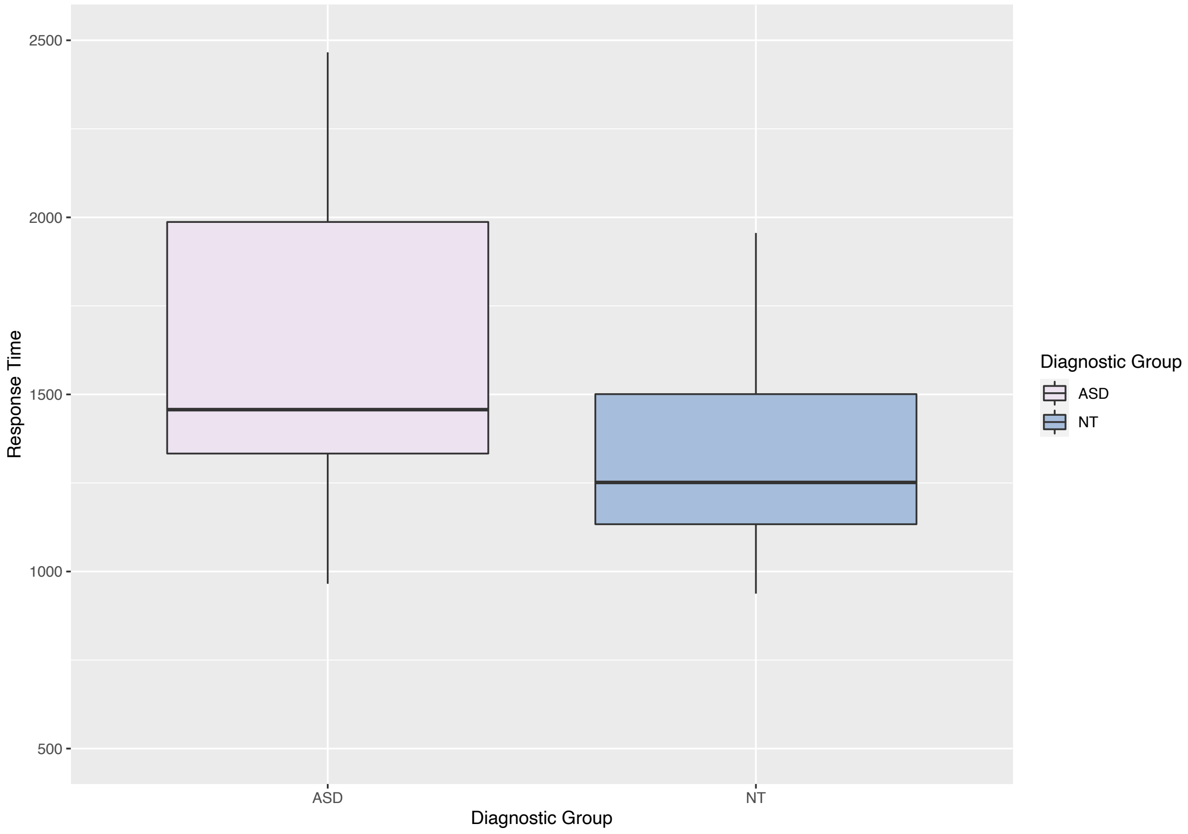


Figure 3. 4. Boxplots of overall mean RT for the ASD and NT groups in the emotion recognition task.

3.3.2.1. Age group × Condition × Emotion

The Age group × Condition × Emotion interaction was examined using linear mixed effects models for each condition, with mean log-transformed RT as the dependent measure. Each model included Age group, Emotion, and their two-way interaction as fixed effects, and with by-subject intercept as random effects. The results of the models are summarised in Table 3.10.

Table 3. 10. Linear mixed effects model results for age group, emotion, and their interactions on mean log-transformed RT for each condition separately.

	Fixed effects	**df**		***F***	***p***	**η^2_p**
Face						
	Age group	2	73	3.02	0.055	0.08
	Emotion	**3**	**219**	**58.48**	**< 0.001**	**0.44**
	Age group × emotion	**6**	**219**	**2.48**	**0.024**	**0.06**
Object						
	Age group	**2**	**72.99**	**10.30**	**< 0.001**	**0.22**
	Emotion	**3**	**218.02**	**26.77**	**< 0.001**	**0.27**
	Age group × emotion	6	218.02	0.92	0.483	0.02
Prosody						
	Age group	2	72.89	0.59	0.559	0.02
	Emotion	**3**	**216.22**	**8.08**	**< 0.001**	**0.10**
	Age group × emotion	6	216.23	1.89	0.084	0.05
Song						
	Age group	2	73.14	2.78	0.069	0.07
	Emotion	**3**	**216.40**	**4.47**	**0.005**	**0.06**
	Age group × emotion	6	216.38	0.27	0.952	0.00

Note. R model equation: lmer (log(Response time) ~ Age Group * Emotion + (1 | Subject)); significant effects and interactions are highlighted using bold font.

Faces

When presented with faces, there was a significant effect of Emotion ($F(3, 219) = 58.48$, $p < 0.001$, $\eta^2_p = 0.44$) and interaction of Age group × Emotion ($F(16, 219) = 2.48$, $p = 0.024$, $\eta^2_p = 0.06$). It was found that adult-level RT was achieved by late childhood for the angry and sad emotions, as no age-related differences were observed for these emotions. These were followed by the recognition of the scared and happy emotions: children showed significantly slower RTs than adults when recognising scared faces (child, M = 2574.37, SD = 1276.23; adult, M = 1672.32, SD = 598.51; $t(44.97) = 3.11$, $p = 0.010$) and happy faces (child, M = 1670.18, SD = 614.87; adult, M = 1277.51, SD = 529.31; $t(51.66) = 2.57$, $p = 0.039$), with adolescents showing no difference to either group.

Examining the performance across emotions for faces, the happy emotion was the fastest recognised emotion in all three age groups: children (happy, M = 1670.18, SD = 614.87; angry, M = 2068.13, SD = 918.03; sad, M = 2091.13, SD = 973.34; scared, M = 2574.37, SD = 1276.23; happy vs. angry: $t(27) = -4.79$, $p < 0.001$; happy vs. sad: $t(27) = -4.75$, $p < 0.001$; happy vs. scared: $t(27) = -7.94$, $p < 0.001$), adolescents (happy, M = 1474.35, SD = 686.16; angry, M = 1769.46, SD = 838.88; scared, M = 1890.78, SD = 856.41; sad, M = 1862.99, SD

= 870.64; happy vs. angry: $t(21) = -3.73$, $p = 0.002$; happy vs. scared: $t(21) = -5.84$, $p < 0.001$; happy vs. sad: $t(21) = -4.02$, $p = 0.002$), and adults (happy, M = 1277.51, SD = 529.31; angry, M = 1611.19, SD = 644.62; scared, M = 1672.32, SD = 598.51; sad, M = 1701.38, SD = 644.29; happy vs. angry: $t(25) = -6.19$, $p < 0.001$; happy vs. scared: $t(25) = -6.72$, $p < 0.001$; happy vs. sad: $t(25) = -8.73$, $p < 0.001$).

The RT pattern for the other emotions was similar in adolescents and adults, which differed from that exhibited by children. That is, whereas the scared emotion was recognised the slowest in children (angry vs. scared: $t(27) = -4.30$, $p < 0.001$; sad vs. scared: $t(17) = -6.31$, $p < 0.001$), the recognition RT of the angry, scared, and sad emotions did not differ in adolescents and adults – see Figure 3.5.

These accuracy patterns among emotions further reinforced the more protracted development for fear: whereas fear was recognised the slowest compared to other emotions in children, the recognition of this emotion reached similar speed levels as that of anger and sadness by adolescence. Happiness remained the fastest recognised emotion across ages.

Objects

When presented with objects, there was a significant effect of Age group ($F(2, 72.99) = 10.30$, $p < 0.001$, $\eta^2_p = 0.22$), with children (M = 1835.28, SD = 724.25) showing slower RT than adolescents (M = 1343.31, SD = 567.26) and adults (M = 1300.12, SD = 517.07), who did not differ from each other (child vs. adolescent: $t(175.81) = 6.68$, $p < 0.001$; children vs. adult: $t(210.68) = 7.84$, $p < 0.001$).

The effect of Emotion was also significant ($F(3, 218.02) = 26.77$, $p < 0.001$, $\eta^2_p = 0.27$). Overall, the happy emotion (M = 1367.64, SD = 555.13) was the fastest recognised emotion from objects, followed by angry (M = 1501.16, SD = 653.13) and sad (M = 1519.74, SD = 594.83), with the scared emotion (M = 1648.15, SD = 800.13) being recognised the slowest (happy vs. angry: $t(74) = -4.87$, $p < 0.001$; happy vs. sad: $t(74) = -6.80$, $p < 0.001$; happy vs.

scared: $t(74) = -9.54$, $p < 0.001$; angry vs. scared: $t(74) = -3.99$, $p < 0.001$; sad vs. scared: $t(74) = -3.24$, $p = 0.002$).

Prosody

When presented with prosody, there was a significant effect of Emotion ($F(3, 216.22) = 8.08$, $p < 0.001$, $\eta^2_p = 0.10$). Across age groups, the happy emotion (M = 1331.09, SD = 486.77) was recognised faster from prosody than the scared (M = 1506.09, SD = 455.72) and sad emotions (M = 1452.00, SD = 437.04), which did not differ from each other (happy vs. scared: $t(72) = -4.48$, $p < 0.001$; happy vs. sad: $t(72) = -3.08$, $p = 0.009$). There was no difference between recognition of the angry emotion and other emotions from prosody.

Song

When presented with song, there was a significant effect of Emotion ($F(3, 216.40) = 4.47$, $p = 0.005$, $\eta^2_p = 0.06$). Across age groups, the scared emotion (M = 1546.39, SD = 602.97) was the slowest recognised emotion from song, with no difference found between the angry (M = 1432.50, SD = 726.65), happy (M = 1362.93, SD = 602.39), and sad emotions (M = 1376.10, SD = 438.59) (scared vs. angry: $t(72) = 2.34$, $p = 0.045$; scared vs. happy: $t(72) = 3.35$, $p = 0.008$; scared vs. sad: $t(72) = 2.48$, $p = 0.045$).

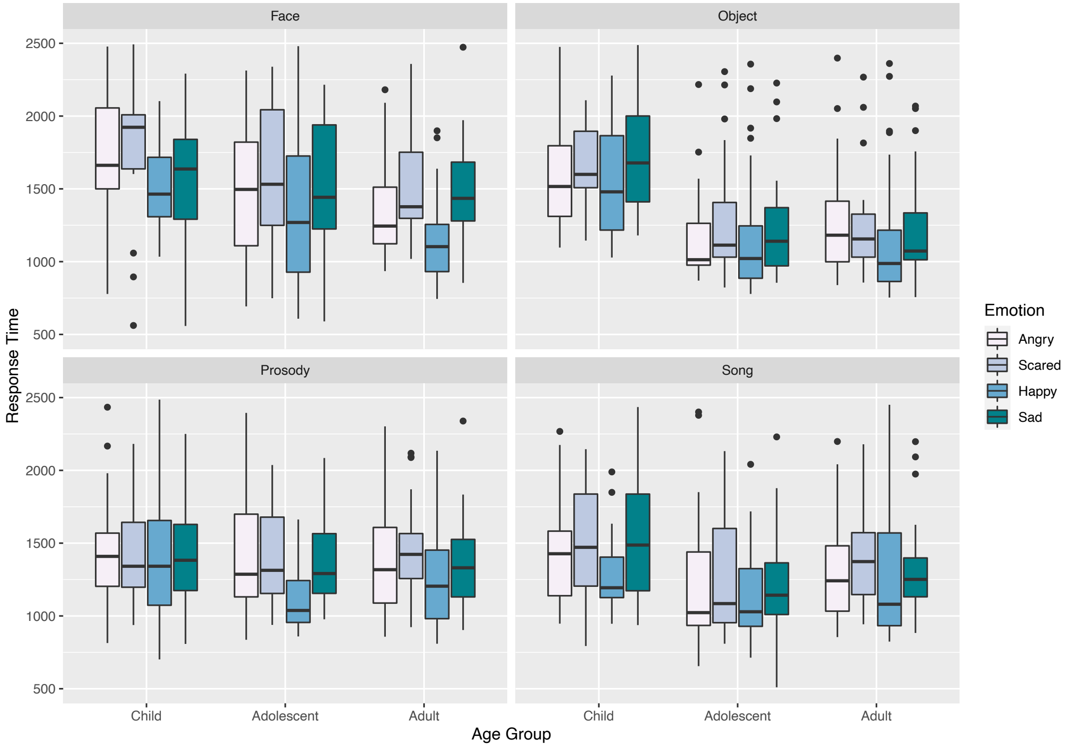


Figure 3. 5. Boxplots of mean RT for children, adolescents, and adults in the emotion recognition task across the four conditions (face, object, prosody, and song) as a function of the four different emotions (angry, scared, happy, and sad).

Summary

Adult-level recognition speed appeared to be well-established by late childhood across emotions for prosody and song and across some emotions for faces (i.e., anger and sadness). The recognition speed continued to improve for the other emotions for faces (i.e., fear and happiness) and across emotions for objects beyond late childhood and reached adult-level performance during adolescence. Adult-level recognition speed appeared to be well-established by late childhood across emotions for prosody and song and across some emotions for faces (i.e., anger and sadness). The recognition speed continued to improve for the other emotions for faces (i.e., fear and happiness) and across emotions for objects beyond late childhood and reached adult-level performance during adolescence. Across age groups, happiness was the fastest recognised emotion for all conditions, whereas fear was the slowest recognised emotion for all conditions. In addition to fear, sadness was also the slowest recognised emotion for speech prosody across age groups.

3.3.3. Speed-accuracy composite score

Table 3. 11. Linear mixed effects model results for diagnostic group, age group, condition, emotion, and their interactions on mean SACS.

Fixed effects	df		*F*	*p*	η^2_p
Diagnostic group	**1**	**69.97**	**10.66**	**0.002**	**0.13**
Age group	**2**	**69.98**	**11.24**	**< 0.001**	**0.24**
Condition	3	70.19	0.01	0.998	0.00
Emotion	**3**	**69.86**	**72.41**	**< 0.001**	**0.76**
Diagnostic group × age group	2	69.98	0.92	0.403	0.03
Diagnostic group × condition	3	70.19	0.39	0.762	0.02
Age group x condition	**6**	**70.20**	**2.28**	**0.045**	**0.16**
Diagnostic group × emotion	3	69.86	0.07	0.974	0.00
Age group × emotion	**6**	**69.87**	**3.12**	**0.009**	**0.21**
Condition × emotion	**9**	**626.48**	**16.82**	**< 0.001**	**0.19**
Diagnostic group × age group × condition	6	70.20	1.30	0.267	0.10
Diagnostic group × age group × emotion	6	69.87	0.60	0.726	0.05
Diagnostic group × condition × emotion	9	626.48	1.50	0.144	0.02
Age group × condition × emotion	**18**	**626.44**	**2.61**	**< 0.001**	**0.07**
Diagnostic group × age group × condition × emotion	18	626.44	0.97	0.487	0.03

Note. R model equation: lmer (Speed-Accuracy Composite Score ~ Diagnostic Group * Age Group * Condition * Emotion + (1 + Condition + Emotion | Subject), control = lmerControl(optimizer = "bobyqa")); significant effects and interactions are highlighted using bold font.

Table 3.11 displays the full model results summary of the linear mixed effects analysis on speed-accuracy composite score data. The analysis showed a significant main effect of Diagnostic group ($F(1, 69.97) = 10.66$, $p = 0.002$, $\eta^2_p = 0.13$), with the ASD group (M = -0.30, SD = 0.97) showing lower SACS than the NT group (M = 0.29, SD = 0.74) (Figure 3.6). This suggests that in general, the ASD group were less efficient than the NT group in the emotion recognition task. The main effects of Age group ($F(2, 69.98) = 11.24$, $p < 0.001$, $\eta^2_p = 0.24$) and Emotion ($F(3, 69.86) = 72.41$, $p < 0.001$, $\eta^2_p = 0.76$), as well as the two-way interactions of Age group × Condition ($F(6, 70.20) = 2.28$, $p = 0.045$, $\eta^2_p = 0.16$), Age group × Emotion ($F(6, 69.87) = 3.12$, $p = 0.009$, $\eta^2_p = 0.21$), and Condition x Emotion ($F(9, 626.48) = 16.82$, $p < 0.001$, $\eta^2_p = 0.19$) were all significant. Most importantly, a three-way interaction of Age group × Condition × Emotion was also significant ($F(18, 626.44) = 2.61$, $p < 0.001$, $\eta^2_p = 0.07$).

No other factors and interactions were significant. I will unpack the three-way interaction in the subsequent subsection.

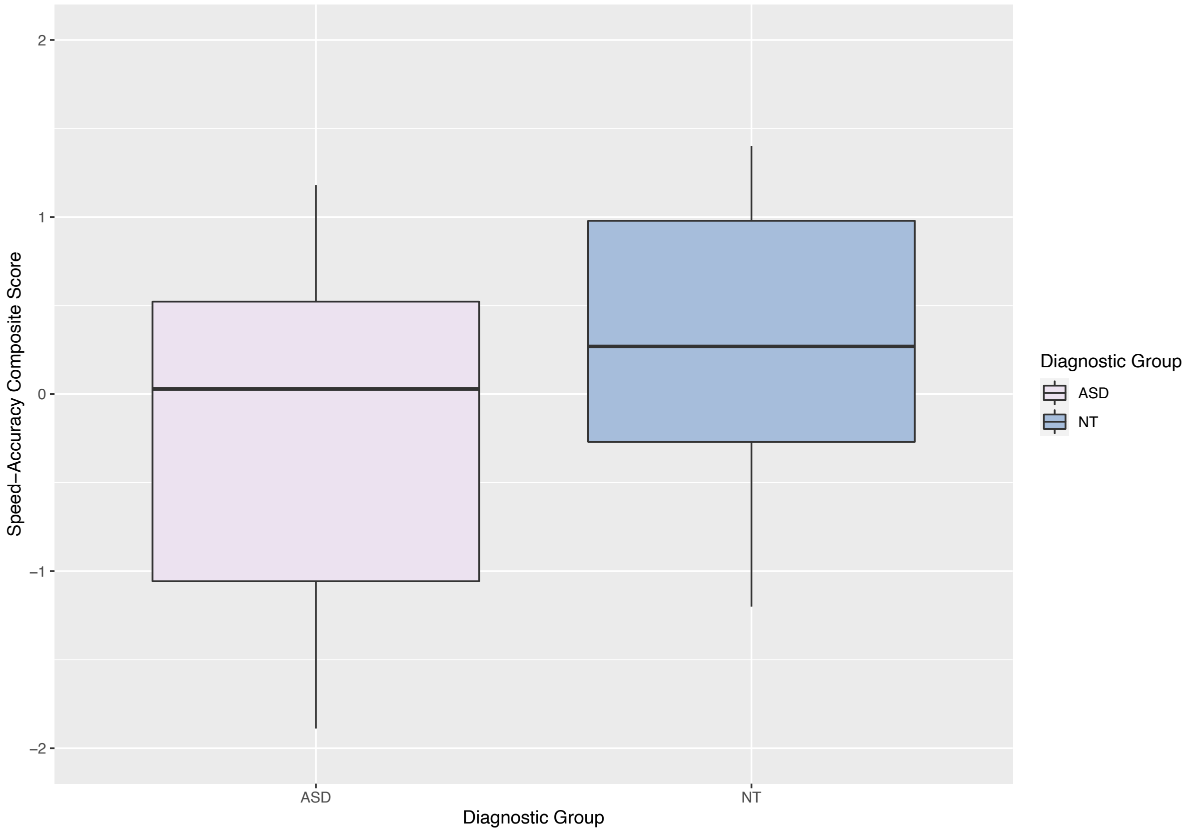


Figure 3. 6. Boxplots of overall mean SACS for the ASD and NT groups in the emotion recognition task.

3.3.3.1. Age group × Condition × Emotion

The Age group × Condition × Emotion interaction was examined using linear mixed effects models for each condition, with mean SACS as the dependent measure. Each model included Age group, Emotion, and their two-way interaction as fixed effects, and with by-subject intercept as random effects. The results of the models are summarised in Table 3.12.

Table 3. 12. Linear mixed effects model results for age group, emotion, and their interactions on mean SACS for each condition separately.

	Fixed effects	df		*F*	*p*	η^2_p
Face						
	Age group	**2**	**73**	**13.54**	**< 0.001**	**0.27**
	Emotion	**3**	**219**	**65.37**	**< 0.001**	**0.47**
	Age group × emotion	**6**	**219**	**7.32**	**< 0.001**	**0.17**

Object						
	Age group	**2**	**73.03**	**6.19**	**0.003**	**0.14**
	Emotion	**3**	**218.11**	**38.88**	**< 0.001**	**0.35**
	Age group × emotion	**6**	**218.11**	**2.86**	**0.010**	**0.07**
Prosody						
	Age group	2	73.23	2.66	0.077	0.07
	Emotion	**3**	**216.80**	**30.57**	**< 0.001**	**0.30**
	Age group × emotion	6	216.82	1.07	0.381	0.03
Song						
	Age group	**2**	**73.13**	**7.52**	**0.001**	**0.17**
	Emotion	**3**	**216.71**	**25.12**	**< 0.001**	**0.26**
	Age group × emotion	6	216.68	2.02	0.065	0.05

Note. R model equation: lmer (Speed-Accuracy Composite Score ~ Age Group * Emotion + (1 | Subject)); significant effects and interactions are highlighted using bold font.

Faces

When presented with faces, there were significant main effects of Age group ($F(2, 73) = 13.54, p < 0.001, \eta^2_p = 0.27$) and Emotion ($F(3, 219) = 65.37, p < 0.001, \eta^2_p = 0.47$), as well as a significant interaction of Age group × Emotion ($F(6, 219) = 7.32, p < 0.001, \eta^2_p = 0.17$). It was found that children showed lower SACS than adults across all emotions for faces: angry (child, M = -0.58, SD = 1.23; adult, M = 0.48, SD = 1.39; $t(50.11) = -2.99, p = 0.013$), scared (child, M = -2.01, SD = 1.64; adult, M = 0.53, SD = 1.00; $t(45.06) = -6.96, p < 0.001$), happy (child, M = 0.77, SD = 0.91, adult: M = 1.55, SD = 0.66; $t(49.20) = -3.65, p = 0.002$), and sad (child, M = -0.76, SD = 1.33; adult, M = 0.11, SD = 1.01; $t(50.09) = -2.72, p = 0.027$). While children showed lower SACS than adolescents (M = -0.37, SD = 1.20) for scared faces ($t(47.79) = -4.10, p < 0.001$), the two groups showed comparable SACS for angry, happy, and sad faces. Moreover, adolescents also showed lower SACS than adults for scared faces ($t(41) = -2.81, p = 0.008$), while no other differences were found between adolescents and adults.

Examining the performance among emotions for faces, all three age groups showed the highest SACS for the happy emotion: children (happy: M = 0.77, SD = 0.91; angry: M = -0.58, SD = 1.23; sad: M = -0.76, SD = 1.33; scared: M = -2.01, SD = 1.64; happy vs. angry: $t(27) = 6.94, p < 0.001$; happy vs. sad: $t(27) = 6.31, p < 0.001$; happy vs. scared: $t(27) = 10.73, p < 0.001$), adolescents (happy, M = 1.12, SD = 1.07; angry, M = -0.03, SD = 1.21; scared, M = -0.37, SD = 1.20; sad, M = -0.61, SD = 1.25; happy vs. angry: $t(21) = 6.16, p < 0.001$; happy vs. scared: $t(21) = 6.38$, $p < 0.001$; happy vs. sad: $t(21) = 8.21$, $p < 0.001$), and adults (happy: M = 1.55, SD = 0.66; angry, M = 0.48, SD = 1.39; scared, M = 0.53, SD = 1.00; sad, M = 0.11, SD = 1.01; happy vs. angry: $t(25) = 4.69$, $p < 0.001$; happy vs. scared: $t(25) = 5.41$, $p < 0.001$; happy vs. sad: $t(25) = 8.09$, $p < 0.001$).

Children, however, showed different SACS patterns for the other emotions compared to adolescents and adults. Specifically, children showed lower SACS for the scared emotion compared to the angry and sad emotions, which did not differ from each other (scared vs. angry: $t(27) = -6.05$, $p < 0.001$; scared vs. sad: $t(27) = -4.60$, $p < 0.001$). Conversely, adolescents and adults showed no difference across the angry, scared, and sad emotions for faces – see Figure 3.7.

These SACS patterns indicate that the recognition efficiency of fear for faces underwent substantial improvement in development and reached similar levels as that of anger and sadness with increasing age. Happiness remained the most efficiently recognised emotion across ages.

Objects

When presented with objects, there were significant main effects of Age group ($F(2, 73.03) = 6.19$, $p = 0.003$, $\eta^2_p = 0.14$) and Emotion ($F(3, 218.11) = 38.88$, $p < 0.001$, $\eta^2_p = 0.35$), as well as a significant interaction of Age group × Emotion ($F(6, 218.11) = 2.86$, $p = 0.010$, $\eta^2_p = 0.07$). It was found that adult-level SACS for the angry emotion with objects was reached by late childhood, as no age-related differences were observed for this emotion. Children, however, showed lower SACS than adults across the other emotions for objects: scared (child, M = -0.84, SD = 1.80; adult, M = 0.37, SD = 1.42; $t(49.09) = -2.73$, $p = 0.013$), happy (child, M = 0.30, SD = 1.05; adult, M = 1.07, SD = 1.40; $t(46.23) = -2.27$, $p = 0.042$), and sad (child, M = -0.78, SD = 1.16; adult, M = 0.59, SD = 1.04; $t(51.93) = -4.57$, $p < 0.001$). While children showed lower SACS than adolescents for scared objects (adolescent, M = 0.52, SD = 1.09; $t(43.71) = -3.25$, $p = 0.007$) and happy objects (adolescent, M = 1.04, SD = 0.88; $t(47.77) =$ -

2.72, $p = 0.027$), the two groups showed no difference in SACS for angry and sad objects. There was also no difference in SACS between adolescents and adults across all emotions for objects.

Examining the performance among emotions for objects, all three age groups showed the highest SACS for the happy emotion: children (happy, M = 0.30, SD = 1.05; angry, M = -1.00, SD = 1.44; scared, M = -0.84, SD = 1.80; sad, M = -0.78, SD = 1.16; happy vs. angry: $t(26) = 6.91$, $p < 0.001$; happy vs. scared: $t(26) = 5.85$, $p < 0.001$; happy vs. sad: $t(26) = 5.54$, $p < 0.001$), adolescents (happy, M = 1.04, SD = 0.88; angry, M = -0.41, SD = 1.20; scared, M = 0.52, SD = 1.09; sad, M = -0.24, SD = 1.59; happy vs angry: $t(21) = 7.63$, $p < 0.001$; happy vs. scared: $t(21) = 3.33$, $p = 0.004$; happy vs. sad: $t(21) = 5.42$, $p < 0.001$), and adults (happy, M = 1.07, SD = 1.40; angry, M = -0.34, SD = 1.37; scared, M = 0.37, SD = 1.42, sad, M = 0.59, SD = 1.04; happy vs. angry: $t(25) = 5.17$, $p < 0.001$; happy vs. scared: $t(25) = 4.22$, $p < 0.001$; happy vs. sad: $t(25) = 2.55$, $p = 0.021$).

Whereas children showed no difference in SACS for angry, scared, and sad objects, adolescents and adults did. Specifically, adolescents showed higher SACS for scared objects compared to angry and sad objects which did not differ from each other (scared vs. angry: $t(21) = 4.83$, $p < 0.001$; scared vs. sad: $t(21) = 3.70$, $p = 0.002$). By contrast, adults showed no difference in SACS for scared and sad objects, where both scared and sad objects had higher SACS than angry objects (scared vs. angry: $t(25) = 2.85$, $p = 0.013$; sad vs. angry: $t(25) = 3.74$, $p = 0.002$) – see Figure 3.7.

These SACS patterns indicate that the recognition efficiency of fear and sadness underwent great improvement in development and excelled that of anger with increasing age. Happiness remained the most efficiently recognised emotion across ages.

Prosody

When presented with prosody, there was a significant effect of Emotion (*F*(3, 216.80) = 30.57, $p < 0.001$, $\eta^2_p = 0.30$). Across age groups, prosody showed the highest SACS when conveying the happy emotion (M = 0.79, SD = 1.40), followed by angry (M = 0.13, SD = 1.34) and sad (M = -0.01, SD = 1.43), with prosody showing the lowest SACS when conveying the scared emotion (M = -0.93, SD = 1.40) (happy vs. angry: *t*(72) = 3.64, $p < 0.001$; happy vs. sad: *t*(72) = 4.58, $p < 0.001$; happy vs. scared: *t*(72) = 9.10, $p < 0.001$; angry vs. scared: *t*(72) = 5.61, $p < 0.001$; sad vs. scared: *t*(72) = 4.85, $p < 0.001$).

Song

When presented with song, there was a significant effect of Age group (*F*(2, 73.13) = 7.52, $p = 0.001$, $\eta^2_p = 0.170$). Overall, children (M = -0.54, SD = 1.70) showed lower SACS for song than adolescents (M = 0.26, SD = 1.26) and adults (M = 0.37, SD = 1.20), who did not differ from each other (child vs. adolescent: *t*(194.70) = -3.82, $p < 0.001$; child vs. adult: *t*(198.21) = -4.56, $p < 0.001$).

The effect of Emotion was also significant (*F*(3, 216.71) = 25.12, $p < 0.001$, $\eta^2_p = 0.26$). Across age groups, song showed the lowest SACS when conveying the scared emotion (M = -0.95, SD = 1.30), with no difference found across the angry (M = 0.29, SD = 1.50), happy (M = 0.35, SD = 1.43), and sad emotions (M = 0.27, SD = 1.27) (scared vs. angry: *t*(72) = -6.97, $p < 0.001$; scared vs. happy: *t*(72) = -6.78, $p < 0.001$; scared vs. sad: *t*(72) = -6.68, $p < 0.001$).

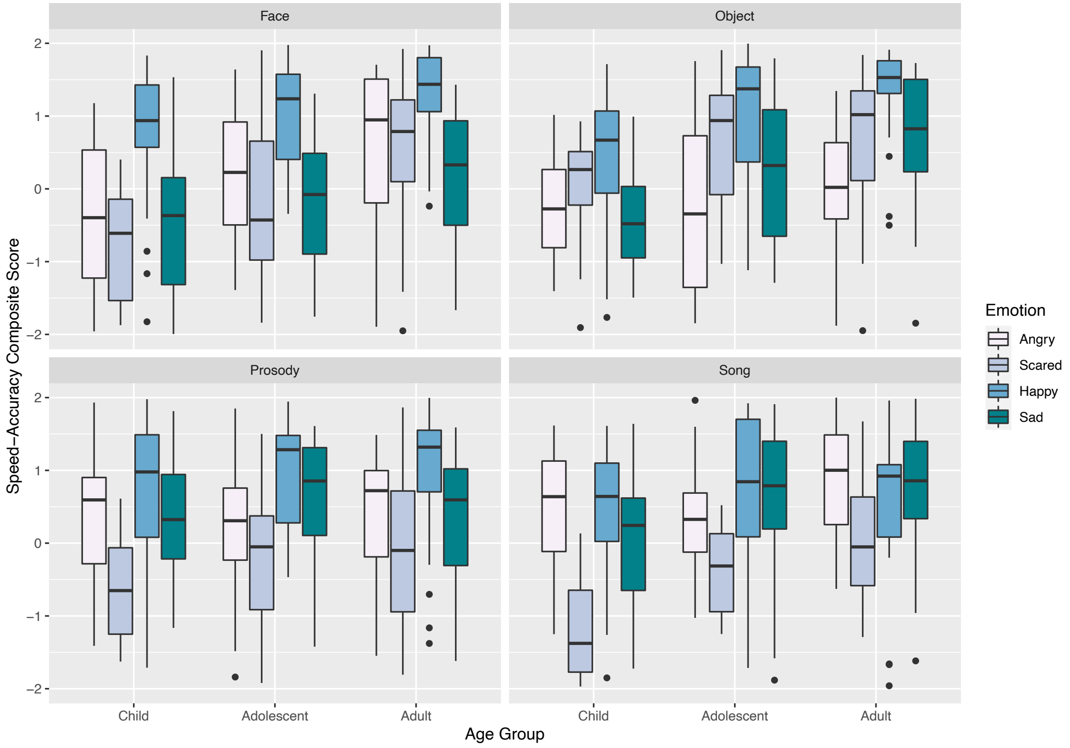


Figure 3. 7. Boxplots of mean SACS for children, adolescents, and adults in the emotion recognition task across the four conditions (face, object, prosody, and song) as a function of the four different emotions (angry, scared, happy, and sad).

Summary

Adult-level recognition efficiency appeared to be well-established by late childhood across emotions for prosody and across some emotions for objects (i.e., anger). The recognition efficiency continued to improve for the other emotions for objects (i.e., fear, happiness, and sadness) and across emotions for both faces and song beyond late childhood and reached adult-level performance during adolescence, though with one exception – the recognition efficiency of fear for faces followed a more protracted trajectory and developed throughout adolescence. Across age groups, happiness was the most efficiently recognised emotion for all conditions. While scared was the least efficiently recognised emotion for faces, prosody, and song, anger and sadness were the least efficiently recognised emotions for objects.

3.4. Discussion

The present study examined whether emotion recognition in human faces, face-like objects, speech prosody, and song differed between autistic and NT individuals across development. First, the findings revealed that the ASD and NT groups did not differ in recognition accuracy across domains and emotions. There were, however, some differences as to how the two groups misidentified emotions conveyed through faces and face-like objects. Importantly, the ASD group was generally slower and less efficient in emotion recognition than the NT group regardless of domain and emotion. Secondly, general age-related improvements were evident across both ASD and NT groups. The trajectories of these improvements, nevertheless, varied by domain and emotion depending on performance index (i.e., accuracy, RT, or SACS). In particular, the recognition of fear for both faces and song reached adult-level performance later in development. Crucially, there was no evidence for differential developmental trajectories of emotion recognition between autistic and NT individuals. Thirdly, across all groups, certain emotions were better recognised than others in the different domains. These findings are further discussed below.

3.4.1. Intact recognition accuracy at the expense of longer response time in ASD across domains and emotions

As discussed in Section 3.1, the literature has accumulated evidence to suggest that emotion recognition impairments in ASD may be specific to domain(s) and emotion(s). However, findings from separate studies could not often be compared due to variations in participant samples and the stimuli used. The present study was carried out to address these issues. It was found that autistic individuals were less efficient in emotion recognition than NT individuals across domains and emotions, as reflected in the speed-accuracy composite measure. This lower efficiency was driven by the slower recognition speed in autistic individuals relative to NT individuals, as recognition accuracy did not differ between the two groups. Contrary to previous reports of emotion-specific impairments in recognition accuracy in ASD (Ashwin et al., 2006; Boraston et al., 2007; Hubbard et al., 2017; S. Wallace et al., 2008), the present study showed no evidence for the effects of specific emotion on recognition accuracy across the different domains.

For human faces, previous studies examining both recognition accuracy and speed in autistic individuals have mostly reported impairments in both measures (Berggren et al., 2016; Eack et al., 2015; Greimel et al., 2014; Kliemann et al., 2010; Sawyer et al., 2012). In the present study, although slower recognition speed with human faces was observed, no difference in recognition accuracy was found between autistic and NT individuals. The apparent contradiction in results for accuracy could be due to differences in the experimental design. For instance, with respect to task demands, Greimel et al. (2014) employed a same-different paradigm, where participants judged whether the test face showed the same or different expression as the preceding target face on each trial. Such tasks would require accurate recognition of both the test and target expressions and/or tolerance to variations in the encoding of emotion between the target and test faces in order to provide a correct response (see Harms et al., 2010 for a discussion). Conversely, the present study employing a forced choice paradigm required accurate recognition of the test expressions only, which may have posed as a less difficult task comparatively. With respect to the emotional set used, previous studies included larger sets of emotions consisting of all six basic emotions (Berggren et al., 2016; Eack et al., 2015; Sawyer et al., 2012), whereas the present emotion set was confined to four basic emotions to ensure representativeness of expressions across all the domains of interest. In addition, given the reports of prominent difficulties with differentiating emotional from neutral faces in autistic individuals (Dalton et al., 2005; Eack et al., 2015), the inclusion of neutrality in the emotional set in previous studies may have resulted in increased task difficulty

(Berggren et al., 2016; Eack et al., 2015; Kliemann et al., 2010). With respect to stimulus presentation time, whereas previous studies restricted the presentation time of facial emotional stimuli (2500ms: Greimel et al., 2014; 150ms: Kliemann et al., 2010), these stimuli were presented until a response was made in the present study. It is plausible that autistic individuals might have used more deliberate and time-consuming processing strategies to compensate for difficulties with emotional understanding. All these factors appear to suggest that the present emotion recognition task likely encompassed lower task demands, where group differences could have been obscured on the accuracy level (Harms et al., 2010). Despite the capability to recognise the four basic emotions through human facial expressions as accurately as NT individuals, autistic individuals needed significantly more time to do so. This is consistent with previous literature that highlighted measures such as RT might be able to reveal more subtle group differences in emotion recognition ability (Uljarevic & Hamilton, 2013; M. Zhang et al., 2021).

For face-like objects, the present finding of intact recognition accuracy in autistic individuals are in line with those reported in prior studies using different types of nonhuman facial stimuli, namely schematic and cartoon faces (Brosnan et al., 2015; Davidson et al., 2019; Rosset et al., 2008). However, the slower recognition observed in the present study contradicts previous findings of comparable recognition speed for nonhuman facial expressions between autistic and NT individuals (Miyahara et al., 2007). This contradiction may be due to differences in the visual complexity of the stimuli used and/or the number and combination of emotions examined between the present and previous studies. Specifically, the face-like object images used in the present study may be more visually complex than still frames of cartoon faces used in the previous study (Miyahara et al., 2007). Moreover, the examination of four basic emotions (one positive, and three negative) in the present study may have prevented the use of elimination method to exclude emotions clearly representing the opposite valence compared to the examination of happiness (positive) and disgust (negative) in the previous study (Miyahara et al., 2007; see Grossman & Tager-Flusberg, 2012 for a discussion). These differences in methodology may have, therefore, contributed to the extended response time observed in the present study, while these speculations need to be scrutinised in future studies.

For speech prosody, the present findings replicated previous studies in demonstrating accurate, yet slower, emotion recognition in autistic individuals relative to NT individuals (Ketelaars et al., 2016; Kujala et al., 2005; Lindström, 2019; Waddington et al., 2018). However, the present study did not provide evidence that reduced perceptual salience obtained by shorter stimulus durations would likely detect group differences in recognition accuracy. This contradicted previous studies that also used short utterances as experimental stimuli and showed impaired recognition accuracy in autistic individuals (Doi et al., 2013; Heaton et al., 2012; Schelinski & von Kriegstein, 2019; Wang & Tsao, 2015). Beyond stimulus duration, there remain other differences between the stimulus sets used in the present study and those in previous studies. One difference relates to the number of syllables within the spoken words: whereas the present study used monosyllabic utterances (i.e., “door”) as experimental stimuli, multisyllabic utterances were used in previous studies (i.e., /su-zu-ki-san/, Doi et al., 2013; three-digit numbers, Heaton et al., 2012; i.e., two-syllable German nouns, Schelinski & von Kriegstein, 2019). Given that syllables at different positions of a word carry emotion specific information (Rao et al., 2013), the use of multisyllabic utterances in the previous studies might have required the integration of global prosodic features across syllables within a short time frame, increasing the difficulty of the task. Secondly, the monosyllabic utterances used in the present study were the final words of the sentence stimuli obtained from the RAVDESS database (Livingstone & Russo, 2018). Conversely, stimuli in the previous studies were specifically recorded as utterances which were not originally embedded in sentences (Doi et al., 2013; Heaton et al., 2012; Schelinski & von Kriegstein, 2019). On a sentence level, words in the final position of sentences have been found to exhibit more emotion discriminative information compared to words presented in other positions (Rao et al., 2013). As such, the stimuli used in the present study may have been richer in emotional content, and hence reducing the task difficulty comparatively. Finally, whereas the present study used the same emotional utterance throughout the experiment, different utterances were used in previous studies (Heaton et al., 2012; Schelinski & von Kriegstein, 2019; Wang & Tsao, 2015). Accordingly, variations in the prosodic structures of individual utterances may have influenced the ease of identifying emotions trial-by-trial. Altogether, the heterogeneity and complexity of the intonation pattern across and within stimuli might have contributed to the disparity in task difficulty and group differences between the present study and previous studies.

For song, the present findings of comparable recognition accuracy between the two groups are in consistent with and extend those reported in prior studies using instrumental music stimuli (Gebauer, Skewes, Westphael, et al., 2014; Heaton et al., 1999; Järvinen et al., 2016; Quintin et al., 2011). However, the present study did not corroborate previous findings that showed comparable emotion recognition speed between autistic and NT individuals with instrumental music (Quintin et al., 2011). It should be noted that in this previous study, participants listened to the music clips for 7s prior to making an emotional judgment (i.e., the point at which response times were obtained; Quintin et al., 2011), whereas response times were taken from stimulus onset in the present study; it is plausible that any differences in emotion recognition during the long period of listening time may have been unnoticed in the previous study. In addition, it remains to be elucidated whether impaired recognition speed in autistic individuals in the present study resulted from the use of shorter stimulus length and/or the musicality of the stimuli (i.e., vocal vs. instrumental).

The overall comparable emotion recognition accuracy between the ASD and NT groups is unsurprising, given that the groups were matched on nonverbal, and especially, verbal ability. Previous studies that matched groups on verbal and/or nonverbal abilities have also tended to find no group differences (Davidson et al., 2019; Fink et al., 2014; J. B. Grossman et al., 2000; Ozonoff et al., 1990; L. J. Taylor et al., 2015). Moreover, as shown in Chapter 2, the implementation of full-scale/verbal/nonverbal IQ matching can substantially weaken the magnitude of observed group differences. It is thought that intellectual ability may constitute a compensatory mechanism for emotion recognition in ASD (Rutherford & Troje, 2012). In support of this, correlations between IQ and emotion recognition ability have been observed in ASD but not in NT (Dyck et al., 2006b; Koldewyn et al., 2010; Tanaka et al., 2012). This implies that emotion recognition may tap into higher-level processes involving more analytical methods in ASD, whereas more intuitive strategies are employed in NT (J. B. Grossman et al., 2000). However, due to a common profile of lower verbal IQ compared to nonverbal IQ in ASD, the absence of matching on verbal IQ may artificially inflate group differences in emotion recognition tasks if the threshold of verbal IQ in the ASD group falls below that in the NT group (Harms et al., 2010). The present study, thus, showed that autistic individuals accurately recognised emotions to similar extents to NT individuals where confounding effects of intellectual ability were eliminated.

Importantly, results showed that autistic individuals were, in fact, able to accurately recognise emotions from different types of stimuli at the expense of longer response time. It is possible that autistic individuals lack the spontaneous decoding of emotional information rather than the capacity to decode this information. Several hypotheses seem plausible in accounting for the reason for this delay in response. The slower recognition speed could be attributed to a more cautious and time-consuming cognitive approach to the task in autistic individuals, as opposed to the more intuitive strategies that would allow for effortless recognition of emotions in NT individuals (J. B. Grossman et al., 2000; Livingston & Happé, 2017; Rutherford & McIntosh, 2007). For instance, autistic individuals may extract facial information from local features, rather than using higher-level configural processing for human facial expressions (Behrmann et al., 2006; Pelphrey et al., 2002). In support of this hypothesis, studies incorporating brief (e.g., < 50ms; Clark et al., 2008; Otsuka et al., 2017) or restricted presentation time (e.g., 300-3000ms; Brewer et al., 2016; Greimel et al., 2014; Griffiths et al., 2019) revealed compromised recognition accuracy in autistic individuals, as opposed to studies with no restriction on presentation time which found preserved recognition accuracy (e.g., Fink et al., 2014; Lacroix et al., 2014). These suggest that accuracy might be affected when the use of alternative strategies is prevented. The applicability of this hypothesis to face-like objects is, however, unclear, given the dissimilar reports of typical configural processing (Akechi et al., 2014; Guillon et al., 2016; Rosset et al., 2008) as well as featural processing of nonhuman faces (Isomura et al., 2014; Riby & Hancock, 2009). In addition, this deliberative processing approach may not accommodate the recognition of emotions from stimuli that are dynamic in nature, such as speech prosody and song. Whereas static visual stimuli provide constant emotional cues throughout presentation and benefit from more detailed attention to specific features, auditory stimuli are inherently dynamic and require listeners to track and integrate acoustic information that develops over time (Schirmer & Adolphs, 2017). Yet, recognition accuracy remained unaffected while slower recognition speed was still observed for speech prosody and song in autistic individuals.

Another plausible explanation of the slower recognition speed could be reflective of the amount of cognitive resources devoted to the processing of emotional stimuli. In relation to this, atypical orientation towards face-like objects (Akechi et al., 2014; Guillon et al., 2016; Pavlova et al., 2017; Ryan et al., 2016) and speech sounds (Čeponienė et al., 2003; Lepistö et al., 2005), as well as reduced sensitivity towards prosodic and musical expressivity (Bhatara et al., 2010; L. S. Brown, 2017; Gebauer, Skewes, Hørlyck, et al., 2014) may contribute to such delays.

Alternatively, the slower recognition speed across domains might be the result of overall slower processing speed in ASD. In line with this, prior studies have reported a fundamental impairment in the speed with which autistic individuals can process information (Haigh et al., 2018; Mayes et al., 2007; Velikonja et al., 2019). Furthermore, the relationship between slower processing speed and measures of social cognition have been illustrated (Haigh et al., 2018; Hedvall et al., 2013; Oliveras-Rentas et al., 2012; Russo-Ponsaran et al., 2015). These possibilities, however, require further investigation to ascertain whether and how processing strategy, cue sensitivity, and processing speed contribute to the accuracy and speed of emotion recognition in ASD.

3.4.2. Differing misclassification patterns for visual but not auditory stimuli between autistic and NT individuals

The similar misattribution patterns between autistic and NT individuals for speech prosody and song indicate that the two groups may employ similar mechanisms to process emotions within the auditory modality. In speech prosody, the analysis of pitch features is thought to contribute to the perception of emotions (R. Gold et al., 2012; Krumhansl & Shepard, 1979). Likewise, pitch also affects the perception of emotions in song, although its expressive power to communicate emotional content in song has been found to be diminished (Hakanpää et al., 2019a, 2019b). As discussed in Section 2.4.3, a previous study found that, similar to NT individuals, recognition of emotional prosody was strongly associated with non-vocal pitch processing in autistic individuals (Globerson et al., 2015). The relationship between non-vocal pitch processing and emotion recognition in song between the two groups, on the other hand, is not well-known. To this end, whether non-vocal pitch processing underlies auditory emotion recognition in autistic and NT individuals similarly will be further explored in Chapter 5.

Autistic individuals exhibited misclassification patterns differing from that by NT individuals with human faces and face-like objects. It was found that for faces, autistic individuals were more likely to misinterpret scared, happy, and sad faces as angry. For face-like objects, autistic individuals were more likely to misinterpret sad objects as happy, but they were less likely to misinterpret angry and scared objects as sad compared to NT individuals. The different misclassification patterns may imply differential ways of decoding facial expressions by the two groups. Inspection of these misclassification patterns highlight that autistic individuals were not only susceptible to within-valence confusions, but also to cross-valence confusions. Indeed, autistic individuals have previously been reported to make both within-valence (Eack et al., 2015; Kennedy & Adolphs, 2012) and cross-valence confusions when recognising emotions from human faces (Whitaker et al., 2017). The present findings, additionally, shed light on the misclassification patterns for face-like objects by the two groups. In fact, over the course of development, NT children have been shown to form superordinate valence-based categories (e.g., positive vs. negative) that slowly differentiate into discrete emotional categories (e.g., angry vs. sad) (Widen, 2013). Upon this gradual maturation, NT children exhibit more systematic errors between emotions of the same valence and rarely confuse emotions of different valences (J. A. Russell & Widen, 2002; Widen, 2013). By contrast, in the present study, autistic individuals exhibited less secure valence-based boundaries, which could disrupt the formation of discrete categories due to persistent cross-valence confusions. In other words, the mental representations for negative expressions (e.g., sad) could be activated by both negative (e.g., angry) and positive expressions (e.g., happy) in autistic individuals. Together, these findings suggest that autistic individuals may have atypical prototypes across emotions, hence reflecting lower sensitivity to boundaries distinguishing between emotions as illustrated in various previous studies (Law Smith et al., 2010; Otsuka et al., 2017; Song & Hakoda, 2018; S. Wang & Adolphs, 2017).

With regards to human faces, one possibility may relate to the atypical processing of these stimuli in autistic individuals as noted in numerous eye-tracking studies (see Harms et al., 2010 for a review). While anger, fear and sadness carry the most salient information in the eye region, happiness carries the most salient information in the mouth region (Bombari et al., 2013; Calder et al., 2000; Eisenbarth & Alpers, 2011). Given the importance of facial features in communicating different emotions, the more frequent misclassifications among these emotions in autistic individuals may, therefore, be attributed to the diminished attention to the core features of the face (Pelphrey et al., 2002), and particularly the eyes when viewing human emotional faces (Corden et al., 2008; D. Neumann et al., 2006; Pelphrey et al., 2002; Reisinger et al., 2020).

On the other hand, previous studies have shown that, despite being less attuned to detect faces in face-like objects in the first instance, autistic individuals process and perceive face-like objects similarly to NT individuals (Akechi et al., 2014; Guillon et al., 2016). Autistic individuals, like NT individuals, showed a preference for face-like objects when displayed in an upright relative to an inverted orientation (Guillon et al., 2016). This finding provides evidence for intact configural processing of face-like objects in autistic individuals (Guillon et al., 2016). Together, it seems unlikely that the different misclassification patterns observed are due to differences in processing strategies employed by the two groups in the case of face-like objects. It is, however, noteworthy that the studies mentioned above provided evidence for the processing of face-like objects in the absence of emotional judgment (Akechi et al., 2014; Guillon et al., 2016). Thus, whether differences in perception of emotional face-like objects between the two groups are unrelated to processing strategies would need to be confirmed by future studies. Another interesting question that remains to be addressed in future studies would be whether processing strategies are shared across human faces and face-like objects in both ASD and typical development. To this end, whether the frequently reported preference for local information processing style (Baron-Cohen, 2008; Happé & Frith, 2006; Mottron et al., 2006) is related to emotion processing differences between autistic and NT individuals across these stimuli will be further explored in Chapter 5.

3.4.3. The functions of emotion type and stimulus domain in the development of emotion recognition

Across both autistic and NT individuals, age-related improvements were observed for all domains and emotions. The patterns of improvement, nevertheless, varied by domain and emotion according to the different performance indices, namely recognition accuracy, speed, and efficiency (i.e., an index of accuracy and speed combined). It was found that emotion-specific trajectories were observed for faces, face-like objects, and song, whereas the recognition of the different emotions developed in parallel for speech prosody.

For human faces, the present study corroborated previous findings in demonstrating that happiness and sadness were earlier-emerging emotion categories, followed by anger, and then fear in terms of accuracy (Durand et al., 2007; L. A. Thomas et al., 2007; Widen & Russell, 2008). Although children recognised happiness and sadness as accurately as adults, the present findings additionally showed that adult-level performance was in fact not fully established for these emotions when taking recognition speed and efficiency into account: children recognised happiness and sadness slower and less efficiently compared to adults. Conversely, although children recognised anger less accurately than adults, they were just as fast as adults at recognising this emotion. The recognition of fear showed the most prominent improvement overall and continued to develop throughout adolescence, which parallels previous findings (Herba et al., 2006). The continued development in the different measures of emotional recognition in faces may be related to the increase in sensitivity to subtle configural changes in facial expressions across emotion boundaries as children get older (Rump et al., 2009; L. A. Thomas et al., 2007). It has been postulated that with age, children develop expertise in configural processing of faces (Mondloch et al., 2002; Schwarzer, 2000) and continue to develop more refined prototypes of each facial expression through learning and experience (Rump et al., 2009). This refinement would allow for more sophisticated and precise comparisons between the expressions encountered and the mental representations of prototypes when making an emotional judgment (Rump et al., 2009). Accordingly, the less rigid mental representations of prototypes may be a plausible explanation for the poorer efficiency in facial emotion recognition in children.

The present findings provide novel evidence for the development of emotion recognition from nonhuman facial stimuli, namely face-like objects, which has not been investigated previously. Results showed that adult-level accuracy was mostly attained by late childhood when recognising anger, fear, and happiness but not sadness. Across emotions, children were generally slower compared to the older groups. Taking into account both accuracy and speed, the developmental trajectories differed by emotion: whereas fear and happiness showed an abrupt improvement in efficiency from childhood to adolescence, sadness improved more gradually across ages. However, no age-related improvement in efficiency was observed for anger. It may be speculated that the speed and efficiency of processing emotions in face-like objects in children might have been impeded by their proficiency in detecting "faces" in these stimuli. A previous study has outlined the crucial role of configural processing in detecting faces in face-like objects (Ichikawa et al., 2011). As mentioned earlier, children have been shown to be less capable of taking advantage of configural information when processing faces compared to adults (Mondloch et al., 2002; M. J. Taylor et al., 2004). It is plausible that the proficiency in face detection in face-like objects is poorer at a younger age, and thus impedes the subsequent processing of emotions. However, based on the current data, it is not possible to discern whether age-related differences in recognition speed and efficiency is due to slower detection of "faces" in these stimuli or simply a maturation process of developing skills relevant for emotion recognition as noted with human faces. It would be interesting to further explore these relationships in future studies to gain greater insights into the associations between stimulus familiarity and emotion processing across development.

For speech prosody, the present study showed that children exhibited adult-level speed and efficiency for all emotions. Although the present study failed to show adult-level accuracy by late childhood as observed in Vidas et al. (2018), it supported previous findings by Chronaki et al. (2015) in showing that emotion recognition of transient prosodic cues continues to develop beyond late childhood. Furthermore, the inclusion of an adolescent group in the present study outlined that the recognition of different emotions in speech prosody continues to improve substantially in parallel beyond late childhood. The discrepancy between studies may be attributed to varying stimulus lengths: stimuli were utterances of three-digit numbers lasting for 2-3s in Vidas et al. (2018), whereas stimuli were short prosodic interjections and monosyllabic words lasting for 700ms and 500ms in Chronaki et al. (2015) and in the present study, respectively. Given the time course of which recognition accuracy rate for prosodic emotions stabilises at around 500-600ms (Pell & Kotz, 2011), the use of prosodic stimuli within this timeframe may be particularly sensitive in demonstrating age-related differences. One reason for the difference between children and adults in their recognition accuracy for speech prosody may relate to their use of cues to infer speakers' emotions. Studies have shown that whereas children tend to rely on semantic content when judging emotions from speech (Aguert et al., 2010; Friend & Bryant, 2000; Morton & Trehub, 2001; Waxer & Morton, 2011), adults make use of prosody as an overriding cue (Aguert et al., 2010). This developmental shift in the use of cues for emotional judgment of speech has been found to begin at approximately 9 years of age (Aguert et al., 2010). This transitional period in development has been suggested to be relatively long, extending to approximately 12 years of age (Le Sourn-Bissaoui et al., 2013). As the prosodic stimuli used in the present study were semantically neutral, poorer accuracy in children and adolescents may reflect this transition towards better proficiency in decoding emotional cues based on prosody.

Results showed that children performed just as fast as the older age groups when recognising emotions from song. In addition, children were also just as accurate as the older age groups at recognising anger, happiness, and sadness from song. The recognition of fear was, however, less accurate in both children and adolescents compared to adults. This suggests a more protracted development of fear recognition with song, which mostly improved throughout adolescence beyond late childhood. Despite the mostly comparable performance seen in the separate performance indices, children did not show the same efficiency demonstrated by adults. These findings, together, partially echoed previous findings of children achieving adult-level accuracy by 10-11 years of age across all four emotions with instrumental music in Vidas et al. (2018). The discrepancy in the findings for fear recognition may have arisen due to the different forms of musicality in Vidas et al. (2018) and in the present study. This particular age-related difference observed for fear recognition may be further explained by the relatively poorer communicative power of this emotion compared to anger, happiness, and sadness through song as shown in previous listener studies (Livingstone & Russo, 2018).

Despite the age differences discussed above, the recognition differences displayed as a function of emotion by domain across the three age groups, nevertheless, showed some similarities. Happiness was the most well-recognised emotion from all domains, whereas fear and sadness were the least well-recognised emotions from faces, anger from objects, and fear from prosody and song. These emotion-specific effects on recognition in the different domains have also been reported in previous studies (Calvo & Lundqvist, 2008; Chronaki et al., 2018; Elfenbein & Ambady, 2002; Gaspar et al., 2011; Juslin & Laukka, 2003; Livingstone & Russo, 2018; Pell, Monetta, et al., 2009). Interestingly, these patterns remained relatively stable across age groups. This suggests that recognition differences as a function of emotion by domain may not be subject to any delays in development (i.e., not due to one emotion developing later in the developmental course compared to other emotions in a given domain).-Rather, this may imply that some emotions provide more salient cues than others in a given domain and that different domains do not merely carry redundant information for conscious emotion processing (App et al., 2011; Elfenbein & Ambady, 2002). This notion may be implicated in evidence demonstrating that combining and integrating multisensory cues can facilitate emotion recognition (Collignon et al., 2008; Klasen et al., 2014; Paulmann & Pell, 2011; Schirmer & Adolphs, 2017).

Importantly, there was no evidence that emotion recognition follows different developmental trajectories between autistic and NT individuals across domains and emotions, though autistic individuals showed slower and less efficient emotion recognition across ages. This finding appears in contrast to those reported in Rump et al. (2009), by which age-related improvement in NT individuals was not observed in autistic individuals during facial emotion recognition. It is noteworthy that in this previous study, the diverging trajectories were found for the recognition of dynamic facial expressions under brief exposure times while controlling for the subtlety of expressions (Rump et al., 2009). In contrast, the present study demonstrated that autistic and NT individuals may indeed show similar developmental trajectories for the recognition of prototypical expressions. The trajectories may, nonetheless, diverge as NT individuals continue to refine their emotional recognition skills for discriminating between more subtle expressions. Future research may be directed to examine whether this is true for domains other than human faces.

3.4.4. Limitations

Some limitations relating to the study design must be considered when interpreting the present findings. First, although attempts were made to closely match stimuli across domains *within* the same modality, there were substantial differences in presentation mode and duration when comparing domains *between* modalities. Specifically, while images of human faces and face-like objects were static and were presented until a response was made, clips of speech prosody and song were dynamic in nature and lasted only for 500ms. Such manipulations, in fact, mirrored those designed for the cross-modal affective priming study, which will be further described in Chapter 4 (Section 4.2). Thus, in addition to examining the proposed research questions of this chapter, the present data provided a baseline to indicate participants were able to adequately recognise emotions from this set of stimuli under the same presentation mode (i.e., static or dynamic) and duration (i.e., restricted or unrestricted) as that in Chapter 4 but in the absence of priming. This is to ensure any effects of priming observed in Chapter 4 were not confounded by poor emotion recognition of these stimuli. The differences in presentation mode and duration between modalities may, nonetheless, result in discrepant levels of difficulty for recognition (e.g., emotion recognition of auditory stimuli may be more difficult given the need to track acoustic information overtime within a short duration). Despite such differences, no effects of domain were observed, substantiating the finding of generalised emotion recognition ability across domains in autistic individuals. The replicability of the present findings can be scrutinised through future studies employing further stimulus matching criteria.

Secondly, the present task required participants to respond by pressing a corresponding key on a response pad that matched the emotional label displayed on the screen. This requirement of coordination might have been particularly demanding for autistic participants, who often present with motor coordination difficulties (Fournier et al., 2010; Fulceri et al., 2019). The possibility of motor (rather than emotion recognition) difficulties resulting in slower performance in autistic participants, therefore, could not be ruled out. It is important that future studies include a control task that measures motor coordination (e.g., word-word matching task which involves participants pressing a matching word on the keyboard in response to the target word presented on the screen; Fink et al., 2014), as well as one that measures non-social processing speed (e.g., a subtest of the Trail Making Test which involves participants drawing a continuous line to connect numbers 1-25 in ascending order; Battery, 1994). The inclusion of different control tasks will allow potential effects of these abilities to be teased apart from emotion recognition ability, and hence provide better interpretation of results arisen between the ASD and NT groups.

Thirdly, it is important to note that there was substantial variability of performance in the current dataset. Although all participants performed at above chance level across conditions with RT outliers of each of their performance removed, variability in performance was visible when the dataset was further broken down by emotion for each condition (as can be seen in the series of boxplots presented in Sections 3.3.1 and 3.3.2). Notably, outliers were predominantly driven by children's performance when data were grouped by diagnostic group. This is not surprising, given the challenges of developing an emotion recognition task to accommodate a wide age range of participants. The variability in performance may partially reflect difficulty of the task for children, which may be overcome by adopting an adaptive paradigm in future work with participants only proceeding to the task when they have achieved a certain level of scores (e.g., Rump et al., 2009). Moreover, while particularly poor performance (e.g., denoted as outliers) may be indicative of inattentiveness on the task, this does not seem to be the case for the current dataset. With close inspection of individuals' performance among the full dataset, outliers often occurred for one particular emotion; for example, one participant scored 75%, 100%, and 63% for angry, happy, and sad emotions but 13% for the scared emotion, which may represent important effects of emotion rather than influences of inattentiveness on items expressing the same emotion that were presented in random order throughout the experiment. Future studies should, however, include catch trials to ensure detection of performance potentially influenced by inattentiveness, in order to provide better clarity on such interpretation.

3.4.5. Chapter summary

In summary, the present study found that autistic individuals were just as accurate as their NT counterparts at recognising prototypical emotions from human faces, face-like objects, speech prosody, and song, when given sufficient processing time. As slower recognition speed was found across domains and emotions, these results contradicted previous views of domain- and/or emotion-specific difficulties in ASD. The generally slower recognition speed could be attributed to a more deliberative processing strategy, poorer orientation to socioemotional significance of stimuli, and slower processing speed, which do not necessarily need to be mutually exclusive. Despite the observation of comparable emotion recognition accuracy for both human faces and face-like objects, there were differences in the misclassification patterns between autistic and NT individuals for these stimuli but not for speech prosody and song. This suggests that autistic individuals may employ different strategies for emotion processing of human faces and face-like objects compared to NT individuals, whereas similar mechanisms may underlie emotion processing of speech prosody and song in the two groups. The present study also adds to the extending literature on emotion recognition not only in ASD, but also in typical development, by exploring the developmental trajectories across different domains and emotions. Age-related differences were observed, which varied by domain, emotion, and performance index, which were independent of diagnostic group differences. This indicates that although improvement in emotion recognition along the developmental course was seen in both groups, the speed of emotion recognition in autistic individuals did not reach the same level as that in NT individuals even with increased age – likely a difference that persists across the lifespan in ASD. As interpersonal communication proceeds in a time-based manner, delayed interpretation of emotion information may hinder individuals' adaptation to a social situation in ASD. Altogether, these results outline that although accuracy and speed of emotion recognition appears to generalise across nonverbal communicative domains, the underlying mechanisms of this process may be specific to modality, or even domain.

Chapter 4: A behavioural study of cross-modal affective priming of speech prosody and song on human faces and face-like objects in ASD and NT

Furthering the investigation of explicit emotion processing in Chapter 3, this chapter examined the domain generality and specificity of implicit emotion processing in autism spectrum disorder (ASD) using a cross-modal affective priming paradigm – specifically, whether implicit emotion priming generalises across the speech prosody and song domains or is specific to either one domain in autistic individuals. The present study involved the same participants who took part in the study presented in Chapter 3, which will allow for more reliable comparisons between performance at the implicit versus the explicit levels.

4.1. Introduction

Emotion processing can operate not only at the explicit/conscious level, but can also occur in an implicit/unconscious mode (Celeghin et al., 2020; Clausi et al., 2017; Habel et al., 2007; Jessen & Grossmann, 2015; Lane, 2008; Liddell et al., 2004; Okon-Singer et al., 2007). Research has shown that implicit emotion processing is distinctive from conscious emotion processing, with the former representing a fast, automatic, and stimulus-driven process and the latter a slow, capacity-limited, attentional demanding process (Bargh & Chartrand, 1999; Birnboim, 2003; K. R. Scherer, 2001; W. Schneider & Chein, 2003; W. Schneider & Shiffrin, 1977). There appears to be a dissociation between the two processes, such that impairment at the explicit level does not necessarily imply impairment at the implicit level (Roux et al., 2010; Wagenbreth et al., 2016; Wieser et al., 2006). Thus, following the investigation of emotion processing at the explicit level in Chapter 3, this chapter focuses on the emotion processing ability at the implicit level in individuals with autism spectrum disorder (ASD) compared to neurotypical (NT) individuals.

As presented in Section 1.2.1.1, previous research using various measurements has mostly noted impairments in implicit emotion processing in autistic individuals. For example, there is evidence from functional magnetic resonance imaging (fMRI) studies which have consistently reported reduced levels of activation in brain regions in autistic individuals compared to NT individuals during implicit emotion processing of facial and body expressions irrelevant to the central tasks (Ciaramidaro et al., 2018; Critchley et al., 2000; Kana et al., 2016). Similarly, evidence from electroencephalogram (EEG) studies have reported diminished neural response to emotionally spoken syllables during passive listening tasks in autistic individuals compared to NT individuals (Fan & Cheng, 2014; Lindström et al., 2018). These findings, thus, appear to suggest that ASD may be associated with impairments in implicit emotion processing on the neural level, but more specifically, the implicit appraisal of emotions and implicit discrimination between emotional expressions.

The implicit processing of emotional information can be examined not only neurologically, but also behaviourally. For example, emotional meaning of different expressions can be implicitly processed to influence one's subsequent behaviour. Where emotion perception often takes place in a multisensory environment in real-world settings (de Gelder & Bertelson, 2003; Lewkowicz & Ghazanfar, 2009), the emotional information induced by a preceding stimulus from one modality can modulate the emotional judgment of a stimulus in another modality. This phenomenon has been demonstrated using the cross-modal affective priming paradigm (Carroll & Young, 2005; Degner, 2011; Hsu & Schütt, 2012; Murphy & Zajonc, 1993). Under this paradigm, auditory and visual emotional cues conveying either congruent or incongruent emotional connotations are presented consecutively. Early explanations of this phenomenon proposed the spreading of activation is an underlying mechanism of affective priming – that is, the preceding prime stimulus is thought to activate emotionally congruent representations by spreading activation throughout the conceptual network. The preactivated representations, thereby, facilitate the encoding of congruent targets (Collins & Loftus, 1975; De Houwer & Randell, 2004; Hermans et al., 1994). In support of these accounts, it has been shown that emotional judgment of target stimulus is faster and more accurate when the target is preceded by a prime of a congruent emotion (e.g., angry-angry) than when the prime and target are of different emotions (e.g., angry-sad) (Carroll & Young, 2005). Importantly, the manipulation of the time interval between the prime and target onsets, known as the stimulus onset asynchrony (SOA), is essential for capturing the early, automatic processing of the prime stimuli in such paradigms (De Houwer et al., 1998; Hermans et al., 2001; Klauer, 1997; Neely, 1977). Specifically, automatic priming can only be assessed in cases of very short SOAs (e.g., < 300ms; Posner & Snyder, 2004).

In the field of autism research, the cross-modal priming paradigm has been less commonly used to study emotion processing in autistic individuals, yet findings have not been consistent. One study by Kamio et al. (2006) investigated the priming effects of emotional faces (subliminally presented for 16ms followed by a dot pattern mask for 484ms) on subsequent liking ratings of Japanese ideographs in individuals with high-functioning pervasive developmental disorders (HFPDD). Despite comparable recognition of the emotional faces at the conscious level compared to NT individuals, the emotional faces only primed subsequent information processing in the NT group but not in the HFPDD group (Kamio et al., 2006). By contrast, Vanmarcke and Wagemans (2017) found that the priming effects of both coarse and fine emotional facial primes (presented for 83ms followed by a perceptual mask of 83ms) on valence judgment of emotional faces did not differ between autistic and NT individuals. Beyond these disparate findings regarding implicit emotion priming of emotional faces, it remains unclear whether implicit emotion priming of auditory emotional cues is impaired in ASD. In a study by Ben-Yosef et al. (2017), autistic individuals failed to exhibit a priming effect of affective vocalisations on subsequent processing of emotional face targets when the cognitive load was high (i.e., when face targets were presented at low frequencies). However, in this study, the target stimuli were presented following the presentation of the prime stimuli, the duration of which lasted for ≥ 1500ms. The SOA in this study appears to exceed the critical timeframe for automatic priming to be captured, as discussed earlier (De Houwer et al., 1998; Hermans et al., 2001; Klauer, 1997; Neely, 1977). Taken together with these inconsistent findings, it is yet to be elucidated whether implicit emotion priming is impaired in ASD, especially in the case of auditory emotional primes.

With specific regards to auditory stimuli, priming effects have been demonstrated with both emotional prosody (Jaywant & Pell, 2012; Pell, 2005; Pell et al., 2011; L. D. Scherer & Larsen, 2011; Schwartz & Pell, 2012) and musical chords (L. Zhou et al., 2019) in neurotypical development. However, the implicit emotion priming of auditory stimuli between autistic and NT individuals has not been previously examined. Evidence from the multisensory integration literature may provide some insights into this. These studies vary in their designs substantially, including those with synchronous presentation of congruent cross-modal stimuli (Charbonneau et al., 2013; Vannetzel et al., 2011; Xavier et al., 2015) and synchronous presentation of in/congruent cross-modal stimuli (O'Connor, 2007; Wagener et al., 2020). Despite the varying study designs, these findings appear to converge on reduced cross-modal integration of vocal stimuli (Charbonneau et al., 2013; O'Connor, 2007; Vannetzel et al., 2011; Xavier et al., 2015), but typical cross-modal influence of musical stimuli in ASD (L. S. Brown, 2017; Wagener et al., 2020). As noted, it is unclear whether the discrepant patterns observed between the two domains are reflective of conscious or automatic processing. With this in mind, the direct comparison between the prosody and music domains within a cross-modal prime-target paradigm using a short SOA will contribute to the understanding of the domain specificity (or otherwise) of implicit emotion priming in ASD and typical development.

The development of implicit emotion processing, particularly in the case of priming, has not been well-explored in either ASD or typical development. In typical development, there is some evidence showing that implicit influence of emotional information on the discrimination between old and new emotional faces improved with age from childhood and adolescence through into adulthood (Mathersul et al., 2009; L. M. Williams et al., 2009). By contrast, age was not a significant correlate of implicit emotion priming of faces on preferential responses in typical development, which was, in fact, also true in ASD (Kamio et al., 2006). Further evidence from separate studies on implicit emotion processing has also demonstrated that impairments in ASD are found across children (Lindström et al., 2018), adolescents (Riby et al., 2012), and adults (Ciaramidaro et al., 2018; Critchley et al., 2000; Fan & Cheng, 2014; Kana et al., 2016) on both neural and physiological levels. These findings, together, suggest that age may not be an important moderator of implicit emotion processing in ASD or typical development.

As discussed in Section 2.1.1.2, there is a body of work on conscious emotion processing suggesting that deficits in ASD may not be fundamental across all emotions but specific to certain emotions. The findings of Kamio et al. (2006), however, provided no evidence for any modulating effects of specific emotion on the impairments observed in implicit emotion priming among autistic individuals. Notably, this study only examined the priming effects of happy and fearful faces; thus, whether emotion-general performance could be replicated when examining additional emotions is yet to be explored. Furthermore, the review's findings from Chapter 2 highlighted emotion-general impairments for human faces but emotion-specific impairments for speech prosody (i.e., anger, happiness, and disgust) and music (i.e., fear and sadness). It is plausible that specific emotions may have different effects depending on the stimulus domain – though results from Chapter 3 showed no emotion-specific impairments. The effects of specific emotion on implicit emotion priming of auditory stimuli will, therefore, be examined in the present study.

To my knowledge, there have been no studies investigating the implicit processing of emotional speech prosody and song through examining their influence on subsequent emotional judgment of human faces and face-like objects between ASD and neurotypical development across ages. Therefore, the present study was carried out to address this issue. By using the same stimulus types as those on the emotion recognition tasks in Chapter 3, it was possible to examine priming between prime-target pairs of higher co-occurrences in the environment (i.e., song-face and particularly speech-face) and prime-target pairs of less frequent co-occurrences (i.e., song-object and speech-object). Namely, whether prime-target pairs that co-occur more frequently in the environment play a privileged role in cross-modal affective priming was examined. In essence, human faces, but not everyday objects, are more frequently encountered with accompanying human voices regardless of whether it is spoken or sung. With regards to this, Carroll & Young (2005) found emotion priming effects for high co-occurrence prime-target pairs (e.g., vocal bursts-faces), but not for low co-occurrence prime-target pairs (e.g., vocal bursts-printed words). Note that this effect was only found in terms of accuracy, whereby priming was found to be equivalent across low and high co-occurrence pairs in terms of response time (Carroll & Young, 2005). Furthermore, this effect has not been previously studied in autistic individuals. It is, therefore, of interest to examine whether higher co-occurrence between primes and targets would strengthen cross-modal affective priming to similar extents in autistic individuals relative and NT individuals.

In this exploratory study, a cross-modal affective priming paradigm was implemented, where an auditory prime (speech prosody or song) expressing either a congruent or incongruent emotion to the target was presented to the participants on each trial. Participants were instructed to judge the emotion expressed in the visual targets (human faces or face-like objects) that were subsequently presented following a short delay of 200ms – an appropriate SOA confirmed to be reflective of automatic, rather than conscious emotion processing (Hermans et al., 1994, 2001). The effects of priming, as indexed by the difference in performance between the congruent and incongruent conditions, was compared between diagnostic groups across age groups. In addition, the domain specificity (or otherwise) of implicit emotion priming within the auditory modality, and the moderating effects of age group, specific emotion, and cooccurrence of cross-modal stimuli between the ASD and NT groups were examined. The specific research questions are:

1. Are there differences in cross-modal emotion priming between the ASD and NT groups?
2. Are there differences in the developmental trajectory of cross-modal affective priming between the ASD and NT groups?
3. Are differences, if any, in cross-modal priming effects modulated by prime type (speech prosody vs. song), target type (face vs. face-like object), and prime emotion (angry vs. scared vs. happy vs. sad)?

Based on previous findings from the multisensory integration literature as discussed earlier, it is anticipated that the ASD group will show weaker priming of speech prosody, but comparable priming of song, relative to the NT group. While age does not appear to moderate implicit emotion priming, the sparse data and infrequent involvement of child and adolescent participants in this area of research give rise to the need for examining cross-modal emotion priming in the two populations from a developmental perspective. Similarly, limited research suggests no moderating effects of specific (prime) emotion on priming; it remains exploratory whether such effects will reveal with a larger set of emotions employed. Finally, it is hypothesised that face targets will induce stronger priming than face-like object targets, given their higher co-occurrences with speech prosody and song. Again, whether the effects of prime-target co-occurrence will moderate priming differently in the ASD and NT groups remains exploratory, given the scarce literature on this topic.

4.2. Methods

4.2.1. Participants

The same participants from the emotion recognition experiment also took part in the present experiment (see Section 3.2.1 for full details on participant recruitment, eligibility, demographics, and ethical considerations). This includes 76 native English speakers, with 38 participants with ASD and 38 NT participants (14 children aged 7-11 years, 11 adolescents aged 12-15 years, and 13 adults aged 16-56 in each group). The two groups were matched on chronological age, gender, receptive vocabulary (ROWPVT-4; Martin & Brownell, 2011), and nonverbal reasoning ability (RSPM; Raven, 1983). An optimal sample size was estimated through an a prior power analysis using G*Power 3.1 (Faul et al., 2007). It was determined that with a medium effect size ($f = 0.25$; Cohen, 1988), an alpha level of 0.05, and power of 0.80, a total sample size of 158 was required to detect the Diagnostic group (ASD vs. NT) × Age group (child vs. adolescent vs. adult) × Prime type (prosody vs. song) × Target type interaction (face vs. object). Yet, challenges with participant recruitment limited this study to be able to only test 76 participants mentioned above.

4.2.2. Design

This study adopted a mixed design, with between-subjects factors Diagnostic group (ASD vs. NT) and Age group (Child vs. Adolescent vs. Adult) and within-subjects factors Prime type (Prosody vs. Song), Target type (Face vs. Face-like object), and Prime emotion (Angry vs. Scared vs. Happy vs. Sad) as independent variables. The dependent variable was the strength of priming effects measured on the cross-modal affective priming task. This was indexed by a difference score between the congruent and incongruent conditions for accuracy, response time, and efficiency, respectively. That is, a larger difference score reflects facilitated emotional judgment of the target due to pre-activation of a congruent prime, and hence stronger priming effects.

4.2.3. Stimuli

The same set of auditory and visual stimuli from the emotion recognition experiment was used in the present experiment (see Section 3.2.3 for full details on stimulus development and validation). A total of 16 prosodic and 16 sung stimuli were chosen to be the auditory primes, with four prosodic and four sung clips representing each of the four emotional categories (angry, scared, happy, and sad). A total of 64 facial and 64 face-like object stimuli were chosen to be the visual targets, with 16 facial and 16 face-like object images representing each of the four emotional categories (angry, scared, happy, and sad).

4.2.4. Apparatus

The experiment was run with the same apparatus and the same key press response as those in the emotion recognition experiment (see Section 3.2.4 for full details). All participants completed both the emotion recognition and priming tasks as well as a battery of background tasks (will be further described in Chapter 5) in a random order, either during one single session or across multiple sessions. The emotion recognition tasks were always administered after the priming task to minimise habituation of the participants to associating images to emotion labels in the priming task that might eliminate the prime-target association to be examined.

4.2.5. Procedure

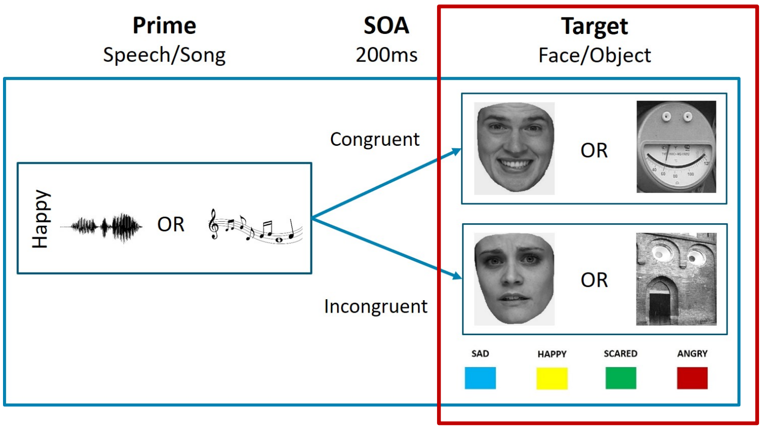


Figure 4. 1. The experimental procedure of the cross-modal affective priming task. On each trial, a prime stimulus (prosody or song) was presented 200ms prior to the presentation of the target stimulus (face or face-like object) which remained on the screen until a response was made to identify the emotion presented in the image.

The priming task assessed participants' emotion recognition in the visual targets after hearing an emotionally congruent or incongruent auditory prime. There was a total of 256 trials, with two blocks of 64 trials presenting a face target and two blocks of 64 trials presenting an object target. The order of the face and face-like object target conditions was counterbalanced between participants and each target condition was preceded by four practice trials. Each auditory prime was paired with a visual target from each of the four emotional categories to create congruent (e.g., angry-angry) and incongruent (e.g., angry-scared, angry-happy, and angry-sad) prime-target pairs. This resulted in 16 congruent and 48 incongruent trials for each of the prime-target pairs (prosody-face, song-face, prosody-object, song-object). Trials were pseudorandomised to ensure that stimuli presenting the same emotion occurred at most twice in a row, stimuli presented by one actor was always followed by stimuli presented by another actor, the same type of auditory input (prosody or song) occurred at most five times in a row, and the same gender of the actor presenting the stimulus occurred at most five times in a row. Two versions of pseudo-randomisation were adopted and counterbalanced between participants.

With a stimulus onset asynchrony (SOA) of 200ms, the prime (prosody or song) was presented 200ms prior to the target onset (human face or face-like object) on each trial. Thus, the prime and target stimuli overlapped for 300ms before the auditory prime offset. Participants were instructed to decide as quickly and as accurately as possible the emotion label (angry, scared, happy, or sad) that best described the expression presented in the image. The target stimulus remained on the screen until a response was made on the response pad by pressing the corresponding key of the chosen emotion. The presentation of the auditory and visual stimuli stopped as soon as the participant responded – see Figure 4.1 for a visual illustration of the task procedure. Responses were considered to be correct if the emotion selected by the participant corresponded to the intended emotion expressed in the visual targets. Accuracy and reaction time (RT) from the target onset were recorded for each trial.

4.2.6. Statistical analyses

All analyses were performed in R (RStudio Team, 2018). To quantify what constitutes a congruency effect due to stimulus priming, a congruency mean difference score was calculated in terms of mean accuracy, mean response time (RT), and mean SACS. As all participants scored above the chance level of 0.25 overall and for each prime-target condition, all accuracy data were retained for analyses. A congruency mean accuracy difference score by prime type and target type was computed by subtracting the mean accuracy for incongruent prime-target pairs from the mean accuracy for congruent prime-target pairs for each participant. Participants' RT on the priming task, measured from target onset, was based on correct responses only. RTs less than 150ms and more than 2.5 SD of the mean of each participant were excluded. The exclusion of incorrect responses (4740 observations; 24% of total observations) and RT outliers (470 observations; 3% of correct response observations), resulted in a dataset of 14245 observations across all participants (73% of the original number of observations). A congruency mean RT difference score by prime type and target type was computed by subtracting the mean RTs for congruent prime-target pairs from the mean RTs for incongruent prime-target pairs. The priming speed-accuracy composite score (SACS) was calculated by subtracting the normalised RTs from the normalised mean accuracy scores [Z(Accuracy) – Z(RT) = SACS] for each participant by prime type, target type, and congruency level. A high composite score reflects efficient performance, and a low composite score reflects poor performance. A congruency mean SACS difference score by prime type and target type was computed by subtracting the mean SACS for incongruent prime-target pairs from the mean SACS for congruent prime-target pairs for each participant.

First, to examine differences in priming effects, three separate linear mixed effects models were constructed using the *lme4* package (Bates et al., 2015), with congruency mean difference scores for accuracy, RT, and SACS as dependent measures, respectively. Each model included Diagnostic group (ASD vs. NT), Age group (child vs. adolescent vs. adult), Prime type (prosody vs. song), Target type (human face vs. face-like object), and all possible interactions as fixed effects. Secondly, in addressing the question of whether priming effects were modulated by specific emotion, separate linear mixed effects models for each prime-target condition (song-face, prosody-face, song-object, and prosody-object) were conducted on the congruency mean difference scores for accuracy, RT, and SACS, respectively. Each model included Diagnostic group (ASD vs. NT), Age group (child vs. adolescent vs. adult), Prime emotion (angry vs. scared vs. happy vs. sad), and all possible interactions as fixed factors. The maximal random effects structure was initially specified with by-subject intercept and by-subject slopes for prime type and target type for all the linear mixed effects models. Due to convergence issues, the number of iterations was increased and derivation calculations were turned off for the congruency mean RT difference model as recommended by Brown (2020). Model specifications are listed in the corresponding results summary tables.

For all the linear mixed effects analyses, the statistical significance of the fixed effects were obtained using the *anova*() function from the *lmerTest* package (Kuznetsova et al., 2017). For effect sizes, partial eta-squared (η^2_p) was computed for each fixed effect using the *eta_squared*() function in the *effectsize* package (Ben-Shachar et al., 2020). A $\eta^2_p \geq 0.01$ was interpreted as a small effect size, a $\eta^2_p \geq 0.09$ as a medium effect size, and a $\eta^2_p \geq 0.25$ as a large effect size (Cohen et al., 2013). Significant effects and interactions emerging from the models were followed up through post-hoc pairwise comparisons using the *pairwise_t_test* function in the *rstatix* package (Kassambara, 2020). Correction of the post-hoc tests for multiple comparisons was performed with Benjamini-Hochberg (false discovery rate) procedure (Benjamini & Hochberg, 1995).

4.3. Results

4.3.1. Are differences in priming effects between diagnostic groups and age groups modulated by prime type and target type?

4.3.1.1. Congruency mean accuracy difference

Table 4.1 displays the full model results summary of the linear mixed effects analysis on the congruency mean accuracy difference score. The analysis revealed a significant main effect of Prime type ($F(1, 70) = 9.09$, $p = 0.004$, $\eta^2_p = 0.11$). Importantly, the four-way interaction of Diagnostic group × Age group × Prime type × Target type was significant ($F(2, 70) = 5.46$, $p = 0.006$, $\eta^2_p = 0.13$), which I will unpack below. No other effects and interactions were found to be significant.

Table 4. 1. Linear mixed effects model results for diagnostic group, age group, prime type, target type, and their interactions on congruency mean accuracy difference score.

Fixed effects	**df**		***F***	***p***	**η^2_p**
Diagnostic group	1	70	1.26	0.266	0.02
Age group	2	70	1.83	0.169	0.05

Prime type	**1**	**70**	**9.09**	**0.004**	**0.11**
Target type	1	70	3.43	0.068	0.05
Diagnostic group × age group	2	70	0.38	0.688	0.01
Diagnostic group × prime type	1	70	1.07	0.304	0.02
Age group × prime type	2	70	0.38	0.684	0.01
Diagnostic group × target type	1	70	0.19	0.667	0.00
Age group × target type	2	70	2.06	0.135	0.06
Prime type × target type	1	70	1.85	0.178	0.03
Diagnostic group × age group × prime type	2	70	0.43	0.651	0.01
Diagnostic group × age group × target type	2	70	0.11	0.897	0.00
Diagnostic group × prime type × target type	1	70	2.58	0.112	0.04
Age group × prime type × target type	2	70	0.87	0.424	0.02
Diagnostic group × age group × prime type × target type	**2**	**70**	**5.46**	**0.006**	**0.13**

Note. R model equation: lmer (Congruency Mean Accuracy Difference ~ Diagnostic Group * Age Group * Prime Type * Target Type + (1 + Prime Type + Target Type | Subject)); significant effects and interactions are highlighted using bold font.

Given that one of the main research questions was to investigate whether implicit emotion priming between autistic and NT individuals across age groups is modulated by domain (i.e., prime type), the four-way interaction of Diagnostic group × Age group × Prime type × Target type was examined using separate linear effects models by Prime Type. Each model, therefore, included Diagnostic group, Age group, Target type, and their possible interactions as fixed effects, and with by-subject intercept as random effects. The results are summarised in Table 4.2.

Table 4. 2. Linear mixed effects model results for diagnostic group, age group, prime type, target type, and their interactions on congruency mean accuracy difference score.

	Fixed effects	**df**		***F***	***p***	**η^2_p**
Prosody						
	Diagnostic group	1	70	1.79	0.185	0.02
	Age group	2	70	1.57	0.215	0.04
	Target type	1	70	1.04	0.311	0.01
	Diagnostic group × age group	2	70	0.40	0.670	0.01
	Diagnostic group × target type	1	70	0.07	0.790	0.00
	Age group × target type	2	70	1.41	0.250	0.04
	Diagnostic group × age group × target type	2	70	1.42	0.248	0.04
Song						
	Diagnostic group	1	70	0.54	0.466	0.00
	Age group	2	70	1.75	0.181	0.05
	Target type	**1**	**70**	**6.08**	**0.016**	**0.08**
	Diagnostic group × age group	2	70	0.36	0.701	0.01
	Diagnostic group × target type	1	70	1.37	0.245	0.02
	Age group × target type	2	70	2.46	0.094	0.07
	Diagnostic group × age group × target type	2	70	0.57	0.567	0.02

Note. R model equation: lmer (Congruency Mean Accuracy Difference ~ Diagnostic Group * Age Group * Target Type + (1 | Subject)); significant effects and interactions are highlighted using bold font.

It was found that when primed by prosody, neither the effects of Diagnostic group, Age group, Target type, nor the interactions between them were significant. This suggests that regardless of target type, visual emotion recognition accuracy was primed by prosodic emotions to similar extents in the ASD and NT groups across age groups.

When primed by song, the effect of Target type was significant ($F(1, 70) = 6.08$, $p = 0.016$, $\eta^2_p = 0.08$), with larger congruency mean accuracy difference found for face targets (M = 0.11, SD = 0.15) than object targets (M = 0.06, SD = 0.16). This suggests that across diagnostic groups and age groups, sung emotions had primed the accuracy of emotion recognition in face targets more strongly than that in object targets.

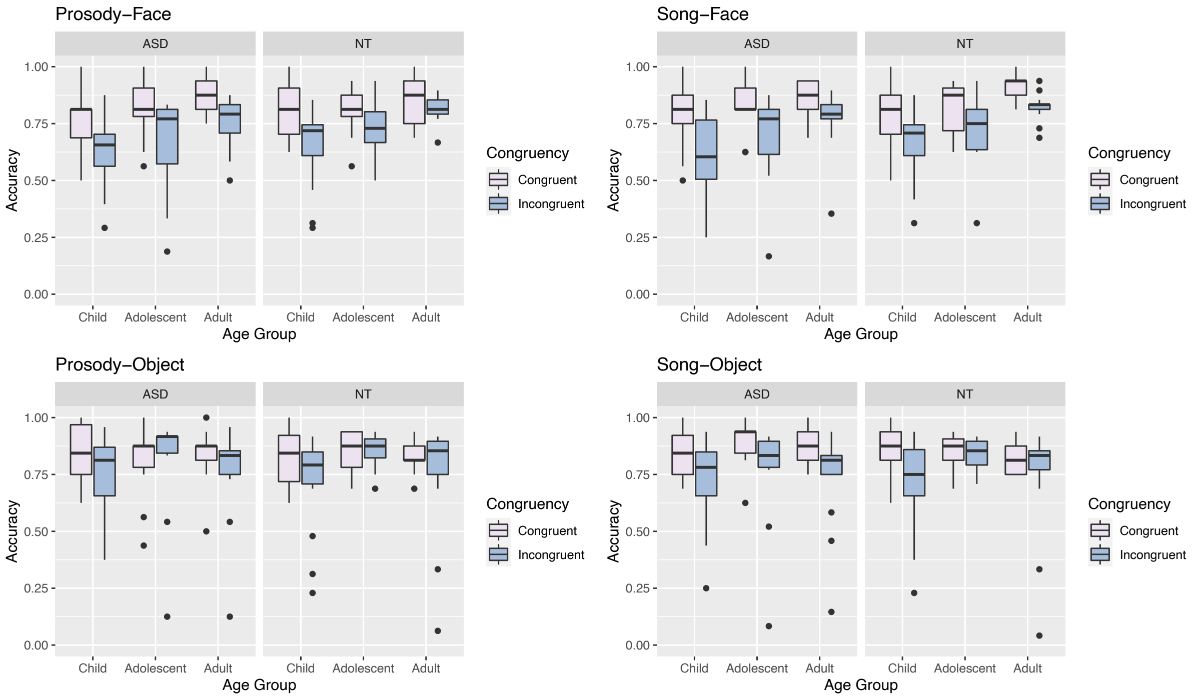


Figure 4. 2. Boxplots of mean accuracy across congruency levels by prime-target condition for each diagnostic by age group.

Further to this, interesting trends in regard to the song-face and prosody-face conditions can be seen in Figure 4.2. There appears to be a more prominent age-related decline in the priming effect of prosody on face targets in the ASD group compared to the NT group. Conversely, there appears to be a more prominent age-related decline in the priming effect of song on face targets in the NT group compared to the ASD group. Age-related differences in the priming effect of prosody and song on object targets were similar across the ASD and NT groups. To better illustrate this emotion priming pattern characterising each group within this four-way interaction, the congruency mean accuracy difference for each diagnostic by age group were compared to the test value of 0 (i.e., no priming) by prime-target condition through a series of one-sample t-tests. Results from these tests would inform whether priming was present across prime-target conditions for each diagnostic and age group.

Table 4. 3. Results summary of one sample t-tests on priming congruency mean accuracy difference against the test value of 0 (i.e., no priming) by diagnostic group, age group and prime-target condition. All p-values reported were adjusted using the Benjamini-Hochberg method.

		Prosody-Face			Song-Face			Prosody-Object			Song-Object		
		M (SD)	*t*	*p*	M (SD)	*t*	*p*	M (SD)	*t*	*p*	M (SD)	*t*	*p*
Child	ASD	**0.22 (0.20)**	**4.16**	**0.003**	**0.13 (0.18)**	**2.82**	**0.024**	**0.12 (0.15)**	**3.01**	**0.030**	0.10 (0.15)	2.56	0.117
	Control	**0.14 (0.17)**	**3.14**	**0.016**	**0.15 (0.17)**	**3.23**	**0.020**	**0.15 (0.15)**	**3.56**	**0.021**	0.11 (0.17)	2.30	0.117
Adolescent	ASD	**0.15 (0.20)**	**2.47**	**0.039**	**0.16 (0.18)**	**2.89**	**0.024**	0.13 (0.21)	2.08	0.096	0.00 (0.12)	0.10	0.921
	Control	**0.11 (0.15)**	**2.48**	**0.039**	0.07 (0.12)	1.87	0.092	0.00 (0.09)	0.13	0.896	0.00 (0.06)	-0.11	0.921
Adult	ASD	0.08 (0.14)	1.99	0.070	**0.11 (0.10)**	**3.97**	**0.011**	0.13 (0.21)	2.28	0.084	0.07 (0.14)	1.82	0.187
	Control	**0.08 (0.06)**	**5.33**	**0.001**	0.05 (0.09)	1.88	0.092	0.09 (0.26)	1.25	0.284	0.08 (0.25)	1.19	0.388

Note. Significant effects are highlighted using bold font.

Table 4.3 displays the full results summary of the one-sample t-tests on congruency mean accuracy difference scores against the test value of 0 for each diagnostic by age group.

One sample t-tests revealed that for the prosody-face condition, the difference score significantly differed from 0 in children ($t(13) = 4.16$, $p = 0.003$) and adolescents ($t(10) = 2.47$, $p = 0.033$) but not in adults within the ASD group. By contrast, the difference score significantly differed across all age groups within the NT group (child: $t(13) = 3.14$, $p = 0.016$; adolescent: $t(10) = 2.48$, $p = 0.039$; adult: $t(12) = 5.33$, $p = 0.001$).

For the song-face condition, the congruency mean accuracy difference score significantly differed from 0 across all age groups within the ASD group (child: $t(13) = 2.82$, $p = 0.024$; adolescent: $t(10) = 2.89$, $p = 0.024$; adult: $t(12) = 3.97$, $p = 0.011$). Conversely, the difference score significantly differed from 0 only in children ($t(13) = 3.23$, $p = 0.020$) but not in adolescents and adults within the NT group.

For the prosody-object condition, the difference score differed significantly from 0 in both children with ASD ($t(13) = 3.01$, $p = 0.030$) and NT children ($t(13) = 3.56$, $p = 0.021$), but not in adolescents and adults across diagnostic groups.

For the song-object condition, the difference score did not differ significantly from 0 across diagnostic groups and age groups.

Taken together, these results indicate that prosodic emotions primed the accuracy of emotion recognition in face targets across all age groups in the NT group but only for children and adolescents in the ASD group. The reverse was seen for sung emotions where they primed the accuracy of emotion recognition in face targets across all age groups in the ASD group but only for children in the NT group. While prosodic emotions primed recognition accuracy of object targets for children but not adolescents and adults across diagnostic groups, sung emotions did not prime recognition accuracy of object targets across diagnostic and age groups.

4.3.1.2. Congruency mean response time difference

Table 4.4 displays the full model results summary of the linear mixed effects analysis on the congruency mean RT difference score. The analysis revealed a significant two-way interaction of Prime type × Target type ($F(1, 140) = 11.59, p < 0.001, \eta^2_p = 0.08$), with no other effects and interactions reaching significance. I will unpack this two-way interaction below.

Table 4. 4. Linear mixed effects model results for diagnostic group, age group, prime type, target type, and their interactions on congruency mean RT difference score.

Fixed effects	**df**		***F***	***p***	**η^2_p**
Diagnostic group	1	70.07	0.62	0.433	0.00
Age group	2	70.07	1.10	0.338	0.03
Prime type	1	78.31	0.86	0.357	0.01
Target type	1	106.28	0.91	0.342	0.00
Diagnostic group × age group	2	70.07	1.09	0.342	0.03
Diagnostic group × prime type	1	78.31	0.01	0.907	0.00

Age group × prime type	2	78.31	1.18	0.314	0.03
Diagnostic group × target type	1	106.28	0.05	0.827	0.00
Age group × target type	2	106.28	1.95	0.148	0.04
Prime type × target type	**1**	**140**	**11.59**	**< 0.001**	**0.08**
Diagnostic group × age group × prime type	2	78.31	1.56	0.217	0.04
Diagnostic group × age group × target type	2	106.28	0.47	0.624	0.00
Diagnostic group × prime type × target type	1	140	0.89	0.347	0.00
Age group × prime type × target type	2	140	1.18	0.311	0.02
Diagnostic group × age group × prime type × target type	2	140	2.76	0.067	0.04

Note. R model equation: lmer (Congruency Mean RT Difference ~ Diagnostic Group * Age Group * Prime Type * Target Type + (1 + Prime Type + Target Type | Subject), control = lmerControl(optCtr = list(maxfun = 1e9), calc.derivs = FALSE)); significant effects and interactions are highlighted using bold font.

Figure 4.3 shows the interaction of Prime type × Target type. Post-hoc analyses revealed a larger congruency mean RT difference when object targets were primed by song (M = 316.09, SD = 481.03) than when they were primed by prosody (M = 121.94, SD = 410.86) (*t*(75) = 2.85, *p* = 0.006). For face targets, there was no significant difference in the congruency mean RT difference between prosodic and sung primes.

In addition, the congruency mean RT difference was larger when object targets were primed by song (M = 316.09, SD = 481.03) than when face targets were primed by song (M = 126.70, SD = 358.98) (*t*(75) = 2.98, *p* = 0.004). There was no significant difference in the congruency mean RT difference between face targets and object targets when primed by prosody.

These results indicate that independent of diagnostic group and age group, the speed of emotion recognition in object targets was particularly primed by sung emotions. By contrast, priming effects on the speed of emotion recognition in face targets were not modulated by prime type across diagnostic groups and age groups.

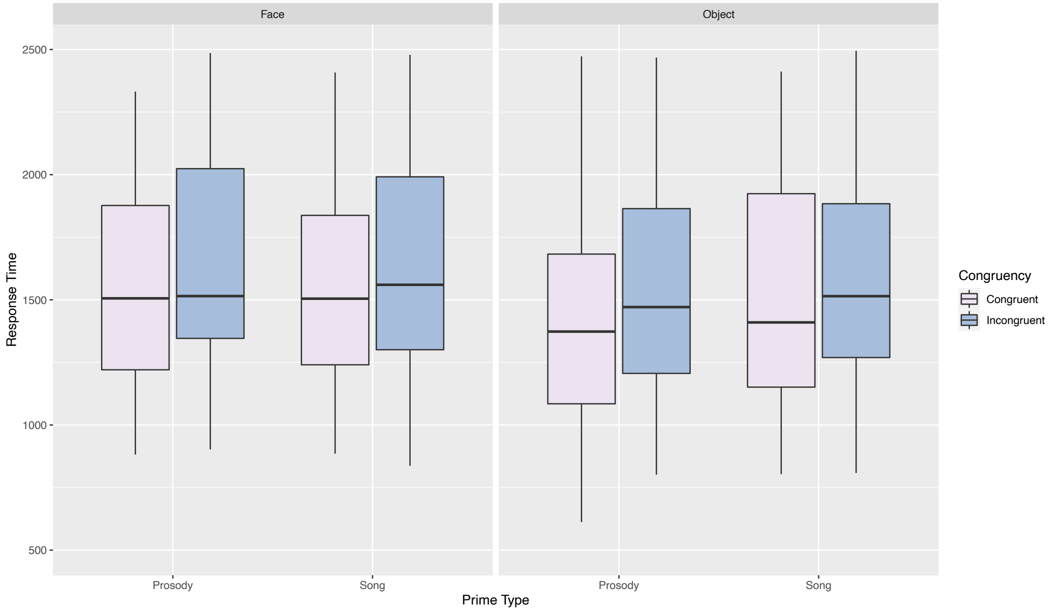


Figure 4. 3. Boxplots of mean RT across congruency levels for each prime type in the face and object conditions, respectively.

4.3.1.3. Congruency mean speed-accuracy composite score difference

Table 4.5 displays the full model results summary of the linear mixed effects analysis on the congruency mean SACS difference score. The linear mixed effects analysis revealed a two-way interaction of Age group × Target type ($F(2, 70) = 3.19$, $p = 0.047$, $\eta^2_{p} = 0.08$), which I will unpack below. No other effects and interactions reached significance.

Table 4. 5. Linear mixed effects model results for diagnostic group, age group, prime type, target type, and their interactions on congruency mean SACS difference score.

Fixed effects	**df**		***F***	***p***	**η^2_p**
Diagnostic group	1	70	1.48	0.228	0.02
Age group	2	70	1.76	0.180	0.05
Prime type	1	70	1.81	0.183	0.03
Target type	1	70	1.42	0.238	0.02
Diagnostic group × age group	2	70	0.66	0.520	0.02
Diagnostic group × prime type	1	70	0.35	0.558	0.00
Age group × prime type	2	70	1.11	0.336	0.03
Diagnostic group × target type	1	70	0.21	0.649	0.00
Age group × target type	**2**	**70**	**3.19**	**0.047**	**0.08**
Prime type × target type	1	70	2.40	0.126	0.03
Diagnostic group × age group × prime type	2	70	0.57	0.566	0.02
Diagnostic group × age group × target type	2	70	0.02	0.976	0.00
Diagnostic group × prime type × target type	1	70	0.30	0.587	0.00
Age group × prime type × target type	2	70	1.76	0.179	0.05
Diagnostic group × age group × prime type × target type	2	70	1.72	0.187	0.05

Note. R model equation: lmer (Congruency Mean SACS Difference ~ Diagnostic Group * Age Group * Prime Type * Target Type + (1 + Prime Type + Target Type | Subject)); significant effects and interactions are highlighted using bold font.

Figure 4.4 shows the interaction of Age group × Target type. Post-hoc analyses revealed that when responding to face targets, the congruency mean SACS difference was larger in children (M = 1.29, SD = 1.08) than in adults (M = 0.55, SD = 0.73) ($t(47.61) = 2.96$, $p = 0.014$), with adolescents showing no difference to either group. By contrast, when responding to object targets, no difference in the congruency mean SACS difference was observed across age groups. These results indicate that regardless of diagnostic group and prime type, priming effects on the efficiency of emotion recognition in face targets were stronger for children than adults, while there were no age-related differences in that in object targets.

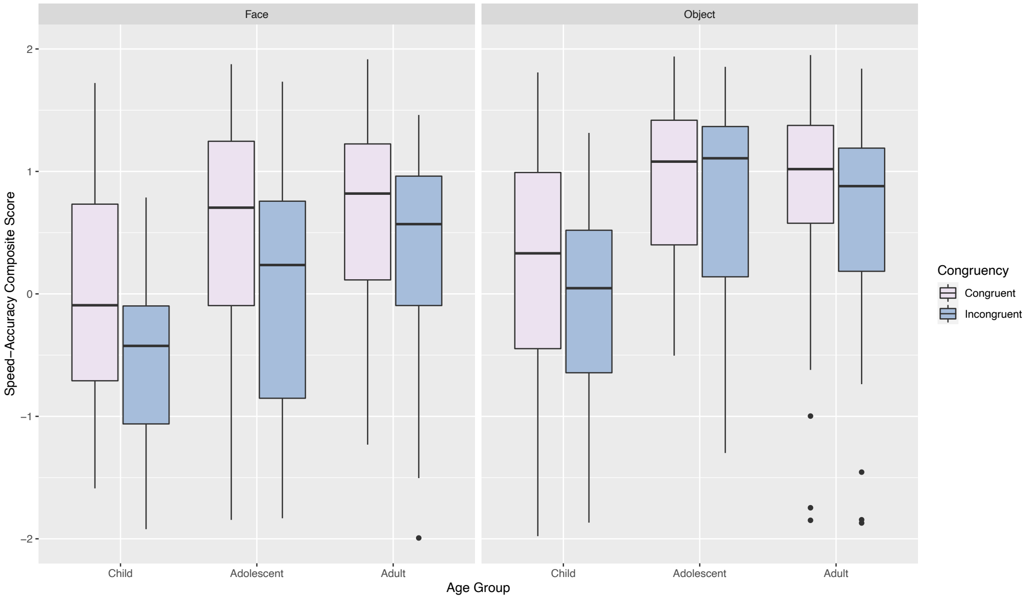


Figure 4. 4. Boxplots of mean SACS across congruency levels for different target types in each age group.

4.3.2. Are differences in priming effects between diagnostic groups and age groups across prime-target conditions modulated by prime emotion?

4.3.2.1. Congruency mean accuracy difference

Table 4.6 displays the full model results summary of the linear mixed effects analysis on the congruency mean accuracy difference score for each prime-target condition.

Table 4. 6. Linear mixed effects model results displayed by prime-target condition for diagnostic group, age group, prime emotion, and their interactions on congruency mean accuracy difference score.

	Fixed effects	**df**		***F***	***p***	**η^2_p**
Prosody-Face						
	Diagnostic group	1	70	1.01	0.317	0.01
	Age group	2	70	2.60	0.082	0.07
	Prime emotion	**3**	**210**	**2.65**	**0.050**	**0.04**
	Diagnostic group × age group	2	70	0.42	0.661	0.01
	Diagnostic group × prime emotion	**3**	**210**	**4.06**	**0.008**	**0.05**
	Age group × prime emotion	**6**	**210**	**3.02**	**0.007**	**0.08**
	Diagnostic group × age group × prime emotion	6	210	1.64	0.138	0.04
Song-Face						
	Diagnostic group	1	70	1.81	0.183	0.03
	Age group	2	70	1.26	0.291	0.03
	Prime emotion	**3**	**210**	**11.41**	**< 0.001**	**0.14**
	Diagnostic group × age group	2	70	0.94	0.396	0.03
	Diagnostic group × prime emotion	3	210	0.81	0.491	0.01
	Age group × prime emotion	6	210	1.75	0.111	0.05
	Diagnostic group × age group × prime emotion	6	210	0.41	0.870	0.01
Prosody-Object						
	Diagnostic group	1	70	1.29	0.259	0.02
	Age group	2	70	0.75	0.474	0.02
	Prime emotion	**3**	**210**	**11.99**	**< 0.001**	**0.15**
	Diagnostic group × age group	2	70	1.01	0.370	0.03
	Diagnostic group × prime emotion	3	210	0.06	0.979	0.00
	Age group × prime emotion	6	210	1.98	0.069	0.05
	Diagnostic group × age group × prime emotion	6	210	0.37	0.898	0.01
Song-Object						
	Diagnostic group	1	70	0.01	0.941	0.00
	Age group	2	70	2.58	0.083	0.07
	Prime emotion	**3**	**210**	**17.59**	**< 0.001**	**0.20**
	Diagnostic group × age group	2	70	0.01	0.987	0.00
	Diagnostic group × prime emotion	3	210	0.27	0.850	0.00
	Age group × prime emotion	6	210	1.38	0.223	0.04
	Diagnostic group × age group × prime emotion	6	210	1.12	0.355	0.03

Note. R model equation for all prime-target conditions: lmer (Congruency Mean Accuracy Difference ~ Diagnostic Group * Age Group * Prime Emotion + (1 | Subject)); significant effects and interactions are highlighted using bold font.

Prosody-face

The linear mixed effects model revealed a significant main effect of Prime emotion ($F(3, 210) = 2.65$, $p = 0.050$, $\eta^2_p = 0.04$), as well as two-way interactions of Diagnostic group × Prime emotion ($F(3, 210) = 4.06$, $p = 0.008$, $\eta^2_p = 0.05$) and Age group × Prime emotion ($F(6, 210) = 3.02$, $p = 0.007$, $\eta^2_p = 0.08$). I will unpack these two two-way interactions below.

Figure 4.5 shows the interaction of Diagnostic group × Prime emotion. Post-hoc analyses showed that when faces were primed by scared prosody, the congruency mean accuracy difference was significantly larger in the ASD group (M = 0.18, SD = 0.27) compared to the NT group (M = 0.00, SD = 0.27) (*t*(74) = 2.84, *p* = 0.006). The two groups, however, did not differ in congruency mean accuracy difference when faces were primed by angry, happy, nor sad prosody.

Within the ASD group, a larger mean accuracy difference was found when faces were primed by happy prosody (M = 0.20, SD = 0.24) than by sad prosody (M = 0.07, SD = 0.23) (*t*(37) = 3.17, *p* = 0.018), while no other differences across emotions of prosodic primes were observed when responding to face targets.

Within the NT group, a larger mean accuracy difference was found when faces were primed by happy prosody (M = 0.18, SD = 0.20) than by scared prosody (M = 0.00, SD = 0.27) (*t*(37) = 3.11, *p* = 0.022), with no other differences across emotions of prosodic primes being observed when responding to face targets.

Together, these results indicate that the ASD group showed stronger priming effects than the NT group when faces were primed by prosody conveying fear. Within each group, different emotions expressed in prosody also primed facial emotion recognition in different ways. That is, whereas face targets were less strongly primed by prosody conveying sadness than happiness in the ASD group, face targets were less strongly primed by prosody conveying fear than happiness in the NT group.

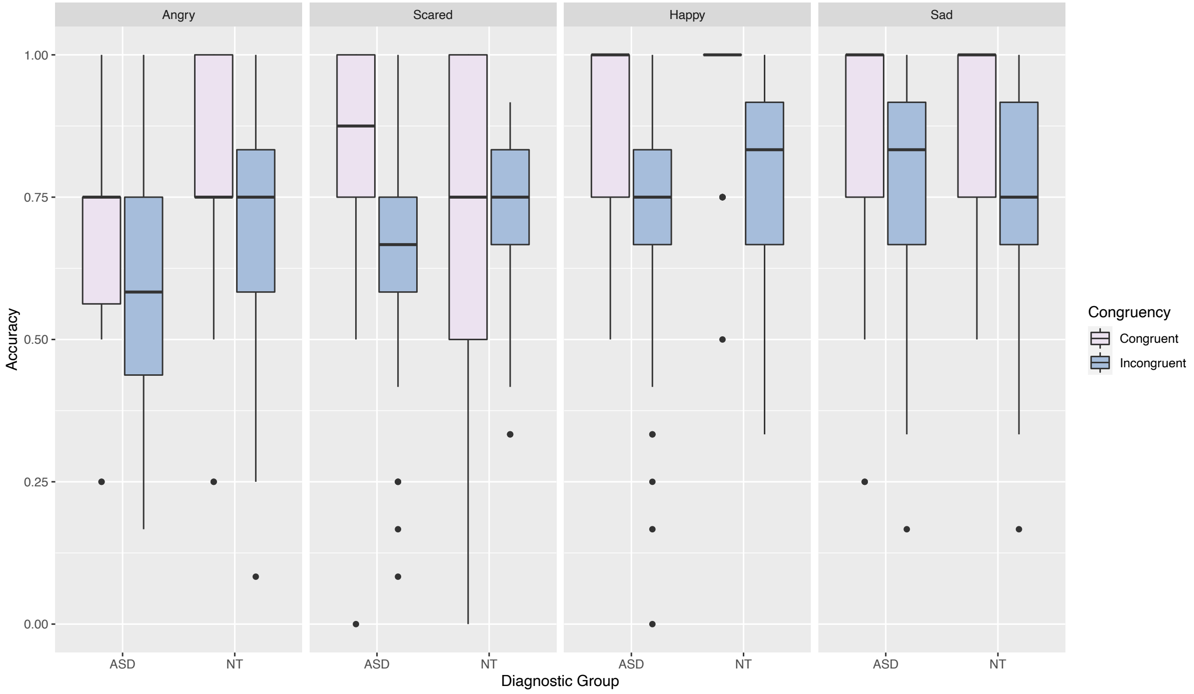


Figure 4. 5. Boxplots of mean accuracy between the ASD and NT groups across congruency levels by prime emotion in the prosody-face condition.

Figure 4.6 shows the interaction of Age group × Prime emotion. Post-hoc analyses showed a larger congruency mean accuracy difference in children than adults when faces were primed by angry prosody (child: M = 0.27, SD = 0.36; adult: M = 0.00, SD = 0.27; $t(49.96) = 3.23$, $p = 0.007$) and sad prosody (child: M = 0.17, SD = 0.20; adult: M = 0.04, SD = 0.16; $t(51.13) = 2.65$, $p = 0.032$), with no difference between children and adults found when faces were primed by happy and scared prosody. Adolescents showed no difference to children and adults across emotions of prosodic primes when responding to face targets.

The priming pattern across prosodic emotions on faces in children also differed from that in adolescents and adults. In children, there was a larger congruency mean accuracy difference when faces were primed by happy (M = 0.23, SD = 0.24) and angry prosody (M = 0.27, SD = 0.36) than by scared prosody (M = 0.04, SD = 0.32) (angry vs. scared: $t(27) = 2.80$,

$p = 0.028$; happy vs. scared: $t(27) = 3.03$, $p = 0.028$), with no other differences across emotions of prosodic primes found when responding to face targets.

Conversely, both adolescents and adults showed no difference in congruency mean accuracy difference across emotions of prosodic primes when responding to face targets.

These results indicate that children showed stronger priming effects on the accuracy of emotion recognition in face targets than adults particularly when prosodic primes conveyed anger and sadness. This partially agrees with that in children, the accuracy of emotion recognition in face targets was more strongly primed by prosody that conveyed anger and happiness compared to fear, whereas no discrepancies among prime emotions were seen in adolescents and adults.

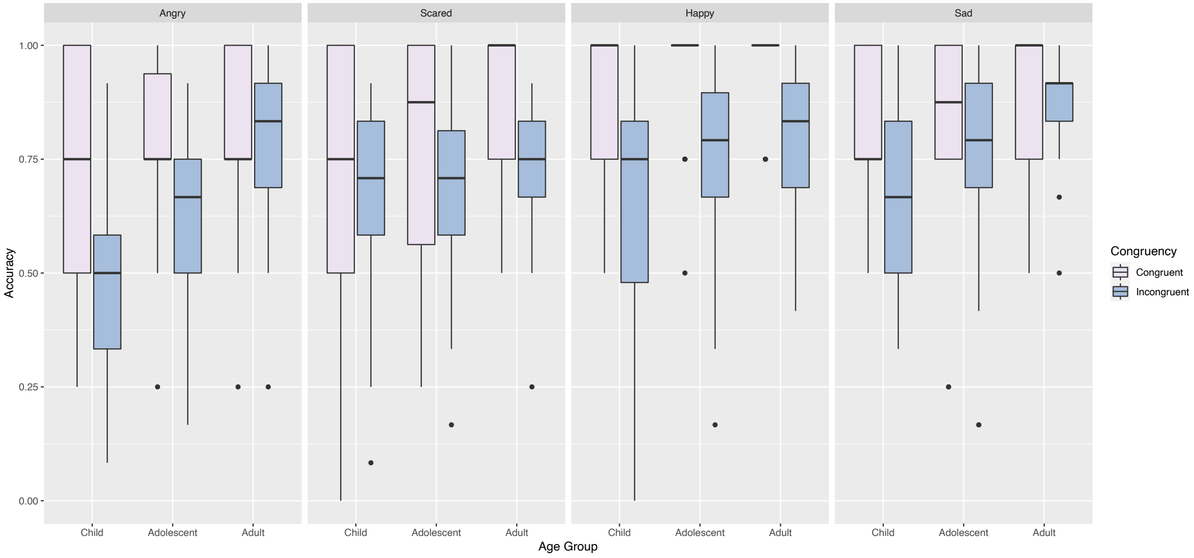


Figure 4. 6. Boxplots of mean accuracy between children, adolescents, and adults across congruency levels by prime emotion in the prosody-face condition.

Song-face

The linear mixed effects model revealed a significant main effect of Prime emotion ($F(3, 210) = 11.41$, $p < 0.001$, $\eta^2_p = 0.14$). Post-hoc analyses showed that the congruency mean accuracy difference was largest when faces were primed by happy song (M = 0.25, SD = 0.23), which significantly differed from that by angry (M = 0.08, SD = 0.29), scared (M = 0.03, SD = 0.27), and sad song (M = 0.09, SD = 0.24) (happy vs. angry: t(75) = 4.12, p < 0.001; happy vs. scared: t(75) = 6.08, $p < 0.001$; happy vs. sad: t(75) = 4.95, $p < 0.001$) (Figure 4.7). These results indicate that across diagnostic groups and age groups, the accuracy of emotion recognition in face targets was most strongly primed by song that conveyed happiness among other emotions.

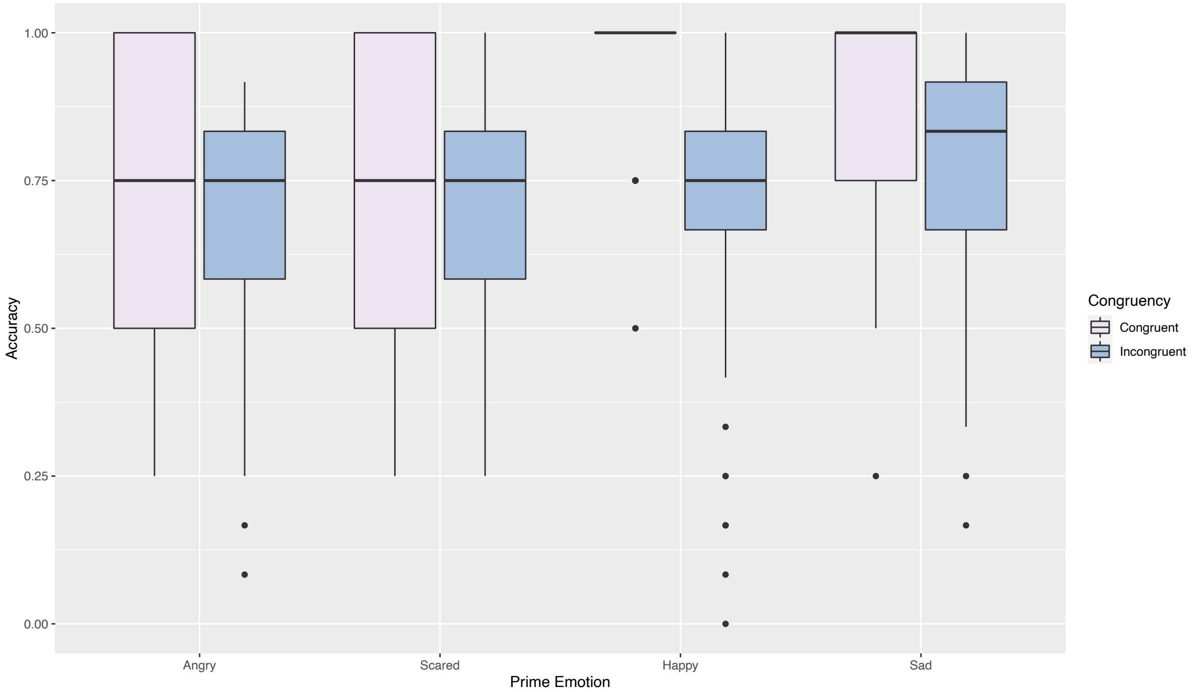


Figure 4. 7. Boxplots of mean accuracy across congruency levels by prime emotion in the song-face condition.

Prosody-object

The linear mixed effects model revealed a significant main effect of Prime emotion ($F(3, 210) = 11.99$, $p < 0.001$, $\eta^2_p = 0.15$). Post-hoc analyses showed that the congruency mean accuracy difference was smallest when objects were primed by angry prosody (M = -0.02, SD = 0.38), which differed significantly from that by scared (M = 0.10, SD = 0.26), happy (M = 0.20, SD = 0.24), and sad prosody (M = 0.15, SD = 0.25) (angry vs. scared: $t(75) = -2.63$, $p = 0.016$; angry vs. happy: $t(75) = -5.76$, $p < 0.001$; angry vs. sad: $t(75) = -4.03$, $p < 0.001$). The congruency mean accuracy difference was also smaller when objects were primed by scared prosody than by happy prosody ($t(75) = -2.78$, $p = 0.014$) (Figure 4.8). These results indicate that across diagnostic groups and age groups, the accuracy of emotion recognition in object targets was more strongly primed by prosody that conveyed happiness but was least primed by prosody that conveyed anger.

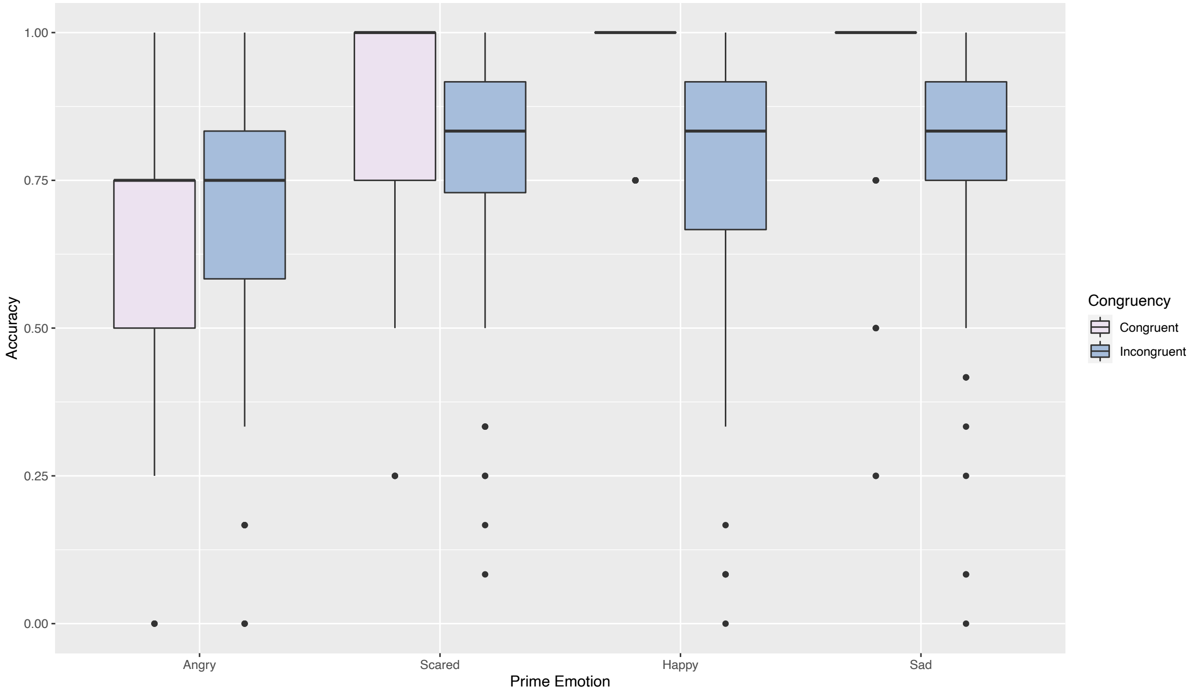


Figure 4. 8. Boxplots of mean accuracy across congruency levels by prime emotion in the prosody-object condition.

Song-object

The linear mixed effects model revealed a significant main effect of Prime emotion ($F(3, 210) = 17.59$, $p < 0.001$, $\eta^2_p = 0.20$). Post-hoc analyses showed that the congruency mean accuracy difference was largest when objects were primed by happy song (M = 0.20, SD = 0.22), followed by sad song (M = 0.10, SD = 0.20), then scared song (M = 0.03, SD = 0.26), with angry song (M = -0.08, SD = 0.36) showing the smallest congruency mean accuracy difference (happy vs. sad: $t(75) = 3.51$, $p = 0.003$; happy vs. scared: $t(75) = 4.39$, $p < 0.001$; happy vs. angry: $t(75) = 7.19$, $p < 0.001$; sad vs. scared: $t(75) = 1.99$, $p = 0.050$; sad vs. angry: $t(75) = 3.98$, $p < 0.001$; scared vs. angry: $t(75) = 2.18$, $p = 0.039$) (Figure 4.9). These results indicate that across diagnostic groups and age groups, the accuracy of emotion recognition in object targets was most strongly primed by song that conveyed happiness, followed by sadness, then fear, with objects being strongly primed by song that conveyed anger the least.

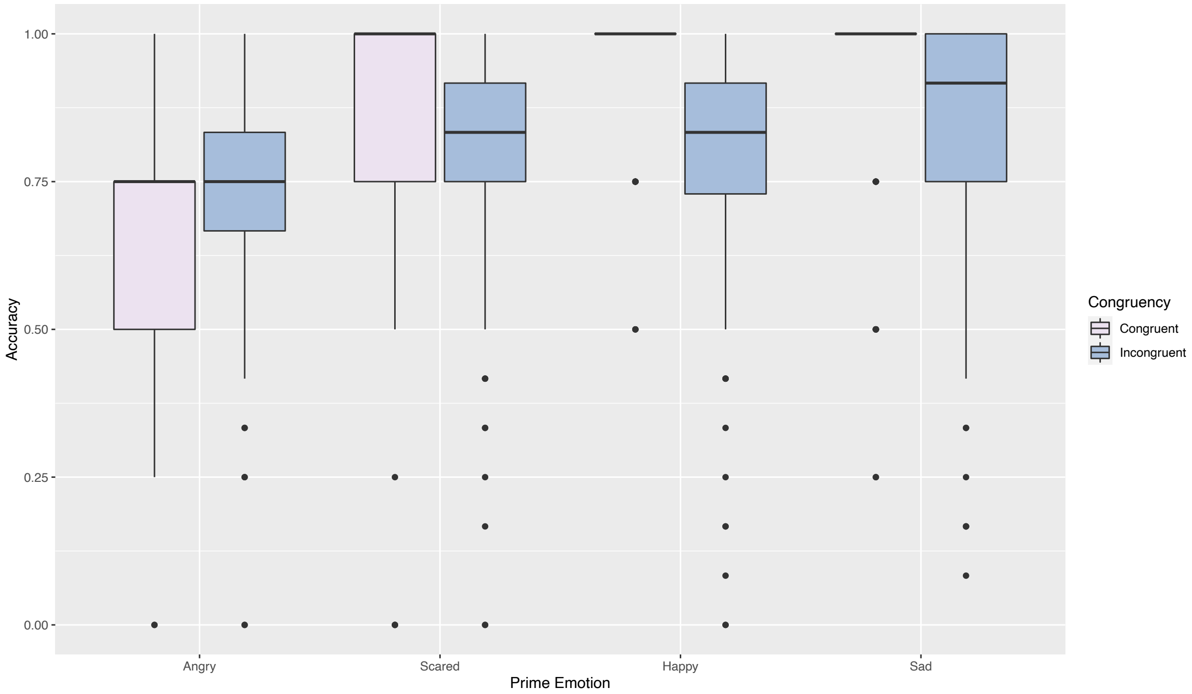


Figure 4. 9. Boxplots of mean accuracy across congruency levels by prime emotion in the song-object condition.

Summary

In general, prime emotion appeared to have modulated the magnitude of cross-modal priming effects in all four prime-target conditions to similar extents across diagnostic groups and age groups. There were, nonetheless, two exceptions: (i) the ASD group showed stronger priming than the NT group when faces were primed by prosody that conveyed fear, and (ii) children showed stronger priming than adults when faces were primed by prosody that conveyed anger and sadness. Overall, visual emotion recognition accuracy was most strongly primed by auditory primes, particularly song, that conveyed happiness. The accuracy of emotion recognition in object targets was relatively less strongly primed by auditory primes that conveyed anger.

4.3.2.2. Congruency mean response time difference

Table 4.7 displays the full model results summary of the linear mixed effects analysis on the congruency mean RT difference score for each prime-target condition.

Table 4. 7. Linear mixed effects model results displayed by prime-target condition for diagnostic group, age group, prime emotion, and their interactions on congruency mean RT difference score.

	Fixed effects	**df**		***F***	***p***	**η^2_p**
Prosody-Face						
	Diagnostic group	1	70.16	0.81	0.372	0.01
	Age group	2	70.20	1.47	0.236	0.04
	Prime emotion	**3**	**207.34**	**7.53**	**< 0.001**	**0.10**
	Diagnostic group × age group	2	70.20	1.66	0.198	0.05
	Diagnostic group × prime emotion	3	207.34	0.80	0.495	0.01
	Age group × prime emotion	6	207.36	0.89	0.507	0.02
	Diagnostic group × age group × prime emotion	6	207.36	0.45	0.843	0.01
Song-Face						
	Diagnostic group	1	279	0.36	0.548	0.00
	Age group	2	279	1.54	0.216	0.01
	Prime emotion	**3**	**279**	**3.24**	**0.022**	**0.03**
	Diagnostic group × age group	2	279	1.64	0.196	0.01
	Diagnostic group × prime emotion	3	279	0.67	0.570	0.01
	Age group × prime emotion	6	279	1.48	0.183	0.03
	Diagnostic group × age group × prime emotion	6	279	0.33	0.921	0.01
Prosody-Object						
	Diagnostic group	1	68.27	0.26	0.609	0.00
	Age group	2	68.24	1.25	0.294	0.04
	Prime emotion	3	203.42	2.30	0.078	0.03
	Diagnostic group × age group	2	68.24	0.54	0.585	0.02
	Diagnostic group × prime emotion	3	203.42	0.39	0.757	0.01
	Age group × prime emotion	6	203.37	1.09	0.367	0.03
	Diagnostic group × age group × prime emotion	6	203.37	0.86	0.528	0.02
Song-Object						
	Diagnostic group	1	69.46	0.28	0.598	0.00
	Age group	2	69.50	1.00	0.374	0.03
	Prime emotion	**3**	**202.01**	**12.74**	**< 0.001**	**0.16**
	Diagnostic group × age group	2	69.50	1.05	0.357	0.03
	Diagnostic group × prime emotion	3	202.01	0.18	0.911	0.00
	Age group × prime emotion	6	202.03	0.69	0.659	0.02
	Diagnostic group × age group × prime emotion	6	202.03	0.56	0.761	0.02

Note. R model equation for Song-Face: lmer (Congruency Mean RT Difference ~ Diagnostic Group * Age Group * Prime Emotion + (1 | Subject), control = lmerControl(calc.derivs = FALSE)); R model equation for Speech-Face, Song-Object, and Speech-Object: lmer (Congruency Mean RT Difference ~ Diagnostic Group * Age Group * Prime Emotion + (1 | Subject)); significant effects and interactions are highlighted using bold font.

Prosody-face

The linear mixed effects model revealed a significant main effect of Prime emotion ($F(3, 207.34) = 7.53$, $p < 0.001$, $\eta^2_p = 0.10$). Post-hoc analyses revealed that the congruency mean RT difference was largest when faces were primed by happy prosody (M = 508.83, SD = 763.60), which differed significantly from that by angry (M = 52.12, SD = 796.44), scared (M = 16.12, SD = 804.96), and sad prosody (M = 211.29, SD = 741.57) (happy vs. angry: $t(72) = 3.89$, $p < 0.001$; happy vs. scared: $t(72) = 4.12$, $p < 0.001$; happy vs. sad: $t(72) = 2.68$, $p = 0.018$) (Figure 4.10). These results indicate that across diagnostic groups and age groups, the

RT of emotion recognition in face targets was most strongly primed by prosody that conveyed happiness among other emotions.

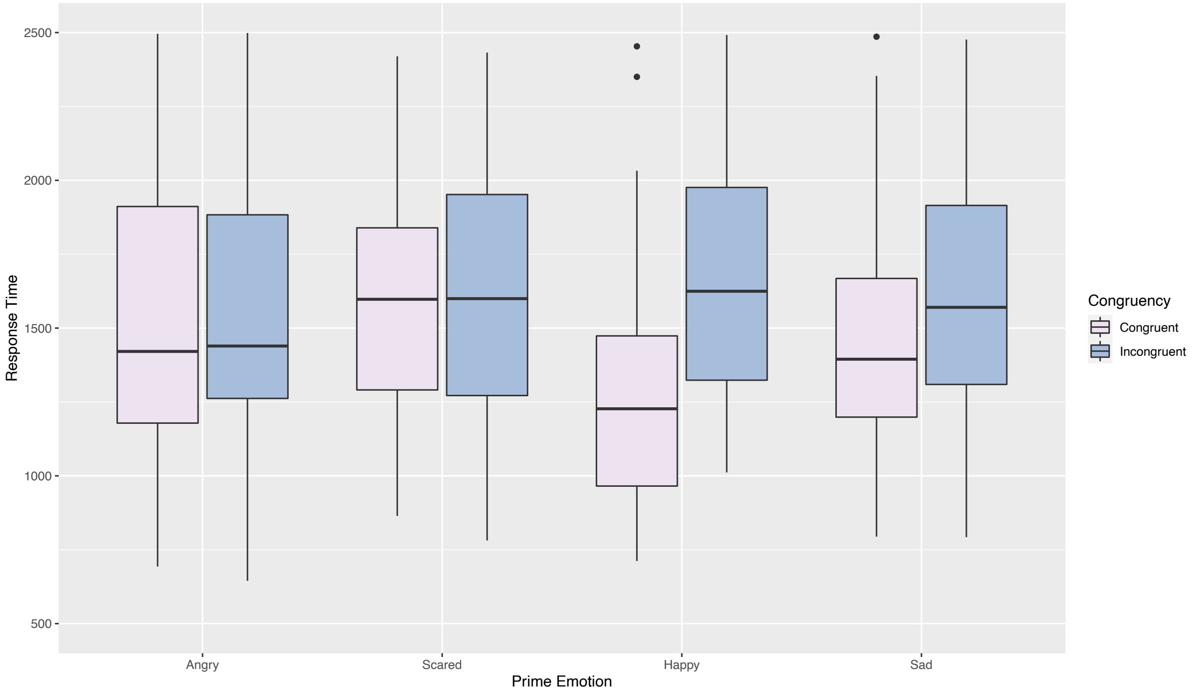


Figure 4. 10. Boxplots of mean RT across congruency levels by prime emotion in the prosody-face condition.

Song-face

The linear mixed effects model revealed a significant main effect of Prime emotion ($F(3, 279) = 3.24$, $p = 0.022$, $\eta^2_p = 0.03$). Post-hoc analyses showed a larger congruency mean RT difference when faces were primed by happy song (M = 284.13, SD = 692.13) than by angry song (M = -98.83, SD = 843.27) ($t(74) = 2.95$, $p = 0.026$), with no other differences across prime emotions observed (Figure 4.11). These results indicate that across diagnostic groups and age groups, the RT of emotion recognition in face targets was more strongly primed by song that conveyed happiness but less strongly primed by that conveyed anger.

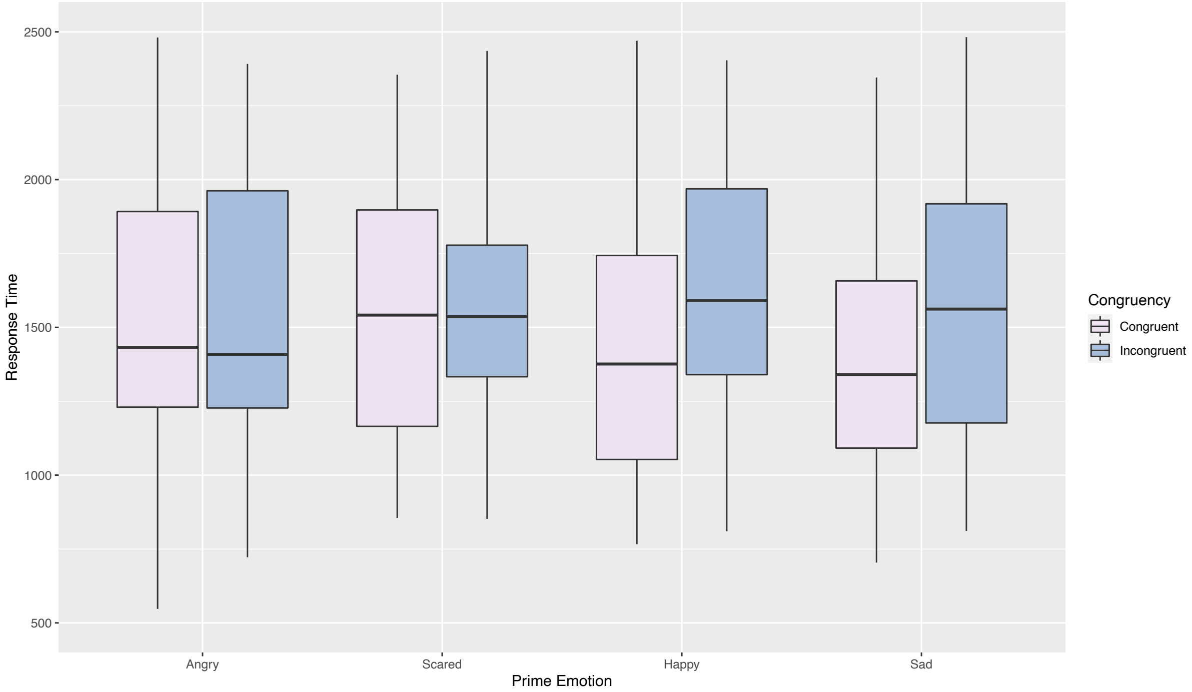


Figure 4. 11. Boxplots of mean RT across congruency levels by prime emotion in the song-face condition.

Prosody-object

The linear mixed effects analysis revealed no significant main effects nor interactions on the congruency mean RT difference across the four factors of interest. These results suggest that prime emotions did not modulate the priming effects of prosody on the RT of emotion recognition in object targets across diagnostic groups and age groups.

Song-object

The linear mixed effects analysis revealed a significant main effect of Prime emotion ($F(3, 202.01) = 12.74$, $p < 0.001$, $\eta^2_p = 0.16$). Post-hoc analyses found that the congruency mean RT difference was largest when objects were primed by happy song (M = 701.52, SD = 718.87), which differed significantly from that by angry (M = 200.31, SD = 1085.16), scared (M = -108.77, SD = 759.43), and sad song (M = 272.21, SD = 691.44) (happy vs. angry: $t(66) = 3.69$, $p < 0.001$; happy vs. sad: $t(66) = 3.91$, $p < 0.001$; happy vs. scared: $t(66) = 6.01$, $p < 0.001$). In addition, the congruency mean RT difference was also larger when objects were primed by sad than scared song ($t(66) = 3.53$, $p = 0.001$). These results indicate that across diagnostic groups and age groups, the RT of emotion recognition in object targets was most strongly primed by song that conveyed happiness but was less strongly primed by song that conveyed fear.

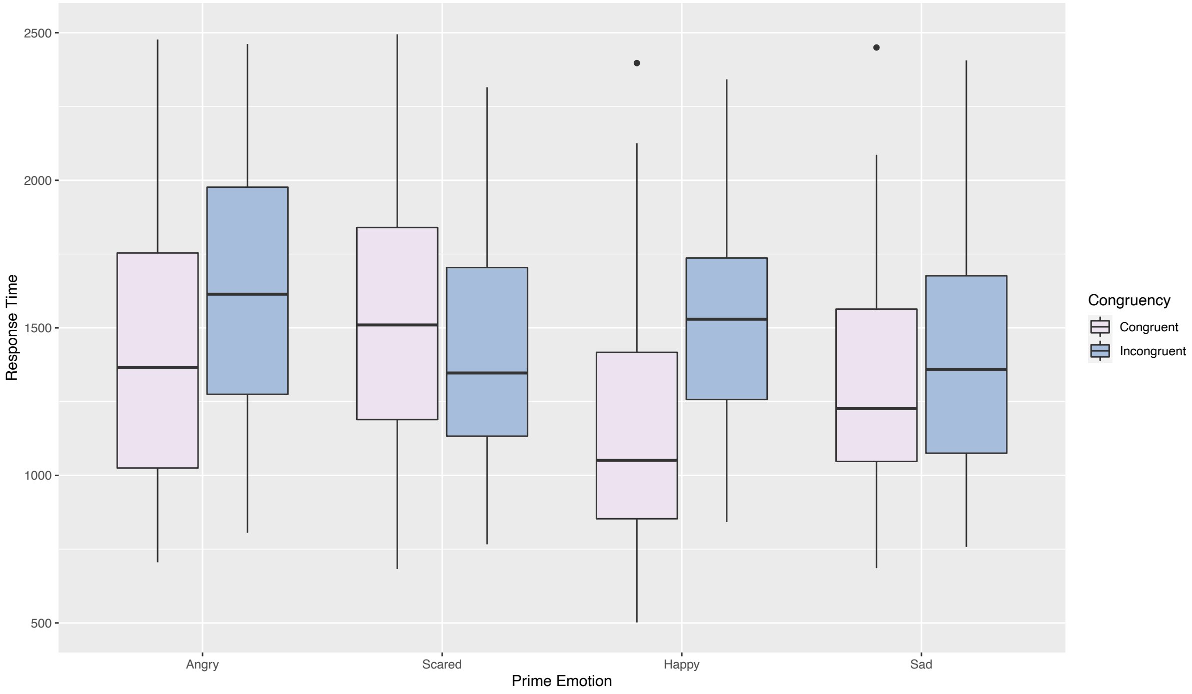


Figure 4. 12. Boxplots of mean RT across congruency levels by prime emotion in the song-object condition.

Summary

Together, prime emotion appeared to have modulated the magnitude of cross-modal priming effects in prosody-face, song-face, and song-object conditions to similar extents across diagnostic groups and age groups. Specifically, the speed of emotion recognition in faces was more strongly primed by prosody and song that conveyed happiness. Similarly, the speed of emotion recognition in objects was most strongly primed by song that conveyed happiness. Prime emotion, however, did not modulate the magnitude of priming effects in the prosody-object condition.

4.3.2.3. Congruency mean speed-accuracy composite score difference

Table 4.8 displays the full model results summary of the linear mixed effects analysis on the congruency mean SACS difference score for each prime-target condition.

Table 4. 8. Linear mixed effects model results displayed by prime-target condition for diagnostic group, age group, prime emotion, and their interactions on congruency mean SACS difference score.

	Fixed effects	**df**		***F***	***p***	**η^2_p**
Prosody-Face						
	Diagnostic group	1	70.16	1.12	0.293	0.02
	Age group	**2**	**70.19**	**3.25**	**0.045**	**0.08**
	Prime emotion	**3**	**207.19**	**6.16**	**< 0.001**	**0.08**
	Diagnostic group × age group	2	70.19	0.10	0.907	0.00
	Diagnostic group × prime emotion	**3**	**207.19**	**3.98**	**0.009**	**0.05**
	Age group × prime emotion	**6**	**207.21**	**2.19**	**0.045**	**0.06**
	Diagnostic group × age group × prime emotion	6	207.21	0.81	0.561	0.02
Song-Face						
	Diagnostic group	1	279	2.62	0.107	0.01
	Age group	2	279	2.34	0.098	0.02
	Prime emotion	**3**	**279**	**10.40**	**< 0.001**	**0.10**
	Diagnostic group × age group	2	279	2.05	0.131	0.01
	Diagnostic group × prime emotion	3	279	0.83	0.476	0.01
	Age group × prime emotion	6	279	2.08	0.055	0.04
	Diagnostic group × age group × prime emotion	6	279	0.40	0.882	0.01
Prosody-Object						
	Diagnostic group	1	64.40	0.58	0.451	0.01
	Age group	2	64.37	1.00	0.375	0.03
	Prime emotion	**3**	**198.76**	**10.99**	**< 0.001**	**0.14**
	Diagnostic group × age group	2	64.37	0.15	0.859	0.00
	Diagnostic group × prime emotion	3	198.76	0.17	0.918	0.00
	Age group × prime emotion	**6**	**198.72**	**2.47**	**0.025**	**0.07**
	Diagnostic group × age group × prime emotion	6	198.72	0.73	0.627	0.02
Song-Object						
	Diagnostic group	1	64.17	0.38	0.541	0.01
	Age group	**2**	**64.20**	**5.25**	**0.008**	**0.14**
	Prime emotion	**3**	**196.20**	**26.37**	**< 0.001**	**0.29**
	Diagnostic group × age group	2	64.20	0.41	0.662	0.01
	Diagnostic group × prime emotion	3	196.20	0.64	0.593	0.01
	Age group × prime emotion	6	196.20	1.37	0.230	0.04
	Diagnostic group × age group × prime emotion	6	196.20	0.42	0.865	0.01

Note. R model equation for Song-Face: lmer (Congruency Mean SACS Difference ~ Diagnostic Group * Age Group * Prime Emotion + (1 | Subject), control = lmerControl(calc.derivs = FALSE)); R model equation for Speech-Face, Song-Object, and Speech-Object: lmer (Congruency Mean SACS Difference ~ Diagnostic Group * Age Group * Prime Emotion + (1 | Subject)); significant effects and interactions are highlighted using bold font.

Prosody-face

The analysis revealed significant main effects of Age group (*F*(2, 70.19) = 3.25, p = 0.045, η^2_p = 0.08) and Prime emotion (*F*(3, 207.19) = 6.16, $p < 0.001$, η^2_p = 0.08), as well as two-way interactions of Diagnostic group × Prime emotion (*F*(3, 207.19) = 3.98, p = 0.009, η^2_p

= 0.05) and Age group × Prime emotion ($F(6, 207.21) = 2.19$, $p = 0.045$, $\eta^2_p = 0.06$). No other main effects and interactions reached significance. I will unpack the two two-way interactions below.

Figure 4.13 shows the interaction of Diagnostic group × Prime emotion. Post-hoc analyses showed that there was a larger congruency mean SACS difference in the ASD group (M = 1.08, SD = 1.82) compared to the TD group (M = 0.03, SD = 1.33) when faces were primed by scared prosody ($t(65.83) = 2.83$, $p = 0.006$). No difference between the ASD and TD groups was found for angry, happy, and sad prosody.

The priming patterns across prosodic emotions on faces also differed between the ASD and TD groups. In the ASD group, a larger congruency mean SACS difference was observed when faces were primed by happy prosody (M = 1.47, SD = 1.43) than that by sad prosody (M = 0.53, SD = 1.51) ($t(36) = 3.49$, $p = 0.008$), with no other difference observed across prime emotions.

In the TD group, a larger congruency mean SACS difference was observed when faces were primed by happy (M = 1.36, SD = 1.51) and sad prosody (M = 0.90, SD = 1.25) than that by scared prosody (M = 0.03, SD = 1.33) (happy vs. scared: $t(35) = 4.09$, $p = 0.001$; sad vs. scared: $t(35) = 2.71$, $p = 0.031$), with no other difference observed across prime emotions.

These results indicate that while the ASD group showed stronger priming effects of prosody that conveyed happiness than sadness on the efficiency of emotion recognition in faces, the TD group showed no difference in the priming effects between these two emotions. Moreover, whereas the efficiency of emotion recognition in faces was equally primed by prosody that conveyed fear and other emotions in the ASD group, the efficiency of emotion recognition in faces was less strongly primed by prosody that conveyed fear compared to happiness and sadness in the TD group.

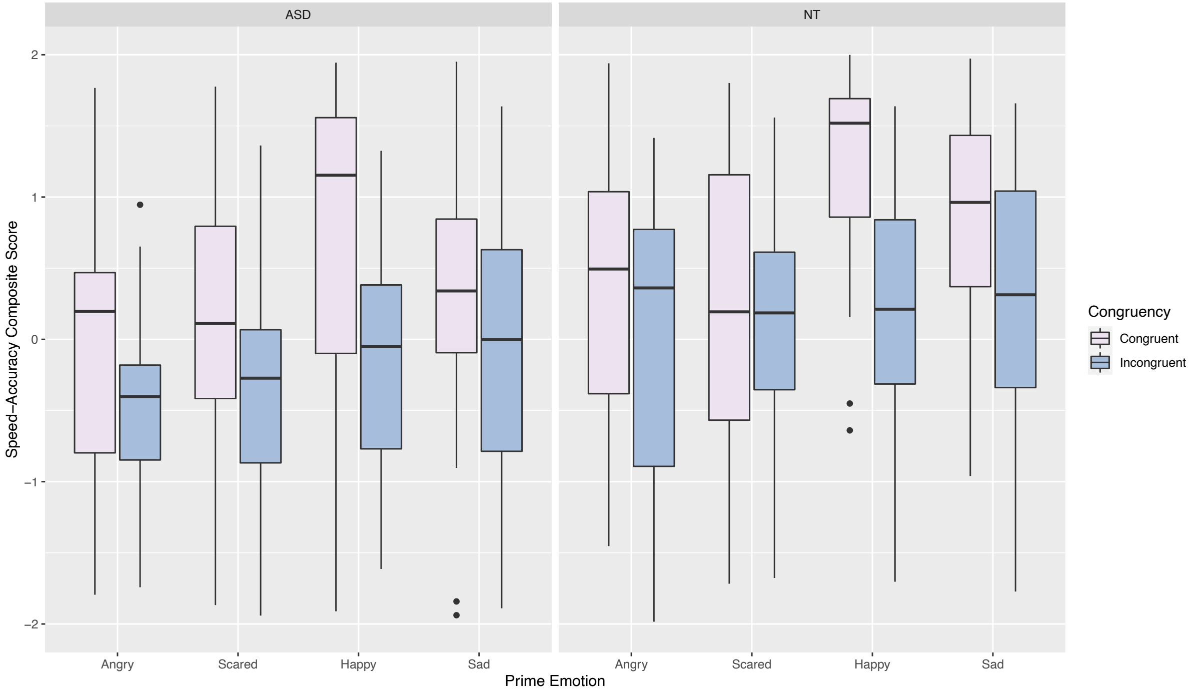


Figure 4. 13. Boxplots of mean SACS across congruency levels by prime emotion for the ASD and TD groups in the prosody-face condition.

Figure 4.14 shows the interaction of Age group × Prime emotion. Post-hoc analyses revealed that children showed larger congruency mean SACS difference than adults when faces were primed by angry prosody (child: M = 1.50, SD = 1.84; adult: M = -0.21, SD = 1.47; $t(50.90) = 3.77$, $p = 0.001$) and sad prosody (child: M = 1.15, SD = 1.19; adult: M = 0.29, SD = 1.23; $t(51.33) = 2.61$, $p = 0.036$), where no difference was observed between the two groups for scared and happy prosody. There was also no difference in congruency mean SACS difference between adolescents and adults across prosodic emotions when responding to face targets.

The priming patterns across prosodic emotions on faces were found to differ across the three age groups. In children, the congruency mean SACS difference was larger when faces were primed by angry (M = 1.50, SD = 1.84) and happy prosody (M = 1.58, SD = 1.40) than that by scared prosody (M = 0.46, SD = 1.75) (angry vs. scared: $t(24) = 2.85$, $p = 0.026$; happy vs. scared: $t(24) = 3.05$, $p = 0.026$), with no other differences observed across prime emotions.

In adolescents, the congruency mean SACS difference did not differ across prime emotions.

In adults, the congruency mean SACS difference was larger when faces were primed by happy prosody (M = 1.23, SD = 1.39) than that by angry prosody (M = -0.21, SD = 1.47) ($t(25) = 3.21$, $p = 0.022$).

These results indicate that children showed stronger priming effects on the efficiency of emotion recognition in face targets than adults particularly when prosodic primes conveyed anger and sadness. This partially agrees with that in children, the efficiency of emotion in face targets was more strongly primed by angry and happy prosody, whereas in adults, the efficiency of emotion in face targets was less strongly primed by anger.

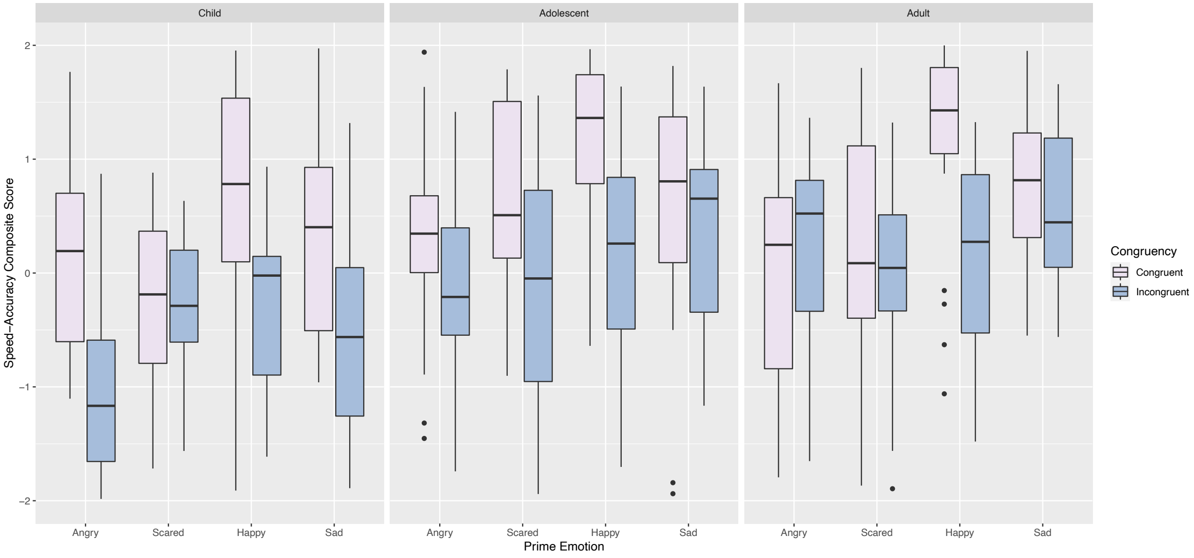


Figure 4. 14. Figure 4.14. Boxplots of mean SACS across congruency levels by prime emotion for children, adolescents, and adults in the prosody-face condition.

Song-face

The linear mixed effects model yielded a significant main effect of Prime emotion ($F(3, 279) = 10.40$, $p < 0.001$, $\eta^2_p = 0.10$). Post-hoc analyses revealed that the congruency mean SACS difference was largest when faces were primed by happy song (M = 1.43, SD = 1.28), which differed significantly from that by angry (M = 0.28, SD = 1.81), scared (M = 0.13, SD = 1.82), and sad song (M = 0.59, SD = 1.39) (happy vs. angry: $t(74) = 4.51$, $p < 0.001$; happy vs. scared: $t(74) = 5.01, p < 0.001$; happy vs. sad: $t(74) = 4.48, p < 0.001$) (Figure 4.15). These results indicate that across diagnostic groups and age groups, the efficiency of emotion recognition in face targets was most strongly primed by song that conveyed happiness among other emotions.

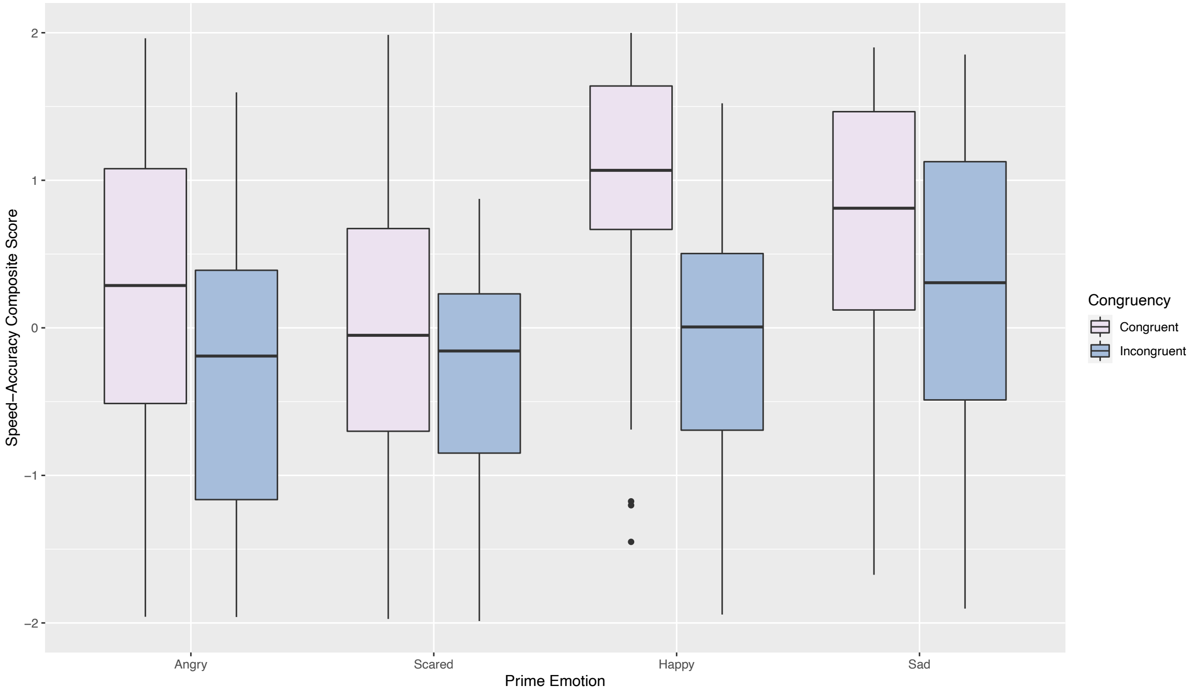


Figure 4. 15. Boxplots of mean SACS across congruency levels by prime emotion in the song-face condition.

Prosody-object

The analysis yielded a significant main effect of Prime emotion ($F(3, 198.76) = 10.99$, $p < 0.001$, $\eta^2_p = 0.14$), as well as a significant interaction of Age group × Prime emotion ($F(6, 198.72) = 2.47$, $p = 0.025$, $\eta^2_p = 0.07$). No other effects and interactions were significant. I will unpack this two-way interaction below.

Figure 4.16 shows the interaction of Age group × Prime emotion. Post-hoc analyses found no difference in congruency mean SACS difference across age groups for the different prime emotions. The priming patterns across prosodic emotions on objects were found to differ across the three age groups.

In children, the congruency mean SACS difference did not differ across prime emotions when responding to object targets.

In adolescents, the congruency mean SACS difference was the largest when objects were primed by happy prosody (M = 1.35, SD = 1.11), which differed significantly from that by angry (M = -0.30, SD = 1.01), scared (M = 0.34, SD = 1.10), and sad prosody (M = 0.58, SD = 1.14) (happy vs. angry: $t(19) = 5.69$, $p < 0.001$; happy vs. scared: $t(19) = 3.52$, $p = 0.005$; happy vs. sad: $t(19) = 4.02$, $p = 0.002$).

In adults, the congruency mean SACS difference did not differ across happy (M = 0.88, SD = 1.41), scared (M = 0.82, SD = 1.23), and sad prosody (M = 0.92, SD = 1.06) when responding to object targets, which were all in turn showing larger congruency mean SACS difference than angry prosody (M = -1.32, SD = 3.63) (happy vs. angry: $t(20) = 3.27$, $p = 0.011$; scared vs. angry: $t(20) = 3.06$, $p = 0.012$; sad vs. angry: $t(20) = 3.36$, $p = 0.011$).

These results indicate that whereas the efficiency of emotion recognition in object targets was not modulated by the emotion of the prosodic primes in children, the efficiency of emotion recognition in objects was more strongly primed by prosody that conveyed happiness in adolescents but was less strongly primed by prosody that conveyed anger in adults.

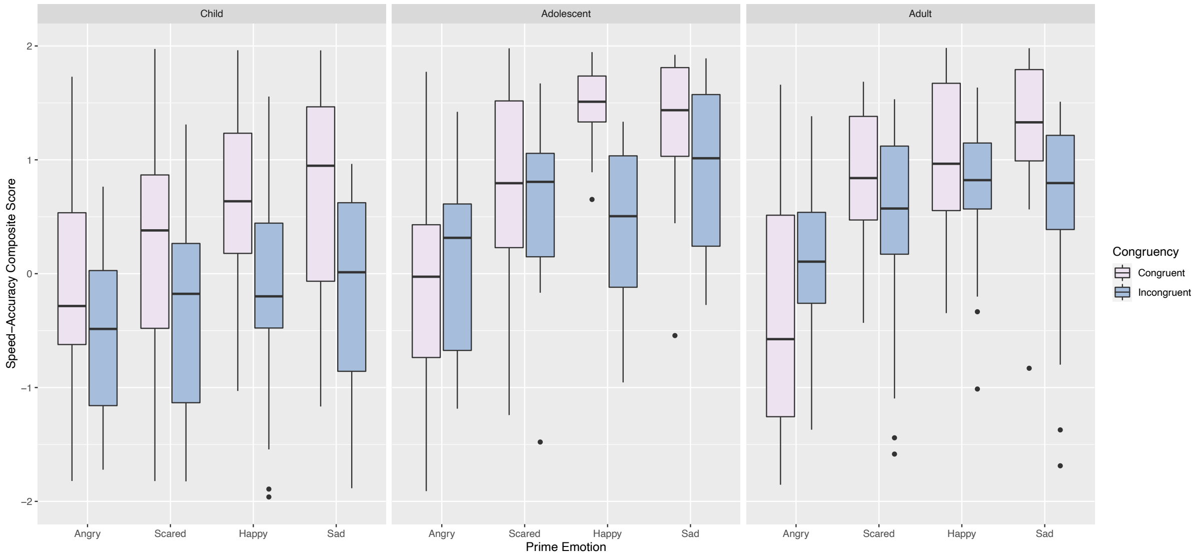


Figure 4. 16. Boxplots of mean SACS across congruency levels by prime emotion for children, adolescents, and adults in the speech-object condition.

Song-object

The linear mixed effects analysis revealed a significant main effect of Age group ($F(2, 64.20) = 5.25$, $p = 0.008$, $\eta^2_p = 0.14$). Post-hoc analyses demonstrated that adolescents (M = 0.15, SD = 0.59) showed a smaller congruency mean SACS difference than children (M = 0.98, SD = 1.02) and adults (M = 0.75, SD = 1.14), who did not differ from each other (adolescent vs. child: $t(44.61) = -3.63$, $p = 0.002$; adolescent vs. adult: $t(38.78) = -2.36$, $p = 0.035$) (Figure 4.17). This suggests that in general, adolescents showed smaller priming effects of sung emotions on the efficiency of emotion recognition in object targets compared to children and adults.

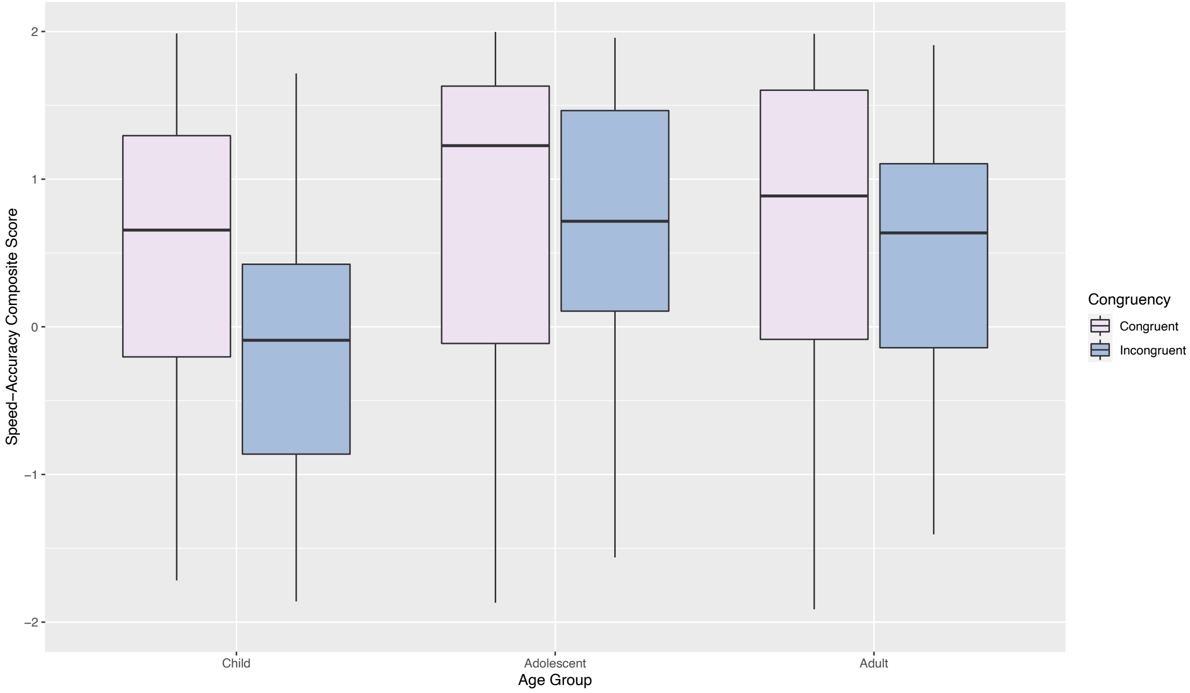


Figure 4. 17. Boxplots of mean SACS across congruency levels by age group in the song-object condition.

A significant main effect of Prime emotion was also depicted ($F(3, 196.20) = 26.37$, $p < 0.001$, $\eta^2_p = 0.29$). Post-hoc analyses revealed that the congruency mean SACS difference was largest when objects were primed by happy song (M = 1.65, SD = 1.39), followed by sad song (M = 0.78, SD = 1.32), with angry (M = -0.14, SD = 1.90) and scared song (M = 0.05, SD = 1.24) showing the least congruency mean SACS difference (happy vs. sad: $t(66) = 4.75$,

$p < 0.001$; happy vs. angry: $t(66) = 7.30$, $p < 0.001$; happy vs. scared: $t(66) = 6.88$, $p < 0.001$; sad vs. angry: $t(66) = 3.55$, $p = 0.001$; sad vs. scared: $t(66) = 2.94$, $p = 0.005$) (Figure 4.18). These results indicate that across diagnostic groups and age groups, the efficiency of emotion recognition in object targets was most strongly primed by song that conveyed happiness, followed by sadness, with anger and fear showing the least priming effects on objects.

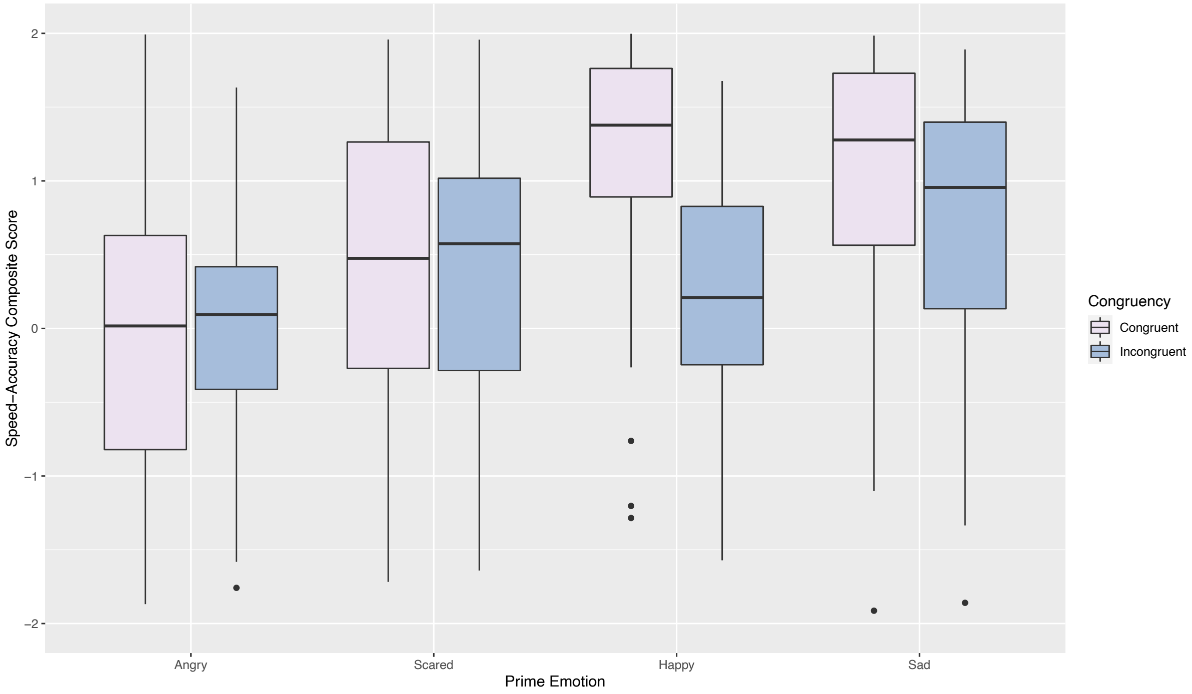


Figure 4. 18. Boxplots of mean SACS across congruency levels by prime emotion in the song-object condition.

Summary

In general, prime emotion appeared to have modulated the magnitude of cross-modal priming effects in all four prime-target conditions to similar extents across diagnostic group and age group. There were, nonetheless, two exceptions: (i) the ASD group showed stronger priming than the TD group when faces were primed by prosody that conveyed fear, and (ii) children showed stronger priming than adults when faces were primed by prosody that conveyed anger and sadness. Overall, visual emotion recognition efficiency was most strongly primed by auditory primes, particularly song, that conveyed happiness. The efficiency of emotion recognition efficiency in object targets was relatively less strongly primed by auditory primes that conveyed anger.

4.4. Discussion

The present study was the first study, to my knowledge, to directly compare the implicit processing of prosodic and sung emotions as well as their influence on subsequent emotional judgment of human faces and face-like objects between autistic and NT individuals across development, via a cross-modal affective priming paradigm. The most important findings are the following. First, the strength of cross-modal affective priming did not differ between the ASD and NT groups, suggesting no impairments in implicit emotion processing across auditory domains in ASD. However, the two groups appear to show different developmental patterns in implicit emotion processing across the auditory domains. That is, in the NT group, prosodic but not sung emotions implicitly primed subsequent emotional judgment of human facial expressions across all age groups; the implicit priming of sung emotions was seen only in NT children but not in NT adolescents and adults. The reverse was observed for the ASD group, where sung but not prosodic emotions implicitly primed subsequent emotional judgment of human facial expressions across all age groups; the implicit priming of prosodic emotions was seen only in children and adolescents with ASD but not in adults with ASD. Secondly, priming effects of both prosodic and sung emotions on emotional judgment of human facial expressions were weaker in adults compared to children, regardless of diagnostic group. Thirdly, prosodic and sung emotions did not prime emotional judgment of face-like object expressions across participants in general, except for children when primed by prosodic emotions. Finally, priming effects appeared stronger when auditory primes conveyed happiness, which was particularly evident for sung cues relative to prosodic cues. These findings will be further unpacked and discussed in the subsequent sections.

4.4.1. The special status of speech prosody during cross-modal affective priming in NT but not in ASD across development

Overall, the present study provided no evidence for impairments in implicit emotion priming in autistic individuals relative to NT individuals. This is demonstrated by the finding of no significant group differences in the extent to which visual emotion recognition was primed by implicitly processed auditory emotions. These findings, however, appear to contradict the numerous studies that reported atypical emotion processing at the implicit level in ASD, particularly for speech prosody (Fan & Cheng, 2014; Lindström et al., 2018). It should be noted that these studies examined implicit emotion processing by measuring participants' neural responses to emotionally spoken syllables during passive listening tasks. Thus, the intact implicit emotion processing on the behavioural level (i.e., its influence on emotional recognition) observed in the present study does not necessarily counter the fact that implicit emotion processing may be impaired in ASD on the neurophysiological level (i.e., influence on neural responses) as observed in these previous studies. With specific regards to implicit emotion priming, the present study corroborated previous findings by Vanmarcke and Wagemans (2017) but contradicted those by Kamio et al. (2006). The discrepancy in findings across the previous and present affective priming studies may relate to the association between the implicitly preactivated concept and the concept required for the target response. For instance, the present study found no impairments in ASD as participants were presented with *emotional* prosody and song as primes then made *emotional* judgments of human facial and face-like object expressions. Likewise, Vanmarcke and Wagemans (2017) also found no impairments in ASD as participants were presented with *emotional* faces (happy vs. sad) as primes then made *valence* judgments (positive vs. negative) of human facial expressions. These two studies, thus, demonstrate a common concept of emotional meaning to be transferred from the prime to the response to the target. By contrast, Kamio et al. (2006) found impaired emotion priming in ASD where participants were presented with *emotional* faces as primes and made *liking* judgments of ideographs, which denotes a less related concept between emotional meaning of the prime and preferential judgments required as responses to targets. Thus, it may be the case where atypicality lies in the extent to which implicitly processed emotional information can guide non-emotionally-related behaviours in a broader context in ASD. Drawing together findings from previous literature and the present study, it opens up new questions as to which aspect(s) of emotion processing at the implicit level is affected in ASD, such as whether it relates to its impact on the neurophysiological versus behavioural levels, implicit appraisal of the emotional significance of incoming cues, and/or the implicit activation of emotional meaning to guide (non)emotionally-related behaviours.

Although no overall impairments in implicit emotion priming were observed across domains in the present study, discrepant priming patterns of prosody and song were observed between the ASD and NT groups across development, mainly in the accuracy measure. Whereas the priming of prosody on human faces was observed across development in the NT group, an age-related decline in the priming of prosody on human faces was observed in the ASD group. In contrast, whereas the priming of song on human faces declined with age in the NT group, it was found across development in the ASD group. These results suggest that for NT individuals across development, emotions conveyed in speech prosody may have a special status during implicit emotion priming when identifying facial expressions. On the other hand, for autistic individuals, particularly adults, prosodic cues may not share the same importance for priming the interpretation of facial expressions as they do for NT individuals, whereas sung cues may be preferred.

The reduced implicit processing bias towards speech prosody in autistic individuals compared to NT individuals has also been demonstrated in the literature on speech perception. Indeed, studies have shown that NT individuals prefer speech over non-speech stimuli from a very early age (Alegria & Noirot, 1978; Vouloumanos et al., 2010). Conversely, autistic individuals do not exhibit such bias (Filipe et al., 2018; Järvinen-Pasley & Heaton, 2007; Klin, 1991, 1992) and show greater visual attention to performers' face and body during singing than story-telling (G. A. Thompson & Abel, 2018). Moreover, findings from ERP studies using a passive listening paradigm provide substantial evidence for atypical early, preattentive processing of speech stimuli in ASD (Čeponienė et al., 2003; Kujala et al., 2005; Whitehouse & Bishop, 2008; J. Zhang et al., 2019). In particular, the attenuated ERPs to speech stimuli in ASD relative to typical development was only observed during the passive listening condition (Whitehouse & Bishop, 2008). This difference, however, disappeared when attention to the sound stream was required (Whitehouse & Bishop, 2008). It was, therefore, suggested that top-down inhibitory processes may play an important role in the atypical initial orientation towards speech sounds in ASD (Whitehouse & Bishop, 2008). In line with this, the social motivation theory hypothesises that autistic individuals have diminished attention to certain stimuli because of its social-cognitive aspects, such as faces and speech (Chevallier et al., 2012). Furthermore, this theory posits that social motivation deficits should precede social cognitive deficits (Chevallier et al., 2012). Based on this theoretical framework, the atypical initial orientation towards speech sounds (Čeponienė et al., 2003; Kujala et al., 2005; Lepistö et al., 2005; Whitehouse & Bishop, 2008) may have resulted in reduced automatic responses to emotional prosodic stimuli, exerting an influence on subsequent cross-modal emotional transfers such as priming in ASD.

The present study noted a lack of emotion priming of sung cues on human faces in NT adolescents and adults, giving rise to the special role of speech in implicit emotion priming. Indeed, using electrodermal activity and electrocardiogram measures, it was shown that lower autonomic reactivity to emotions in music was related to more typical social functioning in NT individuals (Järvinen et al., 2016). Despite this, the present findings are seemingly inconsistent with prior studies which showed affective priming of musical cues on visual emotional words (Goerlich et al., 2012; Steinbeis & Koelsch, 2011) and human faces (L. Zhou et al., 2019) in NT individuals, though mainly at the electrophysiological level. Although these prior studies also showed a lack of cross-modal priming effects of musical stimuli for accuracy at the behavioural level, this was due to ceiling performance on this measure in these studies (Goerlich et al., 2012; Steinbeis & Koelsch, 2011; L. Zhou et al., 2019). One could, therefore, suspect that the accuracy measure lacks sensitivity to capture the prime-target relationship. However, the present study did not observe ceiling performance. This explanation for the lack of emotion priming of sung cues on human faces in NT adults, therefore, seems unlikely.

There are substantial differences between these previous studies examining affective priming of music in typical development and the present study, which may have resulted in the discrepant findings. First, the use of sung stimuli in the present study differs considerably from the instrumental musical stimuli used in previous studies in terms of acoustic characteristics, such as timbre features (Mokhsin et al., 2014). Related to this, a previous study by Zhang and colleagues (2019) showed that there are specialised neural responses to sung music when performing on a cross-modal affective priming task, which were not observed for instrumental music. Specifically, it was found that the integration of emotional information between sung and facial stimuli required sustained, controlled attention allocation, as reflected in a larger P3 and late positive component (LPC) for incongruent than congruent prime-target pairs (Zhang et al., 2019). On this note, the manipulation of SOA to eliminate attention allocation to the prime stimuli in the present study may have limited the priming effects of sung cues on human faces. Secondly, the number of emotions used differed between the previous studies (i.e., happy and sad in Goerlich et al., 2012; happy and angry/scared in Zhou et al., 2019) and the present study (i.e., happy, angry, scared, and sad). Moreover, whereas participants were instructed to respond on the basis of valence to the targets in the above-mentioned studies (Goerlich et al., 2012; L. Zhou et al., 2019), participants responded on the basis of emotion category in the present study. Given the greater task demands and possible ambiguity between emotional categories than valence (Carroll & Young, 2005; Eerola & Vuoskoski, 2011), a potential "leakage effect" may have inhibited priming effects to be observed, where an incongruent emotion may have primed a target (e.g., scared-sad). These possibilities in accounting for the discrepant findings of the present and previous studies should be considered with caution as the priming effects of sung stimuli were nonetheless observed across all age groups in ASD.

One possibility of why the priming effects of sung stimuli remained for the ASD group could be related to the greater activation for these stimuli at the neural level. In support of this, neuroimaging studies using fMRI and diffusion tensor imaging (DTI) demonstrated that relative to NT individuals, activation in the neural system associated with speech and song processing (e.g., the left inferior frontal gyrus) was reduced for speech stimuli, but comparable or greater for sung stimuli in autistic individuals (Lai et al., 2012; Sharda et al., 2015). In addition, fronto-temporal connectivity was found to be greater for song relative to speech in ASD, while these differences were not seen in NT (Lai et al., 2012; Sharda et al., 2015). The greater neural responses to sung over spoken stimuli may imply greater orientation towards sung cues for emotional information to be processed and transferred through priming in ASD. Moreover, the present findings corroborate previous multisensory integration research in showing the reduced cross-modal influence of spoken stimuli (Ben-Yosef et al., 2017; Charbonneau et al., 2013; O'Connor, 2007; Vannetzel et al., 2011; Xavier et al., 2015), but typical cross-modal influence of musical stimuli on facial expressions in ASD (L. S. Brown, 2017; Wagener et al., 2020). Taken together, the cross-modal influence of music on facial expressions appears to occur at both the conscious and automatic levels in ASD, even as a transient cue in a sung format as demonstrated in the present study. Considering the timeframe for capturing automatic processes was controlled for in the present study, the present findings outline the important role of automaticity in response to emotional stimuli as a prerequisite to cross-modal transfers and/or integration of emotion. In particular, prosodic cues appear to be prioritised in implicit emotion priming in typical interpersonal emotional communication.

4.4.2. A gradual age-related decline in the magnitude of cross-modal affective priming

Across both the ASD and NT groups, the magnitude of implicit emotion priming of both prosody and song was less strong in adults compared to children when responding to face targets, with adolescents showing no difference to either age group. This indicates a gradual age-related decline that is most salient between children and adults. This result seems counter to previous findings of no effects of age in implicit emotion priming (Kamio et al., 2006) and age-related improvement in implicit emotion processing (Batty & Taylor, 2006; Mathersul et al., 2009; L. M. Williams et al., 2009), which may be subject to the design and response demands in the present study.

Given that the prime and target stimuli were presented through different modalities in the present study, one possibility of the age-related decline in priming may concern the developmental changes in auditory versus visual dominance in multisensory processing. There is much evidence suggesting a developmental shift from an auditory to visual dominance during multisensory processing (Hirst, Stacey, et al., 2018; Nava & Pavani, 2013; Robinson & Sloutsky, 2004). For example, the Colavita effect, denoting a visual dominance effect demonstrated by participants responding only to the visual element of cross-modally presented stimuli, has been found to be less robust and perhaps reversed (i.e., auditory dominance) in children compared to adults (Wille & Ebersbach, 2016; see also Hirst, Cragg, et al., 2018 for a review). Likewise, with particular regard to emotion processing, a recent study by Ross et al. (2021) showed an auditory dominance towards emotional vocal bursts over body expressions in children that was not seen in adults. In this sense, younger groups may have been more prone to the influence of auditory cues during cross-modal priming than older groups in the present study. However, the developmental transition from auditory to visual dominance has been suggested to occur at approximately 9-10 years of age (Nava & Pavani, 2013; Ross et al., 2021; Wille & Ebersbach, 2016), which does not seem to account for why similar priming between children and adolescents was observed in the present study. Age-related differences in modality dominance may, thus, only somewhat explain the present finding for children, giving rise to the possibility for an alternative explanation.

Considering participants were instructed to respond to the facial expressions, it is plausible that increased age may have contributed to the inhibition of automatic responses to the task-irrelevant information elicited in the auditory primes. Accordingly, the influence of the auditory primes regardless of congruency may have been minimised in adults. In support of this idea, a previous study by Herba et al. (2006) found that with increasing age, children became less distracted by task-irrelevant emotional information when matching emotional faces on the basis of identity. Furthermore, evidence from fMRI studies have reported age-related differences in brain activation during cognitive control over interfering emotional content, though these studies mostly compared between adults and adolescents (Monk et al., 2003; Passarotti et al., 2009; Lihong Wang et al., 2008). For instance, in Passarotti et al. (2009), it was found that when judging the age of emotional faces, adolescents exhibited greater activation in the right amygdala than adults, a finding that has also been noted during incidental processing of task-irrelevant emotional content in previous studies (Guyer et al., 2008; Monk et al., 2003). Moreover, this greater amygdala reactivity in adolescents was found to be coupled with reduced activation in the right ventrolateral prefrontal cortex (Passarotti et al., 2009), a region that has been implicated in self-regulatory control and emotional regulation (Marsh et al., 2006; Pavuluri et al., 2009). Together, the authors interpreted these results as indicating greater automatic responses to emotional content in adolescents compared to adults, despite that emotion processing is not required, which may be modulated by poor control of emotional circuits (Passarotti et al., 2009). In other words, there appears an increased cognitive control exerting over automatic responses to emotional content which may undergo gradual development throughout adolescence (Passarotti et al., 2009; Ravindranath et al., 2020).

It is, therefore, plausible that automatic responses to emotional information emerge from a young age (Lobaugh et al., 2006), but the degree of their influence on subsequent socioemotional judgment may become more dependent on their relevance to task demands with increased age. The evidence from previous studies discussed so far characterises especially the age-related improvement in cognitive control over task-irrelevant emotional content under within-modal contexts, while the present findings extended this to that under cross-modal contexts. More importantly, despite adults maybe being less susceptible to auditory interference when judging emotions in visual stimuli, priming was observed for sung and prosodic cues on facial expressions in adults with ASD and NT adults, respectively. This may further reinforce the diverging preference during implicit emotion processing between the two groups. Specifically, for NT adults, prosodic cues may not be easily inhibited during implicit emotion priming, given their special role in face-to-face conversations.

4.4.3. The modulating effect of prime-target co-occurrence on cross-modal affective priming

So far, the discussion has focused on the cross-modal emotion priming effects of prosody and song on human faces and their differences between the ASD and NT groups across ages. However, in general, cross-modal emotion priming effects of both prosody and song on face-like objects were not seen regardless of diagnostic group and age group, except for children with prosodic cues. It is plausible that the lack of cross-modal affective priming for face-like objects was due to their low co-occurrence with the human voice (in either the spoken or sung form), as opposed to the high co-occurrence between human faces and the human voice (in both the spoken and sung form). There may be a weaker association formed between face-like objects and the human voice, as they are not typically encountered simultaneously in our everyday social interaction. The present findings are consistent with and extend those reported in Carroll & Young (2005), by demonstrating the role of prime-target co-occurrence with different stimuli. Namely, whereas vocal bursts and printed words were used as low co-occurring pairs in Carroll & Young (2005), the present study using the human voice and face-like objects as low co-occurring pairs also found restricted cross-modal emotional transfer for accuracy in most groups.

However, this interpretation cannot explain the existence of cross-modal emotion priming of prosody on face-like objects in children. To the best of my knowledge, no previous studies have examined the role of co-occurrence in emotion priming nor the ability to process emotions from audiovisual stimuli involving nonhuman faces (e.g., a speaking cartoon character) in children. The lack of research in this area makes it challenging to find evidence to interpret this finding. Nevertheless, based on the co-occurrence framework, it is possible that children may encounter pairings of the human voice and nonhuman faces more often compared to adolescents and adults, perhaps through watching cartoons or playing with talking toys that are nonhuman looking. Unfortunately, it is not feasible to extrapolate from the current data whether priming effects on face-like objects in children were confounded by frequent exposure to such pairings during early development. Continued investigation into whether and how exposure to socioemotional stimuli, such as television viewing, contributes to emotion processing would provide greater understanding of how social experience is shaped and refined throughout development.

4.4.4. An advantage for happy prosody and song in cross-modal affective priming

Across participants, implicit emotion priming effects were consistently modulated by prime emotion: the judgment of facial and face-like object emotions was more strongly primed by auditory cues that conveyed happiness. This was particularly evident for sung primes, whereby the happiness advantage was observed across all performance indices regardless of target types. In contrast, the happiness advantage for prosodic primes was less clear, as it did not always show greater priming among other emotions and across target types. This happiness advantage, however, could not be simply explained by participants' recognition of this emotion at the explicit level, when considering the results presented in Chapter 3. For prosody, although happiness was consistently found to be most well-recognised emotion on the explicit emotion recognition task, as mentioned earlier, it did not always show stronger priming than other emotions. Moreover, for song, happiness constantly showed the strongest priming effects compared to other emotions, despite being recognised equally well as other emotions (except for fear which was the least well-recognised emotion) on the explicit emotion recognition task. Alternative explanations of this finding are outlined below.

The happiness advantage has consistently been demonstrated in the human face domain, primarily on visual search tasks, where the detection of happy faces was found to be quicker and more accurate compared to faces portraying other emotions (Calvo et al., 2008; Calvo & Marrero, 2009; Juth et al., 2005). The present finding, therefore, appears to extend previous observations of the happiness advantage in preattentive processing of human faces to that in speech prosody and song. Borrowing from the discussion on the underlying mechanisms of the happiness advantage in the face domain, two factors can be considered: the emotional content of expressions (i.e., emotional salience) and the physical perceptual properties of the expressions conveyed through the auditory primes (i.e., perceptual salience). With respect to the emotional salience account, the happiness advantage arisen from implicit processing seems to be at odds with the adaptive biological account, where a threat-detection advantage is postulated. From a biological point of view, protection and survival must be safeguarded prior to attending to benefit and pleasure (Neuberg et al., 2011; Tooby & Cosmides, 1990). Accordingly, the processing of expressions signalling potential harm (e.g., either as direct threat for the perceiver such as anger; or in various indirect ways such as fear and sadness) are thought to be prioritised over expressions signalling potential benefit (e.g., happiness). A number of studies have provided supporting evidence for this prioritisation (Dolan et al., 2001; Eastwood et al., 2001; Fox et al., 2000; Krysko & Rutherford, 2009; Öhman et al., 2001). The emotional salience account, therefore, appears an unlikely explanation for the happiness advantage outlined by the current data.

In terms of perceptual salience, the distinctiveness of specific facial features is thought to be responsible for facilitating initial orientation during preattentive search processes in the human face domain (Calvo et al., 2008; Calvo & Nummenmaa, 2008; Frischen et al., 2008; Nummenmaa & Calvo, 2015). For instance, a smile with teeth showing in a happy face introduces a high luminance contrast, which may produce a detection advantage (Frischen et al., 2008). This is plausible as studies employing schematic line drawings of faces to control for perceptual variance across realistic photographs have tended to not find the happiness advantage (Eastwood et al., 2001; Fox et al., 2000; Öhman et al., 2001). In addition, the detection advantage of happiness remained the same when the photographs were displayed in the upright and inverted positions (Calvo et al., 2008), further supporting the beneficial role of specific features. Accordingly, the privileged cross-modal priming of happiness may be attributed to specific acoustic cues characterising this emotion in the present study. Among the large number of acoustic features that have been studied, four parameters have emerged as important candidates subserving listeners' ratings of emotional expressions in speech and music: speech rate/tempo, voice intensity/sound level, voice quality/timbre and fundamental frequency (F_0)/pitch (Juslin, 1997, 2000; Juslin & Laukka, 2003; Lieberman & Michaels, 1962; Pralus et al., 2019; Scherer & Oshinsky, 1977). With respect to these acoustic cues, the intensity and duration of the auditory stimuli used in the present study were normalised. In addition, the mean fundamental frequency (F_0) and the mean F_0 excursion size did not differ between emotions within the present speech and song stimulus sets (see Section 3.2.3.1). In spite of the subtle differences between emotions across these acoustic features, the processing advantage for happiness was nonetheless revealed, suggesting that the contribution of specific acoustic features may be unlikely. To this end, it should be considered that emotional information in the voice is transmitted via multiple acoustic features (Banse & Scherer, 1996; Pralus et al., 2019; Sobin & Alpert, 1999). fMRI studies have shown that emotional acoustic information is integrated in regions within the right mid superior temporal gyrus (STG), to form an emotional gestalt prior to higher order cognitive processing (Fecteau et al., 2007; Schirmer & Kotz, 2006). Furthermore, it has been illustrated that hemodynamic responses to emotional information during passive listening was driven by a combination of acoustic features, whereby none of the parameters alone could explain the responsiveness of this brain region (Wiethoff et al., 2008). Taken together, these findings give rise to the emotion-specific patterns derived from multiple acoustic cues in conjunction that determine expression distinctiveness, rather than single distinctive acoustic features per se.

On this note, the distinctiveness of happy speech prosody and song could have been resulted from the general underrepresentation of positive emotions in the present study. As mentioned previously, the potential ambiguity between anger, fear, and sadness as belonging to the same valence may have led to a “leakage effect”, inhibiting priming effects to be observed (Carroll & Young, 2005; Eerola & Vuoskoski, 2011). Thus, the questions of whether expression distinctiveness is responsible for different degrees of cross-modal emotion priming, as well as what contributes to this distinctiveness of auditory expressions are yet to be addressed in future research. As noted earlier, the happiness advantage was less clear for prosodic than sung primes. Given that emotion intensity plays an important role in vocal emotion recognition (Juslin & Laukka, 2001), it is speculated that priming of happy prosody may have been attenuated due to its lower intensity ratings obtained from the validation process compared to other emotions (see Section 3.2.3.3). More importantly, since speech and song are considered overlapping forms of vocal expression and share similar acoustic features (Atmaja & Akagi, 2020; K. R. Scherer et al., 2015; B. Zhang, Essl, et al., 2015; B. Zhang, Provost, Swedberg, Essi, et al., 2015), future research should compare the constituents of the happy expression between the two domains that contributes to stronger cross-modal emotion priming.

Importantly, there was no evidence of any difference between the ASD and NT groups in the magnitude of implicit emotion priming across emotions in the different prime-target conditions, though with one exception where the ASD group showed larger emotion priming of scared prosody on human faces relative to the NT group. These findings, thus, provide novel evidence for the absence of emotion-specific (nor emotion-general) impairments in the implicit emotion priming of prosody and song in autistic individuals.

4.4.5. Limitations

An important limitation of the present study is the uneven proportion of trials representing congruent versus incongruent prime-target relations. As an exploratory investigation of the role of specific emotion from multiple categories, the present study paired each emotional prime with targets expressing the four emotions of interest in turn. This resulted in each emotional prime being paired with a target of a congruent emotion once (e.g., angry-angry) but an incongruent emotion thrice (e.g., angry-scared, angry-happy, angry-sad). Thus, congruent prime-target pairs occurred at 25% while incongruent prime-target pairs occurred at 75% of the time throughout the experiment. The disproportionately lower occurrence of congruent trials could have reduced affective priming effects, as observed in previous priming studies (Klauer et al., 1997; Spruyt et al., 2007). It has been proposed that in incongruent trials, automatic stimulus-response route causes interference and becomes suppressed in the following trial, which may in turn leads to reduced facilitation in the subsequent congruent trials (Kornblum et al., 1990; Kunde, 2003). Such effects may, thus, become particularly problematic as the proportion of incongruent trials increases. One might suspect that the lack of priming effects observed for face-like object targets in the present study might have resulted from this manipulation. Priming effects were, nonetheless, found for face targets, suggesting the effects of congruence proportion could not be the sole explanation for this suspicion. It would be worthwhile for future studies to vary congruence proportions to examine the replicability of the present findings, specifically, to scrutinise whether the lack of group differences was limited by reduced overall priming effects in the NT group.

Despite all participants performing at above chance level across prime-target conditions, variability in the accuracy dataset was, nonetheless, noted. Outliers were particularly observed for object targets when primed by incongruent prosody and song. One may suspect poor recognition of object targets regardless of prime type is related to poor baseline performance (i.e., in the absence of priming). As demonstrated in Chapter 3 (Section 3.2.7), all participants scored above chance level in the simple emotion recognition task across stimulus types including objects. Moreover, this suspicion seems unlikely when cross-checking the baseline performance (data from Chapter 3) of the outlier scorers in the present study, who had all scored well at baseline (> 67%). Alternatively, the variability of data may be explained by interference effects of incongruent primes occurred for these participants, resulting in poor emotion recognition of the object targets – though it is unclear why this might have been the case for some participants. If indeed variability reflects interference effects of incongruent primes, priming effects measured for these participants would have been exacerbated, due to a larger congruency difference score. Nonetheless, a lack of priming effects on objects was observed across groups (Section 4.3.1.1), while the reverse would have been expected due to the exacerbated congruency difference score. This suggests the potential influence of data variability on the overall results may have been minimal.

The finding of larger priming effects in the ASD group relative to the NT group when faces were primed by prosody that conveyed fear was unexpected. On one hand, this finding may be interpreted in line with Sahuquillo-Leal et al. (2022), which found an atypical heightened sensitivity towards threat (i.e., revealed by the percentage of first fixations) in the ASD relative to the NT group during the processing of socio-emotional scenes. In essence, fear may be implicitly preactivated to a greater extent, and hence an increased facilitation of subsequent emotional judgment, in autistic compared to NT individuals in the present study. This interpretation, however, cannot account for the comparable priming observed for song conveying fear between the two groups. An alternative explanation may be that possible differences in the mechanisms may underlie the implicit processing of prosody between the ASD and NT groups. As shown in Section 3.2.4.1, although nonsignificant, fear had a relatively small size of pitch excursions compared to other emotions within the current set of prosody stimuli. Moreover, fear was the only emotion that differed in the size of pitch excursions between the present prosody and song stimulus sets – specifically, fear had a significantly smaller pitch excursion size for prosody than song stimuli. If pitch indeed plays a critical role in typical implicit emotion processing (L. Zhou et al., 2019), it is not surprising that the NT group showed weaker priming of prosody conveying fear (compared to happiness). This gives rise to the possibility that autistic individuals may rely on alternative acoustic cues than pitch during implicit emotion processing, as a result of compensation. The reliance on alternative cues may not have limited implicit processing of fear in prosody in autistic individuals even though pitch information is less salient, resulting in the observed group difference. As outlined in Section 1.2.1.1, an important objective to evaluate emotion processing at the implicit level in addition to the explicit level is to overcome influences of compensation, which may not have been achieved according to this hypothesis. The possible effects of compensation may have undermined differences to be observed between the ASD and NT groups in the present study, which may further explain why the present study did not show atypical implicit emotion processing as did previous neurophysiological studies (Fan & Cheng, 2014; Lindström et al., 2018). Future studies should incorporate neuroimaging methods to quantify the effects of compensation on implicit emotion processing more precisely, which will be informative in determining how differences in processing strategies lead to optimal behavioural performance in ASD.

4.4.6. Chapter summary

In summary, the present study found no differences in implicit emotion priming between autistic individuals relative to NT individuals as assessed on a cross-modal affective task, which is true for different types of auditory primes (i.e., speech prosody and song) and visual targets (i.e., human faces and face-like objects). Despite this, the two groups appeared to show different developmental patterns in implicit emotion processing across the auditory domains. Specifically, speech prosody but not song implicitly primed emotional judgment in human faces in NT individuals across development. This speech-orienting tendency during implicit emotion priming, however, appeared to decline with age in autistic individuals. Regardless of diagnostic group, priming was not observed between pairs of stimuli that had a low co-occurrence in the environment (i.e., song-object and prosody-object). Notably, cross-modal affective priming was found to be stronger in children compared to adults. This was possibly due to greater inhibitory control exerting over task-irrelevant information, and thus the influence of auditory primes may have been attenuated in adults. Despite the lower susceptibility to auditory interference in adults, priming was nevertheless observed for prosody in NT adults and for song in adults with ASD, further reinforcing the divergent preference during implicit emotion processing between the two groups. Interestingly, a priming advantage for happiness was observed across all participants, where the claims of emotion-specific impairments in ASD were not supported by the current data on implicit emotion processing.

Overall, it appears that cross-modal interactions may become more fine-tuned for interpersonal events (i.e., prosody-face) in emotional communication during development in NT, as shown by the greater tendency to orient to spoken over sung cues. Conversely, in line with the social motivation theory, the lack of speech-orientation during interpersonal events may underlie difficulties with social communication in ASD, particularly in situations where facial expressions are used to mask felt emotions which are available through the prosodic expressions. Altogether, these results outline the importance of the delicate weighting between speech- versus song-oriention at an early stage of emotion processing.

Chapter 5: Potential correlates of emotion processing in ASD and NT

This chapter aims to address a couple of discussion points raised in Chapter 3 (Section 3.4.2) with regards to whether individuals with autism spectrum disorder (ASD) and neurotypical (NT) individuals employ similar or different strategies for emotion processing across communicative domains, as well as issues discussed in Chapter 2 (Section 2.4.5) with regards to insufficient research considering the moderating effect of alexithymia on emotion processing. As an attempt to address these, this chapter examines the association between a number of potential correlates and emotion processing in ASD and typical development. These include (a) cognitive processing style, (b) musical training, musical perception abilities, and pitch perception abilities, and (c) autistic and alexithymic traits.

5.1. Introduction

In Chapter 3, the results showed that individuals with autism spectrum disorder (ASD) were able to recognise emotions from prototypical expressions as accurately as neurotypical (NT) individuals across different communicative domains; this was, however, accomplished with extended processing time. In addition, differences in the response patterns between the two groups were observed for human faces and face-like objects. In line with previous studies (Hobson et al., 1988; Tang et al., 2019; Weeks & Hobson, 1987; see also Harms et al., 2010 for a review), these findings highlight that although autistic individuals may show intact emotion recognition to some extent, there may still be differences in the underlying nature of these processes contributing to the different processing speed and response patterns observed between the two groups. Emotion processing encompasses sequential stages of extracting and integrating relevant features of emotional expressions, prior to deriving the emotional meaning of these expressions (Adolphs, 2002; Belin et al., 2004; Pell & Kotz, 2011; Schirmer & Kotz, 2006). It is, therefore, important to explore potential correlates of the proficiency in extracting these relevant properties and how they may differ between autistic and NT individuals. Such investigation will help us heighten understanding of the way autistic individuals process emotions, as well as the strengths and weaknesses that accompany successful recognition under various circumstances.

5.1.1. Cognitive processing style and visual emotion processing

As discussed in Chapter 3, the lack of efficiency in the decoding of emotional information coupled with different response patterns with human faces and face-like objects may be related to the use of alternative strategies to process these stimuli in autistic individuals. Previous literature has shown that whereas NT individuals typically employ global, configural-based strategy to process emotional faces, autistic individuals tend to use local, feature-based strategy to do so (Deruelle et al., 2004; Doi et al., 2013; Isomura, Ogawa, et al., 2014; Kätsyri et al., 2008; Pelphrey et al., 2002; Riby & Hancock, 2009; see Harms et al., 2010 for a detailed discussion). Because typical facial emotion recognition entails global processing, the findings revealed in Chapter 3 may be attributed to the peculiar perceptual processing strategy referred to as weak central coherence (WCC; Frith, 1989; Frith & Happé, 1994) or enhanced perceptual functioning (Mottron et al., 2006) in ASD. These two theoretical accounts, while not explicitly attempting to explain emotion recognition difficulties, proposed that the local, detail-focused processing style is a characteristic of ASD (Happé & Frith, 2006; Mottron et al., 2006). The two accounts, however, differ in their emphasis. The WCC account posits that the local processing style results from reduced top-down influences on global processing which undermines the integration of component parts into a coherent whole (U. Frith, 1989; U. Frith & Happé, 1994). In contrast, the EPF account suggests this local processing style results from overly enhanced low-level perception of local details, while global processing is intact but more optional in ASD (Mottron et al., 2006); the accelerated low-level perception, together with atypical relations between low and high levels of processing, may interrupt the processing of complex materials (Minshew et al., 1997).

Although facial emotion recognition impairments have previously been discussed in the context of local processing bias in ASD (Deruelle et al., 2004; Kätsyri et al., 2008; Pelphrey et al., 2002), only a limited number of studies have directly investigated whether facial emotion recognition was related to the more general perceptual processing style in ASD. In support of this hypothesis, Gross (2005) found that in comparison to NT children, children with ASD had greater difficulty recognising facial expressions of emotions, which was coupled with lower engagement in global processing on a global-local task. Across groups, and in children with ASD particularly, lower engagement in global processing was associated with lower accuracy in recognising emotional expressions in human and canine faces (Gross, 2005). These findings, thus, demonstrated a relationship between cognitive processing style and facial emotion recognition, where poorer facial emotion recognition may be explained by the tendency for local processing style in children with ASD. In another study by Oerlemans et al. (2013), although an association between local processing style and poorer facial emotion recognition was noted in the ASD proband group, it was not observed in the sibling group nor the NT group. Moreover, while children and adolescents with ASD showed impaired facial emotion recognition, this was coupled with an absence of group difference in local processing style (Oerlemans et al., 2013). Thus, despite the within-group association, impaired facial emotion recognition could not be attributed to differences in cognitive processing style. Crucially, correlations between performance by ASD probands and their siblings revealed that the two groups resembled each other significantly on facial emotion recognition but not local processing style (Oerlemans et al., 2013). The authors concluded that local processing style and social cognition (encompassing face processing and facial and prosodic emotion recognition) may be relatively dissociable cognitive constructs, given that local processing performance did not differ between groups and that it appeared not to be familial (Oerlemans et al., 2013). Taken together with these sparse and disparate findings, the relationship between cognitive processing style and facial emotion recognition in ASD remains unclear and more empirical evidence is needed in order to test this hypothesis.

Notably, these two studies examining the association between cognitive processing style and facial emotion recognition in ASD employed global-local tasks that measured different aspects of cognitive processing style. In Gross (2005), participants were presented with target stimuli consisting of a shape or a form made up of smaller objects (e.g., a star shape made from pennies) and then selected one of the three photographs that resembled the target stimulus the most: (a) an object conserving the configuration of the target (e.g., a star) representing a global response, (b) objects conserving the components of the target (e.g., pennies) representing a local response, (c) an object dissimilar to the target (e.g., a plastic fork) representing an unrelated response. The measurement of interest was the endorsement of global responses, which indicated participants' sensitivity to *global* processing on this task. In Oerlemans et al. (2013), participants were presented a target pattern of 3 red and 6 white squares in a 3 × 3 matrix and then detected this pattern in one of the four response options. Among these response options, the distractor patterns looked either (a) very similar to the target pattern where a correct/fast detection represented a local processing style or (b) very dissimilar to the target pattern where a correct/fast detection represented a global processing style. The measurement of interest was the difference between similar and dissimilar trials, with a larger difference representing a more local processing style on this task, which indicated participants' sensitivity to *local* processing. Based on these findings, it appears that sensitivity to global information (i.e., global advantage), as opposed to sensitivity to local information, may be a more sensitive measure that characterises differences in processing style between autistic and NT individuals and in demonstrating its relationship with facial emotion recognition. Another important measure that has also been shown to characterise differences in processing style between autistic and NT individuals is the tendency to be interfered by incongruent local details when processing global information, referred to as the local-to-global interference (Guy et al., 2019; Rinehart et al., 2000; L. Wang et al., 2007). The association between local-to-global interference and facial emotion recognition, however, has not been previously studied. The use of both measures, namely global advantage and local-to-global interference as potential correlates in the present study will shed light on whether slower and less efficient facial emotion recognition in ASD is related to the slower processing of global information that is not due to interference by local details (i.e., a lack of global advantage) and/or disrupted processing of global information due to interference by local details (i.e., higher susceptibility to local-to-global interference).

5.1.2. Musical training, musical perception abilities, pitch perception abilities, and auditory emotion processing

In Chapter 3, similar misattribution patterns were found between autistic and NT individuals when recognising emotions from speech prosody and song, indicating that the two groups may employ similar mechanisms to recognise emotions within the auditory modality. Despite this, slower and less efficient emotion recognition from speech prosody and song was also noted in ASD. The processing strategies employed by autistic individuals for auditory emotion processing have been less explored, and importantly, whether these strategies differ from those used by NT individuals is not well-known. A number of potential correlates of auditory emotion processing as demonstrated in previous research are, therefore, of particular interest, including musical training, musical perception abilities, and pitch perception abilities.

There is compelling evidence not only for the positive association between years of musical training and emotion recognition in music (Castro & Lima, 2014; Lima & Castro, 2011a; Livingstone et al., 2010), but also for the positive effects of musical training (i.e., comparisons between musicians vs. nonmusicians) on emotion recognition in speech prosody (Correia et al., 2020; Farmer et al., 2020; Lima & Castro, 2011b; W. F. Thompson et al., 2004; see Martins et al., 2021 for a review). Beyond the effects of musical training, naturally good musical abilities have also been found to be related to better emotion recognition of nonverbal vocalisations (e.g., laughter, crying) and speech prosody (Correia et al., 2020). For example, a positive association was found between musical perception abilities (i.e., an aggregated variable comprising measures of musical beat perception, pitch discrimination, and duration discrimination) and prosodic emotion recognition across participants, even when musical training was held constant (Correia et al., 2020). Notably, these positive effects of musical training and perception abilities on explicit emotion recognition appear to generalise across auditory domains, which is not surprising given their intricate overlap in the use of acoustic cues to convey emotion (Arbib, 2013; Juslin & Laukka, 2003; Patel, 2010). The association of musical training and perception abilities with implicit emotion processing across auditory domains are, however, less well-known, with a previous study showing that musical training does not seem to facilitate affective priming of musical stimuli (Steinbeis & Koelsch, 2011). This may indicate that the beneficial role of musical training on the explicit level of emotion processing may not extend to that on the implicit level.

Only one study to date has investigated the effects of musical training on auditory emotion recognition in autistic individuals (Quintin et al., 2011). It was found that musicianship was not a significant predictor of emotion recognition in instrumental musical excerpts, both across and within the ASD and NT groups (Quintin et al., 2011). However, the null finding might be due to the classification of musician versus non-musician groups, which was determined by whether participants had received at least two years of musical training. This classification appears to be a relatively low requirement compared with other studies investigating the effects of musical training (e.g., ≥ 13 years; Marques et al., 2007; Parbery-

Clark et al., 2009; Schön et al., 2004; W. F. Thompson et al., 2004). The effects of musical training and musical perception abilities on explicit emotion recognition in the speech prosody and vocal music domains, as well as implicit emotion processing across domains, have not yet been examined in ASD. Therefore, it would be worthy to further explore the relationships between musical training, musical perception abilities, and emotion processing of speech, as well as song, across implicit and explicit levels to gain better understanding of the underlying nature of auditory emotion processing not only in ASD, but also in typical development.

The cross-domain benefits of musical training and effects of musical perception ability on auditory emotion recognition skills may be a consequence of enhancements in aspects of low-level auditory perception (Kannyo & DeLong, 2011; Kraus & Chandrasekaran, 2010; Mankel et al., 2020; Pinheiro et al., 2015). Learning music has been found to gradually improve the perception of pitch (Besson et al., 2007; Magne et al., 2006; Marques et al., 2007; Micheyl et al., 2006; Moreno et al., 2009; Schön et al., 2004). Various features can be extracted from the pitch contour, such as slope, standard deviation, mean, and range, which serve as important cues for deducing emotional meaning from expressions in speech prosody and instrumental and vocal music (Hakanpää et al., 2019b; Hammerschmidt & Jürgens, 2007; Juslin & Laukka, 2003; Monnot et al., 2003). Thus, in order to successfully differentiate between emotional expressions in these domains, the ability to distinguish between high versus low and rising versus falling pitches is thought to be essential, which involves the recognition of large, as well as small, pitch fluctuations (Breitenstein et al., 2001; Lieberman & Michaels, 1962). It is, therefore, not surprising that pitch perception abilities have been found to be related to emotion recognition across auditory domains (Globerson et al., 2013, 2015; Gosselin et al., 2015; Lima et al., 2016).

The literature on pitch processing abilities in autistic individuals has been mixed, with previous studies showing enhanced pitch processing (Bonnel et al., 2003, 2010; Heaton, 2003;

Jiang et al., 2015), specifically in a subset of the ASD group (Heaton, Williams, et al., 2008; Jones et al., 2009), intact pitch processing (Altgassen et al., 2005; Chowdhury et al., 2017; Germain et al., 2019; Globerson et al., 2015; Jones et al., 2009; Schelinski & von Kriegstein, 2019), and impaired pitch processing (Bhatara, Babikian, et al., 2013; Jiang et al., 2015; Kargas et al., 2015; Schelinski & von Kriegstein, 2019; Sota et al., 2018). Furthermore, previous studies examining the relationship between pitch discrimination abilities and prosodic emotion recognition have also yielded inconsistent findings. In a previous study by Globerson et al. (2015), autistic individuals showed intact pitch discrimination and naming abilities for high-low judgment of non-vocal pitches, which were strongly associated with emotional prosody recognition. This association was, in fact, more pronounced in the ASD group compared to the NT group, suggesting that auditory perception abilities play a significant role in emotional prosody recognition in ASD (Globerson et al., 2015). By contrast, Schelinski & von Kriegstein (2019) found impaired pitch discrimination and naming for high-low judgment of vocal pitches in autistic individuals. Moreover, whereas emotional prosody recognition was correlated with vocal pitch perception abilities in NT individuals, no such correlation was observed in autistic individuals (Schelinski & von Kriegstein, 2019). It was postulated that vocal pitch information was perhaps not available for prosodic emotion recognition in autistic individuals to the same extent as it was in NT individuals (Schelinski & von Kriegstein, 2019). Research in this area remains scarce and the discrepant findings regarding the association between pitch perception and prosodic emotion recognition are not surprising, since pitch processing abilities in ASD appear to vary widely across studies. To my knowledge, no previous studies have investigated the association between pitch processing ability and explicit emotion recognition of song, as well as implicit emotion processing of speech prosody and song in both ASD and typical development.

The present study, thus, sought to explore whether and how musical training, musical perception abilities, and pitch perception abilities relate to auditory emotion processing across stimulus domains and levels of processing in autistic and NT individuals. It is also hoped that the investigation of these correlates will complement the findings observed in Chapter 3, for instance, through exploring whether comparable emotion recognition accuracy and response patterns, yet with slower speed, in the ASD group relative to the NT group is related to similarities or differences in auditory-related skills between the two groups.

5.1.3. Alexithymic traits, autistc traits, and emotion processing

As outlined in Chapter 1 (Section 1.3.3), previous research has debated whether impairments in emotion recognition is integral to ASD. Bird and Cook (2013) suggested that upon the strikingly inconsistent findings in the literature, two subgroups may be identified among the ASD population: one showing emotion recognition impairments with an ASD diagnosis, and the other showing no emotion recognition impairments but with an ASD diagnosis. Given that alexithymia can also be found in the NT population (Gündel et al., 2004; Jessimer & Markham, 1997; Prkachin et al., 2009), emotion recognition impairments in the former subgroup have been proposed to be a consequence of co-occurring alexithymia, rather than ASD per se (Bird & Cook, 2013).

Alexithymia, defined by a reduced ability to identify and describe one's own emotions, has been found to result in an impaired ability to recognise others' emotions (Connelly & Denney, 2007; Nemiah & Sifneos, 1970; Poquérusse et al., 2018; Suslow & Donges, 2017; G. J. Taylor & Bagby, 2000). Given its high incidence in ASD with a prevalence rate of approximately 50% (Berthoz & Hill, 2005; Milosavljevic et al., 2016), alexithymia has been brought into the scientific discussion as a plausible explanation for why some autistic individuals have difficulty with emotion recognition, while others perform at typical levels (Bird & Cook, 2013; see Sivathasan et al., 2020 for a recent review). Thus, this theory, while remaining controversial, provides an alternative perspective on why the literature reports highly conflicting results.

Several studies have investigated emotion recognition in autistic and NT individuals with varying degrees of alexithymia (Allen et al., 2013; Cook et al., 2013; Heaton et al., 2012; Keating et al., 2021). While a number of studies have identified alexithymia as accountable for impairments in emotion recognition across facial, vocal, and musical domains in ASD (Allen et al., 2013; Cook et al., 2013; Heaton et al., 2012; Ketelaars et al., 2016; Milosavljevic et al., 2016; Ola & Gullon-Scott, 2020), there is also evidence suggesting alexithymia could not account for emotion recognition impairments in ASD (Keating et al., 2021; Kliemann et al., 2013). Being a transdiagnostic factor, it has been recommended that alexithymia should be considered within psychological research, in order to reduce the conflation of behaviours caused by alexithymia and ASD (Hickman, 2019). One of the aims of the present study was, therefore, to examine the relationship between alexithymia and emotion processing across individuals with and without ASD. In particular, whether co-occurring alexithymia is related to explicit emotion recognition (which was found to be slower and less efficient in the ASD group in Chapter 3) but not implicit emotion priming (which was found to be largely intact in the ASD group in Chapter 4) will be examined.

In addition to examining the potential contribution of alexithymia to emotion processing, the present study evaluated the association between severity of ASD (i.e., autistic traits) and emotion processing. As noted earlier, the ability to recognise emotions seems to vary substantially among individuals with an ASD diagnosis (Bird & Cook, 2013). Since autism is known as a spectrum disorder, it is possible that the presence or the degree of impairment in emotion recognition may be influenced by the divergent symptomatology at the individual level within this group (Happé et al., 2006). Like alexithymia, behaviours associated with ASD can also be found in the NT population (Grinter et al., 2009; Ingersoll, 2010), where milder versions of the traits typical of autism are more common among NT individuals (Baron-Cohen et al., 2001; Ruzich et al., 2015). Previous research has demonstrated the effects of autistic traits on emotion recognition, with higher autistic traits being associated with poorer emotion recognition within and across the ASD and NT populations (McKenzie et al., 2018; Poljac, Poljac, & Wagemans, 2013; Wallace et al., 2011; see Trevisan & Birmingham, 2016 for a review). Thus, impairments in emotion recognition could be driven by the higher autistic traits among autistic individuals. The inclusion of this additional construct will help to elucidate whether emotion processing is better explained by severity of ASD or co-occurring alexithymia.

5.1.4. The present study

The present study sought to expand the lens in order to consider several potential correlates and other, possibly unnoticed, variables that may influence emotion processing across visual and auditory modalities in ASD and NT (as investigated in Chapters 3 and 4). Through examining these correlates, the present study aims to address the following specific research questions:

1. Is there a relationship between cognitive processing style (indexed by global advantage and local-to-global interference) and emotion recognition of human faces and face-like objects in ASD and NT groups? Specifically, could differences in emotion recognition of these stimuli be explained by differences in cognitive processing style between the two groups?
2. Is there a relationship between years of musical training, musical perception abilities, pitch perception abilities, and emotion recognition and priming of speech prosody and song in ASD and NT groups? Specifically, could similarities/differences in emotion recognition of these stimuli between the two groups be explained by these variables?

3. Are alexithymia and autistic traits important predictors of emotion recognition and priming? Specifically, could atypicalities in emotion processing in ASD be attributed to co-occurring alexithymia?

First, the relationship between visual emotion recognition of human faces and face-like objects and cognitive processing style was examined. As discussed earlier, aside from the global advantage measure, the association between local-to-global interference and visual emotion recognition in ASD relative to typical development is not well-known. It is anticipated that a global processing style (i.e., more global advantage and lower local-to-global interference) will be associated with better emotion recognition of human faces and face-like objects, given the importance of global, configural-based strategy for processing these stimuli (Ichikawa et al., 2011; Leder & Carbon, 2006). Because global processing appears to take longer and require more effort in autistic compared to NT individuals (Van der Hallen et al., 2015), it is anticipated that the more effortful global processing will be associated with poorer emotion recognition of human faces and face-like objects in autistic individuals.

Secondly, the relationship between emotion recognition and priming of speech prosody and song and the variables relating to auditory-related skills was examined. These included years of musical training, musical perception abilities, and pitch perception abilities. Given that autistic and NT participants appeared to draw on similar auditory-related mechanisms (as demonstrated by their comparable misattribution patterns in Chapter 3), it is hypothesised that more musical training and better musical and pitch perception abilities will be associated with better emotion recognition of speech and song in both the ASD and NT groups. Drawing on previous evidence reporting impaired affective priming of musical chords in individuals with congenital amusia – a condition defined by impaired musical and pitch perception (L. Zhou et al., 2019), it is hypothesised that more musical training and better musical and pitch perception abilities will also be associated with stronger emotional priming of speech and song in both the ASD and NT groups.

Finally, the contribution of alexithymic and autistic traits to overall emotion recognition and priming was examined. Based on previous literature as discussed in Section 5.1.3, it is hypothesised that while both alexithymic and autistic traits will predict emotion recognition and priming. That is, higher alexithymic and autistic traits will be associated with poorer emotion recognition and weaker emotional priming. In accordance with the alexithymia hypothesis (Bird & Cook, 2013; Cook et al., 2013), the predictive power of alexithymic traits will be over and above that of autistic traits, giving rise to the hypothesis of poorer emotion processing in ASD being explained by co-occurring alexithymia.

5.2. Methods

5.2.1. Participants

The same participants who took part in the emotion recognition and priming experiments also completed a battery of background tasks critical for the investigation of potential correlates of emotion processing in this chapter. This includes 76 native English speakers, with 38 participants with ASD and 38 NT participants matched on chronological age, gender, receptive vocabulary (ROWPVT-4; Martin & Brownell, 2011), and nonverbal reasoning ability (RSPM; Raven, 1983) (see Section 3.2.1 for full details of participant recruitment, eligibility, and demographics).

5.2.2. Design

This study used a correlational design to address the proposed research questions. To address the first research question, variables including the global advantage effect and the local-to-global interference effect measured on the Navon task as an index of cognitive processing style were correlated with performance accuracy, response time (RT), and efficiency for human faces and face-like objects measured on the emotion recognition task in Chapter 3. The Navon's classical paradigm (Navon, 1977) was used as it enables both aspects of cognitive processing style to be assessed. Under this paradigm, participants are presented with compound stimuli (i.e., large letters made up of smaller letters) that consist of congruent or incongruent information across local and global levels. Participants are then asked to detect targets at either the local or global level, while ignoring the other level. The global advantage (i.e., when reactions to global targets are faster than reactions to local targets) and local-to-global interference (i.e., when reactions to global targets are slowed down due to interfering local targets) can both, therefore, be assessed on this task (see Pletzer et al., 2014; L. Wang et al., 2007).

Regarding the second research question, variables including years of musical training, global score measured on the Montreal Battery Evaluation of Amusia (MBEA; for adult participants) or the Montreal Battery of Evaluation of Musical Abilities (MBEMA; for child/adolescent participants) as an index of musical perception ability, and pitch direction discrimination threshold as an index of pitch perception ability were correlated with performance accuracy, RT, and efficiency for speech prosody and song measured on the emotion recognition task in Chapter 3 and congruency mean accuracy, RT, and efficiency difference score for speech prosody and song as an index of priming effects on the cross-modal affective priming task in Chapter 4. In terms of pitch perception abilities, while previous studies examining the relationship of pitch perception abilities and prosodic emotion recognition in ASD employed pitch discrimination naming tasks that targeted high-low judgment of discrete pitches (Globerson et al., 2015; Schelinski & von Kriegstein, 2019), the present study assessed participants' pitch discrimination naming abilities for ascending-descending judgment of gliding pitches, given that emotional connotation can vary depending on rising versus falling pitch contours (Bänziger & Scherer, 2005; Kalinli, 2016; Nwe et al., 2003).

In response to the third research question, the independent variables were scores on the 20-Item Toronto Alexithymia Scale (TAS-20) and the Autism Spectrum Quotient (AQ) as indices of alexithymic and autistic traits, respectively. The dependent variables were overall performance accuracy, RT, and efficiency on the emotion recognition task in Chapter 3 and congruency mean accuracy, RT, and efficiency difference score across conditions as an index of overall priming effects on the cross-modal affective priming task in Chapter 4.

5.2.3. Emotion measures

5.2.3.1. Emotion recognition

The emotion recognition task assessed participants' recognition of four basic emotions (angry, scared, happy, and sad) across four stimulus types (human face, face-like objects, speech prosody, and song) – see Section 3.2.3 for full details on stimulus development and validation. There was a total of 320 trials blocked by stimulus type, with 32 trials in the song, 32 in the prosody, 128 in the face, and 128 in the face-like object conditions. Two versions of pseudo-randomisation were adopted and counterbalanced between participants, with each stimulus presenting twice in the experiment. Two practice trials preceded the start of each condition. On each trial, an auditory stimulus (prosody or song) or a visual stimulus (face or face-like object) was presented through the headphones and on the screen in front, respectively. Participants were instructed to decide as quickly and accurately as possible which of the four emotion labels (angry, scared, happy, or sad) best described the emotion presented in the stimulus. Responses were made by pressing the corresponding key of the chosen emotion on the respond pad. There were no limits to response times and no feedback was given regarding the accuracy of judgment in each trial. The presentation of the stimuli was terminated as soon as a response was made. Mean accuracy, mean response time (RT; RTs less than 150ms or more than 2.5 SD of the mean of each participant for each condition were excluded), and mean speed-accuracy composite score (SACS; a composite measure with accuracy and RT combined denoting recognition efficiency) were computed for each participant by condition – see Section 3.2.6 for full details of these calculations.

5.2.3.2. Cross-modal affective priming

The cross-modal affective priming task assessed participants' emotion recognition in the visual targets (human faces or face-like objects) after hearing an emotionally congruent or incongruent auditory prime (speech prosody or song) following a short delay. The same set of stimuli from the emotion recognition task was used. Each auditory prime was paired with a visual target from each of the four emotion categories to create congruent (e.g., angry-angry) and incongruent (e.g., angry-scared, angry-happy, and angry-sad) prime-target pairs. This resulted in a total of 256 trials blocked by target type, with 16 congruent and 48 incongruent trials for each prime-target condition (song-face, song-object, prosody-face, and prosody-object). Two versions of pseudo-randomisation were adopted and counterbalanced between participants – see full details of pseudo-randomisation process in Section 4.2.4. Four practice trials preceded the start of each block. On each trial, a visual target (face or face-like object) appeared on the screen 200ms following the onset of the auditory prime (prosody or song) presented through the headphones. Participants were instructed to decide as fast and accurately as possible the emotion label (angry, scared, happy, or sad) that best described the expression presented in the visual target. The target stimulus remained on the screen until a response was made on the response pad by pressing the corresponding key of the chosen emotion. The presentation of the auditory and visual stimuli terminated as soon as the participant responded. A congruency mean difference score for accuracy, RT (RTs less than 150ms or more than 2.5 SD of the mean of each participant were excluded), and SACS were computed by prime type (i.e., a score for prosody and song, respectively) and across prime and target types (i.e., overall score). Each congruency mean difference score represented the difference between performance on the congruent versus incongruent conditions, with a larger score indicating greater priming effects – see Section 4.2.5 for full details of these calculations. The congruency mean difference score for each performance index will be respectively referred to emotion priming accuracy, RT, and SACS hereafter for ease of reading.

5.2.4. Cognitive measures

5.2.4.1. Navon

The Navon task assessed participants' cognitive processing style (Navon, 1977) using PsychoPy (Peirce, 2007). Participants were presented with composite letters that comprised of small H or S characters embedded within large H or S characters. The embedded small characters were either congruent (e.g., small Hs within a large H) or incongruent (e.g., small Ss within a large H) with the large characters. During the global processing condition, participants were instructed to respond to the large letters; during the local processing condition, participants were instructed to respond to the small letters. Participants completed a total of 320 trials, with 160 per condition, of which 80 were congruent. Cognitive processing style was indexed by the global advantage and local-to-global interference. Global advantage was calculated by subtracting the RT to global judgment from the RT to local judgment on consistent trials [local consistent RT – global consistent RT]; a larger difference characterises a global advantage. The local-to-global interference was calculated by subtracting the RT to global judgment on inconsistent trials from consistent trials [global inconsistent RT – global consistent RT); a larger difference indicates a higher local-to-global interference as local information dominates over global information.

5.2.5. Auditory measures

5.2.5.1. Musical background survey

A musical background questionnaire was administered as part of the demographic questionnaire. Details about the type(s) of instrument studied (including voice) and the number of years of formal musical training for each instrument learnt were obtained. The years of formal musical training were then summed across all instruments (including voice) offline (Pfordresher & Halpern, 2013).

5.2.5.2. The Montreal Battery of Evaluation of Amusia (MBEA) and The Montreal Battery of Evaluation of Musical Abilities (MBEMA)

The Montreal Battery Evaluation of Amusia (MBEA) and the Montreal Battery of Evaluation of Musical Abilities (MBEMA; an adapted version of MBEA) were administered to adult and child/adolescent participants respectively to assess participants' musical perception and memory (Peretz et al., 2003, 2013). MBEA comprised a total of six subtests, including contour, interval, scale, rhythm, meter, and memory. Each subtest consisted of 2 practice trials (3 on the meter subtest) and 30 experimental trials. MBEMA comprised a total of five subtests, including contour, interval, scale, rhythm, and memory. Each subtest consisted of 2 practice trials and 20 experimental trials. In the first four subtests, participants were presented with pairs of identical or non-identical melodies with respect to their melodic or rhythmic organisation. Participants were instructed to classify whether the melodies within the pair sounded the same or different. In the meter subtest (specific to MBEA), participants were presented with the harmonised melodies, where the melodies were accompanied by chords to outline their binary or ternary structures. Participants were instructed to classify whether the melodies as either a march (i.e., duple meter) or a waltz (i.e., triple meter). In the memory recognition subtest, participants were presented half of the time with melodies which already occurred in the previous subtests, and novel melodies the other half of the time. Participants were instructed to indicate for each melody whether they had heard it before during the previous subtests or whether it was new. A global score was calculated from the performance across all subtests as a measure of general musical perception ability.

5.2.5.3. Pitch direction discrimination

The pitch direction discrimination task was taken from previous studies (F. Liu et al., 2010, 2012). On each trial, participants heard three gliding pitches, with two gliding in the same direction and the other gliding in the opposite direction. Each gliding pitch was 600ms in duration. Participants were required to identify the “odd-one-out” target of the three pitches, which always appeared in the first or final position of the sequence. An adaptive-tracking procedure with a two-down-one-up staircase method and a variable change in step size was used. Starting with six semitones as the excursion size of the first gliding tone, the step size decreased by one semitone, then by 0.1 semitone after four reversals and 0.02 semitones after eight reversals. The threshold was calculated as the mean excursion value of the target glide of the last six reversals, with a lower threshold (i.e., a score close to 0) indicating optimal performance (i.e., higher sensitivity to changes in pitch direction).

5.2.6. Psychometric measures

5.2.6.1. The 20-Item Toronto Alexithymia Scale (TAS-20)

The 20-Item Toronto Alexithymia Scale (TAS-20; Bagby et al., 1994) was used as a measure of alexithymic traits in adult participants only, as its implementation in younger participants is not recommended (J. D. A. Parker et al., 2010). The TAS-20 consisted of three subscales: difficulties describing feelings, difficulty identifying feeling, externally oriented thinking (i.e., tendency to focus attention externally). This questionnaire comprised 20 items that were rated using a 5-point Likert scale, with 1 representing strongly disagree and 5 representing strongly agree. A score of ≤ 51 is considered indicative of non-alexithymia, 52-60 indicates possible alexithymia, and ≥ 61 indicates alexithymia.

5.2.6.2. The Autism Spectrum Quotient Test (AQ)

The Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001) was used as a measure of autistic traits in adult participants. The questionnaire comprised 50 individual questions that make up five subcategories of autistic traits: social skill, attention switching, attention to detail, communication, and imagination. Response options ranged from “definitely agree”, “slightly agree”, “slightly disagree” to “definitely disagree”. Approximately half the questions were worded to elicit an “agree” response from NT individuals, and half to elicit a “disagree” response. Each of the items scored 1 point if the response corresponding autistic-like behaviour either mildly or strongly was chosen. Items were reverse scored as necessary. Higher scores indicated more symptoms of autism. A score of 32 indicated substantial autistic symptoms, with higher scores corresponding to more “autistic-like” behaviour. Note that the AQ was administered to all participants across the three age groups (see Table 3.1 in Section 3.2.1 for summary). However, since one of the main aims of the present study was to compare the role of autistic traits and alexithymic traits in emotion processing to that of alexithymic traits, only the AQ scores for adult participants were used in this study as a comparison with the TAS-20 scores, which were only obtained from adult participants as mentioned above.

5.2.7. Procedure

Prior to the experimental tasks, adult participants or caregivers of child and adolescent participants completed the demographic and music background questionnaire, TAS-20, and AQ. The emotion, cognitive, and pitch tasks were administered to each participant in a random order as part of the full test battery, either during one single session or across multiple sessions.

Participants were allowed to take multiple breaks as needed in between tasks. All tasks were conducted in a sound-proof booth.

5.2.8. Statistical analyses

Statistical analyses were conducted in R (R Core Team, 2019). Separate multiple linear regression models were constructed to assess the main effects of Diagnostic group (ASD vs. NT, dichotomously dummy-coded as 1 or 0) and Chronological age, as well as their interaction, on each of the background measures, namely Navon global advantage, Navon local-to-global interference effect, years of musical training, MBEA/MBEMA global score, and pitch direction discrimination threshold. As the TAS-20 and AQ were only administered to adult participants, the difference in alexithymic traits and autistic traits between adults with ASD and NT adults was assessed using an independent-samples *t*-test.

A series of Kendall's rank correlations and partial correlations (controlling for age) were conducted within and across the ASD and NT groups to examine the correlations between the cognitive/auditory measures (Navon global advantage, Navon local-to-global interference effect, years of musical training, MBEA/MBEMA global score) and performance on the emotion tasks (emotion recognition and priming accuracy, RT, and SACS). To examine the contribution of alexithymic traits and autistic traits to emotion processing, a series of multiple linear regression models were conducted with the total scores on the TAS-20 and AQ as predictors and performance on the emotion tasks (emotion recognition and priming accuracy, RT, and SACS) as dependent measures, within and across the ASD and NT groups. Note that the mean accuracy and RT of emotion recognition performance were arcsine- and log-transformed, respectively, for all relevant analyses.

For effect sizes of each predictor in the multiple linear regression models, partial eta-squared (η^2_p) was computed using the *modelEffectSizes*() function in the *lmSupport* package (Ben-Shachar et al., 2020; Curtin, 2018). A $\eta^2_p \geq 0.01$ was interpreted as a small effect size, a

$\eta^2_p \geq 0.09$ as a medium effect size, and a $\eta^2_p \geq 0.25$ as a large effect size (Cohen et al., 2013). The strength of correlations, denoted as Kendall's tau coefficient (τ), was interpreted with < 0.10 as a very weak correlation, 0.10-0.19 as a weak correlation, 0.20-0.29 as a moderate correlation, and > 0.30 as a strong correlation (Botsch, 2011).

5.3. Results

Summary statistics for each background measure are presented in Table 5.1. While all participants completed the main experiments, it should be noted that there are missing data for some of the background measures due to various reasons (e.g., scheduling constraints and technical issues).

Table 5. 1. Characteristics of the performance on each background task for the ASD and NT groups by age group.

	ASD			**NT**				
	N (M: F)	**M**	**SD**	**N (M: F)**	**M**	**SD**	***t***	***p***
Child								
Navon global advantage (ms)	14 (13:1)	427.98	1296.04	14 (13:1)	39.54	183.40	1.11	0.286
Navon local-to-global interference (ms)	14 (13:1)	141.57	279.51	14 (13:1)	-18.05	129.71	1.94	0.068
Musical training (years)	14 (13:1)	1.18	2.14	14 (13:1)	1.29	1.14	-0.17	0.871
MBEA/MBEMA global score	14 (13:1)	0.75	0.14	14 (13:1)	0.81	0.14	-1.10	0.283
Pitch direction discrimination threshold	11 (10:1)	1.32	0.83	11 (11:0)	1.25	1.92	0.11	0.912
Adolescent								
Navon global advantage (ms)	11 (8:3)	115.45	91.16	11 (8:3)	119.50	71.94	-0.12	0.909
Navon local-to-global interference (ms)	11 (8:3)	36.99	80.86	11 (8:3)	53.29	61.09	-0.53	0.600
Musical training (years)	11 (8:3)	3.09	3.08	11 (8:3)	2.36	3.07	0.55	0.586
MBEA/MBEMA global score	11 (8:3)	0.85	0.13	9 (7:2)	0.89	0.08	-0.95	0.355
Pitch direction discrimination threshold	7 (6:1)	0.48	0.48	11 (8:3)	0.45	0.44	0.09	0.926
Adult								
Navon global advantage (ms)	13 (6:7)	86.23	71.41	13 (6:7)	125.83	40.49	-1.74	0.098
Navon local-to-global interference (ms)	13 (6:7)	27.61	47.07	13 (6:7)	12.30	33.42	0.96	0.350
Musical training (years)	13 (6:7)	5.92	8.47	13 (6:7)	6.96	7.53	-0.33	0.744
MBEA/MBEMA global score	13 (6:7)	0.83	0.12	12 (5:7)	0.86	0.08	-0.78	0.446
Pitch direction discrimination threshold	9 (4:5)	0.37	0.41	12 (6:6)	0.23	0.22	0.91	0.381
TAS-20	**13 (6:7)**	**69.85**	**9.76**	**13 (6:7)**	**45.54**	**16.51**	**4.57**	**< 0.001**
AQ	**13 (6:7)**	**41.46**	**4.29**	**13 (6:7)**	**14.31**	**6.29**	**12.86**	**< 0.001**

Note. MBEMA = Montreal Battery of Evaluation of Musical Ability (Peretz et al., 2003); MBEA = Montreal Battery of Evaluation of Amusia (Peretz et al., 2013); AQ = Autism-Spectrum Quotient (Baron-Cohen et al., 2001); TAS-20 = The 20-Item Toronto Alexithymia Scale (Bagby et al., 1994).

Missing data from participants include 2 adolescents (1 male) and 1 adult (1 male) in the NT group for the MBEA global score; 3 children (all males), 4 adolescents (2 males), and 4 adults (2 males) in the ASD group and 3 children (2 males) and 1 adult (0 males) in the NT group on the pitch direction discrimination task; 1 child (1 male) in the ASD and 1 child (1male) in the NT group on the go/no go task.

Welsh two-sample *t*-tests were used to compare differences between the ASD and NT groups by age group for each background task; significant differences between ASD and NT groups are highlighted using bold font.

5.3.1. Correlations between cognitive processing style and visual emotion processing

Table 5. 2. Multiple linear regression model results for diagnostic group, chronological age, and their interaction on global advantage and local-to-global interference measured on the Navon task.

Predictors	B	SE	df	*t*	*p*	η^2_p
Global advantage						
Diagnostic group	-334.12	221.00	72	-1.51	0.135	0.03
Chronological age	-8.51	6.42	72	-1.33	0.189	0.02
Diagnostic group × chronological age	10.42	9.07	72	1.15	0.255	0.02
Local-to-global interference						
Diagnostic group	-106.71	56.17	72	-1.90	0.061	0.05
Chronological age	-2.10	1.63	72	-1.29	0.202	0.02
Diagnostic group × chronological age	2.40	2.31	72	1.04	0.301	0.01

Table 5. 3. Kendall's rank correlations and partial correlations (controlling for age) between performance on visual emotion recognition tasks and measures of cognitive processing style on the Navon task for ASD, NT and across groups, based on 38 ASD and 38 NT participants.

	Global advantage			Local-to-global interference		
	ASD	NT	All	ASD	NT	All
Without controlling for age						
Face recognition Accuracy	-0.02	0.08	0.03	-0.15	-0.04	-0.09
Face recognition RT	0.07	0.04	0.07	0.08	0.21	0.13
Face recognition SACS	-0.01	-0.02	0.00	-0.14	-0.12	-0.13
Object recognition Accuracy	-0.07	-0.06	-0.05	-0.08	0.06	-0.01
Object recognition RT	-0.05	-0.08	-0.06	0.09	0.04	0.09
Object recognition SACS	0.06	-0.05	0.01	-0.09	0.00	-0.07
After controlling for age						
Face recognition Accuracy	0.03	0.00	0.02	-0.13	-0.06	-0.08
Face recognition RT	0.04	0.11	0.08	0.07	**0.23***	0.12
Face recognition SACS	0.05	-0.13	-0.01	-0.12	-0.15	-0.12
Object recognition Accuracy	-0.07	-0.11	-0.06	-0.08	0.05	0.00
Object recognition RT	-0.09	0.01	-0.05	0.07	0.06	0.09
Object recognition SACS	0.10	-0.14	0.00	-0.07	-0.02	-0.06

Note. Face recognition accuracy was arcsine-transformed; face recognition RT was log-transformed; *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant correlations are highlighted using bold font.

Global advantage

Regression analysis on the Navon global advantage revealed no significant predictors of Diagnostic group, Chronological age, and their interaction (see Table 5.2 for full summary). This indicates that diagnostic group and participants of different ages did not differ in their cognitive processing style for global versus local information.

Correlational analyses showed no significant correlations between global advantage and emotion recognition with faces and objects across measures for both across and within the ASD and NT groups, with and without controlling for age (see Table 5.3). This suggests that an advantage for global processing was not associated with emotion recognition across visual stimuli.

Local-to-global interference

Regression analysis on the Navon local-to-global interference effect revealed no significant predictors of Diagnostic group, Chronological age, and their interaction (see Table 5.2 for full summary). This indicates that diagnostic group and participants' age did not predict their susceptibility to interference by local details when making a global judgment.

Across ASD and NT groups, no significant correlations between local-to-global interference and emotion recognition of faces and objects across measures were observed, both with and without controlling for age (Table 5.3).

Within the NT group, a significant positive correlation was found between local-to-global interference and emotion recognition RT for faces after controlling for age ($\tau = 0.23$, $p = 0.049$). No other correlations regarding local-to-global interference were significant for the NT group, including emotion recognition accuracy and SACS for faces and emotion recognition for objects across performance indices (see Table 5.3).

Within the ASD group, no significant correlations were found between local-to-global interference and emotion recognition from faces and objects across all measures.

Together, the results of the correlational analyses suggest that higher local-to-global interference (i.e., higher susceptibility to interference by local information when processing global information) was associated with slower emotion recognition of faces in NT individuals. This association was, however, not evident in autistic individuals (Figure 5.1). Notably, no correlations between local-to-global interference and emotion recognition in objects were found for both within and across groups.

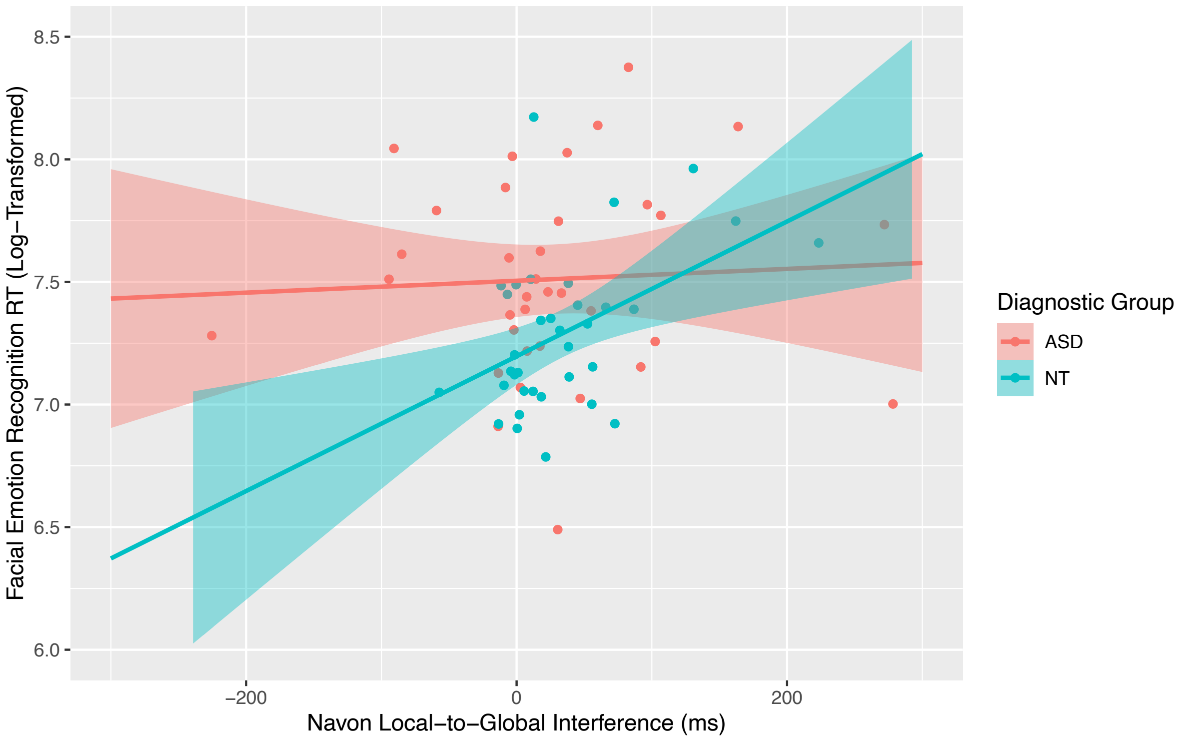


Figure 5. 1. Scatterplot of Navon local-to-global interference (ms) against facial emotion recognition RT (log-transformed). Regression lines were based on linear regressions of paired tasks for each diagnostic group.

5.3.2. Correlations between musical processing, pitch processing, and auditory emotion processing

Table 5. 4. Multiple linear regression model results for diagnostic group, chronological age, and their interaction on years of musical training, musical perception assessed on the MBEA/MBEMA tasks, and pitch perception assessed on the pitch direction discrimination task.

Predictors	B	SE	df	*t*	*p*	η^2_p
Musical training						
Diagnostic group	0.56	1.90	72	0.29	0.772	0.00
Chronological age	**0.19**	**0.06**	**72**	**3.40**	**0.001**	**0.14**
Diagnostic group × chronological age	-0.02	0.08	72	-0.24	0.811	0.00
Musical perception						
Diagnostic group	0.07	0.05	69	1.50	0.139	0.03
Chronological age	0.00	0.00	69	1.61	0.112	0.04
Diagnostic group × chronological age	0.00	0.00	69	-0.77	0.443	0.01
Pitch perception						
Diagnostic group	-0.18	0.43	57	-0.43	0.670	0.00
Chronological age	-0.02	0.01	57	-1.78	0.081	0.05
Diagnostic group × chronological age	0.00	0.02	57	0.08	0.935	0.00

Note. Significant predictors are highlighted using bold font.

Table 5. 5. Kendall’s rank correlations and partial correlations (controlling for age) between performance on auditory emotion recognition and priming tasks and musical processing abilities indexed by the number of years of musical training and performance on the MBEA/MBEMA task for ASD, NT and across groups, based on 38 ASD and 38 NT participants.

	Musical training (years)			**MBEA/MBEMA global score**			**Pitch direction discrimination threshold**		
	ASD	**NT**	**All**	**ASD**	**NT**	**All**	**ASD**	**NT**	**All**
Without controlling for age									
Prosody recognition Accuracy	0.20	0.22	**0.21***	0.20	0.14	**0.20***	-0.17	**-0.39 ****	**-0.34*****
Prosody recognition RT	0.01	-0.10	-0.06	-0.06	0.11	-0.04	0.02	0.12	0.12
Prosody recognition SACS	0.16	0.23	**0.19***	0.22	0.07	**0.17***	-0.18	**-0.37****	**-0.33*****
Song recognition Accuracy	0.16	0.01	0.08	0.11	0.22	0.15	-0.09	**-0.30***	**-0.28****
Song recognition RT	0.05	-0.02	0.00	-0.12	0.01	-0.10	0.20	0.04	0.13
Song recognition SACS	0.08	0.03	0.07	0.11	0.18	0.14	-0.20	**-0.36****	**-0.31*****
Prosody priming Accuracy	-0.18	0.09	-0.01	0.03	0.04	0.07	0.06	-0.10	-0.06
Prosody priming RT	-0.05	-0.17	-0.10	-0.08	0.04	-0.03	0.11	0.06	0.08
Prosody priming SACS	-0.20	0.04	-0.04	-0.02	0.05	0.06	0.10	-0.07	-0.06
Song priming Accuracy	-0.20	0.10	-0.03	-0.06	0.00	0.03	0.16	-0.15	-0.04
Song priming RT	-0.18	-0.07	-0.13	-0.06	0.08	-0.02	0.17	0.04	0.14
Song priming SACS	-0.21	0.07	-0.07	0.00	0.07	0.04	0.23	-0.15	0.01
After controlling for age									
Prosody recognition Accuracy	0.15	0.17	**0.17***	0.14	0.12	0.15	0.10	**-0.34****	**-0.27****
Prosody recognition RT	0.00	-0.07	-0.05	-0.08	0.13	-0.04	0.00	0.08	0.11
Prosody recognition SACS	0.13	0.18	**0.15***	0.20	0.04	0.13	-0.12	**-0.31***	**-0.26****
Song recognition Accuracy	0.06	-0.04	0.02	0.00	0.21	0.10	0.03	**-0.25***	**-0.21***
Song recognition RT	0.07	0.01	0.02	-0.11	0.02	-0.08	0.18	0.00	0.10
Song recognition SACS	0.00	-0.02	0.02	0.03	0.16	0.09	-0.09	**-0.31***	**-0.24****
Prosody priming Accuracy	-0.15	0.09	-0.01	0.08	0.03	0.07	-0.03	-0.09	-0.08
Prosody priming RT	-0.02	-0.16	-0.09	-0.05	0.04	-0.02	0.06	0.04	0.04
Prosody priming SACS	-0.17	0.04	-0.03	0.03	0.05	0.07	0.02	-0.08	-0.08
Song priming Accuracy	-0.18	0.10	-0.02	-0.03	0.00	0.04	0.12	-0.17	-0.07
Song priming RT	-0.20	-0.07	-0.13	-0.07	0.08	-0.03	0.17	0.04	0.13
Song priming SACS	-0.21	0.08	-0.06	0.01	0.08	0.04	0.21	-0.18	-0.02

Note. Prosody and song recognition accuracy were arcsine-transformed; prosody and song recognition RT were log-transformed; *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; significant correlations are highlighted using bold font.

Musical training

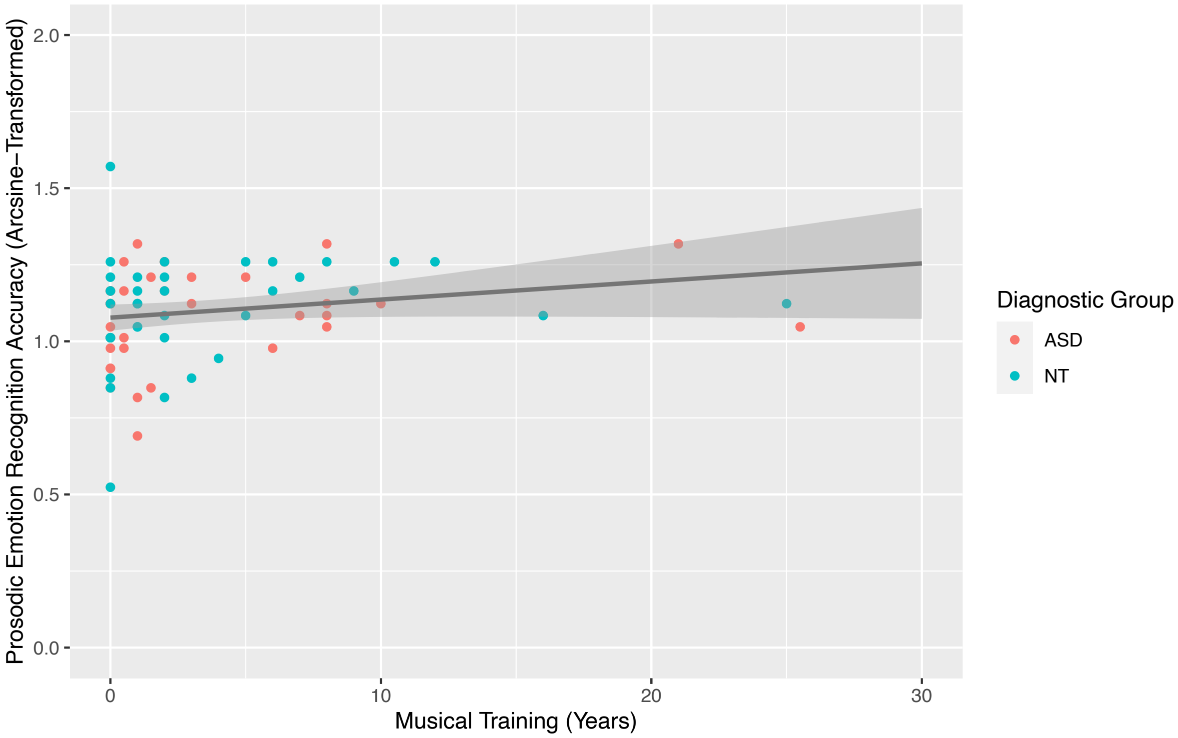
Regression analysis on the years of musical training revealed a significant predictor of Chronological age ($B = 0.19$, SE = 0.06, $t(72) = 3.40$, $p = .001$, $\eta^2_p = 0.14$). This suggests that increased years of musical training was predicted by increased age of participants. The predictors of Diagnostic group and Diagnostic group × Chronological age on years of musical training did not reach significance (see Table 5.4 for full summary).

Across ASD and NT groups, correlational analyses showed significant weak to moderate correlations between years of musical training and emotion recognition accuracy and SACS of prosody (musical training vs. prosody recognition accuracy: $\tau = 0.21$, $p = 0.013$;

musical training vs. prosody recognition SACS: τ = 0.19, p = 0.020), which remained significant after controlling for age (musical training vs. prosody recognition accuracy: τ = 0.17, p = 0.032; musical training vs. prosody recognition SACS: τ = 0.15, p = 0.050). These results suggest that more musical training was associated with more accurate and efficient emotion recognition in speech prosody (Figure 5.2). No other correlations regarding years of musical training were significant across groups, including emotion recognition RT for prosody, emotion recognition in song across performance indices and emotion priming of prosody and song across performance indices (see Table 5.5).

Within the ASD and NT groups, no significant correlations were observed between years of musical training and emotion recognition and priming of prosody and song across performance indices, with and without controlling for age (see Table 5.3).

A)



B)

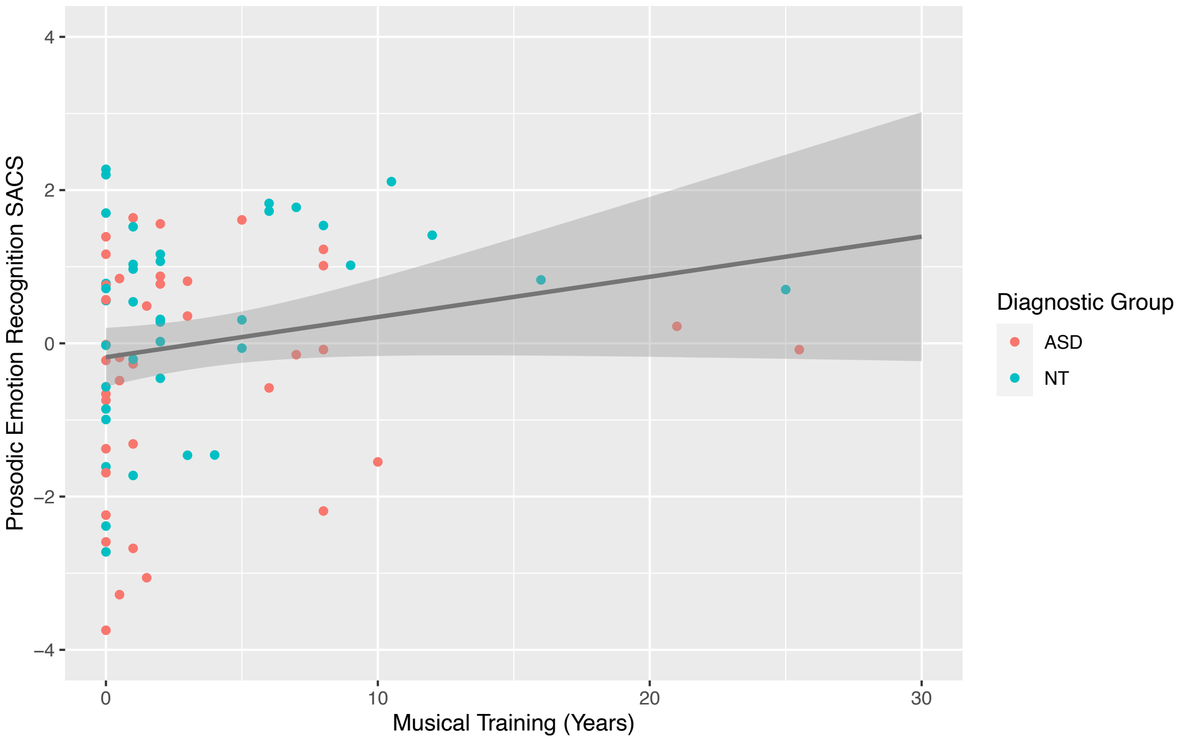


Figure 5. 2. Scatterplot of prosodic emotion recognition A) accuracy (arcsine-transformed) and B) SACS against years of musical training. Regression lines were based on linear regressions of paired measures.

Musical perception abilities

Regression analysis on MBEA/MBEMA global score revealed no significant predictors of Diagnostic group, Chronological age, and their interaction (see Table 5.4 for full summary). This indicates that diagnostic group and participants' age did not predict their overall musical perception abilities.

Across ASD and NT groups, significant weak to moderate correlations between MBEA/MBEMA global score and emotion recognition accuracy and SACS for prosody (global score vs. prosody recognition accuracy: $\tau = 0.20$, $p = 0.019$; global score vs. prosody recognition SACS: $\tau = 0.17$, $p = 0.038$), suggesting better musical perception abilities were associated with more accurate and efficient emotion recognition from speech prosody. However, these correlations became non-significant after controlling for age (global score vs.

speech recognition accuracy: $\tau = 0.15$, $p = 0.060$; global score vs. speech recognition SACS: $\tau = 0.13$, $p = 0.105$). No other correlations regarding MBEA/MBEMA global score were significant across groups, including emotion recognition RT for prosody, emotion recognition from song across performance indices, and emotion priming of prosody and song across performance indices (see Table 5.5).

Within the ASD and NT groups, no significant correlations were observed between MBEA/MBEMA global score and emotion recognition and priming of prosody and song across measures, with and without controlling for age (see Table 5.5).

Pitch processing abilities

Regression analysis on pitch direction discrimination threshold revealed no significant predictors of Diagnostic group and Diagnostic group × Chronological age did not reach significance (see Table 5.4 for sull summary).

Across ASD and NT groups, correlational analyses showed significant moderate to strong correlations between pitch direction discrimination threshold and emotion recognition accuracy and SACS of both prosody and song (pitch threshold vs. prosody recognition accuracy: $\tau = -0.34$, $p < 0.001$; pitch threshold vs. prosody recognition SACS: $\tau = -0.33$, $p < 0.001$; pitch threshold vs. song recognition accuracy: $\tau = -0.28$, $p = 0.002$; pitch threshold vs. song recognition SACS: $\tau = -0.31$, $p < 0.001$). These correlations remained significant after controlling for age (pitch threshold vs. prosody recognition accuracy: $\tau = -0.27$, $p = 0.002$; pitch threshold vs. prosody recognition SACS: $\tau = -0.26$, $p = 0.003$; pitch threshold vs. song recognition accuracy: $\tau = -0.21$, $p = 0.017$; pitch threshold vs. song recognition SACS: $\tau = -0.24$, $p = 0.006$). These results indicate that lower (i.e., better) pitch direction discrimination thresholds were associated with more accurate and efficient emotion recognition from speech prosody and song across groups. No other correlations regarding pitch direction discrimination threshold were significant across groups, including emotion recognition RT for prosody and song and emotion priming of prosody and song across performance indices (see Table 5.5).

Within the NT group, significant strong correlations were found between pitch direction discrimination threshold and emotion recognition accuracy and SACS for both prosody and song (pitch threshold vs. prosody recognition accuracy: $\tau = -0.39$, $p = 0.002$; pitch threshold vs. prosody recognition SACS: $\tau = -0.37$, $p = 0.002$; pitch threshold vs. song recognition accuracy: $\tau = -0.30$, $p = 0.016$; pitch threshold vs. song recognition SACS: $\tau = -0.36$, $p = 0.003$). These correlations remained significant after controlling for age (pitch threshold vs. prosody recognition accuracy: $\tau = -0.34$, $p = 0.006$; pitch threshold vs. prosody recognition SACS: $\tau = -0.31$, $p = 0.011$; pitch threshold vs. song recognition accuracy: $\tau = -0.25$, $p = 0.039$; pitch threshold vs. song recognition SACS: $\tau = -0.31$, $p = 0.010$). These results indicate that lower (i.e., better) pitch direction discrimination thresholds were associated with more accurate and efficient emotion recognition from prosody (Figure 5.3) and song (Figure 5.4) in NT individuals. No other correlations regarding pitch direction discrimination threshold were significant within the NT group, including emotion recognition RT for prosody and song and emotion priming of prosody and song across performance indices (see Table 5.5).

Within the ASD group, no significant correlations were observed between pitch direction discrimination threshold and emotion recognition and priming of prosody and song across measures, with and without controlling for age (see Table 5.5).

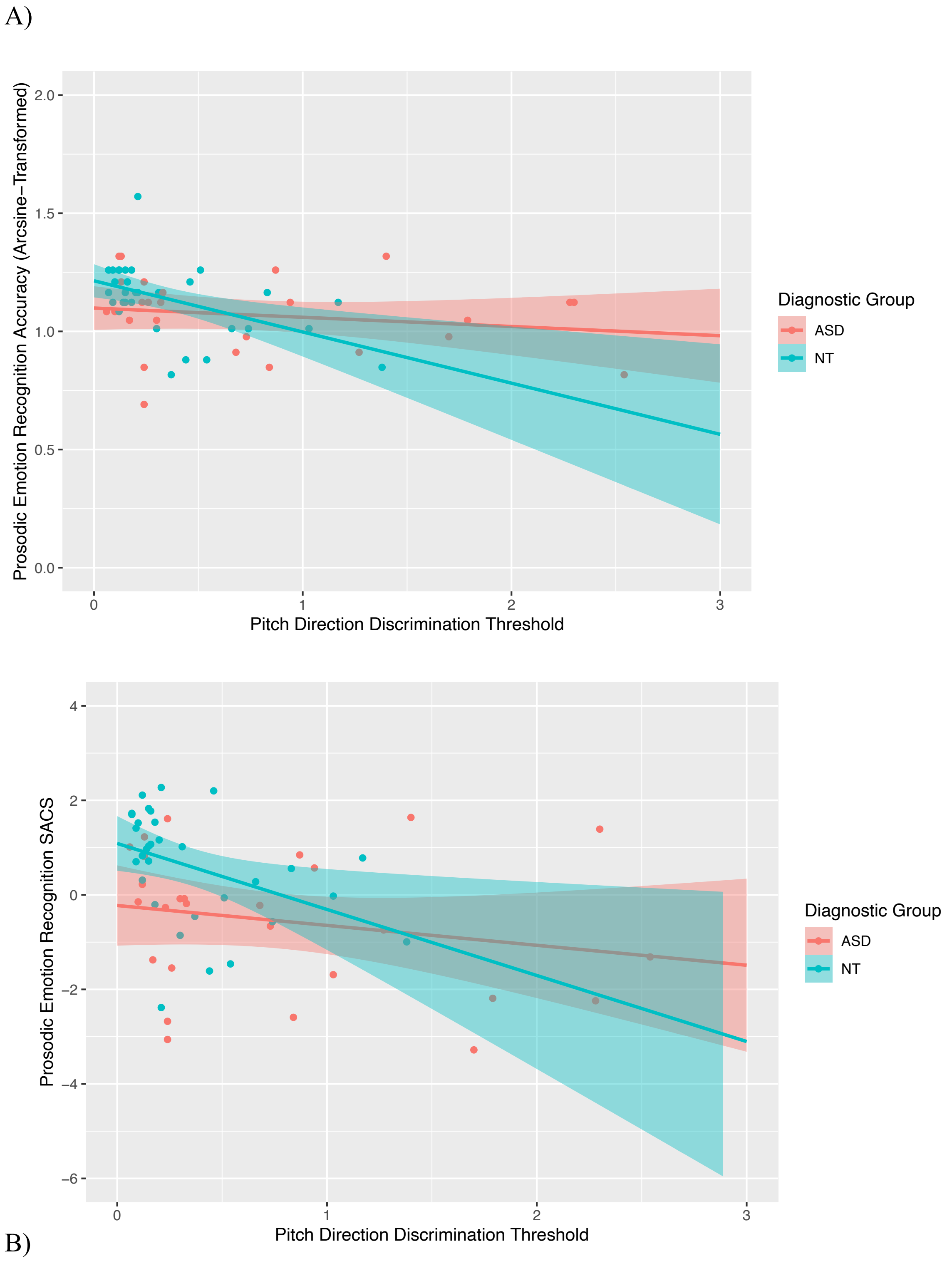
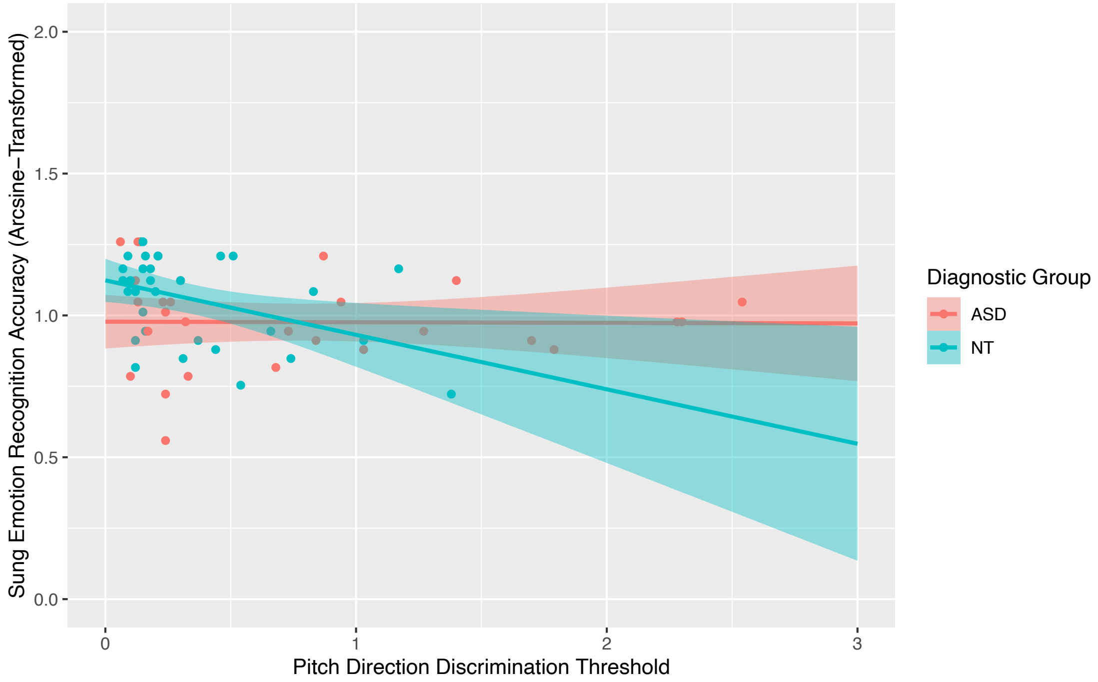


Figure 5. 3. Scatterplot of prosodic emotion recognition A) accuracy (arcsine-transformed) and B) SACS against pitch direction discrimination threshold. Regression lines were based on linear regressions of paired tasks for each diagnostic group.

A)



B)

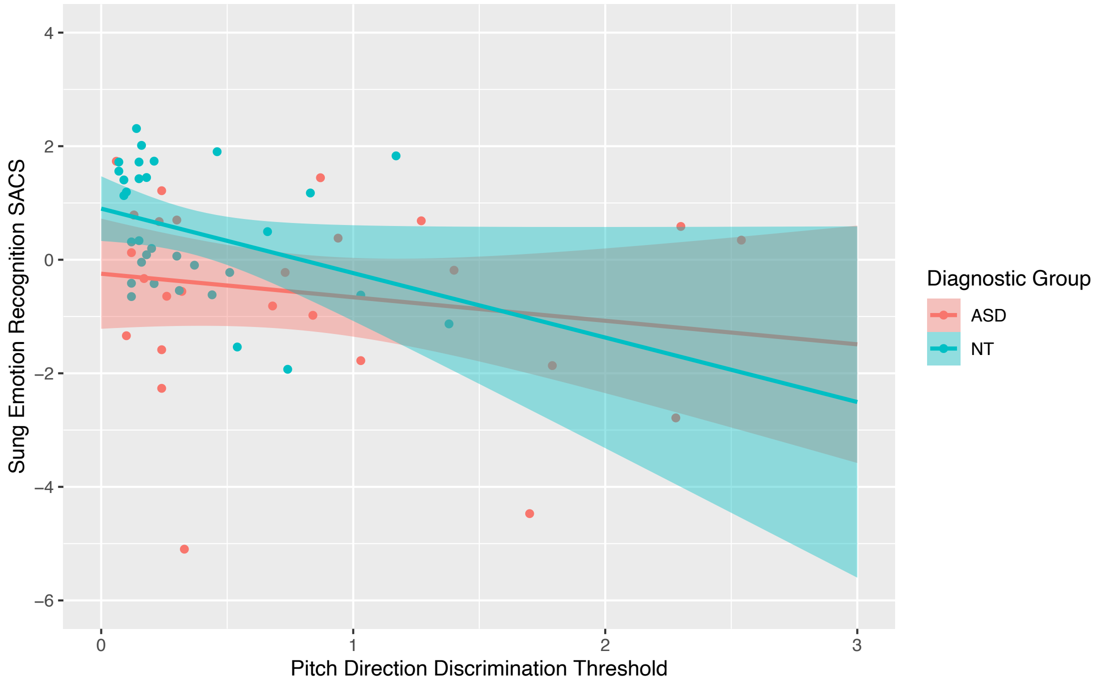


Figure 5. 4. Scatterplot of sung emotion recognition A) accuracy (arcsine-transformed) and B) SACS against pitch direction discrimination threshold. Regression lines were based on linear regressions of paired tasks for each diagnostic group.

5.3.3. Correlations between alexithymic traits, autistic traits, and overall emotion processing

Group differences in alexithymic and autistic traits

An independent-samples *t*-test revealed a significant difference in TAS-20 score between adults with ASD and NT adults ($t(19.48) = 4.57, p < 0.001$), with adults with ASD (M = 69.85, SD = 9.76) showing higher alexithymic traits than NT adults (M = 45.54, SD = 16.51). Specifically, within the ASD group, three participants were presented with possible alexithymia and 10 participants with alexithymia, while no participants obtained a score indicative of non-alexithymia. Within the NT group, nine participants were presented with no alexithymia, one participant with possible alexithymia, and three participants with alexithymia.

An independent-samples *t*-test revealed a significant difference in AQ score between adults with ASD and NT adults ($t(21.19) = 12.86, p < 0.001$), with adults with ASD (M = 41.46, SD = 4.29) showing higher autistic traits than NT adults (M = 14.31, SD = 6.29). Specifically, all participants in the ASD group scored above the cut-off score of 32, which is indicative of clinically significant levels of autistic traits (Baron-Cohen et al., 2001), while all participants in the NT group scored below this cut-off score.

The contribution of alexithymic and autistic traits to emotion processing

Table 5. 6. Multiple linear regression model results for alexithymic traits measured on the TAS-20 and autistic traits measured on the AQ on overall emotion recognition and priming performance within and across ASD and NT adult groups, with 13 participants in each group.

	ASD					**NT**					**All**				
Predictors	***B***	***SE***	***t***	***p***	**η^2_p**	***B***	***SE***	***t***	***p***	**η^2_p**	***B***	***SE***	***t***	***p***	**η^2_p**
Overall emotion recognition accuracy															
AQ	-0.01	0.01	-1.98	0.076	0.28	-0.01	0.00	-1.59	0.142	0.20	0.00	0.00	-0.70	0.293	0.02
TAS-20	0.00	0.00	0.97	0.354	0.09	0.00	0.00	-0.62	0.551	0.04	0.00	0.00	-0.53	0.599	0.02
Overall emotion recognition RT															
AQ	0.00	0.03	0.02	0.987	0.00	**0.03**	**0.01**	**5.06******	**< 0.001**	**0.72**	**0.01**	**0.00**	**2.45***	**0.023**	**0.21**
TAS-20	0.00	0.01	0.19	0.853	0.00	0.00	0.00	-0.97	0.354	0.09	0.00	0.00	-0.21	0.834	0.00
Overall emotion recognition SACS															
AQ	-0.09	0.11	-0.81	0.436	0.06	**-0.11**	**0.03**	**-4.43****	**0.001**	**0.66**	-0.04	0.02	-1.89	0.072	0.13
TAS-20	0.01	0.05	0.26	0.800	0.01	0.00	0.01	-0.03	0.974	0.00	0.00	0.02	-0.19	0.853	0.00
Overall emotion priming accuracy															
AQ	0.00	0.01	0.35	0.736	0.01	0.00	0.01	0.39	0.705	0.02	0.00	0.00	-0.03	0.979	0.00

TAS-20	0.00	0.01	0.62	0.547	0.04	0.00	0.00	0.23	0.825	0.01	0.00	0.00	0.69	0.495	0.02
Overall emotion priming RT															
AQ	0.00	0.01	0.35	0.736	0.01	0.00	0.01	0.39	0.705	0.02	0.00	0.00	-0.03	0.979	0.00
TAS-20	0.00	0.01	0.62	0.547	0.04	0.00	0.00	0.23	0.825	0.01	0.00	0.00	0.69	0.495	0.02
Overall emotion priming SACS															
AQ	0.02	0.10	0.23	0.820	0.01	0.04	0.06	0.47	0.650	0.02	0.00	0.02	0.05	0.960	0.00
TAS-20	0.05	0.05	1.13	0.283	0.11	0.01	0.03	0.25	0.81	0.01	0.02	0.02	0.97	0.344	0.04

Note. Overall emotion recognition accuracy was arcsine-transformed and overall emotion recognition RT was log-transformed; **p* < 0.05, ** *p* < 0.01, *** *p* < 0.001; significant correlations are highlighted using bold font.

Across ASD and NT groups, regression analyses revealed that AQ scores significantly predicted overall emotion recognition RT (B = 0.01, SE = 0.00, $t(23) = 2.45$, $p = 0.023$, η^2_p = 0.21), indicating that higher autistic traits were associated with slower emotion recognition in general (Figure 5.5). AQ scores, however, did not predict other performance indices of emotion recognition nor any performance indices of emotion priming (see Table 5.6). There were no effects of TAS-20 score on any of the emotion recognition and priming performance indices, indicating that alexithymic traits did not contribute to implicit nor explicit emotion processing across groups.

Within the NT group, AQ score significantly predicted overall emotion recognition RT (B = 0.03, SE = 0.01, $t(10) = 5.06$, $p < 0.001$, $\eta^2_p = 0.72$) and overall emotion recognition SACS (B = -0.11, SE = 0.03, $t(10) = -4.43$, $p = 0.001$, $\eta^2_p = 0.66$), indicating that higher autistic traits were associated with slower and less efficient emotion recognition among NT individuals. AQ score, however, did not predict emotion recognition accuracy nor emotion priming across performance indices (see Table 5.6). There were no effects of TAS-20 score on any of the emotion recognition and priming performance indices, indicating that alexithymic traits did not contribute to implicit nor explicit emotion processing among NT individuals.

Within the ASD group, neither AQ or TAS-20 scores predicted emotion recognition and priming across performance indices (see Table 5.6).

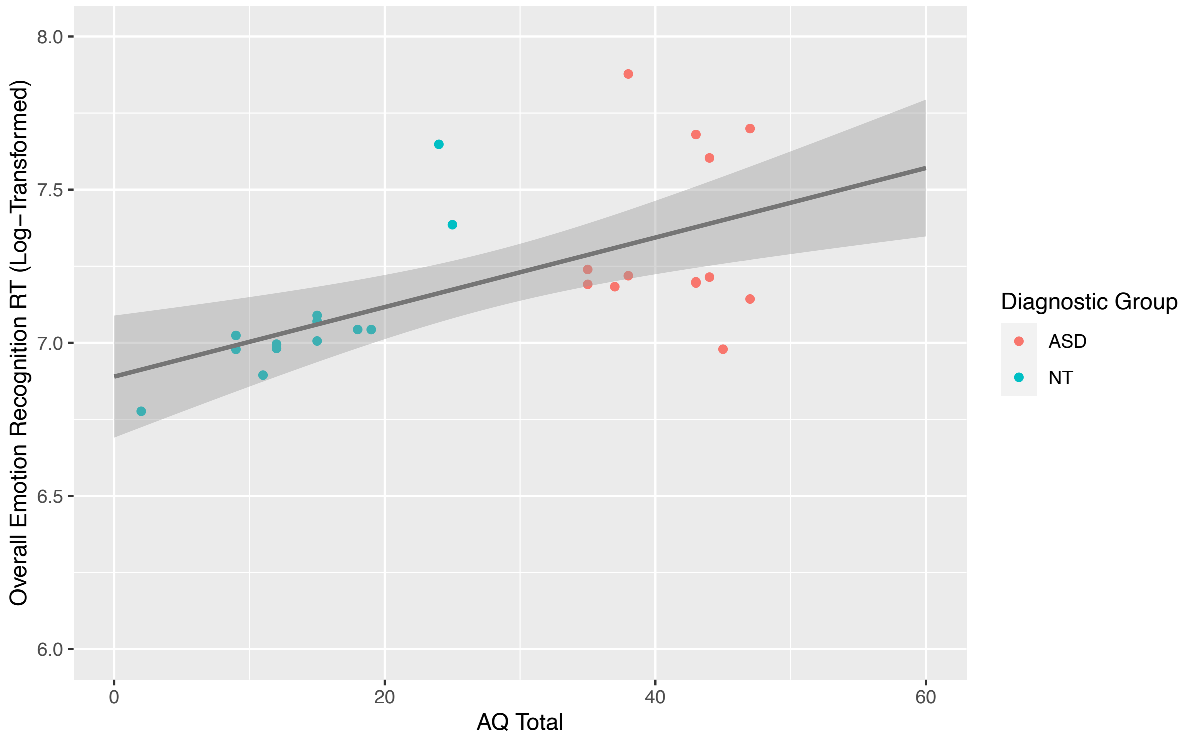


Figure 5. 5. Scatterplot of autistic trait assessed on the AQ against overall emotion recognition RT (log-transformed). Regression lines were based on linear regressions of paired measures across ASD and NT groups.

5.4. Discussion

The present study aimed to expand findings reported in Chapters 3 and 4 by exploring potential correlates relating to visual and auditory emotion processing in autistic and NT individuals. The current data suggest that the two groups may employ different processing strategies during explicit emotion recognition, specifically for human faces, speech prosody, and song. First, it was found that higher susceptibility to local details when making a global judgment was associated with slower recognition of human facial but not face-like object expressions, which was particularly prominent in the NT group. Secondly, across groups, more musical training was related to more accurate and efficient emotion recognition of speech prosody but not song. By contrast, musical perception ability was not associated with any of the emotion recognition measures. Interestingly, lower pitch threshold was related to better prosodic and sung emotion recognition only among NT individuals, while no such associations were observed for autistic individuals. Notably, none of these auditory-related measures correlated with emotion priming. Finally, autistic traits, but not alexithymic traits, contributed to the variability in explicit emotion recognition performance across groups, contradicting the alexithymia hypothesis. Neither of these measures significantly predicted the strength of implicit emotion priming. These findings are further discussed in the following subsections.

5.4.1. The relationship between cognitive processing style and visual emotion recognition

The present study explored whether differences in emotion recognition of human facial and face-like object expressions between autistic and NT individuals (as observed in Chapter 3) were related to their cognitive processing style. Two measures of cognitive processing style were obtained from the Navon task (Navon, 1977), namely the global advantage effect and the local-to-global interference effect. It was found that the two groups did not differ in the global advantage effect, suggesting that both groups were just as readily to identify the global component when presented with a compound stimulus containing both global and local components. These findings while not speaking for the proposal of an impairment in the integration of component parts in ASD as put forward by the WCC account (U. Frith, 1989; U. Frith & Happé, 1994), they replicated previous studies employing the Navon paradigm in showing comparable global advantage effects between the ASD and NT groups (Guy et al., 2019; Juslin & Madison, 2016; Mottron et al., 1999; Ozonoff et al., 1994; Plaisted et al., 1999). Importantly, results showed that the global advantage effect did not correlate with visual emotion recognition across and within the ASD and NT groups. This indicates that the speed of processing global over local features was not related to the ability to process emotions across visual stimuli.

Conversely, the local-to-global interference effect was found to be significantly associated with the speed of emotion recognition of faces in the NT group. That is, higher susceptibility to the slowing down of global processing due to local interference was related to slower emotion recognition in faces. This association, however, did not extend to that of objects in the NT group. Despite a number of physical features in common, recent work has demonstrated faster detection of human faces compared to face-like objects in visual search (Keys et al., 2021). Moreover, magnetoencephalography (MEG) data showed that although the initial neural representation of face-like objects is similar to human faces than matched objects, this ‘face-like’ representation reorganises to be more similar to matched objects than faces within ~250ms (Wardle et al., 2020). Human faces, thus, appear to be somewhat more special in perception, as also evident in an advantage for upright versus inverted orientation when processing human but not cartoon faces in NT individuals (Akdeniz, 2020; Rosset et al., 2010). Accordingly, the recognition of human facial expressions, but not face-like object expressions, may benefit more from interference-free global processing strategy that is used for general information processing, and hence an association observed between the two measures within the NT group.

The association between local-to-global interference and facial emotion recognition, however, was not observed within the ASD group. Compelling evidence suggests qualitative differences in the way autistic individuals process human faces, such as the atypical featural processing for human faces (Deruelle et al., 2004; Hernandez et al., 2009) and/or avoidance of the eyes in autistic individuals (Frazier et al., 2017; Tanaka & Sung, 2016). Moreover, this atypical processing strategy may be specific to human faces only, where autistic children have been shown to exploit typical configural strategy with cartoon faces (Rosset et al., 2008). Accordingly, there may be a reduced tendency to tap into more general perceptual processing mechanisms (e.g., interference-free global processing) when processing emotional faces in ASD, and hence a lack of association observed between the two measures. In fact, a marginally higher local-to-global interference was noted in the ASD relative to the NT group, partially consistent with previous studies demonstrating clear differences in this measure between the two groups (Guy et al., 2019; Rinehart et al., 2000; L. Wang et al., 2007). This observation may give rise to the possibility that slower emotion recognition of human faces in the ASD group was attributed to their marginally higher local-to-global interference. Future research incorporating eye-tracking techniques to measure visual scan paths for processing emotional human faces and control stimuli such as compound letters used in the Navon task may provide clearer insights into these relationships. Importantly, no association was found between cognitive processing style and emotion recognition of face-like objects for either group; it remains uncertain what may be underlying the slower emotion recognition of face-like objects in the ASD group (as observed in Chapter 3).

The investigation of different measures of cognitive processing style and their associations with facial emotion recognition provides implications for the understanding of the processes that may underlie facial emotion recognition. Findings from previous research and the present work appear to converge on the notion that global processing as the default setting in visual perception is particularly relevant for emotion recognition of facial expressions. While previous work by Gross (2005) showed that a preference for global processing in a free-choice task was related to better facial emotion recognition, the present work extended this and showed that global processing interrupted by initial orientation towards local details was related to poorer facial emotion recognition. These findings could be used to explain why autistic individuals are more drawn to particular features of a face which may or may not be relevant for emotional decoding (e.g., an ear, the chin, or region of the hair line) without a tendency to explore core features holistically when processing emotional faces (Pelphrey et al., 2002). This differential processing strategy may, in turn, hamper the recognition of emotions from human faces. By contrast, the ability to process global or local information when directed to a specific level appears to be a less important correlate of facial emotion recognition, as shown in Oerlemans et al. (2013) and in the present study (in the case of global advantage).

5.4.2. The relationship between musical processing, pitch processing, and auditory emotion processing

The ASD and NT groups did not differ in the three auditory perception measures, namely musical training, musical perception, and pitch perception. In particular, the absence of group differences for musical perception and pitch perception are consistent with previous research in showing that these are relatively preserved domains of ability in ASD (Altgassen et al., 2005; Chowdhury et al., 2017; Germain et al., 2019; Globerson et al., 2015; Heaton, 2005; Jamey et al., 2019; Molnar-Szakacs & Heaton, 2012). The three measures, nevertheless, appear to be associated with auditory emotion processing to varying extents.

Across groups, musical training was positively associated with emotional recognition accuracy and efficiency in speech prosody, which was depicted even after controlling for age differences. These results are in line with previous findings (Correia et al., 2020; Farmer et al., 2020; Lima & Castro, 2011b; W. F. Thompson et al., 2004). Notably, this association was only observed when the ASD and NT groups were pulled together, but not within each group separately. This is likely due to the low number of participants with extensive musical training when broken down by groups, with two participants in each group having musical training for more than 13 years (i.e., a criterion used for classifying musician versus non-musician groups in previous studies; Marques et al., 2007; Parbery-Clark et al., 2009; Schön et al., 2004a; W. F. Thompson et al., 2004). Importantly, it is not possible to infer whether the effects of musical training on prosodic emotion recognition are evident to similar extents in the two groups. This undermines the question of whether any benefits of musical training (e.g., enhancement of auditory-related skills) contribute to prosodic emotion recognition similarly in the ASD group relative to the NT group to be addressed and warrants further investigation with a larger sample of participants having more diverged musical training background.

No association between musical training and emotion recognition of song was observed across nor within groups. These findings are in line with those reported in Quintin et al. (2011) which used instrumental excerpts as musical stimuli across ASD and NT groups, while standing in contrast with those reporting significant facilitatory effects of musical training on musical emotion recognition in the NT literature (Castro & Lima, 2014; Lima & Castro, 2011a; Livingstone et al., 2010). Although the relatively low number of years of musical training in both Quintin et al. (2011) and the present study may suffice as a plausible explanation, an association was nonetheless observed for prosodic emotion recognition when the two groups were combined in the present study. This may, therefore, give rise to an alternative possibility. The effects of musical training may depend on the musical stimuli used for emotion recognition and/or the relationship between the musicality of the instrument learnt and the stimuli presented. For instance, in Castro & Lima (2014), participants had received formal training on different types of instruments (i.e., keyboard, strings, woodwinds, and percussion) and recognised emotions from instrumental stimuli developed using digital synthesizers in piano timbre and of long durations (e.g., ~12.4s). In the present study, participants had also received formal training on various types of instruments[1] with only five participants on voice, while the musical stimuli used were sung stimuli obtained from RAVDESS (Livingstone & Russo, 2018) that were segmented and manipulated to last for very short durations (i.e., 500ms). It is possible that musical training, particularly on instrumental music, is less applicable to emotion recognition of the singing voice in very short segments that are not often encountered in participants' musical practice. It remains unclear as to why the effects of musical training on auditory emotion recognition appear to be generalised across to the speech prosody domain,

---

[1] A breakdown of the number of participants who had received formal training on the various types of instruments: brass (n = 12), keyboard (n = 22), percussion (n = 6), plucked string (n = 16), string (n = 7), voice (n = 5), and woodwind (n = 16). Note that 20 of the 50 participants who had received formal musical training learnt multiple musical instruments.

but not to other forms of musicality (i.e., the singing voice) within the musical domain. To further elucidate this, future studies should directly investigate emotional stimuli induced by different timbres within and across domains, involving a larger pool of musicians trained on different instruments including voice.

Results showed that musical perception abilities assessed on the MBEA/MBEMA tasks (Peretz et al., 2003, 2013) did not correlate with any of the auditory emotion recognition measures, after controlling for age differences across and within the ASD and NT groups. These findings appear to contradict previous studies that noted significant correlations between MBEA scores and emotion recognition in speech prosody (Lima et al., 2016; Pralus et al., 2019; Trimmer & Cuddy, 2008), instrumental music (Lévêque et al., 2018), and vocalisations (e.g., crying; Lima et al., 2016). One possibility of this apparent contradiction in findings may be due to musical abilities assessed on the MBEA tasks being more relatable to emotion recognition in auditory stimuli of longer durations, such as sentences (Pralus et al., 2019; Trimmer & Cuddy, 2008), multisyllabic utterances (Lima et al., 2016), and instrumental musical excerpts (Lévêque et al., 2018). By contrast, performance on the MBEA tasks may be less relatable to auditory stimuli of shorter durations, such as the monosyllabic words (i.e., “door) used in the present study. For example, rhythmic perception assessed on the MBEA task might not have been as relevant for emotion recognition in the present study, as rhythmic/speech rate information was not available in the prosodic and sung stimuli employed which were monosyllabic words standardised to last for 500ms. Likewise, emotion recognition of these monosyllabic words might not have required musical memory to the same extent as that for distinguishing between old and new melodies made up of sequence of notes on the musical memory subtest of the MBEA/MBEMA task. The seemingly reduced compatibility of the MBEA measures and characteristics of the stimuli used in the present study may have resulted in the lack of associations observed. However, this speculation seems to be conflicted with the significant associations observed for vocalisations in Lima et al. (2016), which are also stimuli of short durations.

This significant association between MBEA performance and emotion recognition in vocalisations (Lima et al., 2016), as well as some of the other significant associations mentioned above (Lévêque et al., 2018; Pralus et al., 2019), were in fact only observed when data from participants with congenital amusia and control participants were pulled together, but not when these associations were explored separately within each. Here, congenital amusia refers to a neurodevelopmental disorder of musical and pitch perception (Ayotte et al., 2002; Peretz et al., 2002) that is commonly diagnosed using the MBEA task (Peretz et al., 2003). It is, therefore, not surprising that participants with congenital amusia in these studies had significantly lower MBEA scores than control participants, and hence enabling a wider range of musical perception abilities to be correlated with emotion recognition performance in these studies. In this regard, it is possible that the MBEA task may have been less sensitive in capturing more subtle differences within the typical range of musical perception abilities among the present sample, limiting associations between musical perception abilities and auditory emotion recognition to be revealed. Additionally, in previous studies by Globerson et al. (2013, 2015), it was found that pitch tasks that require differentiating and naming pitches (e.g., high vs. low, glide vs. non-glide, ascending vs. descending) were more sensitive predictors of emotional prosody recognition than pitch tasks that require same-different judgments. This observation may partially explain why there was a lack of association between MBEA/MBEMA performance and auditory emotion recognition in the present study, as participants were required to make same-different judgments on the melodic pitch perception subtests. Future studies may re-evaluate these associations by using measures that capture more subtle individual differences in musical perception abilities, to further scrutinise whether an association between musical perception ability and emotion recognition in prosody and song exists.

Although pitch perception ability did not differ between the ASD and NT groups, its association with auditory emotion recognition accuracy and efficiency was only found in the NT group for both speech prosody and song, even after controlling for age differences. The finding of that pitch perception ability was associated with more accurate and efficient emotion recognition of speech prosody in NT individuals is in line with previous findings (Globerson et al., 2013, 2015), while adding novel evidence to the literature for the same association observed for song. In addition, these findings provide supporting evidence that pitch provides crucial emotional information in the two domains (R. Gold et al., 2012; Hakanpää et al., 2019a, 2019b; Krumhansl & Shepard, 1979) and that processing in the two domains may draw on overlapping underlying mechanisms (Merrill et al., 2012; Peretz et al., 2015), which extends to the recognition of emotions (Escoffier et al., 2013).

Conversely, no associations between pitch perception ability and auditory emotion recognition were found in the ASD group across accuracy, speed, and efficiency. While this potentially different reliance on pitch processing (specifically non-vocal gliding tones) may explain why autistic individuals were less efficient in emotion recognition from prosody in song (as shown in Chapter 3), this difference did not seem to affect the recognition accuracy nor response patterns exhibited by autistic individuals relative to NT individuals. This highlights that autistic individuals may make use of alternative acoustic information to decode emotions from auditory stimuli to achieve similar accuracy level as NT individuals, such as voice quality, given that intensity and duration of auditory stimuli were standardised in the present study. This speculation, nevertheless, needs to be further examined in future research. The present findings appear to contradict with previous findings reported in Globerson et al. (2015), which found an association between non-vocal discrete pitch processing and prosodic emotion recognition in both the ASD and NT groups. In this study, although non-vocal discrete pitch processing was intact and associated with prosodic emotion recognition, impaired prosodic emotion recognition was still observed in autistic individuals relative to NT individuals (Globerson et al., 2015). Taken together with these findings and the present findings, it appears that non-vocal pitch processing does not seem to moderate group differences in auditory emotion recognition accuracy – that is, prosodic emotion recognition performance by autistic individuals relative to NT individuals was not determined by whether or not an association between the two variables is found (e.g., a significant association does not imply intact emotion recognition accuracy and a non-significant association does not imply impaired emotion recognition accuracy).

Taking this discussion further, in another study by Schelinski and von Kriegstein (2019), it was found that impairments in prosodic emotion recognition in autistic individuals might be attributed to their impaired vocal discrete pitch processing. It was proposed that impairments in vocal pitch processing might have resulted in vocal pitch information not being available for prosodic emotion recognition in autistic individuals, where this information is available for NT individuals (Schelinski & von Kriegstein, 2019). Thus, it may be the case that vocal pitch processing is a particularly important moderator of group differences in auditory emotion recognition or that the association between non-vocal pitch processing and auditory emotion recognition is mediated by vocal processing specifically. Given the highly variable pitch processing abilities reported in the autism literature (e.g., enhanced/intact non-vocal pitch processing: Altgassen et al., 2005; Bonnel et al., 2003; Chowdhury et al., 2017; Jones et al., 2009; impaired non-vocal pitch processing: Bhatara et al., 2013; Kargas et al., 2015; e.g., enhanced vocal pitch processing: Heaton, Hudry, et al., 2008; e.g., impaired vocal pitch processing: Schelinski & von Kriegstein, 2019), it is not possible to infer whether intact auditory emotion recognition accuracy could be attributed to (intact) vocal pitch processing in the current ASD sample. Nevertheless, the discrepant findings across previous studies and the present study highlight the different contribution of various aspects of pitch information in auditory emotion recognition, which may or may not moderate group differences. Future research should be directed to investigate the relationship of auditory emotion recognition and different aspects of pitch processing for better understanding of the complex picture between the two constructs. Notably, slower auditory emotion recognition in autistic individuals (as demonstrated in Chapter 3) could not be attributed to pitch perception abilities, as neither the ASD nor NT groups showed any associations between the two measures.

The present study provided novel evidence for a lack of association between pitch perception abilities and auditory emotion priming both across and within the ASD and NT groups. In other words, pitch perception ability appeared not to be related to the extent to which visual emotion recognition is facilitated by emotional cues conveyed in the prosodic and sung primes. These findings, however, appear to stand at odds when considering evidence from affective priming studies with individuals with congenital amusia. It has been shown that while priming effects of musical chords on facial emotional judgment were found in NT controls, no such effects were found in individuals with congenital amusia who also exhibited impairments in pitch perception abilities (L. Zhou et al., 2019). One explanation regarding the different findings for the association between pitch perception and auditory emotion priming may relate to how emotions are conveyed through pitch in relation to other acoustic characteristics of the auditory prime stimuli. For instance, given that voice quality also contributes to the differentiation between emotion categories in prosody and song (Hakanpää et al., 2019b; Lugger & Yang, 2007), pitch cues alone may not be sufficient or that it may be outweighed by voice quality cues for emotional meaning to be activated during implicit priming, minimising the reliance on pitch processing ability during this process in the present study. By contrast, emotional connotations were largely determined by the manipulation of the pitch properties within the musical chord stimuli used in Zhou et al. (2019), and hence may have drawn more closely on pitch processing ability for emotional meaning to be activated during implicit priming. Future research is needed to examine these speculations regarding the possibly different weightings of acoustic cues during implicit and explicit auditory emotion processing, as the present findings seem to suggest that pitch cues may be particularly relevant for emotion processing in prosody and song at the explicit level but not at the implicit level.

5.4.3. The relationship between alexithymic traits, autism-like traits, and emotion processing

As previous research suggests that alexithymia can account for atypicalities in emotion recognition (Cook et al., 2013; Heaton et al., 2012; Ketelaars et al., 2016), one of the objectives of this chapter was to evaluate whether alexithymia was related to emotion recognition and priming, and whether differences in emotion processing in ASD could be attributed to co-occurring alexithymia. Multiple linear regression analyses revealed no significant effects of alexithymic traits on overall emotion recognition across performance indices. By contrast, significant effects of autistic traits on overall emotion recognition speed were depicted over and above that of alexithymic traits, with higher autistic traits relating to slower general emotion recognition. It should, however, be noted that the effects of autistic traits across groups were driven by the NT group, as such effects were only seen in the separate analysis for the NT but not ASD group. Nonetheless, the current data provided no evidence in favour of the alexithymia hypothesis, which proposes that alexithymia, but not ASD per se, is the source of emotion recognition impairments (Bird & Cook, 2013; Cook et al., 2013). This notion, nevertheless, remains controversial, given that the current literature in this area reports conflicting findings (Allen et al., 2013; Cook et al., 2013; Heaton et al., 2012; Keating et al., 2021; Ketelaars et al., 2016; Kliemann et al., 2013; Milosavljevic et al., 2016; Ola & Gullon-Scott, 2020). Notably, the effects of alexithymia on emotion recognition were almost exclusively examined with regards to accuracy in these previous studies, with poorer emotion recognition accuracy reported in their ASD groups (Allen et al., 2013; Cook et al., 2013; Heaton et al., 2012; Ketelaars et al., 2016). As no impairments in emotion recognition accuracy were observed for the ASD group in the present work (as demonstrated in Chapter 3), the alexithymia hypothesis could not be reliably attested. It is plausible that the emotion recognition task using prototypical expressions had lower sensitivity to detect subtle group differences on the accuracy level, and hence the effects of alexithymia. The absent effect of alexithymia on emotion recognition speed, in fact, has also been reported in a previous study (Ola & Gullon-Scott, 2020). Thus, whether or not co-occurring alexithymia contributes to emotion recognition impairments in terms of accuracy but not speed (and efficiency, which has not been explored previously) warrants closer examination in future studies with more sensitive measures of emotion recognition. Nonetheless, the significant effect of autistic traits on overall emotion recognition speed complements findings reported in Chapter 3, such that high autistic traits (a distinctive characteristic of the ASD group) were related to slower emotion recognition compared to low autistic traits (a distinctive characteristic of the NT group).

The current data did not provide evidence for the effect of alexithymia on implicit emotion priming, contradicting evidence indicating attenuated (e.g., Rosenberg et al., 2020; Vermeulen et al., 2006; see Donges & Suslow, 2017; Grynberg et al., 2012 for reviews) or elevated implicit emotion processing in alexithymia (J. Parker et al., 1993; Suslow, 1998). On the whole, no significant difference in the strength of emotion priming was observed between the ASD and NT groups in the present work (as demonstrated in Chapter 4). Again, this implies that the alexithymia hypothesis could not be reliably attested. Moreover, no significant effects of autistic traits on implicit emotion priming were observed. Together, these findings align well with those observed in Chapter 4, such that the ASD and NT groups, which are clearly differentiated by their autistic and alexithymic traits, showed no significant difference in the strength of emotion priming, thus the absence of effects of these measures here.

5.4.4. Limitations

A number of limitations need to be considered. First, while the possibility of mistyped responses due to fatigue and boredom over the number of trials on the Navon task cannot be ruled out, another limitation is that the global/local manipulations were completed within-subjects (i.e., participants responded to both global and local targets depending on the designated block). It is possible that participants' observed cognitive processing style was influenced by carryover effects from the within-subjects design, despite the order of global/local blocks being counterbalanced between participants. For example, participants responding to local targets may show more local bias in the subsequent global block. The cost of switching might have limited autistic participants' biases to be reliably measured in the subsequent block, where a particular difficulty in the case of switching attention from a local level to a global level has been noted in previous research (Katagiri et al., 2013; Soriano et al., 2018). In this regard, for autistic participants completing the local block first, their local biases could potentially have been exacerbated. This confounding effect, however, does not seem to be reflected in the current dataset, as the ASD group did not show significantly more local biases in neither the global advantage nor local-to-global interference measures compared to the NT group.

Secondly, as discussed above, there was a relatively small variability in several measures of individual differences captured among the current sample, including years of musical training (and perhaps the types of instruments trained on) and musical perception ability. This limitation might have obscured their associations with emotion processing to be observed, contradicting previous research. Moreover, this has further limited the present study in examining whether individual differences in these factors correlate with emotion processing to similar extents in autistic and NT individuals, which could have important implications for understanding differences in underlying mechanisms of emotion processing between the two groups. A larger sample of participants to capture wider variability in individual differences is warranted to delve further into this area of research.

Thirdly, it should be noted that the recruitment of a representative sample of autistic and NT individuals meant that group analyses according to the presence and absence of alexithymia was not possible. Specifically, the two groups were (almost) distinctively differentiated by their alexithymic traits, with all autistic participants having an alexithymic score indicative of (possible) alexithymia and the majority of the NT participants having no alexithymia with only a few participants with (possible) alexithymia. Hence, there was very little overlap in alexithymic scores between the two groups. It is clear that alexithymic trait is highly conflated with the ASD diagnosis; the present findings should be interpreted with caution as there were no subgroups defined by alexithymic traits within the current ASD sample, which is part of the proposal of the alexithymia hypothesis (Bird & Cook, 2013). To get a clearer picture of the associations between ASD diagnosis, co-occurring alexithymia, and emotion processing impairments, future study designs may incorporate matching the ASD and NT groups on the alexithymic measure in a larger sample.

5.4.5. Chapter summary

In summary, the investigation of various plausible correlates appears to highlight that autistic individuals may employ processing strategies and/or rely on different cues compared to NT individuals during explicit emotion recognition. It was found that local-to-global interference, an index of cognitive processing style, was significantly associated with emotion recognition speed and efficiency in human faces but not face-like objects across groups, which was particularly prominent in the NT group. This suggests that global processing that is free from interference by local details is particularly relevant for quicker and more efficient emotion recognition, but only for human faces – perhaps due to their special status in visual perception. The question of whether the slower and less efficient emotion recognition of human faces in the ASD group (as outlined in Chapter 3) could be attributed to higher local-to-global interference is inconclusive, given that local-to-global interference was only marginally higher in the ASD group relative to the NT group. More musical training, but not musical perception ability, was related to better prosodic emotion recognition across the ASD and NT groups. Non-vocal pitch perception was a prominent correlate of emotion recognition across auditory stimulus types within the NT group. This highlights the cross-domain effects of pitch perception in auditory emotion recognition in typical development, while providing novel evidence for its role in sung emotion recognition. However, the association between non-vocal pitch perception and auditory emotion recognition was not evident within the ASD group. Nevertheless, this lack of association may explain why autistic individuals were less efficient in auditory emotion recognition, but did not result in poorer auditory emotion recognition accuracy, suggesting that autistic individuals may rely on alternative cues to achieve comparable accuracy to NT individuals. The lack of association between pitch perception and auditory emotion recognition speed in both groups suggests that other factors may be more responsible for auditory emotion recognition atypicalities (i.e., slower recognition as observed in Chapter 3) in ASD. None of these auditory perception measures were associated with implicit emotion priming, suggesting that they may play a specific role only in the explicit emotion processing of speech prosody and song. The current data did not provide support for the alexithymia hypothesis, such that alexithymic traits did not contribute to emotion recognition over and above the effects of autistic traits. This suggests that emotion recognition impairments in the ASD group could not be attributed to co-occurring alexithymia, though these findings should be interpreted with caution given the high conflation between alexithymic traits and ASD diagnosis in the current sample. Altogether, these results shed light on the differences in the underlying nature of emotion recognition between autistic and NT individuals which may not be easily depicted from task performance, and thus complementing findings reported in previous chapters.

Chapter 6: General Discussion

When discussing the possible sources of the highly mixed findings in the literature on emotion processing in autism spectrum disorder (ASD), little attention has been paid to the role of communicative domain and level of processing (cf. Harms, Martin, & Wallace, 2010; Lozier, Vanmeter, & Marsh, 2014; Uljarevic & Hamilton, 2013; but see Nuske, Vivanti, & Dissanayake, 2013). The lack of research into domains other than human faces (such as nonhuman faces, speech prosody, and music/song) at both implicit and explicit levels of emotion processing, together with the variations in population characteristics and experimental designs across studies, have undoubtedly led to difficulties in creating a cohesive picture of the *general* emotion processing ability among autistic individuals. This thesis was, therefore, set out to fill these gaps in the literature and to enhance understanding of how, and the extent to which, emotion processing in autistic individuals differs from that in neurotypical (NT) individuals across development from a multi-domain and -processing-level perspective.

Various methodologies were used to address four main research questions in this thesis, with regards to (i) whether emotion processing ability generalises across different communicative domains in ASD, (ii) whether emotion processing ability generalises across different levels of processing in ASD, (iii) whether the developmental trajectory of emotion processing differs between ASD and NT development, and (iv) whether the relationship between emotion processing and several related correlates differed between ASD and NT development.

Here, Section 6.1 summarises and integrates findings from the main studies to address the four research questions outlined above and discusses how these findings fit in the current literature; Section 6.2 evaluates the implications of the results and their contributions to the field of emotion processing in ASD and NT development; Section 6.3 highlights general limitations of this research and outlines future directions of research needs arising from the studies within this thesis; and lastly, Section 6.4 provides concluding remarks of this thesis.

6.1. Summary of findings

6.1.1. The ability to process emotions generalises across communicative domains in ASD

This thesis sought to answer the question of whether emotion processing ability generalises across domains (i.e., domain-general) or whether it is specific to certain domain(s) (i.e., domain-specific) in ASD, when the same participants were involved under the same study paradigms in Chapters 3 and 4, respectively.

In a forced-choice emotion recognition task, Chapter 3 showed that participants with ASD were able to recognise emotions from prototypical expressions in human faces, face-like objects, speech prosody, and song just as accurately as their NT participants. This general ability to accurately recognise emotions across domains, while corroborating previous findings in the nonhuman face and music domains (Heaton et al., 1999; Quintin et al., 2011; Rosset et al., 2008), appears to stand in contrast to the general view of impaired human facial and prosodic emotion recognition accuracy (Doi et al., 2013; Eack et al., 2015; Griffiths et al., 2019; Pelphrey et al., 2002; Schelinski & von Kriegstein, 2019; L. J. Taylor et al., 2015). Several possibilities relating to differences in study methodology may explain this discrepancy, such that the present emotion recognition study (i) did not require tolerance to variations between target and test expressions as it would on same-different tasks (cf. Greimel et al., 2014; see Harms et al., 2010 for a discussion), (ii) used monosyllabic prosodic stimuli extracted from the final position of sentences that could potentially contain richer emotional content and require less integration of global prosodic features compared to multisyllabic utterances, (iii) employed a smaller set of emotions (cf. Berggren et al., 2016; Eack et al., 2015; Sawyer et al., 2012), and/or (iv) allowed unlimited stimulus presentation and response times (cf. Kliemann et al., 2012; Oerlemans et al., 2014), all of which suggests that the present emotion recognition task likely encompasses lower task demands that can obscure differences that might exist between the ASD and NT groups. This limitation of the present study will be further discussed in Section 6.3 to provide a guide for future research.

Despite the intact emotion recognition accuracy, participants with ASD were significantly slower and less efficient in recognising emotions across domains compared to NT participants. This observation although is in line with previous findings in the human face and speech prosody domains (Greimel et al., 2014; Ketelaars et al., 2016; Sawyer et al., 2012), it contradicts with the comparable recognition speed observed between autistic and NT individuals in previous studies on nonhuman faces (namely, cartoon faces) and music (namely, instrumental music) (Miyahara et al., 2007; Quintin et al., 2011). The discrepancy in findings may be explained by the differences in the stimuli used between these previous studies and the present emotion recognition study. For instance, given the moderating effects of perceptual complexity on emotion recognition speed of iconic faces (Kendall et al., 2016), the face-like object emotional stimuli used in the present study, which were perceptually more complex and displayed higher contrasts, may have been a more sensitive medium in detecting group differences in comparison to the cartoon stimuli used in the previous study (Miyahara et al., 2007). With regards to the music domain, the discrepant findings for emotion recognition speed between the previous study (Quintin et al., 2011) and the present emotion recognition study could perhaps be due to differences in (i) stimulus lengths, such that emotions had to be identified after hearing the stimuli for 500ms in the present study compared to those lasting for ≥ 7s in the previous study and (ii) musicality of the stimuli, such that autistic individuals may have been less inclined to orient to sung stimuli in the present study compared to instrumental stimuli in the previous study, given that atypical processing of vocal stimuli has been reported in ASD (Bidet-Caulet et al., 2017; Gervais et al., 2004).

In a cross-modal affective priming task, Chapter 4 showed that the extent to which emotion recognition in human faces and face-like objects was primed by emotions conveyed in speech prosody and song did not differ between participants with ASD and NT participants across accuracy, speed, and efficiency. According to the spreading of activation account, affective priming is thought to be a consequence of the preactivation of emotionally congruent representations in the conceptual network when the preceding prime stimulus is implicitly processed, and thereby facilitating the encoding of subsequent congruent targets (Collins & Loftus, 1975; De Houwer & Randell, 2004; Hermans et al., 1994). As affective priming of speech prosody and song has not been studied previously in ASD, this intact ability to implicitly process emotions from auditory inputs to prime subsequent emotional judgment in ASD is a novel finding, which again was found to generalise across domains within the auditory modality.

Although no domain-specific impairments were observed on the priming task, discrepant patterns of the weighting between speech prosody and song were observed between the ASD and NT groups. Specifically, prosodic but not sung emotions implicitly influenced subsequent emotional judgment of human facial expressions across all age groups in the NT group, while the reverse was noted in the ASD group, with sung but not prosodic emotions implicitly influencing subsequent emotional judgment of human facial expressions across all age groups. Explanations for such discrepant patterns between the ASD and NT groups may pertain to the special role of speech in perception for NT individuals but not for autistic individuals, such that prosodic cues are prioritised over sung cues in the NT group. There is evidence to suggest that autistic individuals show lack of preference for and/or atypical initial orientation towards speech sounds (Čeponienė et al., 2003; Filipe et al., 2018; Kujala et al., 2005). These studies prompted the suggestion that ASD may be associated with atypical speech-specific orientation, which may have resulted in the reduced domain-specificity for speech stimuli during implicit emotion processing among autistic individuals.

In summary, autistic individuals seem able to accurately recognise emotions from human faces, face-like objects, speech prosody, and song, although they do it in a slower and less efficient way than NT individuals. Autistic individuals also showed implicit emotion priming of speech prosody and song on human faces and face-like objects comparable to their NT counterparts. Taken together with these findings, it could be argued that emotion processing ability is domain-general in ASD. That is, the intact recognition accuracy but slower recognition speed and poorer efficiency on the emotion recognition task, together with intact emotion priming across accuracy, speed, and efficiency on the cross-modal affective task, were not specific to particular domains. Although comparable performance was found between autistic and NT individuals, atypical weighting between prosodic and sung cues in implicit emotion priming was depicted in ASD. Specifically, the prioritisation of prosody over song in emotion processing was evident only in the NT group but not in the ASD group.

6.1.2. The ability to process emotions at the implicit and explicit levels is dissociated in ASD

Prior research has shown that the implicit and explicit emotion processing can be dissociated. Considering this is an under-investigated topic in ASD, particularly in the case of multiple domains, Chapters 3 and 4 were used to address the question of whether emotion processing ability generalises across the implicit and explicit levels of processing in ASD.

As outlined earlier, Chapter 4 found no difference between autistic and NT individuals in their emotion priming of prosodic and sung cues on emotion recognition in human faces and face-like objects, suggesting an intact ability to implicitly process emotions in speech prosody and song. In comparison to previous affective priming studies, the present findings are in line with those in Vanmarcke and Wagemans (2017), who found comparable priming of both coarse and fine emotional faces on subsequent valence judgment of faces between autistic and NT individuals, while the present study extended this observation to the priming of speech prosody and song. Conversely, the present findings are different from the findings by Kamio et al. (2006), who reported impaired priming of emotional faces on liking judgments in ideographs in individuals with high-functioning pervasive developmental disorder (HFPDD) relative to NT individuals. The discrepancy across the previous and present findings seem to suggest that implicitly activated emotional meaning can guide subsequent behaviours that are more conceptually related (e.g., emotional/valence judgment) to similar extents between autistic and NT individuals. By contrast, atypicality in ASD may relate to the particular case when subsequent behaviours are less conceptually related (e.g., liking judgment), in line with previous findings showing that autistic individuals rely less on emotional heuristics in decision-making tasks (De Martino et al., 2008).

The present finding of intact implicit emotion processing, however, seems to contradict numerous studies reporting atypical implicit processing of emotional prosody in ASD (Fan & Cheng, 2014; Lindström et al., 2018). Notably, in these previous studies, implicit responses to emotional stimuli were measured using neurophysiological methods during passive listening tasks (Fan & Cheng, 2014; Lindström et al., 2018), whereas implicit emotion processing was measured using behavioural methods through priming in the present study. Thus, the intact implicit emotion processing on the behavioural level in the present study does not necessarily counter the atypical implicit emotion processing on the neurophysiological level in these previous studies in ASD. In fact, this potential mismatch between behaviour and neural underpinnings may indicate the contributing role of compensation in the present behavioural results. Here, compensation in neurodevelopmental disorders is defined as the process of which improves behavioural presentation of the condition, despite persisting core deficit(s) at cognitive and/or neurobiological levels (Livingston & Happé, 2017). To achieve typical task performance, it is proposed that ‘compensated’ participants may recruit additional resources, such as through neural compensation (see Livingston & Happé, 2017 for a detailed discussion). Neural compensation may be evident in additional ‘neural effort’ required from the same neural network used by neurotypicals: for example, although similar neural networks were engaged during emotion processing of music in autistic and NT groups, significantly greater activation in left dorsolateral prefrontal cortex and left rolandic operculum/insula in response to happy contrasted with sad music was found in autistic individuals compared to NT individuals (Gebauer, Skewes, Westphael, et al., 2014). Alternative to the compensation hypothesis, it is also possible that reduced neural activation was not at a threshold for poor behavioural performance: for example, although hypoactivation of the premotor area and the left anterior insula especially in response to happy music excerpts was observed for autistic participants relative to NT participants, behavioural emotional ratings of music pieces overall indicated no differences between the two groups (Caria et al., 2011). These findings highlight the importance of the combined use of behavioural and neurocognitive measures (e.g., EEG, fMRI), which will help to disentangle the disparity between previous electrophysiological findings and the present behavioural results with regards to intact behavioural but atypical neural implicit processing of emotional prosody. To date, no fMRI studies have been conducted to investigate brain activation patterns during implicit emotion processing (priming) of auditory stimuli in ASD. Such investigation will provide insights into whether the present behavioural results were reflective of neural compensation or low sensitivity of task to reflect poor neural encoding.

Findings from Chapter 3 indicated that the ASD group showed comparable accuracy to the NT group in recognising emotions from human faces, face-like objects, speech prosody, and song. However, the ASD group was significantly slower and less efficient than the NT group in doing so. The impaired recognition speed and efficiency in ASD may be attributed to the use of more deliberative processing strategy (J. B. Grossman et al., 2000; Livingston & Happé, 2017; Rutherford & McIntosh, 2007), greater amount of cognitive resources needed for processing emotional stimuli (Bhatara et al., 2010; Čeponienė et al., 2003; Gebauer, Skewes, Hørlyck, et al., 2014; Guillon et al., 2016; Lepistö et al., 2005; Pavlova et al., 2017), and/or generally slower processing speed (Haigh et al., 2018; Hedvall et al., 2013; Oliveras-Rentas et al., 2012; Russo-Ponsaran et al., 2015). Thus, the results presented in Chapter 3 highlighted intact accuracy but impaired speed (i.e., slower) and efficiency (i.e., lower) of explicit emotion recognition in ASD, an observation that has also been noted in previous studies (Ketelaars et al., 2016; Lepistö et al., 2005; Waddington et al., 2018).

Collating findings from the emotion recognition and cross-modal affective priming studies presented in Chapters 3 and 4, ASD appears to be associated with impaired explicit but spared implicit emotion processing. This observation provides supporting evidence for a dissociation between early automatic and late conscious emotion processing (Castner et al., 2007; Kamio et al., 2006; Padovan et al., 2002; Roux et al., 2010; Suslow et al., 2003; Wagenbreth et al., 2016; Wieser et al., 2006), which may reflect the fact that the two levels of processing involve distinct neural networks (B. T. Gold et al., 2006; Habel et al., 2007; Rossell et al., 2003).

6.1.3. The developmental trajectory of emotion processing is comparable between ASD and NT at the explicit but not implicit levels

The systematic review's findings presented in Chapter 2 showed that age was a significant predictor of the severity of impairments in emotion recognition accuracy for the six-emotion composite. Notably, these results represented the age effects where studies investigating different domains were pulled together. Due to insufficient data available, further analyses to examine the age effects on emotion recognition accuracy separately for each domain were precluded. Additionally, the effects of age on emotion processing at the implicit level in ASD have also not been investigated. Thus, Chapters 3 and 4 addressed the scarcity of research by examining the developmental trajectories of emotion processing across domains and processing levels in ASD versus NT development.

Chapter 3 showed that age-related improvements were observed regardless of group, which varied by domain, emotion, and performance index. Importantly, no interacting effects of age group and diagnostic group were observed across domains and emotions, hence providing no evidence for different developmental trajectories between the ASD and NT groups. These findings appear to contradict those by Rump et al. (2009), who reported age-related improvement in facial emotion recognition only in the NT group but not in the ASD group. However, in Rump et al. (2009), the divergent trajectories were found for the recognition of dynamic facial expressions which varied in their subtlety (i.e., from low to high) under brief exposure times, unlike the present study which presented static facial stimuli varying from intermediate to high intensities without restriction on presentation times. It could be the case that the divergent trajectories are reflective of the continued refinement of emotion recognition skills for discriminating between more subtle expressions efficiently throughout adolescence in NT individuals, which is not seen in autistic individuals (Rump et al., 2009). These refined skills might not have been required to the same extent for the recognition of prototypical expressions in the present study, and thus any subtle differences between the older age groups could have been undetected.

In addition, the findings of Chapter 3 do not seem to correspond to the review's findings (Chapter 2), which showed that age was a significant predictor of the magnitude of group differences as aforementioned. It is noteworthy that the age effect observed in the review was only seen for the six-emotion composite but not for the individual emotions. This is perhaps due to the fact that individual-emotion data included in the review came from studies exploring different numbers and/or combinations of emotions. For example, in Akechi et al. (2010), anger was examined alongside one other emotion only (i.e., fear), whereas in Berggren et al. (2016), anger was examined alongside another five emotions (i.e., fear, happiness, sadness, disgust, and surprise). Data contributing to the individual-emotion analyses may, therefore, have represented different levels of task difficulty than those contributing to the composite analyses, which only included data for the recognition of all six emotions combined. Indeed, the inclusion of more emotions has been shown to moderate emotion recognition performance in ASD relative to NT development (M. Zhang et al., 2021). This is likely due to the increased cognitive demands required for distinguishing between all six basic emotions (Hoekert et al., 2007; M. Zhang et al., 2021). In essence, the present emotion recognition study (Chapter 3) examining only four emotions (anger, fear, happiness, and sadness to ensure representativeness of expressions conveyed across domains) may have been easier than other studies examining more emotions (e.g., Heaton et al., 2012; Hobson et al., 1989; Pelphrey et al., 2002; Philip et al., 2010; Rump et al., 2009). Thus, it is plausible that studies including more emotions are more sensitive in detecting a divergence in the developmental trajectory of emotion recognition. In other words, it may be the case where NT individuals develop more sophisticated categorisation skills for distinguishing between a greater number of emotions (e.g., > 4) over the developmental course, which draws apart their performance from that by autistic individuals to a greater extent as age increases.

Chapter 4 showed that, at the implicit level, the priming effects of prosodic and sung emotions on facial emotion recognition efficiency decreased with age, regardless of group. While this may not necessarily imply poorer implicit emotion processing in older participants, it was speculated that these age effects may be due to age-related factors influencing the degree of cross-modal emotional transfers from the auditory primes onto the visual targets. These speculations relate to the possibility of a developmental shift from an auditory to visual dominance during multisensory processing (Hirst, Cragg, et al., 2018; Nava & Pavani, 2013;

Robinson & Sloutsky, 2004) and/or improved inhibition of automatic responses towards task-irrelevant emotional content of the auditory primes as age increases (Herba et al., 2006; Passarotti et al., 2009; Ravindranath et al., 2020). It is plausible that automatic responses to emotional information emerge from a young age (Lobaugh et al., 2006), but the degree of their influence on subsequent emotional judgment becomes more dependent on their relevance to task demands with increased age. These findings, nevertheless, provided novel evidence that the developmental trajectory of implicit emotion priming of prosodic and sung emotions does not seem to differ between autistic and NT individuals.

In summary, results from Chapters 3 and 4 showed that the developmental trajectory of implicit and explicit emotion processing is largely comparable between autistic and their NT counterparts. This is the specific case for the processing of prototypical expressions conveying one of the four basic emotions across human faces, face-like objects, speech prosody, and song as investigated in the present work. At the implicit level, it should be noted that the discrepant patterns for emotion priming of prosody versus song between the two groups became more apparent with increased age (as already discussed in Section 6.1.1). That is, an age-related decline for sung but not prosodic emotional cues was seen during implicit emotion priming in NT individuals, whereas the reverse was seen in autistic individuals. This particularly highlights the special role of prosodic emotional cues in face-to-face conversations in NT individuals, where these cues may not be easily inhibited during implicit emotion priming even if adults are less susceptible to task-irrelevant interference. At the explicit level, emotion recognition was found to improve for both the ASD and NT groups over time, where autistic individuals remained comparatively slower and less efficient than their NT counterparts throughout development (as discussed in Section 6.1.1). Impairments in emotion recognition speed and efficiency may, therefore, not be merely reflective of any developmental delay, rather, these impairments may be persistent across the age span in ASD.

6.1.4. The underlying processes of explicit emotion recognition differ between ASD and NT

As raised in the review presented in Chapter 2, there is little empirical exploration of the processing strategies used to decode emotional expressions, as well as the relationship between co-occurring alexithymia and emotion processing in ASD. Chapter 5 was set out to examine a number of potential correlates and their relationships with explicit emotion recognition (assessed in Chapter 3) and implicit emotion priming (assessed in Chapter 4) to gain further understanding of the similarities and differences in the underlying nature of emotion processing between the two groups.

Focusing on the visual modality, cognitive processing style (indexed by local-to-global interference but not global advantage) was found to be significantly associated with the speed and efficiency of emotion recognition in human faces across the ASD and NT groups. Notably, when examining this association separately for each group, it only remained significant for the NT group but not the ASD group. It is speculated that the lack of association observed for the ASD group may be due to the homogenous datapoints in the local-to-global interference measure, where most datapoints narrowly distributed on the positive end of the measure (indicative of higher local-to-global interference). This narrow distribution of data may have, therefore, minimised its association with facial emotion recognition to be observed in the ASD group alone. Nevertheless, findings from previous research and the present work appear to converge on the notion regarding the relevance of global processing as the default perceptual strategy in facial emotion recognition. Namely, while better facial emotion recognition was related to increased preference for global responses in previous work by Gross (2005), the present work extended this by showing that better facial emotion recognition was related to fewer interruptions to global processing due to initial orientation towards local details. However, it remains inconclusive whether the slower and less efficient facial emotion recognition in the ASD group (as demonstrated in Chapter 3) could be attributed to higher local-to-global interference, given that local-to-global interference was only marginally higher in the ASD group relative to the NT group. Interestingly, no significant associations between cognitive processing style and emotion recognition of face-like object expressions were observed across nor within groups. This could be due to human faces being more special (compared to nonhuman faces) in visual perception (Akdeniz, 2020; Rosset et al., 2010), where interference-free global processing is particularly relevant for the recognition of human facial expressions, but not for face-like object expressions. While these findings provide novel evidence for a lack of association between cognitive processing style and emotion recognition of face-like objects, it remains uncertain what may be underlying the slower and less efficient emotion recognition of face-like objects in the ASD group (see Chapter 3).

Focusing on the auditory modality, musical perception ability was not associated with auditory emotion recognition both across and within the ASD and NT groups. Nevertheless, more musical training was found to be related to more accurate and efficient prosodic emotion recognition when the ASD and NT groups were pulled together, consistent with previous research (Correia et al., 2020; Farmer et al., 2020; Lima & Castro, 2011b; W. F. Thompson et al., 2004). This association was, however, not observed within either group separately, likely due to the low number of participants with extensive musical training when broken down by group. It was, therefore, not possible to infer whether the effects of musical training on prosodic emotion recognition are evident to similar extents between the two groups. Musical training, however, did not correlate with emotion recognition in song both across and within groups. One possible explanation for this nil association may relate to the differences in the musicality of the instrument learnt by musicians and the musicality of the stimuli presented. For instance, the majority of participants who had undertaken musical training in the present study practised on various types of instruments, with only a handful of participants who were trained in singing.

Their musical expertise may, therefore, have been less applicable to emotion recognition of transient cues in the singing voice. This speculation, however, appears to conflict with the association observed between musical training and emotion recognition in speech prosody and warrants further investigation with a larger sample of participants having more diverged musical training background.

With regard to pitch perception ability, it was found to be associated with more accurate and efficient emotion recognition in both prosody and song within the NT group. This observation across domains within the auditory modality provides supporting evidence that emotion processing in the two domains may involve overlapping underlying mechanisms (Escoffier et al., 2013). Conversely, no such associations were seen in the ASD group, which could be interpreted as that pitch information may not be as readily available for autistic individuals during auditory emotion recognition as opposed to NT individuals, in line with previous research (Schelinski & von Kriegstein, 2019; but see Globerson et al., 2015). While this lack of association may explain why emotion recognition efficiency in prosody and song was poorer in the ASD group (as shown in Chapter 3), it remains unclear why emotion recognition was slower in this group. Moreover, despite this lack of association, the ASD group performed just as accurately as the NT group, suggesting that autistic individuals may rely on alternative cues such as timbre/voice quality to achieve similar levels of accuracy in auditory emotion recognition among NT individuals. Taken together, as none of the auditory-related measures correlated with emotion recognition speed for prosody and song, it remains uncertain what may underlie the slower auditory emotion recognition in the ASD group relative to the NT group (as demonstrated in Chapter 3). Furthermore, none of these auditory perception measures were associated with implicit emotion priming of prosody and song in the two groups.

Chapter 5 did not provide supporting evidence for the alexithymia hypothesis (Bird & Cook, 2013). In particular, alexithymic traits did not contribute to emotion recognition impairments (i.e., slower speed) over and above the effects of autistic traits. This suggests that emotion recognition impairments in the ASD group could not be attributed to co-occurring alexithymia. It is, however, important to note that there was a high conflation between alexithymic traits and the ASD diagnosis in the current sample (i.e., the ASD group was distinctively characterised by significantly higher alexithymic traits). Neither alexithymic traits nor autistic traits predicted overall emotion recognition accuracy and the strength of emotion priming across performance indices. These nil associations complement the observation of comparable performance between groups in these aspects as demonstrated in Chapters 3 and 4.

Altogether, the findings presented in Chapter 5 showed that the underlying processing strategies for explicit emotion recognition appeared to differ between autistic and NT individuals, especially for pitch perception relevant for emotion recognition accuracy and efficiency for speech prosody and song, and perhaps also for cognitive processing style relevant for emotion recognition speed and efficiency for human faces. The underlying processes of emotion recognition in face-like objects remains unclear, as it did not correlate with cognitive processing style. Additionally, none of the correlates examined in this study represented an underlying role in implicit emotion priming for either group. Thus, it also remains uncertain whether the strategies underlying emotion processing at the implicit level differ between the ASD and NT groups. Finally, emotion processing differences between the two groups, namely the slower and less efficient emotion recognition in ASD (as shown in Chapter 3), could not be attributed to co-occurring alexithymia based on the current data.

6.2. Implications of findings and their contributions to the field

Evidence from separate studies in the existing literature appeared to show a dissociation of emotion processing ability across domains in ASD – that is, emotion processing impairments seemed to be confined to the human face domain (Lozier et al., 2014; Uljarevic & Hamilton, 2013) and/or speech prosody domain (Baker et al., 2010; Doi et al., 2013; Ketelaars et al., 2016; Schelinski & von Kriegstein, 2019), with a largely preserved ability observed for the nonhuman face domain (Brosnan et al., 2015; Rosset et al., 2008) and music domain (Heaton et al., 1999; Quintin et al., 2011). For the first time, emotion processing of these domains was evaluated simultaneously at the implicit (i.e., speech and song) and explicit levels (i.e., human faces, face-like objects, speech, and song), using the same participant sample of a wide age range (7-56 years), closely matched sets of stimuli, and a fixed number and combination of emotions (i.e., anger, fear, happiness, and sadness). In addition, the use of individual-level matching on age, gender, verbal and nonverbal ability, as well as the use of multiple performance indices (i.e., accuracy, speed, and efficiency), allowed for more reliable and sensitive comparisons to be made between autistic and NT individuals. Exerting greater control over variability across studies and confounding factors within studies, the research presented in this thesis provided evidence suggesting a generalised emotion processing ability across domains and emotions in ASD (i.e., impairments are not specific to certain domain(s) and/or emotion(s)), where there appears to be a dissociation between implicit and explicit levels of processing (i.e., impaired recognition speed and efficiency at the explicit level but spared priming at the implicit level).

The findings presented in this thesis have theoretical implications for the debates over the domain-generality versus domain-specificity of emotion processing in ASD and NT development (Borod et al., 2000; Bowers et al., 1993; Connolly et al., 2020; Lewis et al., 2016; Lima et al., 2013; Nuske et al., 2013; Peelen et al., 2010; K. R. Scherer & Scherer, 2011;

Schlegel et al., 2012). The present findings corroborated and extended previous research in showing a generalised emotion processing ability in ASD not only across the human face and speech prosody domains (Jones et al., 2011; Lerner et al., 2013; Philip et al., 2010), but also in the nonhuman face (in this case face-like objects) and music (in this case song) domains.

Findings from Chapters 3, 4, and 5, together, lent support to the notion that both domain-specific and domain-general factors contribute to emotion perception, with more recent work demonstrating that the two factors are organised within a hierarchical structure (i.e., comprising a supramodal emotion recognition factor superordinating over and above domain-specific factors; Connolly et al., 2020). With this in mind, autistic individuals appear to show impairments at the supramodal domain-general level of the hierarchy (i.e., slower and less efficient emotion recognition across all domains; as demonstrated in Chapter 3), as well as atypicalities in the subordinating domain-specific factors (i.e., a lack of association between pitch processing ability and auditory emotion recognition and perhaps a (marginally) atypical cognitive processing style relevant for facial emotion recognition; as demonstrated in Chapter 5). In addition, the present work provides insights into the plausibility of this emotion perception hierarchy representing the human face, voice and body domains in Connolly et al. (2020) to hold for additional domains such as nonhuman faces and song. The inclusion of additional domains to the hierarchy will enable closer inspection into how the different domains are weighted in this hierarchy, given the present findings suggest an advantage for speech over song in emotion processing in NT individuals that was not seen in autistic individuals (as shown in Chapter 4). It should be noted that it remains uncertain whether and how domain-general and domain-specific factors interact to influence the behavioural manifestation in ASD, provided that the atypicalities in the domain-specific factors did not always correspond to impairments observed at the domain-general level (i.e., intact auditory emotion recognition accuracy in ASD was seen despite its lack of association with pitch perception ability). Nonetheless, the research presented in this thesis went beyond considering emotion processing impairments in ASD as an all-or-nothing phenomenon; any emotion processing difficulties and/or differences in ASD could be understood within the context of this emotion perception hierarchy, while providing evidence that contributes to the domain-specificity versus domain-generality debates.

The research presented in this thesis also has implications for the debates regarding whether emotion processing impairments in ASD are specific to certain emotions. According to the amygdala theory of autism, dysfunction of the amygdala in ASD was hypothesised to result in poorer recognition of fear and other negative emotions specifically (Ashwin et al., 2006; Baron-Cohen et al., 2000; Howard et al., 2000). However, the findings from Chapters 3 and 4 provided evidence against this account and showed no specific impairments for the different emotions, while corroborating previous findings of general ability (or impairment) for all emotions in ASD (Baker et al., 2010; Boggs & Gross, 2010; Lindner & Rosén, 2006; Sawyer et al., 2012; Song et al., 2020; Waddington et al., 2018).

Despite the nil effects of individual emotion on group differences, this thesis advances our knowledge about the role of individual emotion in emotion recognition and their varying developmental trajectories in ASD and NT development. As summarised in Section 3.4.3, some emotions were better recognised than others depending on the domain for both groups, regardless of age. This implies that some emotions may provide more salient cues than others in a given domain (e.g., anger was better recognised through prosody and song, whereas fear was better recognised through faces and object). These findings further support the notion that different domains do not merely carry redundant information for emotion processing (App et al., 2011; Elfenbein & Ambady, 2002), underscoring the importance to flexibly process emotions from different sources to obtain more reliable impressions of others' emotional states. Moreover, the developmental trajectory of emotion recognition was found to vary across emotions depending on the domain (e.g., more protracted trajectories were found for the recognition accuracy of anger and fear in faces, sadness in objects, all emotions in speech, and fear in song as compared to other emotions in their given domain). By implementing a detailed investigation of the interacting effects of age, domain, and emotion, this thesis not only provided novel insights into the development of emotion recognition in ASD, but also in NT development, as, to the best of my knowledge, no previous studies have investigated emotion recognition across these three variables simultaneously in either population.

Some potential practical implications of the research presented in this thesis could also be illustrated. Emotion processing impairments were mainly observed at the explicit level. Recognising both the strengths and weaknesses of the emotion processing ability in ASD, it is important to note that autistic individuals were able to recognise emotions from prototypical expressions just as accurately as their NT counterparts when given sufficient time to process. Undoubtedly, the speed of emotion recognition is just as crucial as accuracy in order to respond to others' emotional states in an appropriate timely manner. Alternatively, without sufficient processing time in naturally dynamic social situations due to brief exposures to emotional cues, autistic individuals would likely produce less accurate emotional judgments as a trade-off. Interventions designed to strengthen the efficiency (comprising both accuracy and speed) of emotion recognition skills in ASD would be most beneficial in the long run. A number of interventions have previously been developed to enhance emotion recognition ability in ASD, with technology-based interventions being shown to significantly improve intervention effects to a greater extent than non-technology-based interventions (Zhi et al., 2021). The majority of these interventions have focused on targeting the human face domain – *FaceSay^{TM}*, *Ucime Emocii* (Learning Emotions), MIX, and *The Transporters* (Rice et al., 2015; Russo-Ponsaran et al., 2016; Vasilevska Petrovska & Trajkovski, 2019; Young & Posselt, 2012) – while less emphasis has been placed on other domains, including speech prosody – an alternative important cue to others' emotions in real-world scenarios (Matsumoto et al., 2012).

Showing that emotion recognition ability is generalised across domains (as demonstrated in Chapter 3), an improvement in one domain could perhaps result in improvement in another. Previous research has shown that autistic individuals have great interest in nonhuman faces such as cartoons (Anthony et al., 2013; Grelotti et al., 2005; Kuo et al., 2014; Spiker et al., 2012) and music such as songs (Allen et al., 2009b, 2009a; Brownell, 2002). Given the non-threatening nature of these stimuli in the sense of direct interpersonal interactions, they could be used as motivational tools for learning emotion recognition skills in a safe, accepting setting that could perhaps benefit skills for the human face and speech prosody domains in return. In combination with the benefits of technology-based interventions, such as reduced social withdrawal behaviours, anxiety, and fear in autistic individuals (Kinsella et al., 2017; Zhi et al., 2021), the additional use of motivating stimuli such as face-like objects and song may especially maximise outcomes of the interventions. Moreover, the use of didactic instructions, repeated practice, and increased presentation speeds together have been shown to be particularly useful for improving emotion recognition not only for accuracy but also for speed (Russo-Ponsaran et al., 2016), which was found to be a particular weakness of emotion recognition in autistic individuals (Chapter 3).

In addition, the intact priming effects of prosodic and sung emotions on emotional judgment of human faces could have potential implications for intervention designs. If playing congruent prosodic and sung emotional primes to enhance facial emotion recognition skills proves effective, it could form the basis of interventions to boost the decoding of facial expressions. In this regard, the positive intervention effects of song on facial emotion recognition have been demonstrated in previous work by Katagiri (2009). As discussed in Chapter 4, the reduced speech-orientation at an early stage of emotion processing may underlie social communication difficulties within interpersonal contexts. Future interventions may, thus, be guided to incorporate speech prosody in addition to song as a facilitating tool for facial emotion recognition, in order to strengthen the cross-modal transfers between the two domains. Taking advantage of the larger priming effects observed at a younger age (as observed in Chapter 4), such intervention would perhaps work best earlier in life – that is, prior to the start of the age-related decline in cross-modal influence of the auditory emotions (particularly of prosodic cues) on facial emotional judgment in ASD. However, this paradigm has only been shown to facilitate emotion recognition of human faces but not of face-like objects. Moreover, whether this paradigm could facilitate emotion recognition of speech prosody and song in reverse remains unknown. Future work examining whether and how these approaches are effective will further help practitioners develop interventions or learning shortcuts to improving the quality of social interactions in ASD.

6.3. Limitations and future directions

The studies presented throughout this thesis took care to explore the research aims of each chapter. However, there are limitations that must be acknowledged in order to develop on work in this area in future. Limitations and recommendations for future research specific to aspects of each study have been discussed throughout this work, and hence the more general points applicable to this thesis as a whole, as well as the challenges of conducting this kind of research, will be discussed below.

As ASD affects only 1-2% of the general population (Baron-Cohen et al., 2009; Brugha et al., 2016; Chiarotti & Venerosi, 2020), a challenging part of conducting research with ASD populations is to reach recruitment goals for an adequate sample size, even with a great amount of effort (Ahmed et al., 2020). Moreover, approximately half of the ASD population have a comorbid intellectual disability (Charman et al., 2011; Postorino et al., 2016), preventing them from participating in the studies of this thesis due to the cognitive demands of the tasks involved. The need to recruit and test an equal number of age-, gender-, verbal- and nonverbal-ability matched controls pose further challenges. The total number of 76 participants (ASD: *N* = 38; NT: *N* = 38) was not sufficient for detecting the Diagnostic group × Age group × Domain × Emotion interaction in the emotion recognition study (Section 3.2.1) nor the Diagnostic group × Age group × Prime type × Target type interaction in the cross-modal affective priming study (Section 4.2.1). According to the a priori power analysis, at least 266 participants (ASD: *N* = 133; NT: *N* = 133) and 158 (ASD: *N* = 79; NT: *N* = 79) would have been required to reach the desired power of 0.80 for the complex designs of the emotion recognition and cross-modal affective priming studies, respectively. This was, unfortunately, not feasible to achieve, given the difficulties associated with recruiting and testing autistic individuals under a limited timeframe. Although the emotion recognition study presented in Chapter 3 may have been limited in power to detect the four-way interaction outlined above, separate three-way interactions of Diagnostic group × Domain × Emotion and Age group × Domain × Emotion were, nonetheless, depicted significant. In addition, despite reduced power, the four-way interaction of Diagnostic group × Age group × Prime type × Target type was found to be significant for the accuracy measure in the cross-modal affective priming study presented in Chapter 4. Effect sizes were presented for all effects and interactions, by which the present results could be treated as preliminary and guide future large-scale studies in this area – particularly to scrutinise the current null finding of the four-way interaction in the emotion recognition task, which would provide further insights into whether the developmental trajectory differs between groups depending on the domain and emotion presented.

As briefly mentioned above, all participants with ASD who took part in the studies of this thesis did not have comorbid intellectual disability, which was determined by their verbal (i.e., standard score of ≥ 70 on the ROWPVT-4; Martin & Brownell, 2011) and nonverbal ability (i.e., percentile of ≥ 5th on the RSPM; Raven, 1983). Hence, the effects that were observed in the present work could not be generalised to individuals on the spectrum with lower verbal and/or nonverbal abilities. Although full-scale IQ, verbal IQ, and nonverbal IQ have been shown to be nonsignificant predictors of the severity of *relative* impairments in ASD (as demonstrated in Chapter 2 and also in Lozier et al., 2014 and Uljarevic & Hamilton, 2013), exploration of whether and how the emotion processing differs across the IQ spectrum among autistic individuals could further provide insights into the heterogeneity of this condition. Future research would benefit from adapting tasks that would appropriately accommodate variations in verbal and nonverbal ability of participants, in order to assess implicit and explicit emotion processing for the different communicative domains in a wider population of individuals across the autism spectrum within the same study.

Given the somewhat exploratory nature of the main experimental studies presented in Chapters 3 and 4 to investigate emotion processing in ASD from a multi-domain perspective, prototypical emotional expressions were used as stimuli that signified relatively comparable levels of recognition difficulty across domains. As such, the emotions expressed in the human face and speech prosody stimuli might have been easier to recognise in the present work compared to those in previous studies (as discussed in Section 6.1.1), in order to match the face-like object and song stimuli (which had not been investigated previously in the autism literature) on the level of recognition difficulty as closely as possible. Specifically, stimuli with moderately high recognition rates and intermediate intensity ratings based on the validation results (see Section 3.2.3) were selected to accommodate the wide age range of participants (i.e., to prevent floor effects in child participants) and to ensure sufficient saliency of these stimuli to be used as prime stimuli in the cross-modal affective priming task. It is possible that the use of prototypical expressions may have reduced the sensitivity of the present tasks in detecting subtle differences between the ASD and NT groups, not only for human faces and speech prosody, but also for face-like objects and song. Thus, largely intact performance was observed on both the recognition and priming tasks, where any group differences that existed might have been obscured. It is likely that as expressions become more subtle (i.e., less prototypical), they become more difficult to categorise – these subtle expressions may, in fact, represent more naturally occurring expressions. Experimental manipulation of the emotional intensities of stimuli across domains may facilitate task sensitivity in detecting subtle group differences, as well as establishing the threshold at which potential group differences emerge, as seen for the human face domain in the literature (Law Smith et al., 2010; Rump et al., 2009; S. Wang & Adolphs, 2017). Future studies should take the present findings on prototypical expressions further by including more subtle expressions to extend understanding of emotion processing ability in ASD.

The use of more subtle expressions (e.g., expressions at low intensity levels) may also increase sensitivity to detect differences in the developmental trajectories of emotion processing between autistic and NT individuals when more sophisticated categorisation skills are needed, as seen with human faces in Rump et al. (2009). On this note, it should be considered that, it was not possible to partition the three age groups further into groups with narrower age ranges, given the small sample size of the present study. Although this thesis attempted to compare the developmental trajectories of emotion processing between ASD and NT development, the use of three age groups comprising relatively wide age ranges may have precluded the detection of subtle differences or delays in the ASD group relative to the NT group across development. The high sample sizes of large-scale studies, as well as future longitudinal studies, are needed to provide more precise developmental changes in emotion processing ability across the age span in autistic individuals relative to NT individuals. This would allow interventions to be effectively implemented during the critical time window of the development of emotion processing in ASD.

One of the main aims of the priming study presented in Chapter 4 was to investigate how implicit, automatic processing of emotional cues manifests behaviourally through priming in ASD and NT development. However, the unexpected finding of stronger priming effects in children than in adults makes it speculative regarding the possibility of attentional influence on the priming effects observed. This seemingly poses an important limitation of the investigation of implicit emotion processing in this study, as under the classical view of executive control and automaticity, an automatic process is one that could not be stopped, altered, or avoided (Posner & Snyder, 1975; Schneider & Shiffrin, 1977). As discussed in Section 4.4.2, some tentative explanations were put forward for this finding. First, the reduced facilitatory effects of the primes in adults may be due to an age-related improvement in cognitive control over automatic responses to emotional content conveyed by the primes, which is irrelevant to the central task (Passarotti et al., 2009; Ravindranath et al., 2020). Secondly, the age-related differences in priming may be related to the different mechanisms recruited in affective priming by children versus adults. According to Posner and Snyder's (1975) two-process theory, automatic aspects of semantic priming lead to facilitation of congruent cues with no inhibition of incongruent cues (i.e., due to interference), whereas conscious anticipatory effects of response priming lead to facilitation of congruent cues as well as inhibition of incongruent cues (i.e., due to interference). Given that children's attention is more easily captured by salient stimuli or events in their surroundings (Farrant & Uddin, 2015; Wainwright & Bryson, 2002), response priming mechanisms (i.e., involving conscious strategic processes) may have been recruited for affective priming in the present study. Conversely, adults may have recruited semantic priming mechanisms (i.e., involving automatic processes), perhaps due to increased experience and/or knowledge that may have strengthened emotional concepts in the associative network for affective priming through spreading activation (Gawronski & Bodenhausen, 2005; Öhman et al., 2012). The priming effects, calculated as the difference between the recognition of congruent versus incongruent targets in this study presented in Chapter 4, may therefore have been larger in children, since both the effects of facilitation and inhibition would have been observed due to response priming. Thirdly, children and adults may both have recruited response priming mechanisms in affective priming, but adults were able to inhibit automatic responses to both congruent and incongruent cues, resulting in smaller priming effects in adults.

The current findings could not, however, confirm any of these postulations, as the implementation of neutral primes would be required to distinguish between facilitatory and inhibition effects, in order to infer which mechanisms had been used for affective priming in this study. More importantly, all three postulations would imply the possibility of attentional influence on the affective priming effects observed in this study to some degree, despite the careful consideration of inducing automatic emotion processing through manipulating the SOA to 200ms, a period that has been confirmed to reflect automatic, rather than conscious, emotion processing (Hermans et al., 1994, 2001; Herring et al., 2013). If indeed attentional influence had acted on the priming effects observed in this study presented in Chapter 4, these findings, nevertheless, may be most compatible with the more recent views under the refined theories of automaticity, particularly the attentional sensitization model of unconscious cognition (Kiefer, 2007, 2012; Kiefer & Martens, 2010; Moors & De Houwer, 2006; Naccache et al., 2002; O. Neumann, 1990). In contrast to the classical views of executive control and automaticity (Moors & De Houwer, 2006; Posner & Snyder, 1975; W. Schneider & Shiffrin, 1977), these refined theories posit that automatic processing depends on a configuration of the cognitive system by attention and task sets. That is, executive control factors such as attention, intention, and task sets amplify corresponding unconscious processing streams toward optimization of task performance by facilitating task-relevant unconscious processes, while attenuating task-irrelevant unconscious processes. Under the postulations of these refined theories, automatic processing is under executive control to some extent (Kiefer & Martens, 2010; Moors & De Houwer, 2006; Naccache et al., 2002; O. Neumann, 1990). Accordingly, the priming effects observed in the study presented in Chapter 4 may be best interpreted as an implicit process that is modulated by executive control. The postulated modulatory role of executive control over emotional interference from the primes (which are task-irrelevant regardless of its congruence to the targets) may prompt further discussion over the interplay between domain-general and domain-specific processes in implicit emotion processing. Related to this, previous neuroimaging studies have revealed inconsistent results regarding the neural systems that mediate emotional interference: some studies reported that a domain-general cognitive control network is engaged in both emotional and nonemotional interference processing (e.g., the dorsal anterior cingulate cortex and the lateral prefrontal cortex; Chechko et al., 2012; Chen et al., 2018; Torres-Quesada et al., 2014; Xu et al., 2016), while other studies reported a domain-specific network during interference processing in emotional but not nonemotional contexts (e.g., the amygdala and rostral anterior cingulate cortex; Etkin et al., 2011; Mitchell et al., 2003). Drawing on the postulation that the perception of socioemotional information entails both domain-general processes (e.g., to respond to salient cues regardless of its social relevance) and domain-specific processes (e.g., to facilitate response to cues given its social relevance) (as discussed in Section 1.3), it is possible that the seemingly conflicting findings of similar versus different brain networks involved in emotion and nonemotion interference processing may reflect a complementary interplay between domain-general and domain-specific processes. Nonetheless, the role of domain-general mechanisms underlying emotion interference processing provides a plausible explanation for why an age-related decline for the implicit influence of emotional primes was seen in the present study, where the reduced smaller priming effects in adults may be due to improved domain-general cognitive control processes across development (Passarotti et al., 2009; Ravindranath et al., 2020). Future studies should address the following in order to better understand the processes of implicit emotion priming in ASD and NT development: (i) to examine whether the present findings were confounded by attentional influences through distinguishing between facilitatory and inhibition effects, (ii) to further scrutinise the age-related differences in affective priming, such as whether automatic mechanisms could be recruited for affective priming in children, as affective priming is not well-understood from a developmental perspective; and (iii) if attentional influences are present under the paradigm with an SOA of 200ms, to control for the amount of attention devoted to affective priming through stricter manipulation of the SOA.

Finally, despite the initial attempts to outline the behavioural manifestations of emotion processing in autistic individuals in this thesis, there is a need for future work to take these findings further by demonstrating real-world applications of the similarities and differences in emotion processing between autistic and NT individuals. An important avenue for future research is to establish whether and to what extent do these differences translate to aversive socio-emotional reciprocity in natural social situations, which could perhaps in turn affect individuals' social well-being. In this sense, it is important to denote what these differences mean for autistic individuals, in order to offer more targeted interventions or learning shortcuts that would make a difference and contribute to more fulfilling life experiences for this population. On this note, an important observation in this thesis is that the underlying processes of emotion recognition, appear to differ between groups, specifically the reliance on pitch cues. Nonetheless, autistic individuals were able achieve similar levels of accuracy to NT individuals in auditory emotion recognition. Future research should also continue to identify these compensatory and/or potentially learned strategies that contribute to positive consequences as such, in hope to provide insights into the optimal contexts for autistic individuals to accomplish successful social interactions in daily life.

6.4. Concluding remarks

This thesis centres on the understanding of implicit and explicit emotion processing across different nonverbal communicative domains, as well as their underlying processes, in autistic and NT individuals throughout development. The studies presented in this thesis demonstrated that (i) emotion processing ability generalises across domains including human faces, face-like objects, speech prosody, and song, (ii) implicit and explicit emotion processing abilities in ASD may be dissociated, with impaired (i.e., slower and less efficient) explicit emotion recognition and spared implicit emotion priming in ASD, (iii) the developmental trajectory of implicit and explicit emotion processing is largely comparable in ASD and NT development, except that the priming of prosodic emotional cues becomes less important with age in ASD but not in NT, (iv) the underlying processes of explicit emotion recognition appear to differ between autistic and NT individuals with respect to pitch perception in prosodic and sung emotion recognition and perhaps cognitive processing style in facial emotion recognition, while those of implicit emotion priming of prosody and song could not be attributed to any of the measures explored in this thesis for either group, including musical training, musical perception, and pitch perception, and (v) emotion processing differences between autistic and NT individuals could not be explained by co-occurring alexithymia above and beyond autistic traits. These findings have shed light on the behavioural profile of emotion processing ability in ASD, which has led to a number of important theoretical and practical implications. The studies presented in this thesis are first of its kind and are thus somewhat exploratory; while these findings provide preliminary evidence for the emotion processing ability of autistic individuals from a multi-domain perspective, prospective research should attempt to address the limitations of the current studies. Importantly, this thesis showed that rather than focusing on emotion processing impairments in ASD as an all-or-nothing phenomenon, it may be useful to understand difficulties as well as differences in ASD related to both general (applicable to all domains) and specific (relevant to specific domains) aspects of emotion processing. It is hoped that the current thesis will help practitioners develop more targeted interventions or learning shortcuts to improving emotion processing skills that may facilitate better function and quality of social interactions in ASD.

References

Abràmoff, M. D., Magalhães, P. J., & Ram, S. J. (2004). Image processing with imageJ. *Biophotonics International*, *11*(7), 36–41. https://doi.org/10.1201/9781420005615.ax4

Adams, J., Hillier-Brown, F. C., Moore, H. J., Lake, A. A., Araujo-Soares, V., White, M., & Summerbell, C. (2016). Searching and synthesising "grey literature" and "grey information" in public health: Critical reflections on three case studies. *Systematic Reviews*, *5*(1), 1–11. https://doi.org/10.1186/s13643-016-0337-y

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews*, *1*(1), 21–62. https://doi.org/10.1177/1534582302001001003

Aguert, M., Laval, V., Le Bigot, L., & Bernicot, J. (2010). Understanding expressive speech acts: The role of prosody and situational context in french-speaking 5- to 9-year-olds. *Journal of Speech, Language, and Hearing Research*, *53*(6), 1629–1641. https://doi.org/10.1044/1092-4388(2010/08-0078)

Ahmed, K. L., Simon, A. R., Dempsey, J. R., Samaco, R. C., & Goin-Kochel, R. P. (2020). Evaluating two common strategies for research participant recruitment into autism studies: Observational study. *Journal of Medical Internet Research*, *22*(9). https://doi.org/10.2196/16752

Aiken, M. (2019). An updated evaluation of Google Translate accuracy. *Studies in Linguistics and Literature*, *3*(3), 253–260. https://doi.org/10.22158/sll.v3n3p253

Akdeniz, G. (2020). Brain activity underlying face and face pareidolia processing: an ERP study. *Neurological Sciences*, *41*(6), 1557–1565. https://doi.org/10.1007/s10072-019-04232-4

Akechi, H., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2014). Neural and behavioural responses to face-likeness of objects in adolescents with autism spectrum disorder. *Scientific Reports*, *4*, 1–7. https://doi.org/10.1038/srep03874

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2010). The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: An ERP study. *Neuropsychologia*, *48*(10), 2841–2851. https://doi.org/10.1016/j.neuropsychologia.2010.05.026

Alegria, J., & Noirot, E. (1978). Neonate orientation behaviour towards human voice. *International Journal of Behavioral Development*, *1*(4), 291–312. https://doi.org/10.1177/016502547800100401

Allen, R., Davis, R., & Hill, E. (2013). The effects of autism and alexithymia on physiological and verbal responsiveness to music. *Journal of Autism and Developmental Disorders*, *43*(2), 432–444. https://doi.org/10.1007/s10803-012-1587-8

Allen, R., Hill, E., & Heaton, P. (2009a). “Hath charms to soothe...”: An exploratory study of how high-functioning adults with ASD experience music. *Autism*, *13*(1), 21–41. https://doi.org/10.1177/1362361307098511

Allen, R., Hill, E., & Heaton, P. (2009b). The subjective experience of music in autism spectrum disorder. *Annals of the New York Academy of Sciences*. https://doi.org/10.1111/j.1749-6632.2009.04772.x

Alonso-Recio, L., Martín, P., Rubio, S., & Serrano, J. M. (2014). Discrimination and categorization of emotional facial expressions and faces in Parkinson’s disease. *Journal of Neuropsychology*, *8*(2), 269–288. https://doi.org/10.1111/jnp.12029

Alshaban, F., Aldosari, M., Al-Shammari, H., El-Hag, S., Ghazal, I., Tolefat, M., Ali, M., Kamal, M., Abdel Aati, N., Abeidah, M., Saad, A. H., Dekair, L., Al Khasawneh, M., Ramsay, K., & Fombonne, E. (2019). Prevalence and correlates of autism spectrum disorder in Qatar: a national study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *60*(12), 1254–1268. https://doi.org/10.1111/jcpp.13066

Altgassen, M., Kliegel, M., & Williams, T. I. (2005). Pitch perception in children with autistic spectrum disorders. *British Journal of Developmental Psychology*, *23*(4), 543–558. https://doi.org/10.1348/026151005X26840

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

Ansari, D., & Karmiloff-Smith, A. (2002). Atypical trajectories of number development: A neuroconstructivist perspective. *Trends in Cognitive Sciences*, *6*(12), 511–516. https://doi.org/10.1016/S1364-6613(02)02040-5

Anthony, L. G., Kenworthy, L., Yerys, B. E., Jankowski, K. F., James, J. D., Harms, M. B., Martin, A., & Wallace, G. L. (2013). Interests in high-functioning autism are more intense, interfering, and idiosyncratic than those in neurotypical development. *Development and Psychopathology*, *25*(3), 643–652. https://doi.org/10.1017/S0954579413000072

App, B., McIntosh, D. N., Reed, C. L., & Hertenstein, M. J. (2011). Nonverbal Channel Use in Communication of Emotion: How May Depend on Why. *Emotion*, *11*(3), 603–617. https://doi.org/10.1037/a0023164

Arbib, M. A. (2013). *Language, music, and the brain: a mysterious relationship* (Vol. 10).

MIT Press.

Argstatter, H. (2016). Perception of basic emotions in music: Culture-specific or multicultural? *Psychology of Music*, *44*(4), 674–690. https://doi.org/10.1177/0305735615589214

Åsberg Johnels, J., Hovey, D., Zürcher, N., Hippolyte, L., Lemonnier, E., Gillberg, C., & Hadjikhani, N. (2017). Autism and emotional face-viewing. *Autism Research*, *10*(5), 901–910. https://doi.org/10.1002/aur.1730

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

Asperger, H. (1943). *Die " Autistischen Psychopathen " im Kindesalter*. 1–64. http://www.autismus-biberach.com/Asperger_Hans-_Autistischen_Psychopathen.pdf

Atkinson, A. P. (2009). Impaired recognition of emotions from body movements is associated with elevated motion coherence thresholds in autism spectrum disorders. *Neuropsychologia*, *47*(13), 3023–3029. https://doi.org/10.1016/j.neuropsychologia.2009.05.019

Atmaja, B. T., & Akagi, M. (2020). On the differences between song and speech emotion recognition: Effect of feature sets, feature types, and classifiers. *IEEE Region 10 Annual International Conference, Proceedings/TENCON*, *2020-Novem*, 968–972. https://doi.org/10.1109/TENCON50793.2020.9293852

Auyeung, B., Allison, C., Wheelwright, S., & Baron-Cohen, S. (2012). Brief report: Development of the adolescent empathy and systemizing quotients. *Journal of Autism and Developmental Disorders*, *42*(10), 2225–2235. https://doi.org/10.1007/s10803-012-1454-7

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The autism spectrum quotient: Children's version (AQ-Child). *Journal of Autism and Developmental Disorders*, *38*(7), 1230–1240. https://doi.org/10.1007/s10803-007-0504-z

Auyeung, B., Wheelwright, S., Allison, C., Atkinson, M., Samarawickrema, N., & Baron-Cohen, S. (2009). The children's empathy quotient and systemizing quotient: Sex differences in typical development and in autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *39*(11), 1509–1521. https://doi.org/10.1007/s10803-009-0772-x

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia. A group study of adults afflicted with a music-specific disorder. *Brain*, *125*(2), 238–251.

https://doi.org/10.1093/brain/awf028

Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The twenty-item Toronto Alexithymia scale-I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research*, *38*(1), 23–32. https://doi.org/10.1016/0022-3999(94)90005-1

Baker, K. F., Montgomery, A. A., & Abramson, R. (2010). Brief report: Perception and lateralization of spoken emotion by youths with high-functioning forms of autism. *Journal of Autism and Developmental Disorders*, *40*(1), 123–129. https://doi.org/10.1007/s10803-009-0841-1

Balkwill, L., & Thompson, W. F. (2016). A Cross-Cultural Investigation of the Perception of Emotion in Music : Psychophysical and Cultural Cues. *Music Perception*, *17*(1), 43–64. https://doi.org/10.2307/40285811

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

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication*, *46*(3–4), 252–267. https://doi.org/10.1016/j.specom.2005.02.016

Barak, B., & Feng, G. (2016). Neurobiology of social behavior abnormalities in autism and Williams syndrome. *Nature Neuroscience*, *19*(6), 647–655. https://doi.org/10.1038/nn.4276

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist*, *54*(7), 462–479. https://doi.org/10.1037/0003-066X.54.7.462

Baron-Cohen, S. (2008). Theories of the autistic mind. *The Psychologist*.

Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The Autism-Spectrum Quotient (AQ) - Adolescent version. *Journal of Autism and Developmental Disorders*, *36*(3), 343–350. https://doi.org/10.1007/s10803-006-0073-6

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 361–374. https://doi.org/10.1098/rstb.2002.1206

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

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *British Journal of Psychiatry*, *194*(6), 500–509. https://doi.org/10.1192/bjp.bp.108.059345

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient (EQ): 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., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The Autism-Spectrum Quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. In *Journal of Autism and Developmental Disorders* (Vol. 31, Issue 1, pp. 5–17). Plenum Publishing Corp. https://doi.org/10.1023/A:1005653411471

Bassili, J. N. (1978). Facial motion in the perception of faces and of emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(3), 373. https://doi.org/10.1037/0096-1523.4.3.373

Bassili, J. N. (1979). Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, *37*(11), 2049. https://doi.org/10.1037/0022-3514.37.11.2049

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Battery, A. I. T. (1994). Army Individual Test Battery. In *Manual of Directions and Scoring* (War Depart).

Batty, M., & Taylor, M. J. (2006). The development of emotional face processing during childhood. *Developmental Science*, *9*(2), 207–220. https://doi.org/10.1111/j.1467-7687.2006.00480.x

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

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, *8*(3), 129–135. https://doi.org/10.1016/j.tics.2004.01.008

Ben-Shachar, M., Makowski, D., & Lüdecke, D. (2020). *Compute and interpret indices of* *effect size. CRAN. R package.*

Ben-Yosef, D., Anaki, D., & Golan, O. (2017). Context processing in adolescents with autism spectrum disorder: How complex could it be? *Autism Research*, *10*(3), 520–530. https://doi.org/10.1002/aur.1676

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate : A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berggren, S., Engström, A. C., & Bölte, S. (2016). Facial affect recognition in autism, ADHD and typical development. *Cognitive Neuropsychiatry*, *21*(3), 213–227. https://doi.org/10.1080/13546805.2016.1171205

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

Berthoz, S., & Hill, E. L. (2005). The validity of using self-reports to assess emotion regulation abilities in adults with autism spectrum disorder. *European Psychiatry*, *20*(3), 291–298. https://doi.org/10.1016/j.eurpsy.2004.06.013

Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*, *25*(3–4), 399–410.

Bhatara, A., Babikian, T., Laugeson, E., Tachdjian, R., & Sininger, T. S. (2013). Impaired timing and frequency discrimination in high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*, 2312–2328. https://doi.org/10.1007/s10803-013-1778-y

Bhatara, A., Quintin, E.-M., Fombonne, E., & Levitin, D. J. (2013). Early sensitivity to sound and musical preferences and enjoyment in adolescents with autism spectrum disorders. *Psychomusicology: Music, Mind, and Brain*, *23*(2), 100–108. https://doi.org/10.1037/a0033754

Bhatara, A., Quintin, E.-M., Levy, B., Bellugi, U., Fombonne, E., & Levitin, D. J. (2010). Perception of Emotion in Musical Performance in Adolescents with Autism Spectrum Disorders. *Autism Research*, *3*(5), 214–225. https://doi.org/10.1002/aur.147

Bidet-Caulet, A., Latinus, M., Roux, S., Malvy, J., Bonnet-Brilhault, F., & Bruneau, N. (2017). Atypical sound discrimination in children with ASD as indicated by cortical ERPs. *Journal of Neurodevelopmental Disorders*, *9*(1), 1–10.

https://doi.org/10.1186/s11689-017-9194-9

Bigand, E., Vieillard, S., Madurell, F., Marozeau, J., & Dacquet, A. (2005). Multidimensional scaling of emotional responses to music: The effect of musical expertise and of the duration of the excerpts. *Cognition and Emotion*, *19*(8), 1113–1139. https://doi.org/10.1080/02699930500204250

Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia to the emotional symptoms of autism. *Translational Psychiatry*, *3*(7), 1–8. https://doi.org/10.1038/tp.2013.61

Birnboim, S. (2003). The Automatic and Controlled Information-Processing Dissociation: Is It Still Relevant? *Neuropsychology Review*, *13*(1), 19–31. https://doi.org/10.1023/A:1022348506064

Bishop, D. V. M. (1997). Cognitive neuropsychology and developmental disorders: Uncomfortable bedfellows. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *50*(4), 899–923. https://doi.org/10.1080/713755740

Blackstock, E. G. (1978). Cerebral asymmetry and the development of early infantile autism. *Journal of Autism and Childhood Schizophrenia*, *8*(3), 339–353. https://doi.org/10.1007/BF01539636

Blasi, A., Mercure, E., Lloyd-Fox, S., Thomson, A., Brammer, M., Sauter, D., Deeley, Q., Barker, G. J., Renvall, V., Deoni, S., Gasston, D., Williams, S. C. R., Johnson, M. H., Simmons, A., & Murphy, D. G. M. (2011). Early specialization for voice and emotion processing in the infant brain. *Current Biology*, *21*(14), 1220–1224. https://doi.org/10.1016/j.cub.2011.06.009

Boersma, P., & Weenink, D. (2004). Praat Version 4.2. 09. *Institute of Phonetic Sciences: Amsterdam*.

Boggs, K. M., & Gross, A. M. (2010). Cue Salience in Face Processing by High Functioning Individuals with Autism Spectrum Disorders. *Journal of Developmental and Physical Disabilities*, *22*(6), 595–613. https://doi.org/10.1007/s10882-010-9198-9

Bombari, D., Schmid, P. C., Schmid Mast, M., Birri, S., Mast, F. W., & Lobmaier, J. S. (2013). Emotion recognition: The role of featural and configural face information. *Quarterly Journal of Experimental Psychology*, *66*(12), 2426–2442. https://doi.org/10.1080/17470218.2013.789065

Bonnel, A., McAdams, S., Smith, B., Berthiaume, C., Bertone, A., Ciocca, V., Burack, J. A., & Mottron, L. (2010). Enhanced pure-tone pitch discrimination among persons with autism but not Asperger syndrome. *Neuropsychologia*, *48*(9), 2465–2475.

https://doi.org/10.1016/j.neuropsychologia.2010.04.020

Bonnel, A., Mottron, L., Peretz, I., Trudel, M., Gallun, E., & Bonnel, A. M. (2003). Enhanced pitch sensitivity in individuals with autism: A signal detection analysis. *Journal of Cognitive Neuroscience*, *15*(2), 226–235. https://doi.org/10.1162/089892903321208169

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

Borenstein, M., & Higgins, J. P. T. (2013). Meta-analysis and subgroups. *Prevention Science*, *14*(2), 134–143. https://doi.org/10.1007/s11121-013-0377-7

Borenstein, M., Higgins, J. P. T., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, *8*(1), 5–18. https://doi.org/10.1002/jrsm.1230

Borod, J. C., Pick, L. H., Hall, S., Sliwinski, M., Madigan, N., Obler, L. K., Welkowitz, J., Canino, E., Erhan, H. M., Goral, M., Morrison, C., & Tabert, M. (2000). Relationships among facial, prosodic, and lexical channels of emotional perceptual processing. *Cognition and Emotion*, *14*(2), 193–211. https://doi.org/10.1080/026999300378932

Botsch, R. E. (2011). *Chapter 12: Significance and measures of association*.

Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *41*(7), 847–857. https://doi.org/10.1111/1469-7610.00672

Bowers, D., Bauer, R. M., & Heilman, K. M. (1993). The nonverbal affect lexicon: theoretical perspectives from neuropsychological studies of affect perception. *Neuropsychology*, *7*(4), 433. https://doi.org/10.1037/0894-4105.7.4.433

Bradley, M. M., & Lang, P. J. (2000). Measuring emotion: Behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), *Cognitive Neuroscience of Emotion* (pp. 242-276). Oxford University Press.

Breitenstein, C., Van Lancker, D., & Daum, I. (2001). The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. *Cognition and Emotion*, *15*(1), 57–79. https://doi.org/10.1080/0269993004200114

Brennand, R., Schepman, A., & Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with Autism Spectrum Disorder. *Research in*

Autism Spectrum Disorders, *5*(4), 1567–1573. https://doi.org/10.1016/j.rasd.2011.03.002

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

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Attention, Perception, and Psychophysics*, *77*(5), 1465–1487. https://doi.org/10.3758/s13414-015-0882-9

Brosch, T., Pourtois, G., & Sander, D. (2010). The perception and categorisation of emotional stimuli: A review. *Cognition and Emotion*, *24*(3), 377–400. https://doi.org/10.1080/02699930902975754

Brosnan, M., Johnson, H., Grawmeyer, B., Chapman, E., & Benton, L. (2015). Emotion recognition in animated compared to human stimuli in adolescents with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(6), 1785–1796. https://doi.org/10.1007/s10803-014-2338-9

Brown, L. S. (2017). The influence of music on facial emotion recognition in children with autism spectrum disorder and neurotypical children. *Journal of Music Therapy*, *54*(1), 55–79. https://doi.org/10.1093/jmt/thw017

Brown, V. A. (2021). An introduction to linear mixed effects modeling in R. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–19. https://doi.org/10.1177/2515245920960351

Brownell, M. D. (2002). Musically adapted Social Stories to modify behaviours in students with Autism: Four case studies. In *Journal of Music Therapy* (Vol. 39, Issue 2, pp. 117–144). https://doi.org/10.1093/jmt/39.2.117

Bruce, V., Campbell, R. N., Doherty-Sneddon, G., Import, A., Langton, S., McAuley, S., & Wright, R. (2000). Testing face processing skills in children. *British Journal of Developmental Psychology*, *18*(3), 319–333. https://doi.org/10.1348/026151000165715

Brugha, T. S., Spiers, N., Bankart, J., Cooper, S. A., McManus, S., Scott, F. J., Smith, J., & Tyrer, F. (2016). Epidemiology of autism in adults across age groups and ability levels. *British Journal of Psychiatry*, *209*(6), 498–503. https://doi.org/10.1192/bjp.bp.115.174649

Bury, S. M., Jellett, R., Spoor, J. R., & Hedley, D. (2020). “It Defines Who I Am” or “It’s Something I Have”: What Language Do [Autistic] Australian Adults [on the Autism

Spectrum] Prefer? *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-020-04425-3

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences*, *14*(12), 534–541. https://doi.org/10.1016/j.tics.2010.09.007

Calder, A. J., Keane, J., Young, A. W., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527–551. https://doi.org/10.1037/0096-1523.26.2.527

Calvo, M. G., & Lundqvist, D. (2008). Facial expressions of emotion (KDEF): Identification under different display-duration conditions. *Behavior Research Methods*, *40*(1), 109–115. https://doi.org/10.3758/BRM.40.1.109

Calvo, M. G., & Marrero, H. (2009). Visual search of emotional faces: The role of affective content and featural distinctiveness. *Cognition and Emotion*, *23*(4), 782–806. https://doi.org/10.1080/02699930802151654

Calvo, M. G., & Nummenmaa, L. (2008). Detection of Emotional Faces: Salient Physical Features Guide Effective Visual Search. *Journal of Experimental Psychology: General*, *137*(3), 471–494. https://doi.org/10.1037/a0012771

Calvo, M. G., Nummenmaa, L., & Avero, P. (2008). Visual search of emotional faces eye-movement assessment of component processes. *Experimental Psychology*, *55*(6), 359–370. https://doi.org/10.1027/1618-3169.55.6.359

Capozzi, F., & Ristic, J. (2020). Attention AND mentalizing? Reframing a debate on social orienting of attention. *Visual Cognition*, *28*(2), 97–105. https://doi.org/10.1080/13506285.2020.1725206

Caria, A., Venuti, P., & De Falco, S. (2011). Functional and dysfunctional brain circuits underlying emotional processing of music in autism spectrum disorders. *Cerebral Cortex*, *21*(12), 2838–2849. https://doi.org/10.1093/cercor/bhr084

Carroll, N. C., & Young, A. W. (2005). Priming of emotion recognition. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *58*(7), 1173–1197. https://doi.org/10.1080/02724980443000539

Carton, J. S., Kessler, E. A., & Pape, C. L. (1999). Nonverbal decoding skills and relationship well-being in adults. *Journal of Nonverbal Behavior*, *23*(1), 91–100. https://doi.org/10.1023/A:1021339410262

CASP, U. K. (2017). Critical appraisal skills programme (CASP). *Qualitative Research Checklist*, *31*(13), 449.

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

Castner, J. E., Chenery, H. J., Copland, D. A., Coyne, T. J., Sinclair, F., & Silburn, P. A. (2007). Semantic and affective priming as a function of stimulation of the subthalamic nucleus in Parkinson's disease. *Brain*, *130*(5), 1395–1407. https://doi.org/10.1093/brain/awm059

Castro, S. L., & Lima, C. F. (2014). Age and musical expertise influence emotion recognition in music. *Music Perception*, *32*(2), 125–143. https://doi.org/10.1525/mp.2014.32.2.125

Celeghin, A., Mazzoni, N., & Mattavelli, G. (2020). Explicit and implicit emotion processing: Neural basis, perceptual and cognitive mechanisms. *Frontiers in Psychology*, *11*, 1–2. https://doi.org/10.1038/35072584

Centers for Disease Control and Prevention. (2020). Prevalence of autism spectrum disorder among children aged 8 Years-Autism and developmental disabilities monitoring network, 11 Sites, United States, 2016. *MMWR Surveillance Summaries*, *69*(4), 1–12. https://doi.org/10.15585/MMWR.SS6904A1

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

Charbonneau, G., Bertone, A., Lepore, F., Nassim, M., Lassonde, M., Mottron, L., & Collignon, O. (2013). Multilevel alterations in the processing of audio-visual emotion expressions in autism spectrum disorders. *Neuropsychologia*, *51*(5), 1002–1010. https://doi.org/10.1016/j.neuropsychologia.2013.02.009

Charman, T., Pickles, A., Simonoff, E., Chandler, S., Loucas, T., & Baird, G. (2011). IQ in children with autism spectrum disorders: Data from the Special Needs and Autism Project (SNAP). *Psychological Medicine*, *41*(3), 619–627. https://doi.org/10.1017/S0033291710000991

Chechko, N., Kellermann, T., Zvyagintsev, M., Augustin, M., Schneider, F., & Habel, U. (2012). Brain circuitries involved in semantic interference by demands of emotional and non-emotional distractors. *PLoS ONE*, *7*(5), 1–9. https://doi.org/10.1371/journal.pone.0038155

Chen, T., Becker, B., Camilleri, J., Wang, L., Yu, S., Eickhoff, S. B., & Feng, C. (2018). A domain-general brain network underlying emotional and cognitive interference processing: evidence from coordinate-based and functional connectivity meta-analyses. *Brain Structure and Function*, *223*(8), 3813–3840. https://doi.org/10.1007/s00429-018-1727-9

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

Chiarotti, F., & Venerosi, A. (2020). Epidemiology of autism spectrum disorders: A review of worldwide prevalence estimates since 2014. *Brain Sciences*, *10*(5). https://doi.org/10.3390/brainsci10050274

Chin, W. C., Li, H. M., Chao, K. Y., & Chang, H. L. (2019). Stressors experienced by children with autism spectrum disorder in Taiwan: Perspectives of children and their parents. *Nursing and Health Sciences*, *21*(2), 206–213. https://doi.org/10.1111/nhs.12583

Chittka, L., Skorupski, P., & Raine, N. E. (2009). Speed-accuracy tradeoffs in animal decision making. *Trends in Ecology and Evolution*, *24*(7), 400–407. https://doi.org/10.1016/j.tree.2009.02.010

Chowdhury, R., Sharda, M., Foster, N. E. V., Germain, E., Tryfon, A., Doyle-Thomas, K., Anagnostou, E., & Hyde, K. L. (2017). Auditory Pitch Perception in Autism Spectrum Disorder Is Associated With Nonverbal Abilities. *Perception*, *46*(11), 1298–1320. https://doi.org/10.1177/0301006617718715

Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., & Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *British Journal of Developmental Psychology*, *33*(2), 218–236. https://doi.org/10.1111/bjdp.12075

Chronaki, G., Wigelsworth, M., Pell, M. D., & Kotz, S. A. (2018). The development of cross-cultural recognition of vocal emotion during childhood and adolescence. *Scientific Reports*, *8*(1), 1–17. https://doi.org/10.1038/s41598-018-26889-1

Ciaramidaro, A., Bölte, S., Schlitt, S., Hainz, D., Poustka, F., Weber, B., Freitag, C., & Walter, H. (2018). Transdiagnostic deviant facial recognition for implicit negative emotion in autism and schizophrenia. *European Neuropsychopharmacology*, *28*(2), 264–275. https://doi.org/10.1016/j.euroneuro.2017.12.005

Clarivate Analytics. (2016). EndNote. *Version X8*.

Clark, T. F., Winkielman, P., & McIntosh, D. N. (2008). Autism and the extraction of emotion from briefly presented facial expressions: Stumbling at the first step of empathy. *Emotion*, *8*(6), 803–809. https://doi.org/10.1037/a0014124

Clausi, S., Iacobacci, C., Lupo, M., Olivito, G., Molinari, M., & Leggio, M. (2017). The role of the cerebellum in unconscious and conscious processing of emotions: A review. *Applied Sciences (Switzerland)*, *7*(5), 21–24. https://doi.org/10.3390/app7050521

Çoğaltay, N., & Karadağ, E. (2015). Introduction to meta-analysis. In E. Karadag (Ed.), *Leadership and Organizational Outcomes: Meta-Analysis of Empirical Studies* (Issue March, pp. 19–28). Springer International Publishing. https://doi.org/10.1007/978-3-319-14908-0

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Routledge). Academic press.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Collignon, O., Girard, S., Gosselin, F., Roy, S., Saint-Amour, D., Lassonde, M., & Lepore, F. (2008). Audio-visual integration of emotion expression. *Brain Research*, *1242*, 126–135. https://doi.org/10.1016/j.brainres.2008.04.023

Collignon, O., Girard, S., Gosselin, F., Saint-Amour, D., Lepore, F., & Lassonde, M. (2010). Women process multisensory emotion expressions more efficiently than men. *Neuropsychologia*, *48*(1), 220–225. https://doi.org/10.1016/j.neuropsychologia.2009.09.007

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Connelly, M., & Denney, D. R. (2007). Regulation of emotions during experimental stress in alexithymia. *Journal of Psychosomatic Research*, *62*(6), 649–656. https://doi.org/10.1016/j.jpsychores.2006.12.008

Connolly, H. L., Lefevre, C. E., Young, A. W., & Lewis, G. J. (2020). Emotion recognition ability: Evidence for a supramodal factor and its links to social cognition. *Cognition*, *197*, 104166. https://doi.org/10.1016/j.cognition.2019.104166

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). Emotional modulation of perception in Asperger's Syndrome. *Journal of Autism and Developmental Disorders*, *38*(6), 1072–1080. https://doi.org/10.1007/s10803-007-0485-y

Correia, A. I., Macgregor, C., Müllensiefen, D., & Schellenberg, E. G. (2020). Enhanced recognition of vocal emotions in individuals with naturally good musical abilities. *Emotion*. https://doi.org/10.1037/emo0000770

Cosmides, L. (1983). Invariances in the acoustic expression of emotion during speech. *Journal of Experimental Psychology: Human Perception and Performance*, *9*(6), 864–881. https://doi.org/10.1037/0096-1523.9.6.864

Cosmides, L., & Tooby, J. (2000). Evolutionary Psychology and the Propositional-attitudes. *Handbook of Emotions*, 2nd ed., pp. 91–115. http://link.springer.com/10.1007/978-94-007-2969-8

Coutinho, E., & Dibben, N. (2013). Psychoacoustic cues to emotion in speech prosody and music. *Cognition and Emotion*, *27*(4), 658–684. https://doi.org/10.1080/02699931.2012.732559

Critchley, H. D., Daly, E. M., Bullmore, E. T., Williams, S. C. R., Amelsvoort, T. Van, Robertson, D. M., Rowe, A., Phillips, M., Mcalonan, G., Howlin, P., Murphy, D. G. M., & Murphy, D. G. M. (2000). The functional neuroanatomy of social behaviour. *Brain*, *123*(11), 2203–2212. https://doi.org/10.1093/brain/123.11.2203

Crocker, A. F., & Smith, S. N. (2019). Person-first language: Are we practicing what we preach? *Journal of Multidisciplinary Healthcare*, *12*, 125–129. https://doi.org/10.2147/JMDH.S140067

Curtin, J. (2018). *lmSupport: Support for linear models (R package Version 2.9. 13)*.

D’Souza, D., Booth, R., Connolly, M., Happé, F., & Karmiloff-Smith, A. (2016). Rethinking the concepts of “local or global processors”: Evidence from Williams syndrome, Down syndrome, and Autism Spectrum Disorders. *Developmental Science*, *19*(3), 452–468. https://doi.org/10.1111/desc.12312

D’Souza, H., & Karmiloff-Smith, A. (2017). Neurodevelopmental disorders. *Wiley Interdisciplinary Reviews: Cognitive Science*, *8*(1–2), 1–10. https://doi.org/10.1002/wcs.1398

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., Andrew, L. A., & Davidson, R. J. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, *8*(4), 519–526. https://doi.org/10.1038/nn1421.Gaze

Darwin, C. (1872). The expression of emotions in man and animals. *London: Murray*, *11*, 1872.

Davidson, D., Hilvert, E., Misiunaite, I., Kerby, K., & Giordano, M. (2019). Recognition of facial emotions on human and canine faces in children with and without autism spectrum disorders. *Motivation and Emotion*, *43*(1), 191–202. https://doi.org/10.1007/s11031-018-9736-9

Davitz, J. R., & Beldoch, M. (1964). *The communication of emotional meaning*. McGraw-Hill.

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, *7*(10), 460–467. https://doi.org/10.1016/j.tics.2003.08.014

De Houwer, J., Hermans, D., & Eelen, P. (1998). Affective and Identity Priming with Episodically Associated Stimuli. *Cognition and Emotion*, *12*(2), 145–169. https://doi.org/10.1080/026999398379691

De Houwer, J., & Randell, T. (2004). Robust affective priming effects in a conditional pronunciation task: Evidence for the semantic representation of evaluative information. *Cognition and Emotion*, *18*(2), 251–264. https://doi.org/10.1080/02699930341000022

De Martino, B., Harrison, N. A., Knafo, S., Bird, G., & Dolan, R. J. (2008). Explaining enhanced logical consistency during decision making in autism. *Journal of Neuroscience*, *28*(42), 10746–10750. https://doi.org/10.1523/JNEUROSCI.2895-08.2008

De Martino, B., Kalisch, R., Rees, G., & Dolan, R. J. (2009). Enhanced processing of threat stimuli under limited attentional resources. *Cerebral Cortex*, *19*(1), 127–133. https://doi.org/10.1093/cercor/bhn062

De Sonneville, L. M. J., Verschoor, C. A., Njiokiktjien, C., Op her Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: Speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology*, *24*(2), 200–213. https://doi.org/10.1076/jcen.24.2.200.989

Deeks, J. J., Altman, D. G., & Bradburn, M. J. (2001). *Chapter 15. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. Systematic reviews in health care: meta-analysis in context*. London: BMJ Publishing Group.

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2019). Analysing data and undertaking meta-analyses. *Cochrane Handbook for Systematic Reviews of Interventions*, 241–284.

Degner, J. (2011). Affective priming with auditory speech stimuli. *Language and Cognitive Processes*, *26*(10), 1710–1735. https://doi.org/10.1080/01690965.2010.532625

DeKosky, S. T., Heilman, K. M., Bowers, D., & Valenstein, E. (1980). Recognition and discrimination of emotional faces and pictures. *Brain and Language*, *9*(2), 206–214.

https://doi.org/10.1016/0093-934X(80)90141-8

Del Re, A. C. (2013). compute. es: Compute effect sizes. *R Package Version*, 0–2.

den Houting, J. (2019). Neurodiversity: An insider's perspective. *Autism*, *23*(2), 271–273. https://doi.org/10.1177/1362361318820762

Derntl, B., Seidel, E. M., Kainz, E., & Carbon, C. C. (2009). Recognition of emotional expressions is affected by inversion and presentation time. *Perception*, *38*(12), 1849–1862. https://doi.org/10.1068/p6448

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

Desimone, R., & Duncan, J. (1995). Neural Mechanisms of Selective Visual Attention. *Annual Review of Neuroscience*, *18*, 193–222. https://doi.org/10.1146/annurev-psych-122414-033400

Desmet, P. (2003). Measuring emotions: Development and application of an instrument to measure emotional responses to products. In *Funology* (Springer, Issue 5). https://doi.org/10.1177/216507990605400503

Dobrescu, A. I., Nussbaumer-Streit, B., Klerings, I., Wagner, G., Persad, E., Sommer, I., Herkner, H., & Gartlehner, G. (2021). Restricting evidence syntheses of interventions to English-language publications is a viable methodological shortcut for most medical topics: a systematic review. *Journal of Clinical Epidemiology*, *137*, 209–217. https://doi.org/10.1016/j.jclinepi.2021.04.012

Doi, H., Fujisawa, T. X., Kanai, C., Ohta, H., Yokoi, H., Iwanami, A., Kato, N., & Shinohara, K. (2013). Recognition of facial expressions and prosodic cues with graded emotional intensities in adults with asperger syndrome. *Journal of Autism and Developmental Disorders*, *43*(9), 2099–2113. https://doi.org/10.1007/s10803-013-1760-8

Dolan, R. J., Morris, J. S., & De Gelder, B. (2001). Crossmodal binding of fear in voice and face. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(17), 10006–10010. https://doi.org/10.1073/pnas.171288598

Donath, J. (2001). Mediated faces. In M. Beynon, C. L. Nehaniv, & K. Dautenhahn (Eds.), *Cognitive Technology: Instruments of Mind Proceedings of the 4th International Conference*. Springer. https://doi.org/10.1007/3-540-44617-6_34

Donges, U.-S., & Suslow, T. (2017). Alexithymia and automatic processing of emotional stimuli: a systematic review. *Reviews in the Neurosciences*, *28*(3), 247–264.

https://doi.org/10.1515/revneuro-2012-0049

Duchenne, B., & Cuthbertson, R. A. (1862). The mechanism of human facial expression or an electro-physiological analysis of the expression of the emotions. In *Ed. and Trans (Original work published 1862)*. Cambridge University Press.

Duggleby, W., Holtslander, L., Kylma, J., Duncan, V., Hammond, C., & Williams, A. (2010). Metasynthesis of the hope experience of family caregivers of persons with chronic illness. *Qualitative Health Research*, *20*(2), 148–158. https://doi.org/10.1177/1049732309358329

Dunn, D. S., & Andrews, E. E. (2015). Person-first and identity-first language: Developing psychologists' cultural competence using disability language. *American Psychologist*, *70*(3), 255–264. https://doi.org/10.1037/a0038636

Durand, K., Gallay, M., Seigneuric, A., Robichon, F., & Baudouin, J. Y. (2007). The development of facial emotion recognition: The role of configural information. *Journal of Experimental Child Psychology*, *97*(1), 14–27. https://doi.org/10.1016/j.jecp.2006.12.001

Duval, S. (2005). The trim and fill method. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments* (pp. 127–144). John Wiley & Sons. https://doi.org/10.1002/0470870168

Duval, S., & Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Dyck, M. J., Piek, J. P., Hay, D., Smith, L., & Hallmayer, J. (2006a). 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

Dyck, M. J., Piek, J. P., Hay, D., Smith, L., & Hallmayer, J. (2006b). 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

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

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, *337*(8746), 867–872. https://doi.org/10.1016/0140-6736(91)90201-Y

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics*, *63*(6), 1004–1013. https://doi.org/10.3758/BF03194519

Ecker, C., Bookheimer, S. Y., & Murphy, D. G. M. (2015). Neuroimaging in autism spectrum disorder: Brain structure and function across the lifespan. *The Lancet Neurology*, *14*(11), 1121–1134. https://doi.org/10.1016/S1474-4422(15)00050-2

Eerola, T., Friberg, A., & Bresin, R. (2013). Emotional expression in music: Contribution, linearity, and additivity of primary musical cues. *Frontiers in Psychology*, *4*, 1–12. https://doi.org/10.3389/fpsyg.2013.00487

Eerola, T., & Vuoskoski, J. K. (2011). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, *39*(1), 18–49. https://doi.org/10.1177/0305735610362821

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.316.7129.469

Eisenbarth, H., & Alpers, G. W. (2011). Happy Mouth and Sad Eyes: Scanning Emotional Facial Expressions. *Emotion*, *11*(4), 860–865. https://doi.org/10.1037/a0022758

Ekman, P. (1992a). An argument for basic emotions. In *Cognition and Emotion* (Vol. 6, Issue 3/4, pp. 169–200). http://www.paulekman.com/wp-content/uploads/2009/02/Universality-Of-Emotional-Expression-A-personal-History.pdf

Ekman, P. (1992b). Are there basic emotions? *Psychological Review1*, *99*(3), 550–553. https://doi.org/10.4081/jear.2011.169

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*(4), 384–392.

Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, *1*(1), 56–75. https://doi.org/10.1007/BF01115465

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

Escoffier, N., Zhong, J., Schirmer, A., & Qiu, A. (2013). Emotional expressions in voice and music: Same code, same effect? *Human Brain Mapping*, *34*(8), 1796–1810. https://doi.org/10.1002/hbm.22029

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition, 44*(3), 227–240. https://doi.org/10.1016/0010-0277(92)90002-Y

Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, *15*(2), 85–93.

https://doi.org/10.1016/j.tics.2010.11.004

Fallshore, M., & Bartholow, J. (2003). Recognition of emotion from inverted schematic drawings of faces. *Perceptual and Motor Skills*, *96*(1), 236–244. https://doi.org/10.2466/pms.2003.96.1.236

Fan, Y. T., & Cheng, Y. (2014). Atypical mismatch negativity in response to emotional voices in people with autism spectrum conditions. *PLoS ONE*, *9*(7), 1–10. https://doi.org/10.1371/journal.pone.0102471

Farmer, E., Jicol, C., & Petrini, K. (2020). Musicianship enhances perception but not feeling of emotion from others' social interaction through speech prosody. *Music Perception*, *37*(4), 323–338. https://doi.org/10.1525/mp.2020.37.4.323

Farrant, K., & Uddin, L. Q. (2015). Asymmetric development of dorsal and ventral attention networks in the human brain. *Developmental Cognitive Neuroscience*, *12*, 165–174. https://doi.org/10.1016/j.dcn.2015.02.001

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fecteau, S., Belin, P., Joanette, Y., & Armony, J. L. (2007). Amygdala responses to nonlinguistic emotional vocalizations. *NeuroImage*, *36*(2), 480–487. https://doi.org/10.1016/j.neuroimage.2007.02.043

Filipe, M. G., Watson, L., Vicente, S. G., & Frota, S. (2018). Atypical preference for infant-directed speech as an early marker of autism spectrum disorders? A literature review and directions for further research. *Clinical Linguistics and Phonetics*, *32*(3), 213–231. https://doi.org/10.1080/02699206.2017.1342694

Fink, E., de Rosnay, M., Wierda, M., Koot, H. M., & Begeer, S. (2014). Brief report: Accuracy and response time for the recognition of facial emotions in a large sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(9), 2363–2368. https://doi.org/10.1007/s10803-014-2084-z

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities*, *18*, 281–294. https://doi.org/10.1111/j.1468-3148.2005.00266.x

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*(6), 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203

Fombonne, E. (2020). Epidemiological controversies in autism. *Swiss Archives of Neurology, Psychiatry and Psychotherapy*, *171*(1), 3–5. https://doi.org/10.4414/sanp.2020.03084

Fournier, K. A., Hass, C. J., Naik, S. K., Lodha, N., & Cauraugh, J. H. (2010). Motor coordination in autism spectrum disorders: A synthesis and meta-analysis. *Journal of Autism and Developmental Disorders*, *40*(10), 1227–1240. https://doi.org/10.1007/s10803-010-0981-3

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion*, *14*(1), 61–92. https://doi.org/10.1080/026999300378996

Frazier, T. W., Georgiades, S., Bishop, S. L., & Hardan, A. Y. (2014). Behavioral and cognitive characteristics of females and males with autism in the simons simplex collection. *Journal of the American Academy of Child and Adolescent Psychiatry*, *53*(3), 329-340.e3. https://doi.org/10.1016/j.jaac.2013.12.004

Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A meta-analysis of gaze differences to social and nonsocial information between Individuals with and without autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(7), 546–555. https://doi.org/10.1016/j.jaac.2017.05.005

Freundlieb, M., Sebanz, N., & Kovács, Á. M. (2017). Out of your sight, out of my mind: Knowledge about another person's visual access modulates spontaneous visuospatial perspective-taking. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(6), 1065–1072. https://doi.org/10.1037/xhp0000379

Frewen, P. A., Dozois, D. J. A., Neufeld, R. W. J., & Lanius, R. A. (2008). *Meta-analysis of alexithymia in posttraumatic stress disorder*. *21*(2), 243–246. https://doi.org/10.1002/jts

Frewen, P. A., Lanius, R. A., Dozois, D. J. A., Neufeld, R. W. J., Pain, C., Hopper, J. W., Densmore, M., & Stevens, T. K. (2008). Clinical and Neural Correlates of Alexithymia in Posttraumatic Stress Disorder. *Journal of Abnormal Psychology*, *117*(1), 171–181. https://doi.org/10.1037/0021-843X.117.1.171

Frick, R. W. (1985). Communicating Emotion. The Role of Prosodic Features. *Psychological Bulletin*, *97*(3), 412–429. https://doi.org/10.1037/0033-2909.97.3.412

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., Baron-Cohen, S., & Golan, O. (2016). Basic and complex emotion recognition in children with autism: Cross-cultural findings. *Molecular Autism*, *7*(1), 52. https://doi.org/10.1186/s13229-016-0113-9

Friend, M., & Bryant, J. B. (2000). A developmental lexical bias in the interpretation of discrepant messages. *Merrill-Palmer Quarterly (Wayne State University. Press)*, *46*(2), 342–369. https://doi.org/10.1126/scisignal.274pe36.Insulin

Frischen, A., Eastwood, J. D., & Smilek, D. (2008). Visual Search for Faces With Emotional Expressions. *Psychological Bulletin*, *134*(5), 662–676. https://doi.org/10.1037/0033-2909.134.5.662

Frith, C. D., & Frith, U. (2008). Implicit and Explicit Processes in Social Cognition. *Neuron*, *60*(3), 503–510. https://doi.org/10.1016/j.neuron.2008.10.032

Frith, U. (1989). *Autism: Explaining the enigma*. Blackwell Publishing.

Frith, U. (2004). Emanuel Miller lecture: Confusions and controversies about Asperger syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(4), 672–686. https://doi.org/10.1111/j.1469-7610.2004.00262.x

Frith, U., & Happé, F. (1994). Autism: beyond "theory of mind." *Cognition*, *50*, 115–132.

Fritz, T. (2009). Emotion investigated with music of variable valence - neurophysiology and cultural influence. *MPI Series in Human Cognitive and Brain Sciences*, *110*, 1–35. http://pubman.mpdl.mpg.de/pubman/item/escidoc:722101:4

Fulceri, F., Grossi, E., Contaldo, A., Narzisi, A., Apicella, F., Parrini, I., Tancredi, R., Calderoni, S., & Muratori, F. (2019). Motor skills as moderators of core symptoms in autism spectrum disorders: Preliminary data from an exploratory analysis with artificial neural networks. *Frontiers in Psychology*, *9*(JAN), 1–12. https://doi.org/10.3389/fpsyg.2018.02683

Gabrielsson, A., & Juslin, P. N. (2003). Emotional expression in music. In R. J. Davidson, K. R. Scherer, & H. H. Goldsmith (Eds.), *Handbook of affective sciences* (pp. 503-534). Oxford University Press.

Gale, C. M., Eikeseth, S., & Klintwall, L. (2019). Children with Autism show Atypical Preference for Non-social Stimuli. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-019-46705-8

Galfano, G., Dalmaso, M., Marzoli, D., Pavan, G., Coricelli, C., & Castelli, L. (2012). Eye gaze cannot be ignored (but neither can arrows). *Quarterly Journal of Experimental Psychology*, *65*(10), 1895–1910. https://doi.org/10.1080/17470218.2012.663765

Gaspar, T., Labor, M., Jurić, I., Dumancić, D., Ilakovac, V., & Heffer, M. (2011). Comparison of emotion recognition from facial expression and music. *Collegium Antropologicum*, *35 Suppl 1*, 163–167. http://www.ncbi.nlm.nih.gov/pubmed/21648329

Gawronski, B., & Bodenhausen, G. V. (2005). Accessibility effects on implicit social cognition: The role of knowledge activation and retrieval experiences. *Journal of Personality and Social Psychology*, *89*(5), 672–685. https://doi.org/10.1037/0022-3514.89.5.672

Gebauer, L., Skewes, J., Hørlyck, L., & Vuust, P. (2014). Atypical perception of affective prosody in Autism Spectrum Disorder. *NeuroImage: Clinical*, *6*, 370–378. https://doi.org/10.1016/j.nicl.2014.08.025

Gebauer, L., Skewes, J., Westphael, G., Heaton, P., & Vuust, P. (2014). Intact brain processing of musical emotions in autism spectrum disorder, but more cognitive load and arousal in happy vs. sad music. *Frontiers in Neuroscience*, *8*, 1–10. https://doi.org/10.3389/fnins.2014.00192

Georgiades, S., Szatmari, P., Boyle, M., Hanna, S., Duku, E., Zwaigenbaum, L., Bryson, S., Fombonne, E., Volden, J., Mirenda, P., Smith, I., Roberts, W., Vaillancourt, T., Waddell, C., Bennett, T., & Thompson, A. (2013). Investigating phenotypic heterogeneity in children with autism spectrum disorder: A factor mixture modeling approach. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *54*(2), 206–215. https://doi.org/10.1111/j.1469-7610.2012.02588.x

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

Germain, E., Foster, N. E. V., Sharda, M., Chowdhury, R., Tryfon, A., Doyle-Thomas, K. A. R., Anagnostou, E., & Hyde, K. L. (2019). Pitch direction ability predicts melodic perception in autism. *Child Neuropsychology*, *25*(4), 445–465. https://doi.org/10.1080/09297049.2018.1488954

Gernsbacher, M. A. (2017). Editorial Perspective: The use of person-first language in scholarly writing may accentuate stigma. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(7), 859–861. https://doi.org/10.1111/jcpp.12706

Gernsbacher, M. A., Dawson, M., & Goldsmith, H. H. (2005). Three reasons not to believe in an autism epidemic. *Current Directions in Psychological Science*, *14*(2), 55–58. https://doi.org/10.1111/j.0963-7214.2005.00334.x

Gernsbacher, M. A., Dawson, M., & Mottron, L. (2006). Autism: Common, heritable, but not harmful. *Behavioral and Brain Sciences*, *29*(4), 413–414. https://doi.org/10.1017/S0140525X06319097

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthélémy, C., Brunelle, F., Samson, Y., & Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience*, *7*(8), 801–802. https://doi.org/10.1038/nn1291

Globerson, E., Amir, N., Golan, O., Kishon-Rabin, L., & Lavidor, M. (2013). Psychoacoustic abilities as predictors of vocal emotion recognition. *Attention, Perception, and*

Psychophysics, *75*(8), 1799–1810. https://doi.org/10.3758/s13414-013-0518-x

Globerson, E., Amir, N., Kishon-Rabin, L., & Golan, O. (2015). Prosody recognition in adults with high-functioning autism spectrum disorders: From psychoacoustics to cognition. *Autism Research*, *8*(2), 153–163. https://doi.org/10.1002/aur.1432

Goerlich, K. S., Witteman, J., Schiller, N. O., Van Heuven, V. J., Aleman, A., & Martens, S. (2012). The nature of affective priming in music and speech. *Journal of Cognitive Neuroscience*, *24*(8), 1725–1741. https://doi.org/10.1162/jocn

Goffaux, V., Hault, B., Michel, C., Vuong, Q. C., & Rossion, B. (2005). The respective role of low and high spatial frequencies in supporting configural and featural processing of faces. *Perception*, *34*(1), 77–86. https://doi.org/10.1068/p5370

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

Gold, B. T., Balota, D. A., Jones, S. J., Powell, D. K., Smith, C. D., & Andersen, A. H. (2006). Dissociation of automatic and strategic lexical-semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience*, *26*(24), 6523–6532. https://doi.org/10.1523/JNEUROSCI.0808-06.2006

Gold, R., Butler, P., Revheim, N., Leitman, D. I., Hansen, J. A., Gur, R. C., Kantrowitz, J. T., Laukka, P., Juslin, P. N., Silipo, G. S., & Javitt, D. C. (2012). Auditory emotion recognition impairments in schizophrenia: Relationship to acoustic features and cognition. *American Journal of Psychiatry*, *169*(4), 424–432. https://doi.org/10.1176/appi.ajp.2011.11081230

Gosselin, N., Paquette, S., & Peretz, I. (2015). Sensitivity to musical emotions in congenital amusia. *Cortex*, *71*, 171–182. https://doi.org/10.1016/j.cortex.2015.06.022

Graby, S. (2015). Neurodiversity: Bridging the gap between the disabled people’s movement and the mental health system survivors’ movement? *Madness, Distress and the Politics of Disablement*, *August*, 231–243. https://doi.org/10.2307/j.ctt1t898sg.21

Graf, W. D., Miller, G., Epstein, L. G., & Rapin, I. (2017). The autism epidemic Ethical, legal, and social issues in a developmental spectrum disorder. *Neurology*, *88*(14), 1371–1380. https://doi.org/10.1212/WNL.0000000000003791

Grandjean, D., Bänziger, T., & Scherer, K. R. (2006). Intonation as an interface between language and affect. *Progress in Brain Research*, *156*(06), 235–247. https://doi.org/10.1016/S0079-6123(06)56012-1

Greimel, E., Schulte-Rüther, M., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., & Konrad, K. (2014). Impairment in face processing in autism spectrum disorder: A developmental perspective. *Journal of Neural Transmission*, *121*(9), 1171–1181. https://doi.org/10.1007/s00702-014-1206-2

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

Griffiths, S., Jarrold, C., Penton-Voak, I. S., Woods, A. T., Skinner, A. L., & Munafò, M. R. (2019). Impaired Recognition of Basic Emotions from Facial Expressions in Young People with Autism Spectrum Disorder: Assessing the Importance of Expression Intensity. *Journal of Autism and Developmental Disorders*, *49*(7), 2768–2778. https://doi.org/10.1007/s10803-017-3091-7

Grinter, E. J., Maybery, M. T., Van Beek, P. L., Pellicano, E., Badcock, J. C., & Badcock, D. R. (2009). Global visual processing and self-rated autistic-like traits. *Journal of Autism and Developmental Disorders*, *39*(9), 1278–1290. https://doi.org/10.1007/s10803-009-0740-5

Gross, T. F. (2005). Global-local precedence in the perception of facial age and emotional expression by children with autism and other developmental disabilities. *Journal of Autism and Developmental Disorders*, *35*(6), 773–785. https://doi.org/10.1007/s10803-005-0023-8

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, and Allied Disciplines*, *41*(3), 369–379. https://doi.org/10.1111/1469-7610.00621

Grossman, R. B., Bemis, R. H., Plesa Skwerer, D., & Tager-Flusberg, H. (2010). Lexical and affective prosody in children with high-functioning autism. *Journal of Speech Language and Hearing Research*, *53*(3), 778–793. https://doi.org/10.1044/1092-4388(2009/08-0127)

Grossman, R. B., & Tager-Flusberg, H. (2012). “who said that?” matching of low- and high-intensity emotional prosody to facial expressions by adolescents with ASD. *Journal of Autism and Developmental Disorders*, *42*(12), 2546–2557. https://doi.org/10.1007/s10803-012-1511-2

Grynberg, D., Chang, B., Corneille, O., Maurage, P., Vermeulen, N., Berthoz, S., & Luminet,

O. (2012). Alexithymia and the processing of emotional facial expressions (EFEs): Systematic review, unanswered questions and further perspectives. *PLoS ONE*, *7*(8). https://doi.org/10.1371/journal.pone.0042429

Guillon, Q., Rogé, B., Afzali, M. H., Baduel, S., Kruck, J., & Hadjikhani, N. (2016). Intact perception but abnormal orientation towards face-like objects in young children with ASD. *Scientific Reports*, *6*, 1–9. https://doi.org/10.1038/srep22119

Gündel, H., López-Sala, A., Ceballos-Baumann, A. O., Deus, J., Cardoner, N., Marten-Mittag, B., Soriano-Mas, C., & Pujol, J. (2004). Alexithymia Correlates with the Size of the Right Anterior Cingulate. *Psychosomatic Medicine*, *66*(1), 132–140. https://doi.org/10.1097/01.PSY.0000097348.45087.96

Guy, J., Mottron, L., Berthiaume, C., & Bertone, A. (2019). A Developmental Perspective of Global and Local Visual Perception in Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *49*(7), 2706–2720. https://doi.org/10.1007/s10803-016-2834-1

Guyer, A. E., Monk, C. S., McClure-Tone, E. B., Nelson, E. E., Roberson-Nay, R., Adler, A. D., Fromm, S. J., Leibenluft, E., Pine, D. S., & Ernst, M. (2008). A developmental examination of amygdala response to facial expressions. *Journal of Cognitive Neuroscience*, *20*(9), 1565–1582. https://doi.org/10.1162/jocn.2008.20114

Habel, U., Windischberger, C., Derntl, B., Robinson, S., Kryspin-Exner, I., Gur, R. C., & Moser, E. (2007). Amygdala activation and facial expressions: Explicit emotion discrimination versus implicit emotion processing. *Neuropsychologia*, *45*(10), 2369–2377. https://doi.org/10.1016/j.neuropsychologia.2007.01.023

Haesen, B., Boets, B., & Wagemans, J. (2011). A review of behavioural and electrophysiological studies on auditory processing and speech perception in autism spectrum disorders. In *Research in Autism Spectrum Disorders*. https://doi.org/10.1016/j.rasd.2010.11.006

Haigh, S. M., Walsh, J. A., Mazefsky, C. A., Minshew, N. J., & Eack, S. M. (2018). Processing Speed is Impaired in Adults with Autism Spectrum Disorder, and Relates to Social Communication Abilities. *Journal of Autism and Developmental Disorders*, *48*(8), 2653–2662. https://doi.org/10.1007/s10803-018-3515-z

Hailstone, J. C., Omar, R., Henley, S. M. D., Frost, C., Kenward, M. G., & Warren, J. D. (2009). It’s not what you play, it’s how you play it: Timbre affects perception of emotion in music. *Quarterly Journal of Experimental Psychology*, *62*(11), 2141–2155. https://doi.org/10.1080/17470210902765957

Hakanpää, T., Waaramaa, T., & Laukkanen, A. M. (2019a). Comparing Contemporary Commercial and Classical Styles: Emotion Expression in Singing. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2019.10.002

Hakanpää, T., Waaramaa, T., & Laukkanen, A. M. (2019b). Emotion Recognition From Singing Voices Using Contemporary Commercial Music and Classical Styles. *Journal of Voice*, *33*(4), 501–509. https://doi.org/10.1016/j.jvoice.2018.01.012

Hallmayer, J., Cleveland, S., Torres, A., Phillips, J., Cohen, B., Torigoe, T., Miller, J., Fedele, A., Collins, J., Smith, K., Lotspeich, L., Croen, L. A., Ozonoff, S., Lajonchere, C., Grether, J. K., & Risch, N. (2011). Genetic heritability and shared environmental factors among twin pairs with autism. *Archives of General Psychiatry*, *68*(11), 1095–1102. https://doi.org/10.1001/archgenpsychiatry.2011.76

Hammerschmidt, K., & Jürgens, U. (2007). Acoustical Correlates of Affective Prosody. *Journal of Voice*, *21*(5), 531–540. https://doi.org/10.1016/j.jvoice.2006.03.002

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

Happé, F., & Frith, U. (2020). Annual Research Review: Looking back to look forward – changes in the concept of autism and implications for future research. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *61*(3), 218–232. https://doi.org/10.1111/jcpp.13176

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, *9*(10), 1218–1220. https://doi.org/10.1038/nn1770

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

Hartling, L., Featherstone, R., Nuspl, M., Shave, K., Dryden, D. M., & Vandermeer, B. (2017). Grey literature in systematic reviews: a cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. *BMC Medical Research Methodology*, *17*(1), 1–11. https://doi.org/10.1186/s12874-017-0347-z

Heaton, P. (2003). Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology & Psychiatry*, *44*(4), 543–551. https://doi.org/10.1111/1469-7610.00143

Heaton, P. (2005). Interval and contour processing in autism. *Journal of Autism and Developmental Disorders*, *35*(6), 787–793. https://doi.org/10.1007/s10803-005-0024-7

Heaton, P. (2009). Assessing musical skills in autistic children who are not savants. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1522), 1443–1447. https://doi.org/10.1098/rstb.2008.0327

Heaton, P., Hermelin, B., & Pring, L. (1999). Can children with autism spectrum disorder perceive affect in music? An experimental investigation. *Psychological Medicine*, *29*(6), 1405–1410. https://doi.org/10.1017/S0033291799001221

Heaton, P., Hudry, K., Ludlow, A., & Hill, E. (2008). Superior descrimination of speech pitch and its relationship to verbal ability in autism spectrum disorder. *Cognitive Neuropsychology*, *25*(6), 771–782. https://doi.org/10.1080/02643290802336277

Heaton, P., Pring, L., & Hermelin, B. (2001). Musical processing in high functioning children with autism. *Annals of the New York Academy of Sciences*, *930*, 443–444. https://doi.org/10.1111/j.1749-6632.2001.tb05765.x

Heaton, P., Reichenbacher, L., Sauter, D., Allen, R., Scott, S., & Hill, E. (2012). Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum disorder and typical development. *Psychological Medicine*, *42*(11), 2453–2459. https://doi.org/10.1017/S0033291712000621

Heaton, P., Williams, K., Cummins, O., & Happé, F. (2008). Autism and pitch processing splinter skills: A group and subgroup analysis. *Autism*, *12*(2), 203–219. https://doi.org/10.1177/1362361307085270

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.3102/10769986006002107

Hedvall, Å., Fernell, E., Holm, A., Åsberg Johnels, J., Gillberg, C., & Billstedt, E. (2013). Autism, processing speed, and adaptive functioning in preschool children. *The Scientific World Journal*, *2013*. https://doi.org/10.1155/2013/158263

Heikkinen, J., Jansson-Verkasalo, E., Toivanen, J., Suominen, K., Väyrynen, E., Moilanen, I., & Seppänen, T. (2010). Perception of basic emotions from speech prosody in adolescents with Asperger's syndrome. *Logopedics Phoniatrics Vocology*, *35*(3), 113–120. https://doi.org/10.3109/14015430903311184

Henson, R. K. (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist*, *34*(5), 601–629. https://doi.org/10.1177/0011000005283558

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

Psychology and Psychiatry and Allied Disciplines, *47*(11), 1098–1106. https://doi.org/10.1111/j.1469-7610.2006.01652.x

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

Hermans, D., De Houwer, J., & Eelen, P. (2001). A time course analysis of the affective priming effect. *Cognition and Emotion*, *15*(2), 143–165. https://doi.org/10.1080/02699930125768

Hermans, D., Houwer, J. De, & Eelen, P. (1994). The affective priming effect: Automatic activation of evaluative information in memory. *Cognition & Emotion*, *8*(6), 515–533. https://doi.org/10.1080/02699939408408957

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

Herring, D. R., White, K. R., Jabeen, L. N., Hinojos, M., Terrazas, G., Reyes, S. M., Taylor, J. H., & Crites, S. L. (2013). On the automatic activation of attitudes: A quarter century of evaluative priming research. *Psychological Bulletin*, *139*(5), 1062–1089. https://doi.org/10.1037/a0031309

Hickman, L. (2019). The importance of assessing alexithymia in psychological research. *Psychiatric News*, *37*(23), 1–1.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hiirola, A., Pirkola, S., Karukivi, M., Markkula, N., Bagby, R. M., Joukamaa, M., Jula, A., Kronholm, E., Saarijärvi, S., Salminen, J. K., Suvisaari, J., Taylor, G., & Mattila, A. K. (2017). An evaluation of the absolute and relative stability of alexithymia over 11 years in a Finnish general population. *Journal of Psychosomatic Research*, *95*, 81–87. https://doi.org/10.1016/j.jpsychores.2017.02.007

Hirst, R. J., Cragg, L., & Allen, H. A. (2018). Vision dominates audition in adults but not children: A meta-analysis of the Colavita effect. *Neuroscience and Biobehavioral Reviews*, *94*, 286–301. https://doi.org/10.1016/j.neubiorev.2018.07.012

Hirst, R. J., Stacey, J. E., Cragg, L., Stacey, P. C., & Allen, H. A. (2018). The threshold for the McGurk effect in audio-visual noise decreases with development. *Scientific Reports*, *8*(1), 1–12. https://doi.org/10.1038/s41598-018-30798-8

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

Hobson, R. P., Ouston, J., & Lee, A. (1989). Naming emotion in faces and voices: Abilities and disabilities in autism and mental retardation. *British Journal of Developmental Psychology*, *7*(3), 237–250. https://doi.org/10.1111/j.2044-835x.1989.tb00803.x

Hoekert, M., Kahn, R. S., Pijnenborg, M., & Aleman, A. (2007). Impaired recognition and expression of emotional prosody in schizophrenia: Review and meta-analysis. *Schizophrenia Research*, *96*(1–3), 135–145. https://doi.org/10.1016/j.schres.2007.07.023

Hong, M., Lee, S. M., Park, S., Yoon, S. J., Kim, Y. E., & Oh, I. H. (2020). Prevalence and Economic Burden of Autism Spectrum Disorder in South Korea Using National Health Insurance Data from 2008 to 2015. *Journal of Autism and Developmental Disorders*, *50*(1), 333–339. https://doi.org/10.1007/s10803-019-04255-y

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. (1998). Practitioner Review: Psychological and Educational Treatments for Autism. *Journal of Child Psychology and Psychiatry*, *39*(3), 307–322. https://doi.org/10.1111/1469-7610.00327

Hsu, N., & Schütt, Z. (2012). *Psychology of priming*. Nova Science Publishers.

Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Research*, *10*(12), 1991–2001. https://doi.org/10.1002/aur.1847

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

Hwang, H., & Matsumoto, D. (2019). Functions of Emotion. In *Noba textbook series: Psychology.* https://doi.org/10.4324/9781315276229-4

Ichikawa, H., Kanazawa, S., & Yamaguchi, M. K. (2011). Finding a face in a face-like object. *Perception*, *40*(4), 500–502. https://doi.org/10.1068/p6926

Ingersoll, B. (2010). Broader autism phenotype and nonverbal sensitivity: Evidence for an association in the general population. *Journal of Autism and Developmental Disorders*, *40*(5), 590–598. https://doi.org/10.1007/s10803-009-0907-0

Ingram, D. G., Takahashi, T. N., & Miles, J. H. (2008). Defining autism subgroups: A taxometric solution. *Journal of Autism and Developmental Disorders*, *38*(5), 950–960. https://doi.org/10.1007/s10803-007-0469-y

IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, *6*(7), e010247. https://doi.org/10.1136/bmjopen-2015-010247

Isomura, T., Ito, H., Ogawa, S., & Masataka, N. (2014). Absence of predispositional attentional sensitivity to angry faces in children with autism spectrum disorders. *Scientific Reports*, *4*, 1–6. https://doi.org/10.1038/srep07525

Isomura, T., Ogawa, S., Yamada, S., Shibasaki, M., & Masataka, N. (2014). Preliminary evidence that different mechanisms underlie the anger superiority effect in children with and without Autism Spectrum Disorders. *Frontiers in Psychology*, *5*, 1–8. https://doi.org/10.3389/fpsyg.2014.00461

Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B., & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*, *12*(1), 18–23. https://doi.org/10.1111/1467-9280.00304

Jack, R. E., Sun, W., Delis, I., Garrod, O. G. B., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*(6), 708–730. https://doi.org/10.1037/xge0000162

Jamey, K., Foster, N. E. V., Sharda, M., Tuerk, C., Nadig, A., & Hyde, K. L. (2019). Evidence for intact melodic and rhythmic perception in children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, *64*, 1–12. https://doi.org/10.1016/j.rasd.2018.11.013

Järvinen-Pasley, A., & Heaton, P. (2007). Evidence for reduced domain-specificity in auditory processing in autism. *Developmental Science*, *10*(6), 786–793. https://doi.org/10.1111/j.1467-7687.2007.00637.x

Järvinen, A., Ng, R., Crivelli, D., Neumann, D., Arnold, A. J., Woo-Vonhoogenstyn, N., Lai, P., Trauner, D., & Bellugi, U. (2016). Social functioning and autonomic nervous system sensitivity across vocal and musical emotion in Williams syndrome and autism spectrum disorder. *Developmental Psychobiology*, *58*(1), 17–26.

https://doi.org/10.1002/dev.21335

Jaywant, A., & Pell, M. D. (2012). Categorical processing of negative emotions from speech prosody. *Speech Communication*, *54*(1), 1–10. https://doi.org/10.1016/j.specom.2011.05.011

Jessen, S., & Grossmann, T. (2015). Neural signatures of conscious and unconscious emotional face processing in human infants. *Cortex*, *64*, 260–270. https://doi.org/10.1016/j.cortex.2014.11.007

Jessimer, M., & Markham, R. (1997). Alexithymia: a right hemisphere dysfunction specific to recognition of certain facial expressions? *Brain and Cognition*, *34*(2), 246–258. https://doi.org/10.1006/brcg.1997.0900

Jiang, J., Liu, F., Wan, X., & Jiang, C. (2015). Perception of melodic contour and intonation in autism spectrum disorder: Evidence from mandarin speakers. *Journal of Autism and Developmental Disorders*, *45*(7), 2067–2075. https://doi.org/10.1007/s10803-015-2370-4

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *38*(5), 527–534. https://doi.org/10.1111/j.1469-7610.1997.tb01539.x

Jones, C. R. G., Happé, F., Baird, G., Simonoff, E., Marsden, A. J. S., Tregay, J., Phillips, R. J., Goswami, U., Thomson, J. M., & Charman, T. (2009). Auditory discrimination and auditory sensory behaviours in autism spectrum disorders. *Neuropsychologia*, *47*(13), 2850–2858. https://doi.org/10.1016/j.neuropsychologia.2009.06.015

Jones, C. R. G., Pickles, A., Falcaro, M., Marsden, A. J. S., 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 and Allied Disciplines*, *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. *J Child Psychol Psychiatry*, *44*(4), 529–542. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12751845

Juslin, P. N. (1997a). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception: An Interdisciplinary Journal*, *14*(4), 383–418. https://doi.org/10.2307/40285731

Juslin, P. N. (1997b). Perceived Emotional Expression in Synthesized Performances of a Short Melody: Capturing the Listener's Judgment Policy. *Musicae Scientiae*, *1*(2), 225–256. https://doi.org/10.1177/102986499700100205

Juslin, P. N. (2000). Cue Utilization in Communication of Emotion in Music Performance: Relating Performance to Perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1797–1813. https://doi.org/10.1037//0096-1523.26.6.1797

Juslin, P. N. (2001). Communicating emotion in music performance: A review and a theoretical framework. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 309-337). Oxford University Press.

Juslin, P. N., & Laukka, P. (2001). Impact of Intended Emotion Intensity on Cue Utilization and Decoding Accuracy in Vocal Expression of Emotion. *Emotion*, *1*(4), 381–412. https://doi.org/10.1037/1528-3542.1.4.381

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*(5), 770–814. https://doi.org/10.1037/0033-2909.129.5.770

Juslin, P. N., Liljeström, S., Västfjäll, D., Barradas, G., & Silva, A. (2008). An Experience Sampling Study of Emotional Reactions to Music: Listener, Music, and Situation. *Emotion*, *8*(5), 668–683. https://doi.org/10.1037/a0013505

Juslin, P. N., & Lindström, E. (2010). Musical expression of emotions: Modelling listeners' judgements of composed and performed features. *Music Analysis*, *29*(1–3), 334–364. https://doi.org/10.1111/j.1468-2249.2011.00323.x

Juslin, P. N., & Madison, G. U. Y. (2016). *The Role of Timing Patterns in Recognition of Emotional Expression from Musical Performance Author (s): Patrik N . Juslin and Guy Madison Published by : University of California Press Stable URL : http://www.jstor.org/stable/40285891 Accessed : 04-05-20. 17*(2), 197–221.

Juslin, P. N., & Sloboda, J. A. (2013). Music and emotion. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 583–645). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-381460-9.00015-8

Just, M. A., Keller, T. A., Malave, V. L., Kana, R. K., & Varma, S. (2012). Autism as a neural systems disorder: A theory of frontal-posterior underconnectivity. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1292–1313. https://doi.org/10.1016/j.neubiorev.2012.02.007

Juth, P., Lundqvist, D., Karlsson, A., & Öhman, A. (2005). Looking for foes and friends:

Perceptual and emotional factors when finding a face in the crowd. *Emotion*, *5*(4), 379–395. https://doi.org/10.1037/1528-3542.5.4.379

Kalinli, O. (2016). Analysis of multi-lingual emotion recognition using auditory attention features. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *08-12-Sept*, 3613–3617. https://doi.org/10.21437/Interspeech.2016-1557

Kallinen, K. (2005). Emotional ratings of music excerpts in the western art music repertoire and their self-organization in the Kohonen neural network. *Psychology of Music*, *33*(4), 373–393. https://doi.org/10.1177/0305735605056147

Kamio, Y., Wolf, J., & Fein, D. (2006). Automatic processing of emotional faces in high-functioning pervasive developmental disorders: An affective priming study. *Journal of Autism and Developmental Disorders*, *36*(2), 155–167. https://doi.org/10.1007/s10803-005-0056-z

Kana, R. K., Patriquin, M. A., Black, B. S., Channell, M. M., & Wicker, B. (2016). Altered Medial Frontal and Superior Temporal Response to Implicit Processing of Emotions in Autism. *Autism Research*, *9*(1), 55–66. https://doi.org/10.1002/aur.1496

Kanner, L. (1943). Autistic distrubances of affective contact. In *Nervous Child* (Vol. 2, pp. 217–250). https://doi.org/10.1111/fwb.12896

Kannyo, I., & DeLong, C. M. (2011). The effect of musical training on auditory perception. *Proceedings of Meetings on Acoustics*, *14*. https://doi.org/10.1121/1.4733850

Kapp, S. K., Gillespie-Lynch, K., Sherman, L. E., & Hutman, T. (2013). Deficit, difference, or both? Autism and neurodiversity. *Developmental Psychology*, *49*(1), 59–71. https://doi.org/10.1037/a0028353

Kargas, N., López, B., Reddy, V., & Morris, P. (2015). The Relationship Between Auditory Processing and Restricted, Repetitive Behaviors in Adults with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *45*(3), 658–668. https://doi.org/10.1007/s10803-014-2219-2

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, *2*(10), 389–398. https://doi.org/10.4324/9781315516691-5

Kassambara, A. (2020). rstatix: pipe-friendly framework for basic statistical tests. R package version 0.4. 0. *Avaliable Online at: Https://Cran. r-Project.Org/Web/Packages/Rstatix/Index*.

Katagiri, M., Kasai, T., Kamio, Y., & Murohashi, H. (2013). Individuals with Asperger's disorder exhibit difficulty in switching attention from a local level to a global level. *Journal of Autism and Developmental Disorders*, *43*(2), 395–403. https://doi.org/10.1007/s10803-012-1578-9

Kätsyri, J., Saalasti, S., Tiippana, K., von Wendt, L., & Sams, M. (2008). Impaired recognition of facial emotions from low-spatial frequencies in Asperger syndrome. *Neuropsychologia*, *46*(7), 1888–1897. https://doi.org/10.1016/j.neuropsychologia.2008.01.005

Keating, C. T., Fraser, D. S., Sowden, S., & Cook, J. L. (2021). 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*, *0123456789*. https://doi.org/10.1007/s10803-021-05083-9

Kellerman, G. R., Fan, J., & Gorman, J. M. (2005). Auditory abnormalities in autism: Toward functional distinctions among findings. *CNS Spectrums*, *10*(9), 748–756. https://doi.org/10.1017/S1092852900019738

Keltner, Dacher, Cordaro, D. T. (2015). Understanding Multimodal Emotional Expression: Recent Advances in Basic Emotion Theory. *Emotion Researcher*, 9–24.

Keltner, D., & Haidt, J. (1999). Social functions of emotions at four levels of analysis. *Cognition and Emotion*, *13*(5), 505–521. https://doi.org/10.1080/026999399379168

Keltner, D., & Haidt, J. (2001). Social functions of emotions. In *Emotions: Currrent issues and future directions.* (pp. 192–213). Guilford Press.

Keltner, D., Tracy, J., Sauter, D., Cordaro, D., & McNeil, G. (2016). Emotional expression. In L. E. Barrett, M. Lewis, & J. Haviland-Jones (Eds.), *Handbook of Emotion* (4th ed.). New York, NY: Guilford Press.

Kendall, L. N., Raffaelli, Q., Kingstone, A., & Todd, R. M. (2016). Iconic faces are not real faces: enhanced emotion detection and altered neural processing as faces become more iconic. *Cognitive Research: Principles and Implications*, *1*(1), 1–14. https://doi.org/10.1186/s41235-016-0021-8

Kennedy, D. P., & Adolphs, R. (2012). Perception of emotions from facial expressions in high-functioning adults with autism. *Neuropsychologia*, *50*(14), 3313–3319. https://doi.org/10.1016/j.neuropsychologia.2012.09.038.

Kenny, L., Hattersley, C., Molins, B., Buckley, C., Povey, C., & Pellicano, E. (2016). Which terms should be used to describe autism? Perspectives from the UK autism community. *Autism*, *20*(4), 442–462. https://doi.org/10.1177/1362361315588200

Ketelaars, M. P., In'T Velt, A., Mol, A., Swaab, H., & Van Rijn, S. (2016). Emotion recognition and alexithymia in high functioning females with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *21*, 51–60. https://doi.org/10.1016/j.rasd.2015.09.006

Keys, R. T., Taubert, J., & Wardle, S. G. (2021). A visual search advantage for illusory faces in objects. *Attention, Perception, and Psychophysics*, *83*(5), 1942–1953. https://doi.org/10.3758/s13414-021-02267-4

Khawar, R., Malik, F., Maqsood, S., Yasmin, T., & Habib, S. (2013). Age and gender differences in emotion recognition. *Journal of Behavioural Sciences*, *23*(3), 53–71. https://doi.org/10.3389/fpsyg.2019.02371

Kiefer, M. (2007). Top-down modulation of unconscious "automatic" processes: A gating framework. *Advances in Cognitive Psychology*, *3*(1–2), 289–306. https://doi.org/10.2478/v10053-008-0031-2

Kiefer, M. (2012). Executive control over unconscious cognition: Attentional sensitization of unconscious information processing. *Frontiers in Human Neuroscience*, *6*, 1–12. https://doi.org/10.3389/fnhum.2012.00061

Kiefer, M., & Martens, U. (2010). Attentional sensitization of unconscious cognition: Task sets modulate subsequent masked semantic priming. *Journal of Experimental Psychology: General*, *139*(3), 464–489. https://doi.org/10.1037/a0019561

Kikuchi, Y., Senju, A., Tojo, Y., Osanai, H., & Hasegawa, T. (2009). Faces do not capture special attention in children with autism spectrum disorder: A change blindness study. *Child Development*, *80*(5), 1421–1433. https://doi.org/10.1111/j.1467-8624.2009.01342.x

Kim, Y. S., Leventhal, B. L., Koh, Y. J., Fombonne, E., Laska, E., Lim, E. C., Cheon, K. A., Kim, S. J., Kim, Y. K., Lee, H. K., Song, D. H., & Grinker, R. R. (2011). Prevalence of autism spectrum disorders in a total population sample. *American Journal of Psychiatry*, *168*(9), 904–912. https://doi.org/10.1176/appi.ajp.2011.10101532

Kinnaird, E., Stewart, C., & Tchanturia, K. (2019). Investigating alexithymia in autism: A systematic review and meta-analysis. *European Psychiatry*, *55*, 80–89. https://doi.org/10.1016/j.eurpsy.2018.09.004

Kinsella, B. G., Chow, S., & Kushki, A. (2017). Evaluating the Usability of a Wearable Social Skills Training Technology for Children with Autism Spectrum Disorder. *Frontiers in Robotics and AI*, *4*(July), 1–9. https://doi.org/10.3389/frobt.2017.00031

Klasen, M., Kreifelts, B., Chen, Y. H., Seubert, J., & Mathiak, K. (2014). Neural processing of emotion in multimodal settings. *Frontiers in Human Neuroscience*, *8*, 8–11.

https://doi.org/10.3389/fnhum.2014.00822

Klauer, K. C. (1997). Affective Priming. *European Review of Social Psychology*, *8*(1), 67–103. https://doi.org/10.1080/14792779643000083

Klauer, K. C., Roßnagel, C., & Musch, J. (1997). List-context effects in evaluative priming. *Journal of Experimental Psychology: Learning Memory and Cognition*, *23*(1), 246–255. https://doi.org/10.1037/0278-7393.23.1.246

Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., & Heekeren, H. R. (2012). The role of the amygdala in atypical gaze on emotional faces in autism spectrum disorders. *Journal of Neuroscience*, *32*(28), 9469–9476. https://doi.org/10.1523/jneurosci.5294-11.2012

Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical reflexive gaze patterns on emotional faces in autism spectrum disorders. *Journal of Neuroscience*, *30*(37), 12281–12287. 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*, 1–13. 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. (1992). Listening Preferences in Regard to Speech in Four Children with Developmental Disabilities. *Journal of Child Psychology and Psychiatry*, *33*(4), 763–769. https://doi.org/10.1111/j.1469-7610.1992.tb00911.x

Klingberg, T. (2000). Limitations in information processing in the human brain: Neuroimaging of dual task performance and working memory tasks. *Progress in Brain Research*, *126*(i), 95–102. https://doi.org/10.1016/S0079-6123(00)26009-3

Koldewyn, K., Whitney, D., & Rivera, S. M. (2010). The psychophysics of visual motion and global form processing in autism. *Brain*, *133*(2), 599–610. https://doi.org/10.1093/brain/awp272

Koolagudi, S. G., & Rao, K. S. (2012a). Emotion recognition from speech: A review. *International Journal of Speech Technology*, *15*(2), 99–117. https://doi.org/10.1007/s10772-011-9125-1

Koolagudi, S. G., & Rao, K. S. (2012b). Emotion recognition from speech using source, system, and prosodic features. *International Journal of Speech Technology*, *15*(2), 265–289. https://doi.org/10.1007/s10772-012-9139-3

Kopec, J., Hillier, A., & Frye, A. (2014). The Valency of Music Has Different Effects on the Emotional Responses of Those with Autism Spectrum Disorders and a Comparison Group. *Music Perception: An Interdisciplinary Journal*, *31*(5), 436–443. https://doi.org/10.1525/mp.2014.31.5.436

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional Overlap: Cognitive Basis for Stimulus-Response Compatibility-A Model and Taxonomy. *Psychological Review*, *97*(2), 253–270. https://doi.org/10.1037/0033-295X.97.2.253

Kostic, A., & Chadee, D. (2014). *The social psychology of nonverbal communication*. Springer.

Kouo, J. L., & Egel, A. L. (2016). The effectiveness of interventions in teaching emotion recognition to children with autism spectrum disorder. *Review Journal of Autism and Developmental Disorders*, *3*(3), 252–265. https://doi.org/10.1007/s40489-016-0081-1

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews Neuroscience*, *11*(8), 599–605. https://doi.org/10.1038/nrn2882

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(4), 579–594. https://doi.org/10.1037/0096-1523.5.4.579

Krysko, K. M., & Rutherford, M. D. (2009). A threat-detection advantage in those with autism spectrum disorders. *Brain and Cognition*, *69*(3), 472–480. https://doi.org/10.1016/j.bandc.2008.10.002

Kujala, T., Lepistö, T., Nieminen-Von Wendt, T., Näätänen, P., & Näätänen, R. (2005). Neurophysiological evidence for cortical discrimination impairment of prosody in Asperger syndrome. *Neuroscience Letters*, *383*(3), 260–265. https://doi.org/10.1016/j.neulet.2005.04.048

Kunde, W. (2003). Sequential modulations of stimulus-response correspondence effects depend on awareness of response conflict. *Psychonomic Bulletin and Review*, *10*(1), 198–205. https://doi.org/10.3758/BF03196485

Kuo, M. H., Orsmond, G. I., Coster, W. J., & Cohn, E. S. (2014). Media use among adolescents with autism spectrum disorder. *Autism*, *18*(8), 914–923. https://doi.org/10.1177/1362361313497832

Kuusikko, S., Haapsamo, H., Jansson-Verkasalo, E., Hurtig, T., Mattila, M. L., Ebeling, H., Jussila, K., Bolte, S., & Moilanen, I. (2009). Emotion recognition in children and adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(6), 938–945. https://doi.org/10.1007/s10803-009-0700-0

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models . *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lacroix, A., Guidetti, M., Rogé, B., & Reilly, J. (2014). Facial emotion recognition in 4- to 8-year-olds with autism spectrum disorder: A developmental trajectory approach. *Research in Autism Spectrum Disorders*, *8*(9), 1146–1154. https://doi.org/10.1016/j.rasd.2014.05.012

Lai, G., Pantazatos, S. P., Schneider, H., & Hirsch, J. (2012). Neural systems for speech and song in autism. *Brain*, *135*(3), 961–975. https://doi.org/10.1093/brain/awr335

Lane, R. D. (2000). Neural correlates of conscious emotional experience. In R. D. Lane & L. Nadel (Eds.), *Cognitive Neuroscience of Emotion* (pp. 345-370). Oxford University Press.

Lane, R. D. (2008). Neural substrates of implicit and explicit emotional processes: A unifying framework for psychosomatic medicine. *Psychosomatic Medicine*, *70*(2), 214–231. https://doi.org/10.1097/PSY.0b013e3181647e44

Laukka, P., Eerola, T., Thingujam, N. S., Yamasaki, T., & Beller, G. (2013). Universal and culture-specific factors in the recognition and performance of musical affect expressions. *Emotion*, *13*(3), 434–449. https://doi.org/10.1037/a0031388

Law Smith, M. J., Montagne, B., Perrett, D. I., 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.neuropsychologia.2010.03.008

Le Sourn-Bissaoui, S., Aguert, M., Girard, P., Chevreuil, C., & Laval, V. (2013). Emotional speech comprehension in children and adolescents with autism spectrum disorders. *Journal of Communication Disorders*, *46*(4), 309–320. https://doi.org/10.1016/j.jcomdis.2013.03.002

Leder, H., & Carbon, C. C. (2006). Face-specific configural processing of relational information. *British Journal of Psychology*, *97*(1), 19–29. https://doi.org/10.1348/000712605X54794

Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., & Näätänen, R. (2005). The discrimination of and orienting to speech and non-speech sounds in children with autism. *Brain Research*, *1066*(1–2), 147–157. https://doi.org/10.1016/j.brainres.2005.10.052

Leppänen, J. M., & Hietanen, J. K. (2007). Is there more in a happy face than just a big smile? *Visual Cognition*, *15*(4), 468–490. https://doi.org/10.1080/13506280600765333

Lerner, M. D., McPartland, J. C., & Morris, J. P. (2013). Multimodal emotion processing in autism spectrum disorders: An event-related potential study. *Developmental Cognitive Neuroscience*, *3*(1), 11–21. https://doi.org/10.1016/j.dcn.2012.08.005

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, *43*(3), 225–251. https://doi.org/10.1016/0010-0277(92)90013-8

Lévêque, Y., Teyssier, P., Bouchet, P., Bigand, E., Caclin, A., & Tillmann, B. (2018). Musical emotions in congenital amusia: Impaired recognition, but preserved emotional intensity. *Neuropsychology*, *32*(7), 880–894. https://doi.org/10.1037/neu0000461

Lewis, G. J., Lefevre, C. E., & Young, A. W. (2016). Functional architecture of visual emotion recognition ability: A latent variable approach. *Journal of Experimental Psychology. General*, *145*(5), 589. https://doi.org/10.1037/xge0000160

Lewkowicz, D. J., & Ghazanfar, A. A. (2009). The emergence of multisensory systems through perceptual narrowing. *Trends in Cognitive Sciences*, *13*(11), 470–478. https://doi.org/10.1016/j.tics.2009.08.004

Liddell, B. J., Williams, L. M., Rathjen, J., Shevrin, H., & Gordon, E. (2004). A temporal dissociation of subliminal versus supraliminal fear perception: An event-related potential study. *Journal of Cognitive Neuroscience*, *16*(3), 479–486. https://doi.org/10.1162/089892904322926809

Lieberman, P., & Michaels, S. B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *The Journal of the Acoustical Society of America*, *34*(7), 922–927. https://doi.org/10.1121/1.1918222

Lima, C. F., Brancatisano, O., Fancourt, A., Müllensiefen, D., Scott, S. K., Warren, J. D., & Stewart, L. (2016). Impaired socio-emotional processing in a developmental music disorder. *Scientific Reports*, *6*(1), 1–13. https://doi.org/10.1038/srep34911

Lima, C. F., & Castro, S. L. (2011a). Emotion recognition in music changes across the adult life span. *Cognition and Emotion*, *25*(4), 585–598. https://doi.org/10.1080/02699931.2010.502449

Lima, C. F., & Castro, S. L. (2011b). Speaking to the trained ear: Musical expertise enhances the recognition of emotions in speech prosody. *Emotion*, *11*(5), 1021–1031. https://doi.org/10.1037/a0024521

Lima, C. F., Garrett, C., & Castro, S. L. (2013). Not all sounds sound the same: Parkinson's disease affects differently emotion processing in music and in speech prosody. *Journal* *of Clinical and Experimental Neuropsychology*, *35*(4), 373–392. https://doi.org/10.1080/13803395.2013.776518

Lin, Y., & Ding, H. (2019). Multisensory integration of emotions in a face-prosody-semantics stroop task. *NeuroManagement and Intelligent Computing Method on Multimodal Interaction, AICMI 2019*. https://doi.org/10.1145/3357160.3357668

Lindner, J. L., & Rosén, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *36*(6), 769–777. https://doi.org/10.1007/s10803-006-0105-2

Lindström, R. (2019). *Neural and perceptual processing of emotional speech prosody in typically developed children and in children with autism spectrum disorder* (Issue February).

Lindström, R., Lepistö-Paisley, T., Makkonen, T., Reinvall, O., Nieminen-von Wendt, T., Alén, R., & Kujala, T. (2018). Atypical perceptual and neural processing of emotional prosodic changes in children with autism spectrum disorders. *Clinical Neurophysiology*, *129*(11), 2411–2420. https://doi.org/10.1016/j.clinph.2018.08.018

Liu, F., Jiang, C., Thompson, W. F., Xu, Y., Yang, Y., & Stewart, L. (2012). The mechanism of speech processing in congenital amusia: Evidence from Mandarin speakers. *PLoS ONE*, *7*(2). https://doi.org/10.1371/journal.pone.0030374

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: Discrimination, identification and imitation. *Brain*. https://doi.org/10.1093/brain/awq089

Liu, H., Hu, Z., Peng, D., Yang, Y., & Li, K. (2010). Common and segregated neural substrates for automatic conceptual and affective priming as revealed by event-related functional magnetic resonance imaging. *Brain and Language*, *112*(2), 121–128. https://doi.org/10.1016/j.bandl.2009.11.001

Liu, T. L., Wang, P. W., Yang, Y. H. C., Shyi, G. C. W., & Yen, C. F. (2019). Association between facial emotion recognition and bullying involvement among adolescents with high-functioning autism spectrum disorder. *International Journal of Environmental Research and Public Health*, *16*(24), 1–14. https://doi.org/10.3390/ijerph16245125

Livingston, L. A., & Happé, F. (2017). Conceptualising compensation in neurodevelopmental disorders: Reflections from autism spectrum disorder. *Neuroscience and Biobehavioral Reviews*, *80*, 729–742. https://doi.org/10.1016/j.neubiorev.2017.06.005

Livingstone, S. R., Muhlberger, R., Brown, A. R., & Thompson, W. F. (2010). Changing musical emotion: A computational rule system for modifying score and performance. *Computer Music Journal*, *34*(1), 41–64. https://doi.org/10.1162/comj.2010.34.1.41

Livingstone, S. R., & Russo, F. A. (2018). The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English. *PLoS ONE*, *13*(5), e0196391. https://doi.org/10.1371/journal.pone.0196391

Lobaugh, N. J., Gibson, E., & Taylor, M. J. (2006). Children recruit distinct neural systems for implicit emotional face processing. *NeuroReport*, *17*(2), 215–219. https://doi.org/10.1097/01.wnr.0000198946.00445.2f

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *The Lancet*, *392*(10146), 508–520. https://doi.org/10.1016/S0140-6736(18)31129-2

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). Autism diagnostic observation schedule, (ADOS-2) modules 1-4. *Los Angeles, California: Western Psychological Services*.

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: How frequent and severe are deficits? *Molecular Autism*, *9*(1), 7. https://doi.org/10.1186/s13229-018-0187-7

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

Lugger, M., & Yang, B. (2007). The relevance of voice quality features in speaker independent emotion recognition. *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, IV–17.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, *18*(2), 199–211. https://doi.org/10.1162/jocn.2006.18.2.199

Mankel, K., Barber, J., & Bidelman, G. M. (2020). Auditory categorical processing for speech is modulated by inherent musical listening skills. *NeuroReport*, 162–166. https://doi.org/10.1097/WNR.0000000000001369

Marchesi, C., Brusamonti, E., & Maggini, C. (2000). Are alexithymia, depression, and anxiety distinct constructs in affective disorders? *Journal of Psychosomatic Research*, *49*(1), 43–49. https://doi.org/10.1016/S0022-3999(00)00084-2

Marks, D. (1997). Models of disability. *Disability and Rehabilitation*, *19*(3), 85–91. https://doi.org/10.3109/09638289709166831

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296–305. https://doi.org/10.1016/j.tics.2005.04.010

Marotta, A., Lupiáñez, J., Martella, D., & Casagrande, M. (2012). Eye gaze versus arrows as spatial cues: Two qualitatively different modes of attentional selection. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(2), 326–335. https://doi.org/10.1037/a0023959

Marotta, A., Román-Caballero, R., & Lupiáñez, J. (2018). Arrows don't look at you: Qualitatively different attentional mechanisms triggered by gaze and arrows. *Psychonomic Bulletin and Review*, *25*(6), 2254–2259. https://doi.org/10.3758/s13423-018-1457-2

Marques, C., Moreno, S., Castro, S. L., & Besson, M. (2007). Musicians detect pitch violation in a foreign language better than nonmusicians: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *19*(9), 1453–1463. https://doi.org/10.1162/jocn.2007.19.9.1453

Marsh, R., Zhu, H., Schultz, R. T., Quackenbush, G., Royal, J., Skudlarski, P., & Peterson, B. S. (2006). A developmental fMRI study of self-regulatory control. *Human Brain Mapping*, *27*(11), 848–863. https://doi.org/10.1002/hbm.20225

Martin, N. A., & Brownell, R. (2011). *ROWPVT-4: Receptive One-Word Picture Vocabulary Test*. Pro-ed an International Publisher.

Martins, M., Pinheiro, A. P., & Lima, C. F. (2021). Does Music Training Improve Emotion Recognition Abilities? A Critical Review. *Emotion Review*, 175407392110220. https://doi.org/10.1177/17540739211022035

Masson, J., & McCarthy, S. (1996). When elephants weep: the emotional lives of animals Random House. *New York*.

Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., Cooper, N., Gordon, E., & Williams, L. M. (2009). Explicit identification and implicit recognition of facial emotions: II. Core domains and relationships with general cognition. *Journal of Clinical and Experimental*

Neuropsychology, *31*(3), 278–291. https://doi.org/10.1080/13803390802043619

Matsumoto, D., Hwang, H. C., & Frank, M. G. (2012). Emotional Language and Political Aggression. *Journal of Language and Social Psychology*, *32*(4), 452–468. https://doi.org/10.1177/0261927X12474654

Mattavelli, G., Barvas, E., Longo, C., Zappini, F., Ottaviani, D., Malaguti, M. C., Pellegrini, M., & Papagno, C. (2021). Facial expressions recognition and discrimination in Parkinson’s disease. *Journal of Neuropsychology*, *15*(1), 46–68. https://doi.org/10.1111/jnp.12209

Maximo, J. O., Cadena, E. J., & Kana, R. K. (2014). The implications of brain connectivity in the neuropsychology of autism. *Neuropsychology Review*, *24*(1), 16–31. https://doi.org/10.1007/s11065-014-9250-0

Mayes, S. D., Calhoun, S. L., Dickerson, S., Learning, S. L. C., Mayes, S. D., & Calhoun, S. L. (2007). Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychology*, *13*(6), 469–493. https://doi.org/10.1080/09297040601112773

Mazefsky, C. A., & Oswald, D. P. (2007). Emotion perception in Asperger’s syndrome and high-functioning autism: The importance of diagnostic criteria and cue intensity. *Journal of Autism and Developmental Disorders*, *37*(6), 1086–1095. https://doi.org/10.1007/s10803-006-0251-6

Mazurek, M. O. (2014). Loneliness, friendship, and well-being in adults with autism spectrum disorders. *Autism*, *18*(3), 223–232. https://doi.org/10.1177/1362361312474121

McKelvie, S. J. (1995). Emotional expression in upside-down faces: Evidence for configurational and componential processing. *British Journal of Social Psychology*, *34*(3), 325–334. https://doi.org/10.1111/j.2044-8309.1995.tb01067.x

McKenzie, K., Murray, A. L., Wilkinson, A., Murray, G. C., Metcalfe, D., O’Donnell, M., & McCarty, K. (2018). The relations between processing style, autistic-like traits, and emotion recognition in individuals with and without Autism Spectrum Disorder. *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2017.08.007

McKone, E., & Robbins, R. (2011). Are faces special?. In A. Calder, G. Ghodes, M. Johnson, & J. Haxby (Eds.), *Oxford Handbook of Face Perception* (pp. 149–176). Oxford University Press, Oxford, UK.

Merrill, J., Sammler, D., Bangert, M., Goldhahn, D., Lohmann, G., Turner, R., & Friederici, A. D. (2012). Perception of words and pitch patterns in song and speech. *Frontiers in Psychology*, *3*(76). https://doi.org/10.3389/fpsyg.2012.00076

Michael, J., & D'Ausilio, A. (2015). Domain-specific and domain-general processes in social perception - A complementary approach. *Consciousness and Cognition*, *36*, 434–437. https://doi.org/10.1016/j.concog.2014.12.009

Micheyl, C., Delhommeau, K., Perrot, X., & Oxenham, A. J. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, *219*(1–2), 36–47. https://doi.org/10.1016/j.heares.2006.05.004

Milosavljevic, B., Carter Leno, V., Simonoff, E., Baird, G., Pickles, A., Jones, C. R. G., Erskine, C., Charman, T., & Happé, F. (2016). Alexithymia in Adolescents with Autism Spectrum Disorder: Its Relationship to Internalising Difficulties, Sensory Modulation and Social Cognition. *Journal of Autism and Developmental Disorders*, *46*(4), 1354–1367. https://doi.org/10.1007/s10803-015-2670-8

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuropsychological Society*, *3*(4), 303–316. https://doi.org/10.1017/s1355617797003032

Mitchell, R. L. C., Elliott, R., Barry, M., Cruttenden, A., & Woodruff, P. W. R. (2003). The neural response to emotional prosody, as revealed by functional magnetic resonance imaging. *Neuropsychologia*, *41*(10), 1410–1421. https://doi.org/10.1016/S0028-3932(03)00017-4

Miyahara, M., Bray, A., Tsujii, M., Fujita, C., & Sugiyama, T. (2007). Reaction time of facial affect recognition in Asperger's disorder for cartoon and real, static and moving faces. *Child Psychiatry and Human Development*, *38*(2), 121–134. https://doi.org/10.1007/s10578-007-0048-7

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J. J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., … Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7). https://doi.org/10.1371/journal.pmed.1000097

Mohn, C., Argstatter, H., & Wilker, F. W. (2011). Perception of six basic emotions in music. *Psychology of Music*, *39*(4), 503–517. https://doi.org/10.1177/0305735610378183

Mokhsin, M. B., Rosli, N. B., Zambri, S., Ahmad, N. D., & Hamidi, S. R. (2014). Automatic music emotion classification using artificial neural network based on vocal and instrumental sound timbres. *Journal of Computer Science*, *10*(12), 2584–2592. https://doi.org/10.3844/jcssp.2014.2584.2592

Molnar-Szakacs, I., & Heaton, P. (2012). Music: A unique window into the world of autism. *Annals of the New York Academy of Sciences*, *1252*(1), 318–324. https://doi.org/10.1111/j.1749-6632.2012.06465.x

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

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

Monk, C. S., McClure, E. B., Nelson, E. E., Zarahn, E., Bilder, R. M., Leibenluft, E., Charney, D. S., Ernst, M., & Pine, D. S. (2003). Adolescent immaturity in attention-related brain engagement to emotional facial expressions. *NeuroImage*, *20*(1), 420–428. https://doi.org/10.1016/S1053-8119(03)00355-0

Monnot, M., Orbelo, D., Riccardo, L., Sikka, S., & Ross, E. (2003). Acoustic Analyses Support Subjective Judgments of Vocal Emotion. *Annals of the New York Academy of Sciences*, *1000*, 288–292. https://doi.org/10.1196/annals.1280.027

Moors, A. (2010). Automatic constructive appraisal as a candidate cause of emotion. *Emotion Review*, *2*(2), 139–156. https://doi.org/10.1177/1754073909351755

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, *19*(3), 712–723. https://doi.org/10.1093/cercor/bhn120

Morrison, A., Polisena, J., Husereau, D., Moulton, K., Clark, M., Fiander, M., Mierzwinski-Urban, M., Clifford, T., Hutton, B., & Rabb, D. (2012). The effect of english-language restriction on systematic review-based meta-analyses: A systematic review of empirical studies. *International Journal of Technology Assessment in Health Care*, *28*(2), 138–144. https://doi.org/10.1017/S0266462312000086

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development*, *72*(3), 834–843. https://doi.org/10.1111/1467-8624.00318

Mottron, L., Burack, J. A., Stauder, J. E. A., & Robaey, P. (1999). Perceptual processing among high-functioning persons with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(2), 203–211. https://doi.org/10.1017/S0021963098003333

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, *36*(1), 27–43. https://doi.org/10.1007/s10803-005-0040-7

Mottron, L., Peretz, I., & Ménard, E. (2000). Local and global processing of music in high-functioning persons with autism: Beyond central coherence? *Journal of Child Psychology and Psychiatry*, *41*(8), 1057–1065. https://doi.org/10.1111/1469-7610.00693

Murphy, S. T., & Zajonc, R. B. (1993). Affect, Cognition, and Awareness: Affective Priming with Optimal and Suboptimal Stimulus Exposures. *Journal of Personality and Social Psychology*, *64*(5), 723–739.

Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, *13*(5), 416–424. https://doi.org/10.1111/1467-9280.00474

Nava, E., & Pavani, F. (2013). Changes in Sensory Dominance During Childhood: Converging Evidence From the Colavita Effect and the Sound-Induced Flash Illusion. *Child Development*, *84*(2), 604–616. https://doi.org/10.1111/j.1467-8624.2012.01856.x

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226–254. https://doi.org/10.1037/0096-3445.106.3.226

Neimann Rasmussen, L., & Montgomery, P. (2018). The prevalence of and factors associated with inclusion of non-English language studies in Campbell systematic reviews: A survey and meta-epidemiological study. *Systematic Reviews*, *7*(1). https://doi.org/10.1186/s13643-018-0786-6

Nemiah, J. C., & Sifneos, P. E. (1970). Psychosomatic illness: a problem in communication. *Psychotherapy and Psychosomatics*, *18*(1–6), 154–160. https://doi.org/10.1159/000286074

Neuberg, S. L., Kenrick, D. T., & Schaller, M. (2011). Human threat management systems: Self-protection and disease avoidance. *Neuroscience and Biobehavioral Reviews*, *35*(4), 1042–1051. https://doi.org/10.1016/j.neubiorev.2010.08.011

Neumann, D., Spezio, M. L., Piven, J., & Adolphs, R. (2006). Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience*, *1*(3), 194–202. https://doi.org/10.1093/scan/nsl030

Neumann, O. (1990). Direct parameter specification and the concept of perception. *Psychological Research*, *52*(2), 207–215.

Nevison, C. D. (2014). A comparison of temporal trends in United States autism prevalence to trends in suspected environmental factors. *Environmental Health: A Global Access Science Source*, *13*(1), 1–16. https://doi.org/10.1186/1476-069X-13-73

Ng, M., de Montigny, J. G., Ofner, M., & Do, M. T. (2017). Environmental factors associated with autism spectrum disorder: A scoping review for the years 2003–2013. *Health Promotion and Chronic Disease Prevention in Canada*, *37*(1), 1–23. https://doi.org/10.24095/hpcdp.37.1.01

Nicolaidis, C. (2012). What can physicians learn from the neurodiversity movement? *Virtual Mentor*, *14*(6), 503–510. https://doi.org/10.1001/virtualmentor.2012.14.6.oped1-1206

Nordström, H., & Laukka, P. (2019). The time course of emotion recognition in speech and music. *The Journal of the Acoustical Society of America*, *145*(5), 3058–3074. https://doi.org/10.1121/1.5108601

Nowell, K. P., Schanding, G. T., Kanne, S. M., & Goin-Kochel, R. P. (2015). Cognitive profiles in youth with autism spectrum disorder: An investigation of base rate discrepancies using the Differential Ability Scales—Second Edition. *Journal of Autism and Developmental Disorders*, *45*(7), 1978–1988. https://doi.org/10.1007/s10803-014-2356-7

Nummenmaa, L., & Calvo, M. G. (2015). Dissociation between recognition and detection advantage for facial expressions: A meta-analysis. *Emotion*, *15*(2), 243–256. https://doi.org/10.1037/emo0000042

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

Nuske, H. J., Vivanti, G., Hudry, K., & Dissanayake, C. (2014). Pupillometry reveals reduced unconscious emotional reactivity in autism. *Biological Psychology*, *101*(1), 24–35. https://doi.org/10.1016/j.biopsycho.2014.07.003

Nwe, T. L., Foo, S. W., & De Silva, L. C. (2003). Speech emotion recognition using hidden Markov models. *Speech Communication*, *41*(4), 603–623. https://doi.org/10.1016/S0167-6393(03)00099-2

O'Connor, K. (2007). Brief report: Impaired identification of discrepancies between expressive faces and voices in adults with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *37*(10), 2008–2013. https://doi.org/10.1007/s10803-006-0345-1

Oerlemans, A. M., Droste, K., Van Steijn, D. J., De Sonneville, L. M. J., Buitelaar, J. K., & Rommelse, N. N. J. (2013). Co-segregation of social cognition, executive function and local processing style in children with ASD, their siblings and normal controls. *Journal of Autism and Developmental Disorders*, *43*(12), 2764–2778. https://doi.org/10.1007/s10803-013-1807-x

Oerlemans, A. M., Van Der Meer, J. M. J., Van Steijn, D. J., De Ruiter, S. W., De Bruijn, Y. G. E., De Sonneville, L. M. J., Buitelaar, J. K., & Rommelse, N. N. J. (2014). Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. *European Child and Adolescent Psychiatry*, *23*(5), 257–271. https://doi.org/10.1007/s00787-013-0446-2

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*(3), 381–396. https://doi.org/10.1037/0022-3514.80.3.381

Öhman, A., Soares, S. C., Juth, P., Lindstrm, B., & Esteves, F. (2012). Evolutionary derived modulations of attention to two common fear stimuli: Serpents and hostile humans. *Journal of Cognitive Psychology*, *24*(1), 17–32. https://doi.org/10.1080/20445911.2011.629603

Okon-Singer, H., Tzelgov, J., & Henik, A. (2007). Distinguishing between automaticity and attention in the processing of emotionally significant stimuli. *Emotion*, *7*(1), 147–157. https://doi.org/10.1037/1528-3542.7.1.147

Ola, L., & Gullon-Scott, F. (2020). Facial emotion recognition in autistic adult females correlates with alexithymia, not autism. *Autism*, *24*(8), 2021–2034. https://doi.org/10.1177/1362361320932727

Oliveras-Rentas, R. E., Kenworthy, L., Roberson, R. B., Martin, A., & Wallace, G. L. (2012). WISC-IV profile in high-functioning autism spectrum disorders: Impaired processing speed is associated with increased autism communication symptoms and decreased adaptive communication abilities. *Journal of Autism and Developmental Disorders*, *42*(5), 655–664. https://doi.org/10.1007/s10803-011-1289-7

Orsini, M., & Smith, M. (2010). Social movements, knowledge and public policy: The case of autism activism in Canada and the US. *Critical Policy Studies*, *4*(1), 38–57. https://doi.org/10.1080/19460171003714989

Otsuka, S., Uono, S., Yoshimura, S., Zhao, S., & Toichi, M. (2017). Emotion perception mediates the predictive relationship between verbal ability and functional outcome in high-functioning adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(4), 1166–1182. https://doi.org/10.1007/s10803-017-3036-1

Ouimet, T., Foster, N. E. V. V, Tryfon, A., & Hyde, K. L. (2012). Auditory-musical processing in autism spectrum disorders: A review of behavioral and brain imaging studies. *Annals of the New York Academy of Sciences*, *1252*(1), 325–331. https://doi.org/10.1111/j.1749-6632.2012.06453.x

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1990). Are there emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry*, *31*(3), 343–361. https://doi.org/10.1111/j.1469-7610.1990.tb01574.x

Ozonoff, S., Strayer, D. L., McMahon, W. M., & Filloux, F. (1994). Executive Function Abilities in Autism and Tourette Syndrome: An Information Processing Approach. *Journal of Child Psychology and Psychiatry*, *35*(6), 1015–1032. https://doi.org/10.1111/j.1469-7610.1994.tb01807.x

Padovan, C., Versace, R., Thomas-Antérion, C., & Laurent, B. (2002). Evidence for a selective deficit in automatic activation of positive information in patients with Alzheimer’s disease in an affective priming paradigm. *Neuropsychologia*, *40*(3), 335–339. https://doi.org/10.1016/S0028-3932(01)00101-4

Paez, A. (2017). Gray literature: An important resource in systematic reviews. *Journal of Evidence-Based Medicine*, *10*(3), 233–240. https://doi.org/10.1111/jebm.12266

Page, M. J., Higgins, J. P. T., & Sterne, J. A. C. (2019). Assessing risk of bias due to missing results in a synthesis. *Cochrane Handbook for Systematic Reviews of Interventions*, 349–374.

Palermo, R., O’Connor, K. B., Davis, J. M., Irons, J., & McKone, E. (2013). New tests to measure individual differences in matching and labelling facial expressions of emotion, and their association with ability to recognise vocal emotions and facial identity. *PLoS ONE*, *8*(6). https://doi.org/10.1371/journal.pone.0068126

Paquette, S., Peretz, I., & Belin, P. (2013). The “Musical Emotional Bursts”: A validated set of musical affect bursts to investigate auditory affective processing. *Frontiers in Psychology*, *4*, 1–7. https://doi.org/10.3389/fpsyg.2013.00509

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-In-noise. *Ear and Hearing*, *30*(6), 653–661.

https://doi.org/10.1097/AUD.0b013e3181b412e9

Parker, J. D. A., Eastabrook, J. M., Keefer, K. V., & Wood, L. M. (2010). Can Alexithymia Be Assessed in Adolescents? Psychometric Properties of the 20-Item Toronto Alexithymia Scale in Younger, Middle, and Older Adolescents. *Psychological Assessment*, *22*(4), 798–808. https://doi.org/10.1037/a0020256

Parker, J. D. A., Keefer, K. V., Taylor, G. J., & Bagby, R. M. (2008). Latent Structure of the Alexithymia Construct: A Taxometric Investigation. *Psychological Assessment*, *20*(4), 385–396. https://doi.org/10.1037/a0014262

Parker, J., Taylor, G., & Bagby, M. (1993). Alexithymia and the processing of emotional stimuli: An experimental study. *New Trends in Experimental and Clinical Psychiatry*, *9*(1–2), 9–14.

Parkinson, B. (2005). Do facial movements express emotions or communicate motives? *Personality and Social Psychology Review*, *9*(4), 278–311. https://doi.org/10.1207/s15327957pspr0904_1

Parr, L. A., Waller, B. M., & Fugate, J. (2005). Emotional communication in primates: Implications for neurobiology. *Current Opinion in Neurobiology*, *15*(6), 716–720. https://doi.org/10.1016/j.conb.2005.10.017

Parron, C., Da Fonseca, D., Santos, A., Moore, D. G., Monfardini, E., & Deruelle, C. (2008). Recognition of biological motion in children with autistic spectrum disorders. *Autism*, *12*(3), 261–274. https://doi.org/10.1177/1362361307089520

Passarotti, A. M., Sweeney, J. A., & Pavuluri, M. N. (2009). Neural correlates of incidental and directed facial emotion processing in adolescents and adults. *Social Cognitive and Affective Neuroscience*, *4*(4), 387–398. https://doi.org/10.1093/scan/nsp029

Patel, A. D. (2010). *Music, language, and the brain*. Oxford university press.

Paulmann, S., & Kotz, S. A. (2008). Early emotional prosody perception based on different speaker voices. *NeuroReport*, *19*(2), 209–213. https://doi.org/10.1097/WNR.0b013e3282f454db

Paulmann, S., & Pell, M. D. (2011). Is there an advantage for recognizing multi-modal emotional stimuli? *Motivation and Emotion*, *35*(2), 192–201. https://doi.org/10.1007/s11031-011-9206-0

Pavlova, M. A., Guerreschi, M., Tagliavento, L., Gitti, F., Sokolov, A. N., Fallgatter, A. J., & Fazzi, E. (2017). Social cognition in autism: Face tuning. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-02790-1

Pavuluri, M. N., Passarotti, A. M., Harral, E. M., & Sweeney, J. A. (2009). An fMRI study of the neural correlates of incidental versus directed emotion processing in pediatric bipolar disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *48*(3), 308–319. https://doi.org/10.1097/CHI.0b013e3181948fc7

Peelen, M. V., Atkinson, A. P., & Vuilleumier, P. (2010). Supramodal representations of perceived emotions in the human brain. *Journal of Neuroscience*, *30*(30), 10127–10134. https://doi.org/10.1523/JNEUROSCI.2161-10.2010

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pell, M. D. (2001). Influence of emotion and focus location on prosody in matched statements and questions. *The Journal of the Acoustical Society of America*, *109*(4), 1668–1680. https://doi.org/10.1121/1.1352088

Pell, M. D. (2005). Nonverbal emotion priming: Evidence from the “Facial Affect Decision Task.” *Journal of Nonverbal Behavior*, *29*(1), 45–73. https://doi.org/10.1007/s10919-004-0889-8

Pell, M. D., Jaywant, A., Monetta, L., & Kotz, S. A. (2011). Emotional speech processing: Disentangling the effects of prosody and semantic cues. *Cognition and Emotion*, *25*(5), 834–853. https://doi.org/10.1080/02699931.2010.516915

Pell, M. D., & Kotz, S. A. (2011). On the time course of vocal emotion recognition. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0027256

Pell, M. D., Monetta, L., Paulmann, S., & Kotz, S. A. (2009). Recognizing Emotions in a Foreign Language. *Journal of Nonverbal Behavior*, *33*(2), 107–120. https://doi.org/10.1007/s10919-008-0065-7

Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics*, *37*(4), 417–435. https://doi.org/10.1016/j.wocn.2009.07.005

Pell, M. D., & Skorup, V. (2008). Implicit processing of emotional prosody in a foreign versus native language. *Speech Communication*, *50*(6), 519–530. https://doi.org/10.1016/j.specom.2008.03.006

Pellicano, E., & den Houting, J. (2021). Annual Research Review: Shifting from ‘normal science’ to neurodiversity in autism science. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. https://doi.org/10.1111/jcpp.13534

Pelphrey, K., Goldman, B. D., Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, *32*(4), 249–261. https://doi.org/10.1023/A

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron*, *33*, 185–191. https://doi.org/10.1016/B978-0-444-62630-1.00033-0

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of Musical Disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*(1), 58–75. https://doi.org/10.1196/annals.1284.006

Peretz, I., Gagnon, L., & Bouchard, B. (1998). Music and emotion: Perceptual determinants, immediacy, and isolation after brain damage. *Cognition*, *68*(2), 111–141. https://doi.org/10.1016/S0010-0277(98)00043-2

Peretz, I., Gosselin, N., Nan, Y., Caron-Caplette, E., Trehub, S. E., & Béland, R. (2013). A novel tool for evaluating children's musical abilities across age and culture. *Frontiers in Systems Neuroscience*, *7*, 1–10. https://doi.org/10.3389/fnsys.2013.00030

Peretz, I., Vuvan, D., Lagrois, M.-E., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140090–20140090. https://doi.org/10.1098/rstb.2014.0090

Pfordresher, P. Q., & Halpern, A. R. (2013). Auditory imagery and the poor-pitch singer. *Psychonomic Bulletin and Review*, *20*(4), 747–753. https://doi.org/10.3758/s13423-013-0401-8

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, *16*(2), 331–348. https://doi.org/10.1006/nimg.2002.1087

Philip, R. C. M., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., Atkinson, A. P., Calder, A. J., Johnstone, E. C., Lawrie, S. M., & 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

Pinheiro, A. P., Vasconcelos, M., Dias, M., Arrais, N., & Gonçalves, Ó. F. (2015). The music of language: An ERP investigation of the effects of musical training on emotional prosody processing. *Brain and Language*, *140*, 24–34. https://doi.org/10.1016/j.bandl.2014.10.009

Plaisted, K., Swettenham, J., & Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(5), 733–742. https://doi.org/10.1017/S0021963099004102

Pletzer, B., Petasis, O., & Cahill, L. (2014). Switching between forest and trees: Opposite relationship of progesterone and testosterone to global-local processing. *Hormones and Behavior*, *66*(2), 257–266. https://doi.org/10.1016/j.yhbeh.2014.05.004

Poljac, E., Poljac, E., & Wagemans, J. (2013). Reduced accuracy and sensitivity in the perception of emotional facial expressions in individuals with high autism spectrum traits. *Autism*, *17*(6), 668–680. https://doi.org/10.1177/1362361312455703

Pons, F., Harris, P. L., & de Rosnay, M. (2004). Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. *European Journal of Developmental Psychology*, *1*(2), 127–152. https://doi.org/10.1080/17405620344000022

Poquérusse, J., Pastore, L., Dellantonio, S., & Esposito, G. (2018). Alexithymia and autism spectrum disorder: A complex relationship. *Frontiers in Psychology*, *9*, 1–10. https://doi.org/10.3389/fpsyg.2018.01196

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. *Attention and Performance V*, 669–682. papers2://publication/uuid/5E57ECE7-3B3B-4C34-9C4F-CDF5620D8711

Posner, M. I., & Snyder, C. R. R. (2004). Attention and Cognitive Control. In *Cognitive psychology: Key readings.* Psychology Press.

Postorino, V., Fatta, L. M., Sanges, V., Giovagnoli, G., De Peppo, L., Vicari, S., & Mazzone, L. (2016). Intellectual disability in Autism Spectrum Disorder: Investigation of prevalence in an Italian sample of children and adolescents. *Research in Developmental Disabilities*, *48*, 193–201. https://doi.org/10.1016/j.ridd.2015.10.020

Pralus, A., Fornoni, L., Bouet, R., Gomot, M., Bhatara, A., Tillmann, B., & Caclin, A. (2019). Emotional prosody in congenital amusia: Impaired and spared processes. *Neuropsychologia*, *134*, 107234. https://doi.org/10.1016/j.neuropsychologia.2019.107234

Prinz, J. (2004). Which emotions are basic? *Emotion, Evolution, and Rationality*, 1–19. https://doi.org/10.1093/acprof:oso/9780198528975.003.0004

Prkachin, G. C. (2003). The effects of orientation on detection and identification of facial expressions of emotion. *British Journal of Psychology*, *94*(1), 45–62. https://doi.org/10.1348/000712603762842093

Prkachin, G. C., Casey, C., & Prkachin, K. M. (2009). Alexithymia and perception of facial expressions of emotion. *Personality and Individual Differences*, *46*(4), 412–417. https://doi.org/10.1016/j.paid.2008.11.010

Quam, C., & Swingley, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development*, *83*(1), 236–250. https://doi.org/10.1111/j.1467-8624.2011.01700.x

Quintin, E. M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2011). Emotion perception in music in high-functioning adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*(9), 1240–1255. https://doi.org/10.1007/s10803-010-1146-0

Quintin, E. M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2013). Processing of musical structure by high-functioning adolescents with autism spectrum disorders. *Child Neuropsychology*, *19*(3), 250–275. https://doi.org/10.1080/09297049.2011.653540

Quinto, L., Thompson, W. F., & Taylor, A. (2014). The contributions of compositional structure and performance expression to the communication of emotion in music. *Psychology of Music*, *42*(4), 503–524. https://doi.org/10.1177/0305735613482023

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230–236. https://doi.org/10.1016/j.cognition.2010.08.003

R Core Team. (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. *URL Https://Www. R-Project. Org*.

Rao, K. S., Koolagudi, S. G., & Vempada, R. R. (2013). Emotion recognition from speech using global and local prosodic features. *International Journal of Speech Technology*, *16*(2), 143–160. https://doi.org/10.1007/s10772-012-9172-2

Raven, J. C. (1983). Manual for Raven's progressive matrices and vocabulary scales. *Standard Progressive Matrices*.

Ravindranath, O., Ordaz, S. J., Padmanabhan, A., Foran, W., Jalbrzikowski, M., Calabro, F. J., & Luna, B. (2020). Influences of affective context on amygdala functional connectivity during cognitive control from adolescence through adulthood. *Developmental Cognitive Neuroscience*, *45*, 100836. https://doi.org/10.1016/j.dcn.2020.100836

Reisinger, D. L., Shaffer, R. C., Horn, P. S., Hong, M. P., Pedapati, E. V., Dominick, K. C., & Erickson, C. A. (2020). Atypical Social Attention and Emotional Face Processing in Autism Spectrum Disorder: Insights From Face Scanning and Pupillometry. *Frontiers in Integrative Neuroscience*, *13*, 1–14. https://doi.org/10.3389/fnint.2019.00076

Resnicow, J. E., Salovey, P., & Repp, B. H. (2004). Is recognition of emotion in music performance an aspect of emotional intelligence? *Music Perception: An Interdisciplinary Journal*, *22*(1), 145–158. https://doi.org/10.1525/mp.2004.22.1.145

Rhodes, G., Burton, N., Jeffery, L., Read, A., Taylor, L., & Ewing, L. (2018). Facial expression coding in children and adolescents with autism : Reduced adaptability but intact norm-based coding. *British Journal of Psychology*, *109*(2), 204–218. https://doi.org/10.1111/bjop.12257

Riby, D. M., & Hancock, P. J. B. (2009). Looking at movies and cartoons: Eye-tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability Research*, *53*(2), 169–181. https://doi.org/10.1111/j.1365-2788.2008.01142.x

Riby, D. M., Whittle, L., & Doherty-Sneddon, G. (2012). Physiological reactivity to faces via live and video-mediated communication in typical and atypical development. *Journal of Clinical and Experimental Neuropsychology*, *34*(4), 385–395. https://doi.org/10.1080/13803395.2011.645019

Rice, L. M., Wall, C. A., Fogel, A., & Shic, F. (2015). Computer-Assisted Face Processing Instruction Improves Emotion Recognition, Mentalizing, and Social Skills in Students with ASD. *Journal of Autism and Developmental Disorders*, *45*(7), 2176–2186. https://doi.org/10.1007/s10803-015-2380-2

Rinehart, N. J., Bradshaw, J. L., Moss, S. A., Brereton, A. V., & Tonge, B. J. (2000). Atypical interference of local detail on global processing in high-functioning autism and Asperger's disorder. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(6), 769–778. https://doi.org/10.1017/S002196309900596X

Robert, F., & Robert, J. (2000). *Faces*. Chronicle books.

Robertson, S. M. (2010). Neurodiversity, quality of life, and autistic adults: shifting research and professional focuses onto real-life challenges. *Disability Studies Quarterly*, *30*(1).

Robinson, C. W., & Sloutsky, V. M. (2004). Auditory dominance and its change in the course of development. *Child Development*, *75*(5), 1387–1401. https://doi.org/10.1111/j.1467-8624.2004.00747.x

Rommelse, N., Langerak, I., Van Der Meer, J., De Bruijn, Y., Staal, W., Oerlemans, A., & Buitelaar, J. (2015). Intelligence may moderate the cognitive profile of patients with ASD. *PLoS ONE*, *10*(10), 1–17. https://doi.org/10.1371/journal.pone.0138698

Rosen, N. E., Lord, C., & Volkmar, F. R. (2021). The Diagnosis of Autism: From Kanner to DSM-III to DSM-5 and Beyond. *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-021-04904-1

Rosenberg, N., Ihme, K., Lichev, V., Sacher, J., Rufer, M., Grabe, H. J., Kugel, H., Pampel, A., Lepsien, J., Kersting, A., Villringer, A., & Suslow, T. (2020). Alexithymia and automatic processing of facial emotions: Behavioral and neural findings. *BMC Neuroscience*, *21*(1), 1–13. https://doi.org/10.1186/s12868-020-00572-6

Ross, P., Atkins, B., Allison, L., Simpson, H., Duffell, C., Williams, M., & Ermolina, O. (2021). Children cannot ignore what they hear: Incongruent emotional information leads to an auditory dominance in children. *Journal of Experimental Child Psychology*, *204*, 105068. https://doi.org/10.1016/j.jecp.2020.105068

Rossell, S. L., Price, C. J., & Nobre, A. C. (2003). The anatomy and time course of semantic priming investigated by fMRI and ERPs. *Neuropsychologia*, *41*(5), 550–564. https://doi.org/10.1016/S0028-3932(02)00181-1

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

Rosset, D. B., Santos, A., Da Fonseca, D., Poinso, F., O'Connor, K., & Deruelle, C. (2010). Do children perceive features of real and cartoon faces in the same way? Evidence from typical development and autism. *Journal of Clinical and Experimental Neuropsychology*, *32*(2), 212–218. https://doi.org/10.1080/13803390902971123

Roux, P., Christophe, A., & Passerieux, C. (2010). The emotional paradox: Dissociation between explicit and implicit processing of emotional prosody in schizophrenia. *Neuropsychologia*, *48*(12), 3642–3649. https://doi.org/10.1016/j.neuropsychologia.2010.08.021

RStudio Team. (2018). *RStudio: integrated development for R. RStudio, Inc, Boston, MA*.

Ruesch, J. (1948). The infantile personality; the core problem of psychosomatic medicine. *Psychosomatic Medicine, 10,* 134-144. https://doi.org/10.1097/00006842-194805000-00002

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

Russell, G., Stapley, S., Newlove-Delgado, T., Salmon, A., White, R., Warren, F., Pearson, A., & Ford, T. (2021). Time trends in autism diagnosis over 20 years: a UK population-based cohort study. *Journal of Child Psychology and Psychiatry*. https://doi.org/10.1111/jcpp.13505

Russell, J. A. (1994). Is There Universal Recognition of Emotion From Facial Expression? A

Review of the Cross-Cultural Studies. *Psychological Bulletin*, *115*(1), 102–141. https://doi.org/10.1037//0033-2909.115.1.102

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Russell, J. A., Bachorowski, J.-A., & Fernández-Dols, J.-M. (2003). Facial and Vocal Expressions of Emotion. *Annual Review of Psychology*, *54*(1), 329–349. https://doi.org/10.1146/annurev.psych.54.101601.145102

Russell, J. A., & Widen, S. C. (2002). A label superiority effect in children's categorization of facial expressions. *Social Development*, *11*(1), 30–52. https://doi.org/10.1111/1467-9507.00185

Russo-Ponsaran, N. M., Evans-Smith, B., Johnson, J., Russo, J., & McKown, C. (2016). Efficacy of a Facial Emotion Training Program for Children and Adolescents with Autism Spectrum Disorders. *Journal of Nonverbal Behavior*, *40*(1), 13–38. https://doi.org/10.1007/s10919-015-0217-5

Russo-Ponsaran, N. M., Mckown, C., Johnson, J. K., Allen, A. W., Evans-Smith, B., & Fogg, L. (2015). Social-emotional correlates of early stage social information processing skills in children with and without autism spectrum disorder. *Autism Research*, *8*(5), 486–496. https://doi.org/10.1002/aur.1463

Rutherford, M. D., & McIntosh, D. N. (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. 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

Rutter, M., Couteur, A. Le, & Lord, C. (2003). *The autism diagnostic interview - revised (ADI-R)*. Western Psychological Services.

Ruzich, E., Allison, C., Smith, P., Watson, P., AuYeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism*, *6*(1), 1–12. https://doi.org/10.1186/2040-2392-6-2

Ryan, C., Stafford, M., & King, R. J. (2016). Brief Report: Seeing the Man in the Moon: Do Children with Autism Perceive Pareidolic Faces? A Pilot Study. *Journal of Autism and*

Developmental Disorders, *46*(12), 3838–3843. https://doi.org/10.1007/s10803-016-2927-x

Safar, K., Vandewouw, M. M., & Taylor, M. J. (2021). Atypical development of emotional face processing networks in autism spectrum disorder from childhood through to adulthood. *Developmental Cognitive Neuroscience*, *51*, 101003. https://doi.org/10.1016/j.dcn.2021.101003

Sahuquillo-Leal, R., Navalón, P., Moreno-Giménez, A., Almansa, B., Vento, M., & García-Blanco, A. (2022). Attentional biases towards emotional scenes in autism spectrum condition: An eye-tracking study. *Research in Developmental Disabilities*, *120*(November 2021), 1–13. https://doi.org/10.1016/j.ridd.2021.104124

Salminen, J. K., Saarijärvi, S., Äärelä, E., Toikka, T., & Kauhanen, J. (1999). Prevalence of alexithymia and its association with sociodemographic variables in the general population of Finland. *Journal of Psychosomatic Research*, *46*(1), 75–82. https://doi.org/10.1016/S0022-3999(98)00053-1

Salomone, E., Bulgarelli, D., Thommen, E., Rossini, E., & Molina, P. (2019). Role of age and IQ in emotion understanding in Autism Spectrum Disorder: implications for educational interventions. *European Journal of Special Needs Education*, *34*(3), 383–392. https://doi.org/10.1080/08856257.2018.1451292

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their Way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(5), 1255–1266. https://doi.org/10.1037/a0018729

Santiesteban, I., Catmur, C., Hopkins, S. C., & Bird, G. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 929–937. https://doi.org/10.1037/a0035175

Santos, A., Rosset, D., & Deruelle, C. (2009). Human versus non-human face processing: Evidence from williams syndrome. *Journal of Autism and Developmental Disorders*, *39*(11), 1552–1559. https://doi.org/10.1007/s10803-009-0789-1

Sasson, N. J., Shasteen, J. R., & Pinkham, A. E. (2016). Brief report: Reduced prioritization of facial threat in adults with autism. *Journal of Autism and Developmental Disorders*, *46*(4), 1471–1476. https://doi.org/10.1007/s10803-015-2664-6

Sawyer, A. C. P., Williamson, P., & Young, R. L. (2012). Can gaze avoidance explain why individuals with Asperger's syndrome can't recognise emotions from facial expressions? *Journal of Autism and Developmental Disorders*, *42*(4), 606–618.

https://doi.org/10.1007/s10803-011-1283-0

Schelinski, S., & von Kriegstein, K. (2019). The relation between vocal pitch and vocal emotion recognition abilities in people with autism spectrum disorder and typical development. *Journal of Autism and Developmental Disorders*, *49*(1), 68–82. https://doi.org/10.1007/s10803-018-3681-z

Schellenberg, E. G., Krysciak, A. M., & Campbell, R. J. (2000). Perceiving emotion in melody : Interactive effects of pitch and rhythm. *Music Perception: An Interdisciplinary Journal*, *18*(2), 155–171. https://doi.org/10.2307/40285907

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*(2), 143–165. https://doi.org/10.1037/0033-2909.99.2.143

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. *Appraisal Processes in Emotion: Theory, Methods, Research*, *92*(120), 57.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*(1–2), 227–256. https://doi.org/10.1016/S0167-6393(02)00084-5

Scherer, K. R. (2009). Emotions are emergent processes: They require a dynamic computational architecture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3459–3474. https://doi.org/10.1098/rstb.2009.0141

Scherer, K. R., Ladd, D. R., & Silverman, K. E. A. (1984). Vocal cues to speaker affect: Testing two models. *The Journal of the Acoustical Society of America*, *76*(5), 1346–1356. https://doi.org/10.1121/1.391450

Scherer, K. R., & Oshinsky, J. S. (1977). Cue utilization in emotion attribution from auditory stimuli. *Motivation and Emotion*, *1*(4), 331–346. https://doi.org/10.1007/BF00992539

Scherer, K. R., & Scherer, U. (2011). Assessing the Ability to Recognize Facial and Vocal Expressions of Emotion: Construction and Validation of the Emotion Recognition Index. *Journal of Nonverbal Behavior*, *35*(4), 305–326. https://doi.org/10.1007/s10919-011-0115-4

Scherer, K. R., Sundberg, J., Tamarit, L., & Salomão, G. L. (2015). Comparing the acoustic expression of emotion in the speaking and the singing voice. *Computer Speech and Language*, *29*(1), 218–235. https://doi.org/10.1016/j.csl.2013.10.002

Scherer, L. D., & Larsen, R. J. (2011). Cross-Modal Evaluative Priming: Emotional Sounds Influence the Processing of Emotion Words. *Emotion*, *11*(1), 203–208. https://doi.org/10.1037/a0022588

Schirmer, A., & Adolphs, R. (2017). Emotion Perception from Face, Voice, and Touch:

Comparisons and Convergence. *Trends in Cognitive Sciences*, *21*(3), 216–228. https://doi.org/10.1016/j.tics.2017.01.001

Schirmer, A., & Kotz, S. A. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *Trends in Cognitive Sciences*, *10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Schlegel, K., Grandjean, D., & Scherer, K. R. (2012). Emotion recognition: Unidimensional ability or a set of modality- and emotion-specific skills? *Personality and Individual Differences*, *53*(1), 16–21. https://doi.org/10.1016/j.paid.2012.01.026

Schmucker, C. M., Blümle, A., Schell, L. K., Schwarzer, G., Oeller, P., Cabrera, L., Von Elm, E., Briel, M., & Meerpohl, J. J. (2017). Systematic review finds that study data not published in full text articles have unclear impact on meta-analyses results in medical research. *PLoS ONE*, *12*(4), 1–16. https://doi.org/10.1371/journal.pone.0176210

Schneider, E., & Zuccoloto, A. (2007). E-prime 2.0 [Computer software]. *Pittsburg, PA: Psychological Software Tools*.

Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science*, *27*(3), 525–559. https://doi.org/10.1016/S0364-0213(03)00011-9

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66. https://doi.org/10.1037/0033-295X.84.1.1

Schön, D., Magne, C., & Besson, M. (2004). The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology*, *41*(3), 341–349. https://doi.org/10.1111/1469-8986.00172.x

Schubert, E. (2004). Modeling perceived emotion with continuous musical features. *Music Perception: An Interdisciplinary Journal*, *21*(4), 561–585. https://doi.org/10.1525/mp.2004.21.4.561

Schünemann, H. J., Vist, G. E., Higgins, J. P., Santesso, N., Deeks, J. J., Glasziou, P., Akl, E. A., & Guyatt, G. H. (2019). Interpreting results and drawing conclusions. *Cochrane Handbook for Systematic Reviews of Interventions*, 403–431. https://doi.org/10.1002/9781119536604.ch15

Schwartz, R., & Pell, M. D. (2012). Emotional Speech Processing at the Intersection of Prosody and Semantics. *PLoS ONE*, *7*(10), 1–9. https://doi.org/10.1371/journal.pone.0047279

Schwarzer, G. (2000). Development of face processing: The effect of face inversion. *Child*

Development, *71*(2), 391–401. https://doi.org/10.1111/1467-8624.00152

Seltzer, M. M., Shattuck, P., Abbeduto, L., & Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *10*(4), 234–247. https://doi.org/10.1002/mrdd.20038

Shafritz, K. M., Bregman, J. D., Ikuta, T., & Szeszko, P. R. (2015). Neural systems mediating decision-making and response inhibition for social and nonsocial stimuli in autism. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *60*, 112–120. https://doi.org/10.1016/j.pnpbp.2015.03.001

Shah, A., & Frith, U. (1993). Why Do Autistic Individuals Show Superior Performance on the Block Design Task? *Journal of Child Psychology and Psychiatry*, *34*(8), 1351–1364. https://doi.org/10.1111/j.1469-7610.1993.tb02095.x

Shakes, P., & Cashin, A. (2019). Identifying Language for People on the Autism Spectrum: A Scoping Review. *Issues in Mental Health Nursing*, *40*(4), 317–325. https://doi.org/10.1080/01612840.2018.1522400

Sharda, M., Midha, R., Malik, S., Mukerji, S., & Singh, N. C. (2015). Fronto-Temporal connectivity is preserved during sung but not spoken word listening, across the autism spectrum. *Autism Research*, *8*(2), 174–186. https://doi.org/10.1002/aur.1437

Shariff, A. F., & Tracy, J. L. (2011). What are emotion expressions for? *Current Directions in Psychological Science*, *20*(6), 395–399. https://doi.org/10.1177/0963721411424739

Shattuck, P. T., Seltzer, M. M., Greenberg, J. S., Orsmond, G. I., Bolt, D., Kring, S., Lounds, J., & Lord, C. (2007). Change in autism symptoms and maladaptive behaviors in adolescents and adults with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *37*(9), 1735–1747. https://doi.org/10.1007/s10803-006-0307-7

Shen, L., Liu, X. K., Zhang, H., Lin, J., Feng, C., & Iqbal, J. (2020). Biomarkers in autism spectrum disorders: Current progress. *Clinica Chimica Acta*, *502*, 41–54. https://doi.org/10.1016/j.cca.2019.12.009

Sifneos, P. E. (1973). The prevalence of 'alexithymic' characteristics in psychosomatic patients. *Psychotherapy and Psychosomatics*, *22*(2–6), 255–262. https://doi.org/10.1159/000286529

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, *49*(22), 2705–2739. https://doi.org/10.1016/j.visres.2009.08.005

Sirois, S., Spratling, M., Thomas, M. S. ., Westermann, G., Mareschal, D., & Johnson, M. . (2008). Precis of neuroconstructivism : How the brain constructs cognition. *Behavioral and Brain Sciences*, *31*, 321–356. https://doi.org/10.1017/S0140525X0800407X

Sivathasan, S., Fernandes, T. P., Burack, J. A., & Quintin, E. M. (2020). Emotion processing and autism spectrum disorder: A review of the relative contributions of alexithymia and verbal IQ. *Research in Autism Spectrum Disorders*, *77*, 101608. https://doi.org/10.1016/j.rasd.2020.101608

Snowdon, C. T., Zimmermann, E., & Altenmüller, E. (2015). Music evolution and neuroscience. *Progress in Brain Research*, *217*, 17–34. https://doi.org/10.1016/bs.pbr.2014.11.019

Sobin, C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes fo fear, anger, sadness, and joy. *Journal of Psycholinguistic Research, 28*(4), 347–365. https://doi.org/10.1023/A:1023237014909

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), 1–9. https://doi.org/10.1007/s10803-017-3428-2

Song, Y., Zhong, J., Jia, Z., & Liang, D. (2020). Emotional prosody recognition in children with high-functioning autism under the influence of emotional intensity: Based on the perspective of emotional dimension theory. *Journal of Communication Disorders*, *88*, 106032. https://doi.org/10.1016/j.jcomdis.2020.106032

Soriano, M. F., Ibáñez-Molina, A. J., Paredes, N., & Macizo, P. (2018). Autism: Hard to Switch from Details to the Whole. *Journal of Abnormal Child Psychology*, *46*(6), 1359–1371. https://doi.org/10.1007/s10802-017-0384-4

Sota, S., Hatada, S., Honjyo, K., Takatsuka, T., & Honer, W. G. (2018). Musical disability in children with autism spectrum disorder. *Psychiatry Research*, *267*, 354–359. https://doi.org/10.1016/j.psychres.2018.05.078

Spiker, M. A., Lin, C. E., Van Dyke, M., & Wood, J. J. (2012). Restricted interests and anxiety in children with autism. *Autism*, *16*(3), 306–320. https://doi.org/10.1177/1362361311401763

Spineli, L. M., & Pandis, N. (2020). Prediction interval in random-effects meta-analysis. *American Journal of Orthodontics and Dentofacial Orthopedics*, *157*(4), 586–588. https://doi.org/10.1016/j.ajodo.2019.12.011

Spruyt, A., Hermans, D., De Houwer, J., Vandromme, H., & Eelen, P. (2007). On the nature of the affective priming effect: Effects of stimulus onset asynchrony and congruency proportion in naming and evaluative categorization. *Memory and Cognition*, *35*(1), 95–106. https://doi.org/10.3758/BF03195946

Steinbeis, N., & Koelsch, S. (2011). Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience*, *23*(3), 604–621. https://doi.org/10.1162/jocn.2009.21383

Stephenson, K. G., Luke, S. G., & South, M. (2019). Separate contributions of autistic traits and anxious apprehension, but not alexithymia, to emotion processing in faces. *Autism*, *23*(7), 1830–1842. https://doi.org/10.1177/1362361319830090

Sterne, J. A. C., Egger, M., & Moher, D. (2011). Chapter 10: Addressing reporting biases In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors),. In *Cochrane Handbook for Systematic Reviews of Interventions* (version 5.). Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Stewart, M. E., McAdam, C., Ota, M., Peppé, S., & Cleland, J. (2013). Emotional recognition in autism spectrum conditions from voices and faces. *Autism*, *17*(1), 6–14. https://doi.org/10.1177/1362361311424572

Sun, X., Allison, C., Wei, L., Matthews, F. E., Auyeung, B., Wu, Y. Y., Griffiths, S., Zhang, J., Baron-Cohen, S., & Brayne, C. (2019). Autism prevalence in China is comparable to Western prevalence. *Molecular Autism*, *10*(1), 1–19. https://doi.org/10.1186/s13229-018-0246-0

Suslow, T. (1998). Alexithymia and automatic affective processing. *European Journal of Personality*, *12*(6), 433–443. https://doi.org/10.1002/(SICI)1099-0984(199811/12)12:6<433::AID-PER320>3.0.CO;2-O

Suslow, T., & Donges, U. S. (2017). Alexithymia components are differentially related to explicit negative affect but not associated with explicit positive affect or implicit affectivity. *Frontiers in Psychology*, *8*(OCT), 1–8. https://doi.org/10.3389/fpsyg.2017.01758

Suslow, T., Kugel, H., Ohrmann, P., Stuhrmann, A., Grotegerd, D., Redlich, R., Bauer, J., & Dannlowski, U. (2013). Neural correlates of affective priming effects based on masked facial emotion: An fMRI study. *Psychiatry Research - Neuroimaging*, *211*(3), 239–245. https://doi.org/10.1016/j.pscychresns.2012.09.008

Suslow, T., Roestel, C., Droste, T., & Arolt, V. (2003). Automatic processing of verbal emotion stimuli in schizophrenia. *Psychiatry Research*, *120*(2), 131–144. https://doi.org/10.1016/S0165-1781(03)00173-2

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 303–314. https://doi.org/10.1098/rstb.2002.1198

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

Tanaka, J. W., & Sengco, J. A. (1997). Features and their configuration in face recognition. *Memory and Cognition*, *25*(5), 583–592. https://doi.org/10.3758/BF03211301

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

Tanaka, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., Brown, C., Stahl, S. S., South, M., McPartland, J. C., Kaiser, M. D., & Schultz, R. T. (2012). The perception and identification of facial emotions in individuals with autism spectrum disorders using the Let's Face It! Emotion Skills Battery. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(12), 1259–1267. https://doi.org/10.1111/j.1469-7610.2012.02571.x

Tang, J. S. Y., Chen, N. T. M., 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

Taylor, G. J., & Bagby, R. M. (2000). An overview of the alexithymia construct. In R. Bar-On & J. D. A. Parker (Eds.), *The handbook of emotional intelligence: Theory, development, assessment, and application at home, school, and in the workplace* (pp. 40-67). Jossey-Bass.

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. O. (2015). Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. *International Journal of Language and Communication Disorders*, *50*(4), 452–466. https://doi.org/10.1111/1460-6984.12146

Taylor, M. J., Batty, M., & Itier, R. J. (2004). The faces of development: A review of early face processing over childhood. *Journal of Cognitive Neuroscience*, *16*(8), 1426–1442. https://doi.org/10.1162/0898929042304732

Tell, D., Davidson, D., & Camras, L. A. (2014). Recognition of emotion from facial expressions with direct or averted eye gaze and varying expression intensities in children with autism disorder and typically developing children. *Autism Research and Treatment*, *2014*, 1–11. https://doi.org/10.1155/2014/816137

Teufel, C., Alexis, D. M., Clayton, N., & Davis, G. (2010). Mental-state attribution drives rapid, reflexive gaze following. *Attention, Perception & Psychophysics*, *72*(3), 695–107. https://doi.org/10.3758/APP

Thomas, L. A., De Bellis, M. D., Graham, R., & LaBar, K. S. (2007). Development of emotional facial recognition in late childhood and adolescence. *Developmental Science*, *10*(5), 547–558. https://doi.org/10.1111/j.1467-7687.2007.00614.x

Thomas, M. S. C., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*, *52*(2), 336–358. https://doi.org/10.1044/1092-4388(2009/07-0144)

Thompson, G. A., & Abel, L. A. (2018). Fostering Spontaneous Visual Attention in Children on the Autism Spectrum: A Proof-of-Concept Study Comparing Singing and Speech. *Autism Research*, *11*(5), 732–737. https://doi.org/10.1002/aur.1930

Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*(11), 1559–1573. https://doi.org/10.1002/sim.1187

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2004). Decoding Speech Prosody: Do Music Lessons Help? *Emotion*, *4*(1), 46–64. https://doi.org/10.1037/1528-3542.4.1.46

Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijsdijk, F. (2016). Heritability of autism spectrum disorders: A meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *57*(5), 585–595. https://doi.org/10.1111/jcpp.12499

Tobin, M. C., Drager, K. D. R., & Richardson, L. F. (2014). A systematic review of social participation for adults with autism spectrum disorders: Support, social functioning, and quality of life. *Research in Autism Spectrum Disorders*, *8*(3), 214–229. https://doi.org/10.1016/j.rasd.2013.12.002

Tong, F., Nakayama, K., Moscovitch, M., Weinrib, O., & Kanwisher, N. (2000). Response properties of the human fusiform face area. *Cognitive Neuropsychology*, *17*(1–3), 257–280. https://doi.org/10.1080/026432900380607

Tonks, J., Williams, W. H., Frampton, I., Yates, P., & Slater, A. (2007). Assessing emotion recognition in 9-15-years olds: Preliminary analysis of abilities in reading emotion from faces, voices and eyes. *Brain Injury*, *21*(6), 623–629. https://doi.org/10.1080/02699050701426865

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology*, *11*(4–5), 375–424. https://doi.org/10.1016/0162-3095(90)90017-Z

Torres-Quesada, M., Korb, F. M., Funes, M. J., Lupiáñez, J., & Egner, T. (2014). Comparing neural substrates of emotional vs. non-emotional conflict modulation by global control context. *Frontiers in Human Neuroscience*, *8*(1 FEB), 1–14. https://doi.org/10.3389/fnhum.2014.00066

Tosto, M. G., Momi, S. K., Asherson, P., & Malki, K. (2015). A systematic review of attention deficit hyperactivity disorder (ADHD) and mathematical ability: Current findings and future implications. *BMC Medicine*, *13*(1). https://doi.org/10.1186/s12916-015-0414-4

Tracy, J. L., Robins, R. W., Schriber, R. A., & Solomon, M. (2011). Is emotion recognition impaired in individuals with autism spectrum disorders? *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-010-1030-y

Trainor, L. (2010). The emotional origins of music. *Physics of Life Reviews*, *7*(1), 44–45. https://doi.org/10.1016/j.plrev.2010.01.010

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

Trimmer, C. G., & Cuddy, L. L. (2008). Emotional Intelligence, Not Music Training, Predicts Recognition of Emotional Speech Prosody. *Emotion*, *8*(6), 838–849. https://doi.org/10.1037/a0014080

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

Umemoto, T. (1990). The Psychological Structure of Music. *Music Perception: An Interdisciplinary Journal*, *8*(2), 115–127. https://doi.org/10.2307/40285492

Uono, S., Sato, W., & Toichi, M. (2011). The specific impairment of fearful expression recognition and its atypical development in pervasive developmental disorder. *Social Neuroscience*, *6*(5–6), 452–463. https://doi.org/10.1080/17470919.2011.605593

Uono, S., Sato, W., & Toichi, M. (2013). Common and unique impairments in facial-expression recognition in pervasive developmental disorder-not otherwise specified and Asperger's disorder. *Research in Autism Spectrum Disorders*, *7*(2), 361–368. https://doi.org/10.1016/j.rasd.2012.10.007

van 't Wout, M., Aleman, A., Bermond, B., & Kahn, R. S. (2007). No words for feelings: alexithymia in schizophrenia patients and first-degree relatives. *Comprehensive Psychiatry*, *48*(1), 27–33. https://doi.org/10.1016/j.comppsych.2006.07.003

Van der Donck, S., Dzhelyova, M., Vettori, S., Mahdi, S. S., Claes, P., Steyaert, J., & Boets, B. (2020). Rapid neural categorization of angry and fearful faces is specifically impaired in boys with autism spectrum disorder. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *61*(9), 1019–1029. https://doi.org/10.1111/jcpp.13201

Van der Donck, S., Dzhelyova, M., Vettori, S., Thielen, H., Steyaert, J., Rossion, B., & Boets, B. (2019). Fast Periodic Visual Stimulation EEG Reveals Reduced Neural Sensitivity to Fearful Faces in Children with Autism. *Journal of Autism and Developmental Disorders*, *49*(11), 4658–4673. https://doi.org/10.1007/s10803-019-04172-0

Van der Donck, S., Vettori, S., Dzhelyova, M., Mahdi, S. S., Claes, P., Steyaert, J., & Boets, B. (2021). Investigating automatic emotion processing in boys with autism via eye tracking and facial mimicry recordings. *Autism Research*, *14*(7), 1404–1420. https://doi.org/10.1002/aur.2490

Van Der Geest, J. N., Kemner, C., Camfferman, G., & Verbaten, M. N. (2002). Looking at images with human figures: Comparison between autistic and normal children. *Journal of Autism and Developmental Disorders*, *32*(2). https://doi.org/10.1023/A:1014832420206

Van Der Geest, J. N., Kemner, C., Verbaten, M. N., & Van Engeland, H. (2002). Gaze behavior of children with pervasive developmental disorder toward human faces: A fixation time study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*(5), 669–678. https://doi.org/10.1111/1469-7610.00055

Van der Hallen, R., Evers, K., Brewaeys, K., & Noortgate, W. Van Den. (2015). Global processing takes time: A meta-analysis on local-global visual processing in ASD. *Psychological Bulletin*, *141*(3), 549–573. https://doi.org/10.1037/bul0000004

Van Lancker, D., Cornelius, C., & Kreiman, J. (1989). Recognition of emotional-prosodic meanings in speech by autistic, schizophrenic, and normal children. *Developmental Neuropsychology*, *5*(2–3), 207–226. https://doi.org/10.1080/87565648909540433

Vanmarcke, S., & Wagemans, J. (2017). Priming Facial Gender and Emotional Valence: The Influence of Spatial Frequency on Face Perception in ASD. *Journal of Autism and Developmental Disorders*, *47*(4), 927–946. https://doi.org/10.1007/s10803-016-3017-9

Vannetzel, L., Chaby, L., Cautru, F., Cohen, D., & Plaza, M. (2011). Neutral versus emotional human stimuli processing in children with pervasive developmental disorders not otherwise specified. *Research in Autism Spectrum Disorders*, *5*(2), 775–783. https://doi.org/10.1016/j.rasd.2010.09.005

Vasilevska Petrovska, I., & Trajkovski, V. (2019). Effects of a Computer-Based Intervention on Emotion Understanding in Children with Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*, *49*(10), 4244–4255. https://doi.org/10.1007/s10803-019-04135-5

Velikonja, T., Fett, A. K., & Velthorst, E. (2019). Patterns of Nonsocial and Social Cognitive Functioning in Adults with Autism Spectrum Disorder: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *76*(2), 135–151. https://doi.org/10.1001/jamapsychiatry.2018.3645

Vermeulen, N., Luminet, O., & Corneille, O. (2006). Alexithymia and the automatic processing of affective information: Evidence from the affective priming paradigm. *Cognition and Emotion*, *20*(1), 64–91. https://doi.org/10.1080/02699930500304654

Vesterinen, H. M., Sena, E. S., Egan, K. J., Hirst, T. C., Churolov, L., Currie, G. L., Antonic, A., Howells, D. W., & Macleod, M. R. (2014). Meta-analysis of data from animal studies: A practical guide. *Journal of Neuroscience Methods*, *221*, 92–102. https://doi.org/10.1016/j.jneumeth.2013.09.010

Vidas, D., Dingle, G. A., & Nelson, N. L. (2018). Children's recognition of emotion in music and speech. *Music & Science*, *1*, 1–10. https://doi.org/10.1177/2059204318762650

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Vieillard, S., Peretz, I., Gosselin, N., Khalfa, S., Gagnon, L., & Bouchard, B. (2008). Happy, sad, scary and peaceful musical excerpts for research on emotions. *Cognition and Emotion*, *22*(4), 720–752. https://doi.org/10.1080/02699930701503567

Vogan, V. M., Leung, R. C., Safar, K., Martinussen, R., Smith, M. Lou, & Taylor, M. J. (2018). Longitudinal examination of everyday executive functioning in children with ASD: Relations with social, emotional, and behavioral functioning over time. *Frontiers in Psychology*, *9*, 1–11. https://doi.org/10.3389/fpsyg.2018.01774

Vorst, H. C. M., & Bermond, B. (2001). Validity and reliability of the Bermond-Vorst Alexithymia Questionnaire. *Personality and Individual Differences*, *30*(3), 413–434. https://doi.org/10.1016/S0191-8869(00)00033-7

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child Development*, *81*(2), 517–527.

https://doi.org/10.1111/j.1467-8624.2009.01412.x

Waddington, F., Hartman, C., de Bruijn, Y., Lappenschaar, M., Oerlemans, A., Buitelaar, J., Franke, B., & Rommelse, N. (2018). Visual and auditory emotion recognition problems as familial cross-disorder phenomenon in ASD and ADHD. *European Neuropsychopharmacology*, *28*(9), 994–1005. https://doi.org/10.1016/j.euroneuro.2018.06.009

Wagenbreth, C., Wattenberg, L., Heinze, H. J., & Zaehle, T. (2016). Implicit and explicit processing of emotional facial expressions in Parkinson's disease. *Behavioural Brain Research*, *303*, 182–190. https://doi.org/10.1016/j.bbr.2016.01.059

Wagener, G. L., Berning, M., Costa, A. P., Steffgen, G., & Melzer, A. (2020). Effects of Emotional Music on Facial Emotion Recognition in Children with Autism Spectrum Disorder (ASD). *Journal of Autism and Developmental Disorders*, *2011*. https://doi.org/10.1007/s10803-020-04781-0

Wainwright, A., & Bryson, S. E. (2002). The development of exogenous orienting: Mechanisms of control. *Journal of Experimental Child Psychology*, *82*(2), 141–155. https://doi.org/10.1016/S0022-0965(02)00002-4

Walker, N., & Raymaker, D. M. (2021). Toward a Neuroqueer Future: An Interview with Nick Walker. *Autism in Adulthood*, *3*(1), 5–10. https://doi.org/10.1089/aut.2020.29014.njw

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

Wallace, S., Coleman, M., & Bailey, A. (2008). An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition and Emotion*, *22*(7), 1353–1380. https://doi.org/10.1080/02699930701782153

Wang, J. E., & Tsao, F. M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities*, *37*(1), 162–170. https://doi.org/10.1016/j.ridd.2014.11.013

Wang, L., Mottron, L., Peng, D., Berthiaume, C., & Dawson, M. (2007). Local bias and local-to-global interference without global deficit: A robust finding in autism under various conditions of attention, exposure time, and visual angle. *Cognitive Neuropsychology*, *24*(5), 550–574. https://doi.org/10.1080/13546800701417096

Wang, Lihong, Huettel, S., & De Bellis, M. D. (2008). Neural substrates for processing task-irrelevant sad images in adolescents: REPORT. *Developmental Science*, *11*(1), 23–32. https://doi.org/10.1111/j.1467-7687.2007.00661.x

Wang, S., & Adolphs, R. (2017). Reduced specificity in emotion judgment in people with autism spectrum disorder. *Neuropsychologia*, *99*, 286–295. https://doi.org/10.1016/j.neuropsychologia.2017.03.024

Wardle, S. G., Taubert, J., Teichmann, L., & Baker, C. I. (2020). Rapid and dynamic processing of face pareidolia in the human brain. *Nature Communications*, *11*(1), 1–14. https://doi.org/10.1038/s41467-020-18325-8

Waxer, M., & Morton, J. B. (2011). Children's judgments of emotion from conflicting cues in speech: Why 6-year-olds are so inflexible. *Child Development*, *82*(5), 1648–1660. https://doi.org/10.1111/j.1467-8624.2011.01624.x

Wazana, A., Bresnahan, M., & Kline, J. (2007). The autism epidemic: Fact or artifact? *Journal of the American Academy of Child and Adolescent Psychiatry*, *46*(6), 721–730. https://doi.org/10.1097/chi.0b013e31804a7f3b

Weeks, S. J., & Hobson, R. P. (1987). The salience of Facial Expression for Autistic Children. *Journal of Child Psychology and Psychiatry*, *28*(1), 137–152. https://doi.org/10.1111/j.1469-7610.1987.tb00658.x

Wegrzyn, M., Vogt, M., Kireclioglu, B., Schneider, J., & Kissler, J. (2017). Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS ONE*, *12*(5), 1–15. https://doi.org/10.1371/journal.pone.0177239

Westwood, H., Kerr-Gaffney, J., Stahl, D., & Tchanturia, K. (2017). Alexithymia in eating disorders: Systematic review and meta-analyses of studies using the Toronto Alexithymia Scale. *Journal of Psychosomatic Research*, *99*, 66–81. https://doi.org/10.1016/j.jpsychores.2017.06.007

Westwood, H., Stahl, D., Mandy, W., & Tchanturia, K. (2016). The set-shifting profiles of anorexia nervosa and autism spectrum disorder using the Wisconsin Card Sorting Test: A systematic review and meta-analysis. *Psychological Medicine*, *46*(9), 1809–1827. https://doi.org/10.1017/S0033291716000581

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

Whitehouse, A. J. O., & Bishop, D. V. M. (2008). Do children with autism "switch off" to speech sounds? An investigation using event-related potentials. *Developmental Science*, *11*(4), 516–524. https://doi.org/10.1111/j.1467-7687.2008.00697.x

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, *5*(1), 72–77. https://doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, *23*(2), 291–312. https://doi.org/10.1016/j.cogdev.2008.01.002

Wiese, E., Wykowska, A., Zwickel, J., & Müller, H. J. (2012). I See What You Mean: How Attentional Selection Is Shaped by Ascribing Intentions to Others. *PLoS ONE*, *7*(9), 1–7. https://doi.org/10.1371/journal.pone.0045391

Wieser, M. J., Mühlberger, A., Alpers, G. W., Macht, M., Ellgring, H., & Pauli, P. (2006). Emotion processing in Parkinson's disease: Dissociation between early neuronal processing and explicit ratings. *Clinical Neurophysiology*, *117*(1), 94–102. https://doi.org/10.1016/j.clinph.2005.09.009

Wiethoff, S., Wildgruber, D., Kreifelts, B., Becker, H., Herbert, C., Grodd, W., & Ethofer, T. (2008). Cerebral processing of emotional prosody-influence of acoustic parameters and arousal. *NeuroImage*, *39*(2), 885–893. https://doi.org/10.1016/j.neuroimage.2007.09.028

Wille, C., & Ebersbach, M. (2016). Semantic congruency and the (reversed) Colavita effect in children and adults. *Journal of Experimental Child Psychology*, *141*, 23–33. https://doi.org/10.1016/j.jecp.2015.07.015

Willenbockel, V., Sadr, J., Fiset, D., Horne, G. O., Gosselin, F., & Tanaka, J. W. (2010). Controlling low-level image properties: The SHINE toolbox. *Behavior Research Methods*, *42*(3), 671–684. https://doi.org/10.3758/BRM.42.3.671

Williams, K., Woolfenden, S., Roberts, J., Rodger, S., Bartak, L., & Prior, M. (2014). Autism in context 1: Classification, counting and causes. *Journal of Paediatrics and Child Health*, *50*(5), 335–340. https://doi.org/10.1111/jpc.12451

Williams, L. M., Mathersul, D., Palmer, D. M., Gur, R. C., Gur, R. E., & Gordon, E. (2009). Explicit identification and implicit recognition of facial emotions: I. Age effects in males and females across 10 decades. *Journal of Clinical and Experimental Neuropsychology*, *31*(3), 257–277. https://doi.org/10.1080/13803390802255635

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

World Health Organization. (2018). *ICD-11 for mortality and morbidity statistics (2018).*

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

Xavier, J., Vignaud, V., Ruggiero, R., Bodeau, N., Cohen, D., & Chaby, L. (2015). A multidimensional approach to the study of emotion recognition in autism spectrum disorders. *Frontiers in Psychology*, *6*(1954), 1–9. https://doi.org/10.3389/fpsyg.2015.01954

Xu, M., Xu, G., & Yang, Y. (2016). Neural systems underlying emotional and non-emotional interference processing: An ALE meta-analysis of functional neuroimaging studies. *Frontiers in Behavioral Neuroscience*, *10*(NOV), 1–15. https://doi.org/10.3389/fnbeh.2016.00220

Yeung, M. K., Lee, T. L., & Chan, A. S. (2020). Impaired Recognition of Negative Facial Expressions is Partly Related to Facial Perception Deficits in Adolescents with High-Functioning Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *50*(5), 1596–1606. https://doi.org/10.1007/s10803-019-03915-3

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion*, *24*(1), 3–47. https://doi.org/10.1080/02699930903205698

Young, R. L., & Posselt, M. (2012). Using The Transporters DVD as a learning tool for children with Autism Spectrum Disorders (ASD). *Journal of Autism and Developmental Disorders*, *42*(6), 984–991. https://doi.org/10.1007/s10803-011-1328-4

Zhang, B., Essl, G., & Provost, E. M. (2015). Recognizing emotion from singing and speaking using shared models. *2015 International Conference on Affective Computing and Intelligent Interaction, ACII 2015*, 139–145. https://doi.org/10.1109/ACII.2015.7344563

Zhang, B., Provost, E. M., Swedberg, R., & Essi, G. (2015). Predicting emotion perception across domains: A study of singing and speaking. *Proceedings of the National Conference on Artificial Intelligence*, *2*(1), 1328–1334.

Zhang, B., Provost, E. M., Swedberg, R., Essi, G., & Essl, G. (2015). Predicting emotion perception across domains: A study of singing and speaking. *Proceedings of the National Conference on Artificial Intelligence*, *2*(1), 1328–1334.

Zhang, J., Meng, Y., Wu, C., Xiang, Y. T., & Yuan, Z. (2019). Non-speech and speech pitch perception among Cantonese-speaking children with autism spectrum disorder: An ERP study. *Neuroscience Letters*, *703*, 205–212. https://doi.org/10.1016/j.neulet.2019.03.021

Zhang, M., Xu, S., Chen, Y., Lin, Y., Ding, H., & Zhang, Y. (2021). Recognition of affective prosody in autism spectrum conditions: A systematic review and meta-analysis. *Autism*. https://doi.org/10.1177/1362361321995725

Zhang, W., Liu, F., Zhou, L., Wang, W., Jiang, H., & Jiang, C. (2019). The Effects of Timbre on Neural Responses to Musical Emotion. In *Music Perception* (Vol. 37, Issue 2). https://doi.org/10.1525/mp.2019.37.2.134

Zhi, W., Cheong, L. S., & Jing, T. (2021). Meta-Analysis of Emotion Recognition Intervention Effects and Influencing Factors in Autism Spectrum Disorder. *Education and Training in Autism and Developmental Disabilities*, *56*(4), 2021.

Zhou, L. F., & Meng, M. (2020). Do you see the “face”? Individual differences in face pareidolia. *Journal of Pacific Rim Psychology*, *14*. https://doi.org/10.1017/prp.2019.27

Zhou, L., Liu, F., Jiang, J., & Jiang, C. (2019). Impaired emotional processing of chords in congenital amusia: Electrophysiological and behavioral evidence. *Brain and Cognition*, *135*, 103577. https://doi.org/10.1016/j.bandc.2019.06.001

Appendix A.

Full references of all studies included in the systematic review and meta-analysis presented in Chapter 2.

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2009). Does gaze direction modulate facial expression processing in children with autism spectrum disorder? *Child Development*, *80*(4), 1134–1146. https://doi.org/10.1111/j.1467-8624.2009.01321.x

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2010). The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: An ERP study. *Neuropsychologia*, *48*(10), 2841–2851. https://doi.org/10.1016/j.neuropsychologia.2010.05.026

Baker, K. F., Montgomery, A. A., & Abramson, R. (2010). Brief report: Perception and lateralization of spoken emotion by youths with high-functioning forms of autism. *Journal of Autism and Developmental Disorders*, *40*(1), 123–129. https://doi.org/10.1007/s10803-009-0841-1

Bast, N., Banaschewski, T., Dziobek, I., Brandeis, D., Poustka, L., & Freitag, C. M. (2019). Pupil Dilation Progression Modulates Aberrant Social Cognition in Autism Spectrum Disorder. *Autism Research*, *12*(11), 1680–1692. https://doi.org/10.1002/aur.2178

Berggren, S., Engström, A. C., & Bölte, S. (2016). Facial affect recognition in autism, ADHD and typical development. *Cognitive Neuropsychiatry*, *21*(3), 213–227. https://doi.org/10.1080/13546805.2016.1171205

Boggs, K. M., & Gross, A. M. (2010). Cue Salience in Face Processing by High Functioning Individuals with Autism Spectrum Disorders. *Journal of Developmental and Physical Disabilities*, *22*(6), 595–613. https://doi.org/10.1007/s10882-010-9198-9

Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, *41*(7), 847–857. https://doi.org/10.1111/1469-7610.00672

Brennand, R., Schepman, A., & Rodway, P. (2011). Vocal emotion perception in pseudo-sentences by secondary-school children with Autism Spectrum Disorder. *Research in Autism Spectrum Disorders*, *5*(4), 1567–1573. https://doi.org/10.1016/j.rasd.2011.03.002

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

Campbell, R., Lawrence, K., Mandy, W., Mitra, C., Jeyakuma, L., & Skuse, D. (2006). Meanings in motion and faces: Developmental associations between the processing of intention from geometrical animations and gaze detection accuracy. *Development and Psychopathology*, *18*(1), 99–118. https://doi.org/10.1017/S0954579406060068

Ciaramidaro, A., Bölte, S., Schlitt, S., Hainz, D., Poustka, F., Weber, B., … Walter, H. (2018). Transdiagnostic deviant facial recognition for implicit negative emotion in autism and schizophrenia. *European Neuropsychopharmacology*, *28*(2), 264–275. https://doi.org/10.1016/j.euroneuro.2017.12.005

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

Couture, S. M., Penn, D. L., Losh, M., Adolphs, R., Hurley, R., & Piven, J. (2010). Comparison of social cognitive functioning in schizophrenia and high functioning autism: More convergence than divergence. *Psychological Medicine*, *40*(4), 569–579. https://doi.org/10.1017/S003329170999078X

Davidson, D., Hilvert, E., Misiunaite, I., Kerby, K., & Giordano, M. (2019). Recognition of facial emotions on human and canine faces in children with and without autism spectrum disorders. *Motivation and Emotion*, *43*(1), 191–202. https://doi.org/10.1007/s11031-018-9736-9

Doi, H., Fujisawa, T. X., Kanai, C., Ohta, H., Yokoi, H., Iwanami, A., … Shinohara, K. (2013). Recognition of facial expressions and prosodic cues with graded emotional intensities in adults with asperger syndrome. *Journal of Autism and Developmental Disorders*, *43*(9), 2099–2113. https://doi.org/10.1007/s10803-013-1760-8

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

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*, 1–8. https://doi.org/10.1155/2014/345878

Fink, E., de Rosnay, M., Wierda, M., Koot, H. M., & Begeer, S. (2014). Brief report: Accuracy and response time for the recognition of facial emotions in a large sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(9), 2363–2368. https://doi.org/10.1007/s10803-014-2084-z

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., … Golan, O. (2016). Basic and complex emotion recognition in children with autism: Cross-cultural findings. *Molecular Autism*, *7*(1), 52. https://doi.org/10.1186/s13229-016-0113-9

Greimel, E., Schulte-Rüther, M., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., & Konrad, K. (2014). Impairment in face processing in autism spectrum disorder: A developmental perspective. *Journal of Neural Transmission*, *121*(9), 1171–1181. https://doi.org/10.1007/s00702-014-1206-2

Griffiths, S., Jarrold, C., Penton-Voak, I. S., Woods, A. T., Skinner, A. L., & Munafò, M. R. (2017). Impaired recognition of basic emotions from facial expressions in young people with autism spectrum disorder: Assessing the importance of expression intensity. *Journal of Autism and Developmental Disorders*, 1–11. https://doi.org/10.1007/s10803-017-3091-7

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, and Allied Disciplines*, *41*(3), 369–379. https://doi.org/10.1111/1469-7610.00621

He, Y., Su, Q., Wang, L., He, W., Tan, C., Zhang, H., … Chen, Y. (2019). The characteristics of intelligence profile and eye gaze in facial emotion recognition in mild and moderate preschoolers with autism spectrum disorder. *Frontiers in Psychiatry*, *10*(JUN), 1–15. https://doi.org/10.3389/fpsyt.2019.00402

Heaton, P., Reichenbacher, L., Sauter, D., Allen, R., Scott, S., & Hill, E. (2012). Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum disorder and typical development. *Psychological Medicine*, *42*(11), 2453–2459. https://doi.org/10.1017/S0033291712000621

Heikkinen, J., Jansson-Verkasalo, E., Toivanen, J., Suominen, K., Väyrynen, E., Moilanen, I., & Seppänen, T. (2010). Perception of basic emotions from speech prosody in adolescents with Asperger's syndrome. *Logopedics Phoniatrics Vocology*, *35*(3), 113–120. https://doi.org/10.3109/14015430903311184

Hubbard, D. J., Faso, D. J., Assmann, P. F., & Sasson, N. J. (2017). Production and perception of emotional prosody by adults with autism spectrum disorder. *Autism Research*, *10*(12), 1991–2001. https://doi.org/10.1002/aur.1847

Isomura, T., Ito, H., Ogawa, S., & Masataka, N. (2014). Absence of predispositional attentional sensitivity to angry faces in children with autism spectrum disorders. *Scientific Reports*, *4*, 1–6. https://doi.org/10.1038/srep07525

Isomura, T., Ogawa, S., Yamada, S., Shibasaki, M., & Masataka, N. (2014). Preliminary evidence that different mechanisms underlie the anger superiority effect in children with and without Autism Spectrum Disorders. *Frontiers in Psychology*, *5*, 1–8. https://doi.org/10.3389/fpsyg.2014.00461

Järvinen, A., Ng, R., Crivelli, D., Neumann, D., Arnold, A. J., Woo-Vonhoogenstyn, N., … Bellugi, U. (2016). Social functioning and autonomic nervous system sensitivity across vocal and musical emotion in Williams syndrome and autism spectrum disorder. *Developmental Psychobiology*, *58*(1), 17–26. https://doi.org/10.1002/dev.21335

Ketelaars, M. P., In'T Velt, A., Mol, A., Swaab, H., & Van Rijn, S. (2016). Emotion recognition and alexithymia in high functioning females with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *21*, 51–60. https://doi.org/10.1016/j.rasd.2015.09.006

Kim, K., Rosenthal, M. Z., Gwaltney, M., Jarrold, W., Hatt, N., McIntyre, N., … Mundy, P. (2015). A virtual joy-stick study of emotional responses and social motivation in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(12), 3891–3899. https://doi.org/10.1007/s10803-014-2036-7

Kliemann, D., Dziobek, I., Hatri, A., Baudewig, J., & Heekeren, H. R. (2012). The role of the amygdala in atypical gaze on emotional faces in autism spectrum disorders. *Journal of Neuroscience*, *32*(28), 9469–9476. https://doi.org/10.1523/jneurosci.5294-11.2012

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

Król, M. E., & Król, M. (2019). A novel machine learning analysis of eye-tracking data reveals suboptimal visual information extraction from facial stimuli in individuals with autism. *Neuropsychologia*, *129*(April), 397–406. https://doi.org/10.1016/j.neuropsychologia.2019.04.022

Kujala, T., Lepistö, T., Nieminen-Von Wendt, T., Näätänen, P., & Näätänen, R. (2005). Neurophysiological evidence for cortical discrimination impairment of prosody in Asperger syndrome. *Neuroscience Letters*, *383*(3), 260–265. https://doi.org/10.1016/j.neulet.2005.04.048

Law Smith, M. J., Montagne, B., Perrett, D. I., 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.neuropsychologia.2010.03.008

Li, Y., Fang, H., Zheng, W., Qian, L., Xiao, Y., Wu, Q., … Ke, X. (2017). A fiber tractography study of social-emotional related fiber tracts in children and adolescents with autism spectrum disorder. *Neuroscience Bulletin*, *33*(6), 722–730. https://doi.org/10.1007/s12264-017-0155-9

Lindström, R., Lepistö-Paisley, T., Makkonen, T., Reinvall, O., Nieminen-von Wendt, T., Alén, R., & Kujala, T. (2018). Atypical perceptual and neural processing of emotional prosodic changes in children with autism spectrum disorders. *Clinical Neurophysiology*, *129*(11), 2411–2420. https://doi.org/10.1016/j.clinph.2018.08.018

Loth, E., Garrido, L., Ahmad, J., Watson, E., Duff, A., & Duchaine, B. (2018). Facial expression recognition as a candidate marker for autism spectrum disorder: How frequent and severe are deficits? *Molecular Autism*, *9*(1), 7. https://doi.org/10.1186/s13229-018-0187-7

O'Connor, K. (2007). Brief report: Impaired identification of discrepancies between expressive faces and voices in adults with Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *37*(10), 2008–2013. https://doi.org/10.1007/s10803-006-0345-1

Oerlemans, A. M., Van Der Meer, J. M. J., Van Steijn, D. J., De Ruiter, S. W., De Bruijn, Y. G. E., De Sonneville, L. M. J., … Rommelse, N. N. J. (2014). Recognition of facial emotion and affective prosody in children with ASD (+ADHD) and their unaffected siblings. *European Child and Adolescent Psychiatry*, *23*(5), 257–271. https://doi.org/10.1007/s00787-013-0446-2

Otsuka, S., Uono, S., Yoshimura, S., Zhao, S., & Toichi, M. (2017). Emotion perception mediates the predictive relationship between verbal ability and functional outcome in high-functioning adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *47*(4), 1166–1182. https://doi.org/10.1007/s10803-017-3036-1

Pelphrey, K., Goldman, B. D., Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., … Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders*, *32*(4), 249–261. https://doi.org/10.1023/A

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

Quintin, E.-M., Bhatara, A., Poissant, H., Fombonne, E., & Levitin, D. J. (2011). Emotion perception in music in high-functioning adolescents with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *41*(9), 1240–1255. https://doi.org/10.1007/s10803-010-1146-0

Rhodes, G., Burton, N., Jeffery, L., Read, A., Taylor, L., & Ewing, L. (2018). Facial expression coding in children and adolescents with autism : Reduced adaptability but intact norm-based coding. *British Journal of Psychology*, *109*(2), 204–218. https://doi.org/10.1111/bjop.12257

Rigby, S. N., Stoesz, B. M., & Jakobson, L. S. (2018). Empathy and face processing in adults with and without autism spectrum disorder. *Autism Research*, *11*(6), 942–955. https://doi.org/10.1002/aur.1948

Sasson, N. J., Pinkham, A. E., Weittenhiller, L. P., Faso, D. J., & Simpson, C. (2016). Context effects on facial affect recognition in schizophrenia and autism: Behavioral and eye-tracking evidence. *Schizophrenia Bulletin*, *42*(3), 675–683. https://doi.org/10.1093/schbul/sbv176

Sasson, N. J., Shasteen, J. R., & Pinkham, A. E. (2016). Brief report: Reduced prioritization of facial threat in adults with autism. *Journal of Autism and Developmental Disorders*, *46*(4), 1471–1476. https://doi.org/10.1007/s10803-015-2664-6

Sawyer, A. C. P., Williamson, P., & Young, R. L. (2012). Can gaze avoidance explain why individuals with Asperger's syndrome can't recognise emotions from facial expressions? *Journal of Autism and Developmental Disorders*, *42*(4), 606–618. https://doi.org/10.1007/s10803-011-1283-0

Schaller, U. M., & Rauh, R. (2017). What difference does it make? Implicit, explicit and complex social cognition in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *47*(4), 961–979. https://doi.org/10.1007/s10803-016-3008-x

Schelinski, S., & von Kriegstein, K. (2019). The relation between vocal pitch and vocal emotion recognition abilities in people with autism spectrum disorder and typical development. *Journal of Autism and Developmental Disorders*, *49*(1), 68–82. https://doi.org/10.1007/s10803-018-3681-z

Shafritz, K. M., Bregman, J. D., Ikuta, T., & Szeszko, P. R. (2015). Neural systems mediating decision-making and response inhibition for social and nonsocial stimuli in autism. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *60*, 112–120. https://doi.org/10.1016/j.pnpbp.2015.03.001

Shanok, N. A., Jones, N. A., & Lucas, N. N. (2019). The nature of facial emotion recognition impairments in children on the autism spectrum. *Child Psychiatry and Human Development*, *50*(4), 661–667. https://doi.org/10.1007/s10578-019-00870-z

Sinzig, J., Morsch, D., & Lehmkuhl, G. (2008). Do hyperactivity, impulsivity and inattention have an impact on the ability of facial affect recognition in children with autism and ADHD? *European Child and Adolescent Psychiatry*, *17*(2), 63–72. https://doi.org/10.1007/s00787-007-0637-9

Stephenson, K. G., Luke, S. G., & South, M. (2019). Separate contributions of autistic traits and anxious apprehension, but not alexithymia, to emotion processing in faces. *Autism*, *23*(7), 1830–1842. https://doi.org/10.1177/1362361319830090

Stewart, M. E., Mcadam, C., Ota, M., Peppé, S., & Cleland, J. (2013). Emotional recognition in autism spectrum conditions from voices and faces. *Autism*, *17*(1), 6–14. https://doi.org/10.1177/1362361311424572

Tanaka, J. W., Wolf, J. M., Klaiman, C., Koenig, K., Cockburn, J., Herlihy, L., … Schultz, R. T. (2012). The perception and identification of facial emotions in individuals with autism spectrum disorders using the Let's Face It! Emotion Skills Battery. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *53*(12), 1259–1267. https://doi.org/10.1111/j.1469-7610.2012.02571.x

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. O. (2015). Evidence for shared deficits in identifying emotions from faces and from voices in autism spectrum disorders and specific language impairment. *International Journal of Language and Communication Disorders*, *50*(4), 452–466. https://doi.org/10.1111/1460-6984.12146

Tell, D., Davidson, D., & Camras, L. A. (2014). Recognition of emotion from facial expressions with direct or averted eye gaze and varying expression intensities in children with autism disorder and typically developing children. *Autism Research and Treatment*, *2014*, 1–11. https://doi.org/10.1155/2014/816137

Tottenham, N., Hertzig, M. E., Gillespie-Lynch, K., Gilhooly, T., Millner, A. J., & Casey, B. J. (2014). Elevated amygdala response to faces and gaze aversion in autism spectrum disorder. *Social Cognitive and Affective Neuroscience*, *9*(1), 106–117. https://doi.org/10.1093/scan/nst050

Uono, S., Sato, W., & Toichi, M. (2013). Common and unique impairments in facial-expression recognition in pervasive developmental disorder-not otherwise specified and Asperger's disorder. *Research in Autism Spectrum Disorders*, *7*(2), 361–368. https://doi.org/10.1016/j.rasd.2012.10.007

Uono, S., Sato, W., & Toichi, M. (2011). The specific impairment of fearful expression recognition and its atypical development in pervasive developmental disorder. *Social Neuroscience*, *6*(5–6), 452–463. https://doi.org/10.1080/17470919.2011.605593

Vannetzel, L., Chaby, L., Cautru, F., Cohen, D., & Plaza, M. (2011). Neutral versus emotional human stimuli processing in children with pervasive developmental disorders not otherwise specified. *Research in Autism Spectrum Disorders*, *5*(2), 775–783. https://doi.org/10.1016/j.rasd.2010.09.005

Waddington, F., Hartman, C., de Bruijn, Y., Lappenschaar, M., Oerlemans, A., Buitelaar, J., … Rommelse, N. (2018). Visual and auditory emotion recognition problems as familial cross-disorder phenomenon in ASD and ADHD. *European Neuropsychopharmacology*, *28*(9), 994–1005. https://doi.org/10.1016/j.euroneuro.2018.06.009

Wallace, S., Coleman, M., & Bailey, A. (2008). An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition and Emotion*, *22*(7), 1353–1380. https://doi.org/10.1080/02699930701782153

Wang, J. E., & Tsao, F. M. (2015). Emotional prosody perception and its association with pragmatic language in school-aged children with high-function autism. *Research in Developmental Disabilities*, *37*(1), 162–170. https://doi.org/10.1016/j.ridd.2014.11.013

Wang, S., & Adolphs, R. (2017). Reduced specificity in emotion judgment in people with autism spectrum disorder. *Neuropsychologia*, *99*, 286–295. https://doi.org/10.1016/j.neuropsychologia.2017.03.024

Wingenbach, T. S. H., Ashwin, C., & Brosnan, M. (2017). Diminished sensitivity and specificity at recognising facial emotional expressions of varying intensity underlie emotion-specific recognition deficits in autism spectrum disorders. *Research in Autism Spectrum Disorders*, *34*, 52–61. https://doi.org/10.1016/j.rasd.2016.11.003

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 Hum Dev*, *43*(5), 775–794. https://doi.org/10.1007/s10578-012-0296-z

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

Yeung, M. K., Han, Y. M. Y., Sze, S. L., & Chan, A. S. (2014). Altered right frontal cortical connectivity during facial emotion recognition in children with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *8*(11), 1567–1577. https://doi.org/10.1016/j.rasd.2014.08.013

Appendix B.

The Critical Appraisal Skills Programme (CASP) case control study checklist used for quality assessment in the systematic review and meta-analysis presented in Chapter 2.

CASP
Critical Appraisal
Skills Programme

www.casp-uk.net
info@casp-uk.net
Summertown Pavilion, Middle
Way Oxford OX2 7LG

CASP Checklist: 11 questions to help you make sense of a Case Control Study

How to use this appraisal tool: Three broad issues need to be considered when appraising a case control study:

- Are the results of the study valid? (Section A)
- What are the results? (Section B)
- Will the results help locally? (Section C)

The 11 questions on the following pages are designed to help you think about these issues systematically. The first three questions are screening questions and can be answered quickly. If the answer to both is "yes", it is worth proceeding with the remaining questions. There is some degree of overlap between the questions, you are asked to record a "yes", "no" or "can't tell" to most of the questions. A number of italicised prompts are given after each question. These are designed to remind you why the question is important. Record your reasons for your answers in the spaces provided.

About: These checklists were designed to be used as educational pedagogic tools, as part of a workshop setting, therefore we do not suggest a scoring system. The core CASP checklists (randomised controlled trial & systematic review) were based on JAMA 'Users' guides to the medical literature 1994 (adapted from Guyatt GH, Sackett DL, and Cook DJ), and piloted with health care practitioners.

For each new checklist, a group of experts were assembled to develop and pilot the checklist and the workshop format with which it would be used. Over the years overall adjustments have been made to the format, but a recent survey of checklist users reiterated that the basic format continues to be useful and appropriate.

Referencing: we recommend using the Harvard style citation, i.e.: *Critical Appraisal Skills Programme (2018). CASP (insert name of checklist i.e. Case Control Study) Checklist. [online] Available at: URL. Accessed: Date Accessed.*

Paper for appraisal and reference:..

Section A: Are the results of the trial valid?

1. Did the study address a clearly focused issue?

Yes ☐
Can't Tell ☐
No ☐

HINT: An issue can be 'focused' In terms of
- the population studied
- Whether the study tried to detect a beneficial or harmful effect
- the risk factors studied

Comments:

2. Did the authors use an appropriate method to answer their question?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider
- Is a case control study an appropriate way of answering the question under the circumstances
- Did it address the study question

Comments:

Is it worth continuing?

3. Were the cases recruited in an acceptable way?

Yes ☐
Can't Tell ☐
No ☐

Comments:

3 if diagnostic instrument (e.g., DSM/ICD) or other available tools (e.g., ADOS/ADI-R/AQ) were reported
2 if diagnosis was simply stated as made by clinicians without specifying what diagnostic instrument was used

HINT: We are looking for selection bias which might compromise validity of the findings

- are the cases defined precisely
- were the cases representative of a defined population (geographically and/or temporally)
- was there an established reliable system for selecting all the cases
- are they incident or prevalent
- is there something special about the cases
- is the time frame of the study relevant to disease/exposure
- was there a sufficient number of cases selected
- was there a power calculation

4. Were the controls selected in an acceptable way?

Yes ☐
Can't Tell ☐
No ☐

Comments:

3 if groups were matched on age, gender, and IQ
2 if groups were matched on one of these criteria
1 if groups were not matched at all

HINT: We are looking for selection bias which might compromise the generalisability of the findings

- were the controls representative of the defined population (geographically and/or temporally)
- was there something special about the controls
- was the non-response high, could non-respondents be different in any way
- are they matched, population based or randomly selected
- was there a sufficient number of controls selected

5. Was the exposure accurately measured to minimise bias?

Yes ☐
Can't Tell ☐
No ☐

Comments:

HINT: We are looking for measurement, recall or classification bias

- was the exposure clearly defined and accurately measured
- did the authors use subjective or objective measurements
- do the measures truly reflect what they are supposed to measure (have they been validated)
- were the measurement methods similar in the cases and controls
- did the study incorporate blinding where feasible
- is the temporal relation correct (does the exposure of interest precede the outcome)

6. (a) Aside from the experimental intervention, were the groups treated equally?

HINT: List the ones you think might be important, that the author may have missed

- genetic
- environmental
- socio-economic

List:

6. (b) Have the authors taken account of the potential confounding factors in the design and/or in their analysis?

Yes ☐
Can't Tell ☐
No ☐

HINT: Look for

- restriction in design, and techniques e.g. modelling, stratified-, regression-, or sensitivity analysis to correct, control or adjust for confounding factors

Comments:

Section B: What are the results?

7. How large was the treatment effect?

HINT: Consider
- what are the bottom line results
- is the analysis appropriate to the design
- how strong is the association between exposure and outcome (look at the odds ratio)
- are the results adjusted for confounding, and might confounding still explain the association
- has adjustment made a big difference to the OR

Comments:

Pay attention to whether effect sizes were reported:
3 if effect sizes were reported
2 if effect sizes were not reported
Do not score based on the size of effect

8. How precise was the estimate of the treatment effect?

HINT: Consider
- size of the p-value
- size of the confidence intervals
- have the authors considered all the important variables
- how was the effect of subjects refusing to participate evaluated

Comments:

Pay attention to how p-values were reported:
3 if exact p-values (and $p < .001$) were reported
2 if approximate p-values were reported (e.g., $p < .05$ or $p < .01$)
1 if no p-values were reported

9. Do you believe the results?

Yes ☐
No ☐

HINT: Consider
- big effect is hard to ignore!
- Can it be due to chance, bias, or confounding
- are the design and methods of this study sufficiently flawed to make the results unreliable
- consider Bradford Hills criteria (e.g. time sequence, does-response gradient, strength, biological plausibility)

Comments:

Section C: Will the results help locally?

10. Can the results be applied to the local population?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider whether
- the subjects covered in the study could be sufficiently different from your population to cause concern
- your local setting is likely to differ much from that of the study
- can you quantify the local benefits and harms

Comments:
3 if paper included both genders and age groups
2 if paper included one gender but both age groups or one age group but both genders
1 if paper included only one gender and one age group

11. Do the results of this study fit with other available evidence?

Yes ☐
Can't Tell ☐
No ☐

HINT: Consider
- all the available evidence from RCT's Systematic Reviews, Cohort Studies, and Case Control Studies as well, for consistency

Comments:

Score down to 2 if paper found no significant group differences but did not back this up in the discussion

Remember One observational study rarely provides sufficiently robust evidence to recommend changes to clinical practice or within health policy decision making. However, for certain questions observational studies provide the only evidence. Recommendations from observational studies are always stronger when supported by other evidence.