



Ainsworth, Kirsty (2016) *Facial expression recognition and the autism spectrum*. PhD thesis.

<http://theses.gla.ac.uk/8287/>

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten:Theses  
<http://theses.gla.ac.uk/>  
theses@gla.ac.uk

# **Facial Expression Recognition and the Autism Spectrum**

**Kirsty Ainsworth**  
M.Sc., B.A. (Hons.)

Submitted in fulfilment of the requirements for the degree of

**Doctor of Philosophy**

School of Psychology  
College of Science and Engineering  
University of Glasgow

December 2016

## Abstract

An atypical recognition of facial expressions of emotion is thought to be part of the characteristics associated with an autism spectrum disorder diagnosis (*DSM-5*, 2013). However, despite over three decades of experimental research into facial expression recognition (FER) in autism spectrum disorder (ASD), conflicting results are still reported (Harms, Martin, & Wallace, 2010). The thesis presented here aims to explore FER in ASD using novel techniques, as well as assessing the contribution of a co-occurring emotion-blindness condition (alexithymia) and autism-like personality traits.

Chapter 1 provides a review of the current literature surrounding emotion perception in ASD, focussing specifically on evidence for, and against, atypical recognition of facial expressions of emotion in ASD.

The experimental chapters presented in this thesis (Chapters 2, 3 and 4) explore FER in adults with ASD, children with ASD and in the wider, typical population. In Chapter 2, a novel psychophysics method is presented along with its use in assessing FER in individuals with ASD. Chapter 2 also presents a research experiment in adults with ASD, indicating that FER is similar compared to typically developed (TD) adults in terms of the facial muscle components (action units; AUs), the intensity levels and the timing components utilised from the stimuli. In addition to this, individual differences within groups are shown, indicating that better FER ability is associated with lower levels of ASD symptoms in adults with ASD (measured using the ADOS; Lord et al. (2000)) and lower levels of autism-like personality traits in TD adults (measured using the Autism-Spectrum Quotient; (S. Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001)). Similarly, Chapter 3 indicates that children with ASD are not significantly different from TD children in their perception of facial expressions of emotion as assessed using AU, intensity and timing components. Chapter 4 assesses the contribution of alexithymia and autism-like personality traits (AQ) to FER ability in a sample of individuals from the typical population. This chapter provides evidence against the idea that alexithymia levels predict FER ability over and above AQ levels.

The importance of the aforementioned results are discussed in Chapter 5 in the context of previous research in the field, and in relation to established theoretical approaches to FER in ASD. In particular, arguments are made that FER cannot be conceptualised under an

‘all-or-nothing’ framework, which has been implied for a number of years (Harms et al., 2010). Instead it is proposed that FER is a multifaceted skill in individuals with ASD, which varies according to an individual’s skillset. Lastly, limitations of the research presented in this thesis are discussed in addition to suggestions for future research.

# Table of Contents

<b>Abstract</b> .....	<b>i</b>
<b>List of Figures</b> .....	<b>v</b>
<b>List of Tables</b> .....	<b>x</b>
<b>Acknowledgements</b> .....	<b>xi</b>
<b>Author’s Declaration</b> .....	<b>xii</b>
<b>List of Abbreviations</b> .....	<b>xiii</b>
<b>Chapter 1 General Introduction</b> .....	<b>1</b>
<b>1.1 What is Autism Spectrum Disorder (ASD)?</b> .....	<b>1</b>
<b>1.2 Issues associated with ASD research</b> .....	<b>4</b>
<b>1.3 Typical perception of facial expressions of emotion</b> .....	<b>5</b>
<b>1.4 The ‘enigma’ of facial expression recognition (FER) in ASD: assessing conflicting literatures</b> .....	<b>6</b>
1.4.1 Early research.....	7
1.4.2 Components of FER processing.....	10
1.4.3 Contribution of participant sample characteristics.....	14
1.4.4 Experimental methods used to assess FER in individuals with ASD.....	18
1.4.5 Summary of FER research .....	20
<b>1.5 Role of the ‘Autism Spectrum Quotient’ in FER ASD research</b> .....	<b>20</b>
<b>1.6 Theoretical context</b> .....	<b>21</b>
<b>1.7 Aims of the thesis</b> .....	<b>23</b>
<b>1.8 Overview of chapters</b> .....	<b>23</b>
<b>Chapter 2 Facial Expression Recognition in Adults with ASD</b> .....	<b>25</b>
<b>2.1 Introduction</b> .....	<b>25</b>
<b>2.2 Method</b> .....	<b>27</b>
2.2.1 Instruments for characterising participants.....	27
2.2.2 Participants .....	28
2.2.3 Stimulus generation .....	32
2.2.4 Procedure.....	34
2.2.5 Analysis .....	37
<b>2.3 Results</b> .....	<b>40</b>
2.3.1 Response patterns .....	40
2.3.2 Reverse correlation: modelling individual differences.....	49
<b>2.4 Discussion</b> .....	<b>71</b>
2.4.1 Limitations.....	73
2.4.2 Conclusion.....	75
<b>Chapter 3 Facial Expression Recognition in Children with ASD</b> .....	<b>76</b>
<b>3.1 Introduction</b> .....	<b>76</b>
<b>3.2 Methods</b> .....	<b>78</b>
3.2.1 Participants .....	78
3.2.2 Experimental set-up .....	80
3.2.3 Stimuli .....	82
3.2.4 Procedure.....	84
3.2.5 Analysis .....	85
<b>3.3 Results</b> .....	<b>87</b>
3.3.1 Response characteristics.....	87

3.3.2	ASD and TD decompositions derived from NMF.....	93
<b>3.4</b>	<b>Discussion.....</b>	<b>107</b>
3.4.1	Limitations.....	109
3.4.2	Conclusion.....	110
<b>Chapter 4</b>	<b>FER in Relation to Alexithymia and Autism-like Traits .....</b>	<b>111</b>
<b>4.1</b>	<b>Introduction.....</b>	<b>111</b>
<b>4.2</b>	<b>Methods .....</b>	<b>114</b>
4.2.1	Participants .....	114
4.2.2	Psychometric measures .....	115
4.2.3	Design.....	117
4.2.4	Stimuli .....	118
4.2.5	Procedure.....	121
4.2.6	Analysis .....	125
<b>4.3</b>	<b>Results .....</b>	<b>126</b>
4.3.1	Descriptive information .....	126
4.3.2	Pearson’s correlation .....	128
4.3.3	Multiple regression analysis.....	133
4.3.4	Slope: highest and lowest 10%.....	136
<b>4.4</b>	<b>Point of Subjective Equivalence (PSE).....</b>	<b>137</b>
<b>4.5</b>	<b>Discussion .....</b>	<b>141</b>
<b>Chapter 5</b>	<b>General Discussion.....</b>	<b>145</b>
<b>5.1</b>	<b>Overview of general discussion .....</b>	<b>145</b>
5.1.1	Summary of findings .....	145
5.1.2	Condensed summary of findings.....	149
<b>5.2</b>	<b>Contributions to the field .....</b>	<b>149</b>
<b>5.3</b>	<b>Theoretical perspectives .....</b>	<b>150</b>
<b>5.4</b>	<b>Implications of these findings .....</b>	<b>150</b>
<b>5.5</b>	<b>General limitations .....</b>	<b>151</b>
<b>5.6</b>	<b>Future directions.....</b>	<b>152</b>
<b>5.7</b>	<b>Conclusions.....</b>	<b>153</b>
<b>Appendices</b>	<b>.....</b>	<b>154</b>
<b>References</b>	<b>.....</b>	<b>155</b>

## List of Figures

Figure 2–1. Random facial expression synthesis. Figure taken from Yu, Garrod and Schyns (2012) p153.....	34
Figure 2–2. Online instructions at the beginning of the experiment. Please note these instructions were given in addition to an information sheet and verbal direction from the experimenter. ....	35
Figure 2–3. Snapshot of dynamic stimulus as viewed by the participants, which was displayed in the central viewing field for 1.25 seconds (snapshot taken at the middle of the time course). Participants viewed the stimuli 60cm away from the screen.....	36
Figure 2–4. Snapshot of the categorization options as viewed by the participant. The stimulus (neutral) stayed on the screen until the participant selected their option. ....	37
Figure 2–5. Snapshot of the break instructions as viewed by the participant after periods of 50 stimuli. ....	37
Figure 2–6. Experimental steps from stimuli generation to reverse correlation analysis (Image source: Yu, Garrod and Schyns, 2012, p. 154).....	39
Figure 2–7 Distribution of responses (emotion button presses) to stimuli (2400 trials per participant).....	40
Figure 2–8 Distribution of responses (intensity button presses) to stimuli (2400 trials per participant).....	41
Figure 2–10. TD Group Part 2: Frequency of stimuli responses per TD participant for emotion and intensity (2400 trials per participant).....	44
Figure 2–11 TD Group Part 3: Frequency of stimuli responses per TD participant for emotion and intensity (2400 trials per participant).....	45
Figure 2–12 ASD Group Part 1: Frequency of stimuli responses per ASD participant for emotion and intensity (2400 trials per participant).....	46
Figure 2–13 ASD Group Part 2: Frequency of stimuli responses per ASD participant for emotion and intensity (2400 trials per participant).....	47
Figure 2–14 ASD Group Part 3: Frequency of stimuli responses per ASD participant for emotion and intensity (2400 trials per participant).....	48
Figure 2–15 Action Unit numbers and their respective names.....	50
Figure 2–16 Six temporal parameters associated with each action unit.....	51
Figure 2–17 From left to right: 1) Single participant data (ASD participant 11 (ID 2984)) of AUs that emerge from the data (lighter blue indicates higher correlation coefficient (rho values), and hence lower p values) for all emotions (happy, surprise, fear, disgust, anger sadness) 2) This is the same as the image 1 but only the rho values that are significant are displayed (i.e. $p \leq 0.05$ ) in red and 3) Action Units indicated as being ‘core’ action units from Ekman and Friesen 1978.....	52

Figure 2–18 Example: ASD participant 11 (ID 2884)’s model at the midpoint of the dynamic videos at 3 intersections of intensity.....	53
Figure 2–19 Dissimilarity matrix of hamming distance (i.e. percentage of binary values that differ) between ASD (n = 11) and TD (n = 11) for all 6 emotions. Cool colours (e.g. dark blue) indicate low hamming distance values and hot colours (e.g. dark red) indicate high hamming distance values. ....	55
Figure 2–20 Reference data (n = 15): comparison of models within group.....	56
Figure 2–21 Hamming distance between AU composition of models for the ASD group compared to reference dataset .....	57
Figure 2–22 Hamming distance between AU composition of models for the TD group compared to reference dataset .....	57
Figure 2–23. ASD vs. template hamming distance per individual means and variance .....	59
Figure 2–24. TD vs. template hamming distance per individual means and variance .....	59
Figure 2–25. ASD vs. template hamming per emotion distance means and variance.....	60
Figure 2–26 TD vs. template hamming distance per emotion means and variance .....	60
Figure 2–27 Within-group comparison of AU composition of models using Hamming distance .....	61
Figure 2–28 Within-group comparison of AU composition of models using Hamming distance .....	61
Figure 2–29 Multiple distribution plot denoting mean (red cross), median (green box) and distribution of intensity values of TD models .....	63
Figure 2–30 Multiple distribution plot denoting mean (red cross), median (green box) and distribution of intensity values of ASD models.....	63
Figure 2–31 Intensity values of TD models for eye AUs only.....	65
Figure 2–32 Intensity values of ASD models for eye AUs only .....	65
Figure 2–33 Intensity values of TD models for mouth AUs only .....	66
Figure 2–34 Intensity values of ASD models for mouth AUs only .....	67
Figure 2–35 Peak temporal values for TD models .....	68
Figure 2–36 Peak temporal values for ASD models .....	69
Figure 2–37 Peak temporal values for TD models: eye region only .....	69
Figure 2–38 Peak temporal values for ASD models: eye region only .....	70
Figure 2–39 Peak temporal values for TD models: mouth region only .....	70
Figure 2–40 Peak temporal values for ASD models: mouth region only.....	71

Figure 3–1 Stimulus Generation using the Generative Face Grammar (Yu et al. 2012). For details see text (section 3.2.3).....	83
Figure 3–2 Example of two trials within the experimental task. Random facial expressions lasted 1.25 seconds (displayed here as a snapshot at 0.75 seconds). The facial expression trials were interleaved with cartoon and motivational messages. ....	85
Figure 3–3 Reaction time trimmed (<3000ms) for all trials.....	88
Figure 3–4 Response distribution of children with ASD in the happy condition.....	89
Figure 3–5 Response distribution of children with ASD in the angry condition. ....	90
Figure 3–6 Response distribution of TD group in the happy condition. ....	90
Figure 3–7 Response distribution of TD group in the angry condition. ....	90
Figure 3–8 Percentage of ‘yes’ and ‘no’ responses from ASD participants in the happy condition i.e. ‘Is this person happy?’ .....	91
Figure 3–9 Percentage of ‘yes’ and ‘no’ responses from ASD participants in the angry condition i.e. ‘Is this person angry?’ .....	92
Figure 3–10 Percentage of ‘yes’ and ‘no’ responses from TD participants in the happy condition i.e. ‘Is this person happy?’ .....	92
Figure 3–11 Percentage of ‘yes’ and ‘no’ responses from TD participants in the angry condition i.e. ‘Is this person angry?’ .....	93
Figure 3–12 Hamming distance between ASD and TD decompositions. ....	94
Figure 3–13 Amplitude values (0:1) of ASD and TD Decompositions for the happy condition. ....	97
Figure 3–14 Amplitude values (0:1) of ASD and TD Decompositions for the angry condition. ....	97
Figure 3–15 Mean amplitude (intensity) values for all decompositions for each group/condition.....	98
Figure 3–16 Sum of the ‘on’ amplitude (intensity) values (i.e. values > 0.2) for all decompositions .....	99
Figure 3–17 ASD happy Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 22 (AU 13: ‘Cheek Puffer’) and AU number 33 (AU23: ‘Lip Tightener’). ....	100
Figure 3–18 ASD happy Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 20 (AU12L: ‘Lip Corner Puller Left’).....	101
Figure 3–19 ASD happy Decomposition 3: Activation of AUs across time and amplitude (intensity) Highest amplitude AU is AU number 18 (AU 12 Lip Corner Puller) .....	101

Figure 3–20 TD happy Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 20 (AU12L: ‘Lip Corner Puller Left’) and AU number 39 (AU 39 ‘Nostril Dialator’).	102
Figure 3–21 TD happy Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 1 (AU1: ‘Inner Brow Raiser) and AU number 30 (AU20L ‘Lip Stretch Left’).	102
Figure 3–22 TD happy Decomposition 3: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 25 (AU 14R ‘Dimpler Right’).	103
Figure 3–23 ASD angry Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 32 (AU 22 ‘Lip Funneler’).	103
Figure 3–24 ASD angry Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 37 (AU 27 ‘Mouth Stretch’).	104
Figure 3–25 ASD angry Decomposition 3: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 12 (AU 9 Nose Wrinkler) and AU number 36 (AU 26 ‘Jaw Drop’).	104
Figure 3–26 TD angry Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 12 (AU 9 ‘Nose Wrinkler’).	105
Figure 3–27 TD angry Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 39 (AU 39 ‘Nostril Compressor’).	105
Figure 3–28 TD angry Decomposition 3: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 16 (AU 11L ‘Nasolabial Deepener Left’).	106
Figure 3–29 Timepoint (seconds) of AUs when taken together (mean average)	107
Figure 4–1 Endpoint images used in stimuli generation for the happy/sad condition	118
Figure 4–2 Endpoint images used in stimuli generation for the disgust/anger condition	119
Figure 4–3 Endpoint images used in stimulus generation for the surprise/fear condition	119
Figure 4–4 Stimulus Generation Example for disgust/angry Condition	120
Figure 4–5 Example of one participant’s data for the disgust/anger condition (blue dots) and the Cumulative Normal Function fitted to the data (red line)	125
Figure 4–6 Pearson's correlation between slope values and AQ score (happy/sad condition)	129
Figure 4–7 Pearson's correlation between slope values and AQ score (disgust/anger condition)	129
Figure 4–8 Pearson's correlation between slope values and AQ score (surprise/fear condition)	130

Figure 4–9 Pearson's correlation between slope values and TAS score (happy/sad condition).....	131
Figure 4–10 Pearson's correlation between slope values and TAS score (disgust/anger condition).....	131
Figure 4–11 Pearson's correlation between slope values and TAS score (surprise/fear condition).....	132
Figure 4–12 Pearson's correlation between autism-like traits (AQ score) and alexithymia (TAS score).....	132
Figure 4–13 Pearson's correlation between PSE values and AQ score (disgust/anger condition).....	138
Figure 4–14 Pearson's correlation between PSE values and AQ score (surprise/fear condition).....	139
Figure 4–15 Pearson's correlation between PSE values and AQ score (happy/sad condition).....	139
Figure 4–16 Pearson's correlation between PSE values and TAS score (happy/sad condition).....	140
Figure 4–17 Pearson's correlation between PSE values and TAS score (disgust/anger condition).....	140
Figure 4–18 Pearson's correlation between PSE values and TAS score (surprise/fear condition).....	141

## List of Tables

Table 2-1 Demographics of ASD and TD participants (n = 22) *p < 0.01 .....	30
Table 2-2 Autism Diagnostic Observation Schedule (2 <sup>nd</sup> Edition) Score for ASD Participants .....	31
Table 2-3 Autism Spectrum Quotient Score (AQ) and Toronto Alexithymia Scale Score (TAS) .....	32
Table 3-1 Demographic information from children with ASD and TD children .....	80
Table 3-2 Trial Numbers of Aggregate Group Data .....	82
Table 3-3 Percentage of responses for each condition .....	89
Table 3-4 Action Unit Number and Associated AU Name .....	96
Table 4-1 Demographic Information .....	127
Table 4-2 Descriptive information about the psychometric measures .....	128
Table 4-3 Average scores for dependent variables for all participants, separated by location .....	128

## Acknowledgements

I would like to take this opportunity to thank all of the people who have supported me throughout my PhD. Firstly, I would like to thank Dr. David Simmons for his patience, hard work and guidance throughout the PhD as my first supervisor. I would like to thank Prof. Philippe Schyns for his support throughout my PhD as my second supervisor and the excellent opportunities he has provided me. I would also like to thank Prof. James Tanaka for welcoming me into his lab in Canada and for inspiring me with the excellent work he does with children with autism. I thank Prof. Karen Dobkins for her guidance and support, and for making me feel welcome in her lab in California. My thanks also go to Prof. Ralph Adolphs and his team at Caltech for their fantastic help and support. I would like to thank Prof. Alex Clark for his never-ending wisdom about the PhD journey. I would also like to thank Dr. Craig Melville for his support and guidance during the last stages of my PhD.

I would like to send my thanks to all of the people who participated in my research and to the volunteers who helped me with ‘Face Lab’ and ‘Brain Lab’ – your enthusiasm and trust was second to none.

I would like to thank the Economic and Social Research Council for funding my PhD and for providing me with excellent opportunities and support over the past four years. I thank the University of Glasgow Graduate School who have provided me with invaluable opportunities for learning and development. I would also like to thank the Glasgow Science Centre for providing me with fantastic facilities to collect data.

I would like to thank my family for their enduring support and love. In particular I would like to thank my Dad, George and my Brother, Richard for their kindness and faith in me. To my partner, Colin, for his unconditional support and for taking on the unenviable task of keeping me happy throughout the PhD. To my friends, Lauren and Sarah -for being there though good times, and through difficult times- and to Emmy, Claire and Taylor for their amazing generosity and kindness during my time as a PhD student.

Lastly, I would like to thank my late Mum, Jennifer, whose examples of strength and resilience have helped me unendingly, and to whom this PhD is dedicated.

## Author's Declaration

I declare that this thesis, submitted to the University of Glasgow for the degree of Doctor of Philosophy, is the result of my own research, except where otherwise acknowledged, and that this thesis has not been submitted for a higher degree to any other university or institution.

Signed .....

(Kirsty Ainsworth)

Date: .....

## List of Abbreviations

ADOS	Autism Diagnostic Observation Schedule
ADI	Autism Diagnostic Interview
AQ	Autism Spectrum Quotient
ASD	Autism Spectrum Disorder
AU	Action Unit
DSM	Diagnostic and Statistical Manual for Mental Disorders
DV	Dependent Variable
FER	Facial Expression Recognition
IQ	Intelligence Quotient
IRI	Interpersonal Reactivity Index
IV	Independent Variable
MRA	Multiple Regression Analysis
NMF	Non-negative Matrix Factorization
PCA	Principle Components Analysis
PDD-NOS	Pervasive Developmental Disorder – Not Otherwise Specified
PIQ	Performance Intelligence Quotient
RT	Reaction Time
TAS	Toronto Alexithymia Scale
TD	Typically Developing
VIQ	Verbal Intelligence Quotient
VQ	Vector Quantization

# Chapter 1 General Introduction

Autism Spectrum Disorder (ASD) is a common neurodevelopmental disorder that is characterised in part by difficulties with social communication and interaction. It has long been thought that a core part of these social difficulties is the recognition of one's own emotions and those of others. A particular focus of the research in this field has been on the recognition of facial expressions of emotion in other people. However, the scientific literature is equivocal about the nature of how, and the extent to which, facial expression recognition (FER) in individuals with ASD differs from that of typically developed (TD) individuals. Hence, there is a demand for a better understanding of exactly how FER manifests in individuals with ASD.

This introductory chapter will provide a description of ASD and the particular issues that are intrinsic to ASD research. Then, facial expression recognition, and how this skill manifests itself in individuals with ASD, will be discussed with reference to literature in the field. Next, the 'Autism Spectrum Quotient' (AQ) will be described and its role in ASD research will be evaluated, followed by a discussion of the theories of facial expression recognition in ASD. Lastly, an overview of the experimental chapters of this thesis will be outlined.

## 1.1 What is Autism Spectrum Disorder (ASD)?

According to the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM 5), ASD is characterised by a dyad of 'impairments': 1) 'impairments' in social communication and interaction and 2) restricted, repetitive patterns of behaviour, interests, or activities (American Psychiatric Association, 2013). As suggested by its name, ASD is a spectrum disorder, meaning that within the diagnosis of ASD there is large variability in the 'severity' of the symptoms from one individual to the next (Wing, 1988).

Changes in the diagnostic criteria for ASD have, over recent years, changed how we think about the definition of ASD. The DSM moved from its 4<sup>th</sup> edition to its 5<sup>th</sup> in 2013, bringing with it several changes to the diagnostic criteria for ASD. Under DSM-IV, Pervasive Developmental Disorder (PDD) was an umbrella term used to denote the sub-categories: Autistic Disorder, Asperger's Disorder, Childhood Disintegrative Disorder (CDD), Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) and Rett's Disorder. During this time Autistic Disorder, Asperger's Disorder and PDD-NOS

were also collectively referred to using the informal term ‘Autism Spectrum Disorders’. The DSM updated its classification from Pervasive Developmental Disorder to the all-encompassing term Autism Spectrum Disorder in its 5<sup>th</sup> edition. In addition to this the DSM-5 introduced a new classification of ‘Social (Pragmatic) Communication Disorder’ (SCD), which applies to those who do not meet all necessary criteria for a diagnosis of ASD. The removal of Asperger’s Disorder from DSM criteria has, however, been met with mixed reactions from the autism community. One major criticism has been that individuals who would have previously been diagnosed with Asperger’s Disorder may no longer meet requirements for ASD *or* SCD and so miss out on the support services they may have otherwise been entitled to (Volkmar & Reichow, 2013).

Despite contentions that have emerged out of recent changes in the DSM diagnostic criteria of ASD, it is important to note that the DSM is only one of two major internationally recognised diagnostic criteria manuals. The DSM-5 is used predominantly in the USA, whereas the International Statistical Classification of Diseases and Related Health Problems, 10<sup>th</sup> edition (ICD-10) is used in the rest of the world. This becomes problematic as the ICD-10, although quite well harmonised with DSM-IV, uses different diagnostic criteria for ASD compared to the DSM-5. For example, the ICD-10 still recognises the separate conditions Asperger’s Disorder and Childhood Autism, although an update to the ICD diagnostic criteria (ICD-11) is expected to be published in 2018 (see [www.who.int/classifications/icd/revision/en/](http://www.who.int/classifications/icd/revision/en/)).

The DSM and ICD provide detailed classifications that are central to ASD diagnostic procedures, largely administered using standardised behavioural observation techniques. The Autism Diagnostic Observation Schedule (ADOS) and the Autism Diagnostic Interview (ADI) are generally considered ‘gold standard’ for diagnosis when administered by a trained clinician (Falkmer, Anderson, Falkmer, & Horlin, 2013). There are, however, several accepted practices in research for obtaining a confirmation of a participant’s diagnosis of ASD. The procedure that is considered to be the best in ASD research is a ‘re-diagnosis’ of ASD using the ADOS and/or ADI, administered by a qualified individual. This allows the researcher not only to confirm that the participant reaches the required cut-off score for ASD, but also to obtain sub-scores on different components of the ADOS and ADI that might not be accessible from the individual’s medical records. Often researchers work alongside clinicians, in which case ADOS and ADI assessments can be obtained as part of a multidisciplinary team. In cases where there is no multidisciplinary team, members of a research team can be trained on the ADOS or ADI, but they must meet

‘research accreditation’ standard before they are allowed to administer the ADOS or ADI as part of a study that may go on to be published in a scientific journal. In cases where it is not feasible to re-administer the ADOS to ADI for diagnosis verification, the original ADOS or ADI documentation can be referred to in order to verify that the participant has a diagnosis of ASD.

A topic of current debate is the large increase in prevalence of ASD over the past twenty-five years. The term ‘autism’ was only first recognised under DSM diagnostic criteria in 1980 (DSM III), previously termed as ‘schizophrenic reaction, childhood type’ (DSM I, 1952) and ‘schizophrenia, childhood type’ (DSM II, 1968). However, the first study to assess the prevalence of autism-like behaviour in children was conducted in the UK in 1966 (Lotter, 1966), reporting a prevalence rate of 4.5 per 10,000. Subsequent studies reported a prevalence rate of between 4 and 5 per 10,000 (Brask, 1972; Wing, Yeates, Brierley, & Gould, 1976; Wing and Gould 1979). In the 1980s thirteen studies reported prevalence rates which varied from 2.5 to 16 per 10,000 (for full review see Wing (1993)). In the 1990s prevalence rates continued to increase, through the 2000s and 2010s reaching a high of 147 per 10,000 in 2014, as reported by U.S. Department of Health and Human Services, Centers for Disease Control and Prevention (Wingate et al., 2014). Exactly why there has been such a large increase in autism diagnosis over the past 25 years is not fully understood. In some cases the rise in prevalence of autism has been termed an ‘autism epidemic’ (Cave, 2008; Gillberg, Cederlund, Lamberg, & Zeijlon, 2006; Leonard, Annaz, Karmiloff-Smith, & Johnson, 2011; Liu, King, & Bearman, 2010). Many researchers reject the notion of an ‘epidemic’, proposing that there is no ‘real’ increase in autism prevalence at all, only changes in understanding, awareness and better diagnostic tools (Gernsbacher, Dawson, & Goldsmith, 2005; Wazana, Bresnahan, & Kline, 2007).

Lastly, another topic that has prompted much discussion in the field of ASD is the greater number of males diagnosed with ASD compared to females. ASD has typically been regarded as affecting four times as many males as females (Rutter, 1978), with more recent reports suggesting that it may affect five times as many males as females (Wingate et al., 2014). There has been significant debate about this, a consistent argument being that ASD presents differently in females and, therefore, many females with ASD are overlooked or misdiagnosed, contributing to the apparent lopsided male to female ratio (Frazier, Georgiades, Bishop, & Hardan, 2014; Head, McGillivray, & Stokes, 2014).

The complexities associated with the diagnosis of ASD (as outlined above) have implications for how research is conducted in this field, as well as the way in which findings from research studies can be interpreted. The next section will outline some prominent issues intrinsic to ASD research.

## 1.2 Issues associated with ASD research

There is current debate over the terms that are used to describe individuals with ASD. The majority of recent, peer-reviewed publications have adopted the terms ‘individuals with autism spectrum disorder’ or ‘individuals with autism spectrum condition’. However, evidence suggests that individuals with ASD may prefer the terms ‘autistic’ or ‘person on the spectrum’ (Kenny et al., 2016). The person-first perspective of ‘individuals with ASD’ was, however, found to be a term in which *several* stakeholders (i.e. professionals, parents, caregivers and people on the spectrum) agreed to represent the condition. Henceforth, the terms ‘individuals with ASD’, ‘children with ASD’ and ‘adults with ASD’ will be used to describe the participants who took part in the research presented in this thesis.

Additionally, in certain sections of data analysis, participants with ASD may be described collectively as the ‘ASD group’ in order to make parallels with typically developed control participants who will be described as the ‘TD group’. It is hoped that using these terms will be acceptable to all relevant stakeholders, and that these terms are deemed appropriate, insofar as they are descriptive and clear in the context of this thesis.

Another contested issue in ASD research is the proposition that the majority of ASD research reflects only a minority of individuals with ASD, largely from the ‘higher’ end of the spectrum (i.e. average or above average IQ, with no co-morbid learning disabilities). The key reason for this is that most ASD research is designed for individuals who have typical IQ levels, rather than those who have IQs in the Learning Disabled (LD) range (< 70; sometimes called ‘low functioning’ though this term has been soundly rejected by the ASD community (Kenny et al., 2016)). Research with lower IQ groups is complicated by a combination of practical and ethical issues. Practical issues include ensuring that the task is not distressing to the participants and that the demands of the task are appropriate to the IQ level of the participant group. Ethical issues include ensuring that the participant is able to give truly informed consent and that they are willing to take part in the research.

Consequently, studies that examine complex cognitive skills are restricted by who is appropriate to take part in the experiment. This means that certain areas of ASD research,

such as complex cognition, often only reflect the abilities of a subset of the wider ASD population.

As well as being mindful of the issues described thus far, a feature that is essential for successful ASD research is the provision of an adequate control group. A control group is necessary in ASD research in order to have a reference with which participants with ASD can be compared. Several characteristics must be taken into account when selecting controls for ASD research, in particular: age, gender and level of education. It is also common for research studies to screen their control participants for autism-like personality traits, ensuring that all participants have no signs of having ASD. FER studies similar to those presented in this thesis (but in the typical population) have suggested that there is a difference in FER ability relative to ethnicity (Jack, Caldara, & Schyns, 2012b; Jack, Garrod, & Schyns, 2014). Hence, although recording information about ethnicity is not always routine in ASD research (Pierce et al., 2014) in the studies presented in this thesis, ethnicity was also recorded.

### **1.3 Typical perception of facial expressions of emotion**

The human face contains highly mobile features that allow an individual to send many social signals such as gestures (e.g. nodding), indications of where the person is directing their attention and patterns of gaze and movement that regulate turn taking in face-to-face interactions. The most highly researched of these signals, however, is facial expressions of emotion, regarded as a fundamental source of information about a person's current emotional state (Ekman, 1992). The earliest scientific exploration into facial expressions of emotion was by Duchenne in 1862 who found that it was possible to stimulate the independent movement of individual facial muscles on a participant who had lost the ability to move the muscles in his face due to nerve damage. Subsequently, Darwin (1872) hypothesised that certain facial expressions of emotion were universally recognizable, i.e. across all countries in the world facial expressions of emotion are universally recognized and easily interpreted through specific facial movements. Ekman et al (1992) built on this, providing a more specific case for 'basic facial expressions of emotion' suggesting that happy, surprise, fear, disgust, anger and sadness are universally represented using the same combination of facial movements in across the world. There have been arguments against this theory, however. For example, Russell (1994) questioned the reliability of the basic emotion concept, concluding that the association of facial expressions and labels is likely to vary across cultures. These sentiments are backed up by Jack, Caldara, and Schyns

(2012a) who found differences in the perception of happy, surprise, fear, disgust, anger and sadness in Eastern Asian participants compared to Western Caucasian participants.

Research has also suggested that the six emotions (claimed by Ekman et al (1992) to be the 'basic' facial expression of emotion) do not provide a sufficiently exhaustive explanation of basic emotion perception (Jack, Sun, Delis, Garrod, & Schyns, 2016). Jack et al. (2016) suggest instead that, because of an overlap in components of emotion perception in surprise and fear, and disgust and anger, four facial expression of emotion are the true 'basics'.

Typical development of facial expression recognition begins in early infancy and develops rapidly throughout the first year of life (Grossmann, 2010). Although some evidence of facial expression perception has been reported in neonates (i.e. < 4 weeks old; Farroni, Menon, Rigato, and Johnson (2007), there is a general consensus in the research literature that infants begin to recognise facial expressions around 7 months of age (Caron, Caron, & Myers, 1982; Nelson, 1987; WalkerAndrews, 1997). Given this, it is difficult to surmise whether reactions to facial expression at this young age reflect *recognition* of emotion per se or whether reactions to facial expression stimuli are merely recognition of affect in general. Nelson, 1987 notes that, even by the age of 5, the ability to correctly identify a range of facial expressions of emotion is only rudimentary. The neural processes involved in the perception of emotional faces are thought to develop in a staggered fashion from childhood to adulthood (Batty and Taylor, 2006). ERPs recorded in 82 children aged 4 to 15 years during an implicit processing task with emotional faces revealed that as participants increased with age, ERPs signals changed. For example, ERPs of P1 were found in the youngest participants and slowly changed with increasing age until N170 (associated with FER in adulthood) emerged in adolescence. Note, this was a cross-sectional study not a longitudinal study so individual differences may have had an impact on the results. It is, nevertheless, apparent that FER in typical individuals matures as individuals transition from childhood to adulthood.

## **1.4 The 'enigma' of facial expression recognition (FER) in ASD: assessing conflicting literatures**

It was a sustained 'impairment' of social communication and interaction with others that first led Kanner (1943) to define autism as a "disturbance of affective contact". However, the cause and presentation of the characteristics associated with this continues to puzzle researchers to this day. Hence, ASD has been described as an 'enigma' (Uta Frith, 1989),

and despite huge advances in an evidence-based understanding of ASD, many aspects of the condition are still not well understood.

Visual perception is one area of ASD that, despite being extensively researched, presents very few universally agreed findings. Results range from impairments in some aspects, through to superior abilities in others, but often no differences are found compared to TD individuals (Dakin & Frith, 2005; Simmons et al., 2009). The perception of faces is generally thought to be ‘special’ to the typical visual system, as compared to other classes of objects (Gould, 1994). Because of this and the social difficulties associated with ASD, researchers have long been interested in the perception of faces in individuals with ASD, with several reviews published on this topic alone (Golarai, Grill-Spector, & Reiss, 2006; Jemel, Mottron, & Dawson, 2006; Weigelt, Koldewyn, & Kanwisher, 2012). Much research has suggested that, in individuals with ASD, faces do not have the same ‘special’ status as that found in the visual system of the typical population. This claim is not, however, universally accepted in the field and still continues to be debated (Simmons et al., 2009).

A specific aspect of visual/facial perception research in ASD that has sustained debate for several decades is the perception of facial expressions of emotion (Harms et al., 2010; Simmons et al., 2009; Uljarevic & Hamilton, 2013). This area of research has given rise to over 140 research papers varying greatly in the method, design and participant variables. However, all are essentially asking the same question: are individuals with ASD ‘impaired’ at recognising facial expressions of emotion? The aim of this section is to assess evidence of an atypical manifestation of facial expression recognition (FER) in individuals with ASD.

### **1.4.1 Early research**

The first major research study to assess FER in individuals with ASD used schematic drawings and photographs of facial expressions of emotion and asked children to pair these with videotapes of emotional gestures, showing the body only (Hobson (1986a)). Hobson (1986a) reported that children who were ‘autistic’ (at this point, the term ‘ASD’ was not yet established) were significantly less accurate than mental-age-matched controls at recognising what expressions corresponded to the gestures. Hence, he concluded that ‘autistic’ children were impaired in the recognition of emotions and subsequently ratified this finding by repeating the study with schematic drawings of gestures (Hobson, 1986b).

Subsequent to this, Hobson and colleagues found that adolescents and young adults with ASD were worse than verbal-age matched controls at matching facial expressions of emotion, despite having a superior ability in matching photographs of upside-down faces (Hobson, Ouston, & Lee, 1988). Following on from these seminal studies, a succession of research papers were published using similar methods as seen in (Hobson, 1986a) using static images of facial expressions (Braverman, Fein, Lucci, & Waterhouse, 1989; Ozonoff, Pennington, & Rogers, 1990). In particular, stimuli of black and white static facial expressions (now regarded as ‘classic’ stimuli) from the resource ‘Pictures of Facial Affect’ were used (POFA; P. Ekman and W. V. Friesen (1976); (Macdonald et al., 1989; Tantam, Monaghan, Nicholson, & Stirling, 1989)). Note, the POFA stimuli depict only what Ekman and Friesen (1978) regard as the six ‘basic’ facial expressions of emotion: happy, surprise, fear, disgust, anger and sadness. Research on FER in ASD flourished in the 1990s, with several studies ratifying the indications of the seminal papers that FER is atypical in individuals with ASD (Bormannkischkel, Vilsmeier, & Baude, 1995; Capps, Yirmiya, & Sigman, 1992; Celani, Battacchi, & Arcidiacono, 1999; Davies, Bishop, Manstead, & Tantam, 1994; Feldman, Mcgee, Mann, & Strain, 1993). However, a significant number of papers indicated *no* FER atypicality in individuals with ASD as compared to TD individuals (S. Baron-Cohen, Wheelwright, & Jolliffe, 1997; Buitelaar, Van der Wees, Swaab-Barneveld, & Van der Gaag, 1999; Fein, Lucci, Braverman, & Waterhouse, 1992; Loveland et al., 1997; Prior, Dahlstrom, & Squires, 1990). Hence, this phase of research gave rise to conflicting conclusions about whether individuals with ASD are impaired in FER, which has reverberated throughout the research field to the current date.

After 1990, over 100 studies were published in this field, all of which were aiming to ascertain whether FER in ASD is fundamentally different from the typical population. Uljarevic and Hamilton (2013) aimed to address the confusion within the FER ASD literature by conducting a meta-analysis across a large number of studies in this area. Meta-analysis is a tool used to obtain information about the effect size of a particular phenomenon by evaluating data from a series of research studies (Cooper, Hedges, & Valentine, 2009). Uljarevic and Hamilton (2013) performed a systematic literature search for all studies published on FER in ASD before December 2011. They performed a meta-analysis on a selection of data from this period, aiming to report an overall effect size of the results. They revealed a large negative effect size between groups (-0.80), indicating that participants with ASD were, on the whole, less accurate at FER than TD participants. They also report that there was a large amount of heterogeneity across studies, indicating

that despite finding an *overall* negative effect size, the effect sizes of each study are considerably variable. In addition to this, they assessed the contribution of IQ and age on FER scores, finding no overall contribution of either construct. In addition, they indicated that task (i.e. which task was used in the study) did not contribute to effect size. This finding is interesting given the argument from Harms et al (2010) that task demands were particularly liable to influence FER outcomes. Uljarevic and Hamilton (2013) also assessed differences between individual emotions and reported that the recognition of surprise, fear, disgust, anger and sadness were all significantly poorer in ASD individuals. They reported an effect for happy, however they indicated that this is marginal, and so not reliably different between the ASD and TD groups. However, the data used for the meta-analysis conducted by Uljarevic and Hamilton only reflect a proportion of the data available. For example, if participants completed more than one FER task, Uljarevic and Hamilton chose only one task result to include in the meta-analysis and discarded the remaining data. Their reasons were understandable insofar as they did not want to introduce multiple comparison effects due to using the same sample of participants more than once. However, the downside to this is that a large proportion of the data, which was potentially meaningful to the meta-analysis, was excluded. On top of this, Uljarevic and Hamilton indicated that they actively chose certain studies over others: studies with static stimuli were chosen over those with dynamic stimuli and studies assessing basic emotion labelling tasks were chosen over other types of task. Selecting specific tasks in this way may have introduced an element of bias into the meta-analysis. No further analysis was conducted with the discarded tasks, hence, Uljarevic and Hamilton provide no confirmation to the reader that the meta-analysis results would have remained the same given the inclusion of the other data. The inclusion of certain participant groups was also selective in the meta-analysis because, when two or more control groups were used (for example one group matched on verbal IQ and another group matched on non-verbal IQ), they were selective to the data they chose, opting for data that was compared to verbal IQ controls over non-verbal IQ controls. Another area of concern in this meta-analysis was the bias toward static stimuli. For example, in studies that reported both dynamic and static FER tasks, Uljarevic and Hamilton excluded the dynamic task that is arguably more ecologically valid. Since differences in static and dynamic emotion processing has been reported in this field (e.g. Philip et al. (2010)) it would have been useful to have a clearer picture of the influence of dynamic stimuli of effects size. Lastly, the data included in this meta-analysis inclusive of studies published until December 2011 and so do not reflect the advances in methodology presented in this area of research over past 5 years. The pertinent questions raised by the early research, as discussed in this section, is further reviewed below.

## 1.4.2 Components of FER processing

### 1.4.2.1 Timing: comparison of FER in static and dynamic stimuli

Since the seminal research papers on FER in ASD were published, researchers have aimed to increase the ecological validity of facial expression stimuli. However, results remain mixed with regard to whether, when stimuli are more ‘life-like’, clearer differences in FER will emerge. Apart from efforts to use videotaped facial expressions in Hobson (1986a), Loveland et al. (1997) conducted the first study to use dynamic stimuli (as opposed to photographs or drawings) to assess FER in individuals with ASD. They presented children with video clips of actors conveying happy, sad and angry facial expressions. The study offered a complex assessment of FER ability in children with ASD because they tested accuracy of FER both verbally and non-verbally, as well as assessing FER based on an emotive or neutral face (i.e. the actor producing the facial expression) paired with an emotive or ‘flat’ sentence (i.e. the actor saying, for example, “I am happy I am going to the zoo”). They found that there was no difference between children with ASD and TD children in their accuracy at labelling facial expressions of emotion in all conditions. However, it could be argued that because children were provided with a ‘trigger’ word while the stimuli were presented (i.e. in the sentence vocalised by the actor) the children may have been relying more on verbal content of the utterance than visual content of the face and hence it is possible that the recognition of facial expressions of emotion *per se* may not have been directly measured using this method. Gepner, Deruelle, and Grynfeldt (2001) built upon this initial dynamic stimuli study by assessing FER using two-second videos of actors performing facial expressions of surprise, happy, sadness and disgust, which were presented to children with ASD and mental-age matched controls. As in Loveland et al. (1997), results revealed no difference between children with ASD and TD children. However, Gepner et al. (2001) asserted that task performance improved with chronological age in the TD children, while task performance remained unchanged in the children with ASD. Hence, although overall findings suggested no significant difference between children with ASD and TD children, slight ‘improvements’ in FER could be found in the TD group as a function of age. Taken broadly, the findings of these initial studies of FER using dynamic stimuli gave rise to the notion that dynamic stimuli are somewhat more difficult for individuals with ASD to recognise than static stimuli.

Advances in methodology that were implemented after this period (as discussed further in section 1.4.4) gave rise to greater sensitivity to measuring FER ability. For example, Philip et al. (2010) assessed FER, measuring the accuracy of participants’ ability to recognise

emotions from photographs of facial expression of emotion (P. Ekman & W. V. Friesen, 1976) that morphed from neutral expressions to a full emotion expression. Participants were required use the emotion labels ‘happy’, ‘surprise’, ‘fear’, ‘disgust’, ‘anger’ and ‘sadness’ to categorise stimuli that differed in the extent to which they express the emotion. Philip et al. (2010) found a significant difference in FER accuracy between adults with ASD and TD adults. A succession of papers reflected similar results. For example, when ‘morphed’ facial expression stimuli were presented at varying speeds to children with ASD, Sato, Uono, and Toichi (2013) found that children with ASD were significantly more likely than typical children to rate slow-moving morphs as ‘natural’ looking. Sato et al. (2013) suggested from these data that dynamics are not only part and parcel of effective FER, they are what underlie poor FER in individuals with ASD. In addition to this, Lerner, McPartland, and Morris (2013) indicated that ‘social information processing speed’ underpinned emotion recognition from faces and voices. It is, therefore, tempting to conclude that it is a difficulty with processing the temporal aspects of facial expressions of emotion that is integral to understanding FER abilities in individuals with ASD. Lerner et al. (2013) found that children with ASD who were less accurate in a basic FER labelling task had slower N170 latency (the event-related potential response associated with face perception). However, this result has not always been replicated. For example, Gold et al (2013) found that adults with ASD were equally proficient at detecting facial expressions of emotions from static faces as compared to dynamic faces. Further to this, Enticott et al. (2014) argued against a ‘broad advantage of using dynamic faces’ after finding that adults with ASD had *improved* accuracy at labelling disgust and angry faces when stimuli were dynamic, as compared to static. Thus, taken together, the majority of evidence suggests that individuals with ASD are hindered in FER when stimuli are dynamic and arguably more ‘life-like,’ but further research is needed in order to convince all researchers in the field.

#### **1.4.2.2 Amplitude: Do ‘intense’ stimuli improve FER accuracy in individuals with ASD?**

Another component of FER that has been attributed to FER ability in individuals with ASD is intensity of the facial expression. For example, Law Smith, Montagne, Perrett, Gill, and Gallagher (2010) found that, for the emotions ‘surprise’ and ‘fear’, FER accuracy was directly linked to the intensity of the stimulus, where individuals with ASD were significantly less accurate than TD individuals at detecting fear and surprise when the stimuli were less intense. Similarly, G. L. Wallace et al. (2011) found that adolescents with ASD required more intense stimuli in order to accurately recognise basic facial expressions

of emotion. Tell, Davidson, and Camras (2014), on the other hand, indicated that intensity did not make a significant difference to FER accuracy. Thus, while few studies have assessed the contribution of intensity to FER ability, it is clear that mixed conclusions have arisen from the evidence presented in the literature thus far.

#### **1.4.2.3 Are difficulties with FER specific to the eye region?**

An atypicality in eye-contact is a behaviour that is indicative of the ASD diagnostic criterion ‘deficits in nonverbal communicative behaviors used for social interaction’ (*DSM-5*, 2013). In a lab setting, this atypicality has been ratified, with research indicating an avoidance of the eye region when viewing images of social scenes (Klin, Jones, Schultz, Volkmar, & Cohen, 2002) and faces (Joseph & Tanaka, 2003; Pelphrey et al., 2002). Research has reported that the avoidance of the eye region is also present when viewing facial expressions of emotion (Gross, 2008). Crawford, Moss, Anderson, Oliver, and McCleery (2015) conducted an experiment in which adults with ASD, adults with Fragile-X syndrome (FXS; a genetic disorder similar in symptomology to ASD) and TD adults passively viewed images of neutral, happy and angry faces (note, there were always two faces on the screen and one of these was always neutral). They used eye tracking to gauge where the participants ‘spontaneously looked’, finding that the ASD and FXS participants were relatively similar to TD, but the FXS participants spent less time looking at the eyes. However, the extent to which this helps explain FER in individuals with ASD is still unclear. For example, the first study to assess the *independent* contribution of the eye and mouth regions to FER accuracy in individuals with ASD indicated that the whole face was more informative than the eyes or mouth alone for both ASD and TD groups (S. Baron-Cohen et al., 1997). This was, however only the case for basic facial expressions of emotion, and S. Baron-Cohen et al. (1997) reported poorer FER in adults with ASD when participants were asked to label ‘complex’ facial expressions of emotions (for example ‘admiration’ and ‘thoughtfulness’). Similarly, Rutherford and Towns (2008), found that both ASD and TD groups looked significantly longer at the eyes than the mouth when asked to label photographs of basic emotions (as measured by eye-tracking equipment), and neither overall ‘looking time at eyes’ nor ‘first fixation on the face’ was found to distinguish the groups. Nevertheless, they did report a significant lack of fixation on the eye area in individuals with ASD compared to the control group when viewing *complex* emotions. Peterson, Slaughter, and Brownell (2015) found, on the other hand, that children with ASD were significantly worse at basic *and* complex versions of the ‘Reading The Mind In The Eyes’ task (S. Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001)

compared to TD children. Research has also assessed the relative contribution of the mouth region, hypothesizing that individuals with ASD will pay *more* attention to the mouth region in FER (Evers, Kerkhof, Steyaert, Noens, & Wagemans, 2014). Hence, despite there being some evidence that atypical use of the eye information influences FER abilities, this does not appear to be a *core* indicator of an FER difficulty. Hence, further research is needed to more accurately understand how information obtained from the eye regions of faces influences FER in individuals with ASD.

#### **1.4.2.4 Are difficulties with FER specific to certain emotions?**

The majority of studies in ASD FER have focussed on the so-called ‘basic’ emotions: happy, surprise, fear, disgust, anger and sadness (Ekman and Friesen, 1976). Overall, there is evidence to suggest that ‘happy’ is the emotion that individuals with ASD have least trouble with as compared to the other emotions. For example, Uljarevic and Hamilton’s 2013 meta-analysis of 48 studies in ASD FER literature reported an overall negative effect size (i.e. the ASD group were less accurate than the TD group) for FER of surprise, fear, disgust, anger and sadness but not happy. These overall findings have led some to propose that individuals with ASD have a general difficulty with negative emotions. However, there is mixed evidence as to whether FER ability is less accurate in individuals with ASD for specific emotions only. For example, research studies have indicated there may be a particular difficulty with fear (Corden, Chilvers, & Skuse, 2008; Humphreys, Minschew, Leonard, & Behrmann, 2007), disgust (S. Wallace, Coleman, & Bailey, 2008), sadness (Boraston, Blakemore, Chilvers, & Skuse, 2007), anger (Ashwin, Chapman, Colle, & Baron-Cohen, 2006) and surprise (Baroncohen, Spitz, & Cross, 1993). On the other hand, studies have also reported no difficulty with negative emotions (e.g Lacroix, Guidetti, Roge, and Reilly (2009) and Piggot et al. (2004)) and some have found differences between ASD and TD groups in FER of happy too (Philip, 2010, Humphreys et al 2007). Hence, evidence for particular emotions being disproportionately difficult for individuals with ASD is currently somewhat mixed.

#### **1.4.2.5 Conclusion about components of FER processing**

The research reviewed thus far suggests that, despite a wide range of research on FER in ASD, the specific cause of atypical FER in individuals with ASD, should it exist at all, remains unclear. It is notable from the research that, although individuals with ASD may perform in a similar manner to control individuals at recognising facial expressions from the 1-dimensional components of the face (i.e. the presentation of a full expression given in

a static photograph), greater difficulties appear to arise with other components of facial expressions, such as timing and intensity. Hence, these research findings give rise to a demand for a greater understanding about the multi-dimensional components of facial expressions and how these contribute to FER differences between ASD and TD populations. Despite this thirst for knowledge, and a mandate to assess these components of FER, few studies have extended methodological designs past 1-dimensional static photographs of facial expression and hence the methodology commonly used in this area of research has not yet caught up with the questions that are being asked (discussed further in section 1.4.4). Further to this, there are complications in this area of research that contribute to the complex findings reported in the literature, namely the heterogeneity within participants, which will now be discussed.

### **1.4.3 Contribution of participant sample characteristics**

#### **1.4.3.1 Are individual differences and ‘sub-groups’ overlooked in FER research in ASD?**

The studies described in section 1.4 have indicated that there are conflicting results in the FER ASD literature, however it could be argued that the majority of these papers describe results in terms of *between group* analyses only, neglecting the possibility of *within group* differences. As described in section 1.1, ASD is, by its very nature, a heterogeneous disorder and so it is likely that differences in FER could be apparent within a group of participants with ASD. The existence of sub-groups within participant samples of individuals with ASD could, therefore explain the conflicting findings present in the FER ASD literature because, by grouping all individuals with ASD together at analysis stage, findings may reflect broad differences compared to TD individuals in *some* cases and no difference in others. For example, Lerner et al. (2013) found an overall poorer FER accuracy in children with ASD (as compared to typical children) but typical FER accuracy in a subgroup of children with ASD. Back and Jordan (2014) argue for ‘individual differences’ as an important factor in FER in the typical population. Although this study does not relate specifically to individuals with ASD, results revealed large individual differences in labelling dynamic facial expressions of emotion. Hence, Back and Jordan (2014) provide a ‘cautionary note’ to researchers conducting studies in FER, advocating that individual differences may play a bigger role than is currently assumed in the FER literature. For example, individual differences in language ability, verbal IQ, intellectual ability, or perhaps the ‘severity’ of ASD characteristics themselves may play an important part in the ability to accurately recognise facial expression of emotion. There is however, a

need for more evidence to support this argument, which is compounded by limitations in terms of methodological possibilities (discussed further in section 1.4.4). These data provide useful insight into the contribution of individual differences in FER, which, despite being understudied, may be particularly important for accurate understanding of FER in ASD due to the heterogeneous nature of the condition. This argument is complicated by the fact that the majority of research in the field of FER in ASD samples only from individuals with average or above average IQ, a complication that is now discussed in more detail.

### 1.4.3.2 Is FER linked to ‘severity’ of ASD symptoms?

Research has suggested that scores on ASD diagnostic tests (such as the ADOS (Lord et al., 2000)) may predict outcomes on FER tasks, indicating that individuals who have higher levels of ASD symptoms (and hence have more profound difficulties and complex needs) are less accurate at FER. For example, Uono, Sato, and Toichi (2013) found a significant correlation between social dysfunction and FER ability in adults with ASD, as measured by accuracy scores of labelling photographs of facial expressions. However, this was only apparent for the expression of fear (not happy, surprise, disgust, anger or sadness). It is also evident here that there may be some circularity to the argument that social dysfunction level predicts FER ability, i.e. poor FER may also have a *causal* effect on social dysfunction level. This finding was also reflected in a study by Lerner et al. (2013), which assessed event-related potential (ERP) activity in children with ASD in relation to their ability to recognise facial expressions in a behavioural photo-labelling task. Results revealed that children with ASD made significantly more errors in FER, which correlated significantly with latency of the N170 ERP. However, they reported large variability within the ASD group in terms of their FER ability, indicating that the top quartile performed at least a standard deviation above the mean of the control group and the bottom quartile performed at least 3 standard deviations below the control mean. B. T. Williams and Gray (2013) reported that autism ‘severity’ (as assessed using the ADOS) was linked to the ability to recognise emotions in a basic FER labelling task. B. T. Williams and Gray (2013) did, however, find this connection only for expressions of anger and fear (not happy, surprise, disgust or sadness) and only in an emotion *recognition* task (no connection was found in an emotion *matching* task with the same children). Hence, the number of studies assessing individual differences is very small in the FER ASD literature and of those that assess it, results are somewhat divergent.

There are two factors which further compound the issues related to individual differences in FER ASD research: 1) very few studies report results on individuals with below-average IQ, learning disabilities or complex needs (Harms et al., 2010) and so there is a bias in the literature towards participants who have higher cognitive ability and 2) individuals with ASD with higher cognitive ability may process facial expressions of emotion in an atypical way but perform well in lab-based FER tasks because they have developed rule-based strategies to ‘work-out’ what the components of the face represent, for example ‘an upturned mouth = happy’ or ‘a downturned mouth = sad’ (Rutherford & McIntosh, 2007; Walsh, Vida, & Rutherford, 2014). In light of the points raised in this section, it could be argued that mixed findings emerge from this area of literature, which are further compounded by a biased sampling from individuals with ASD who have higher cognitive ability. This, in turn, is complicated by the potential use of rule-based strategies to decode emotions.

#### **1.4.3.3 FER in the typical population: the case for Alexithymia**

Alexithymia is a condition which was first described as a ‘relative constriction in emotional functioning, poverty of fantasy life and inability to find appropriate words to describe emotions’ (Sifneos, 1973). The term alexithymia derives from the Greek terms ‘a’ (meaning lack), ‘lexis’ (meaning word) and ‘thymos’ (meaning emotion; Taylor (1984)). Hence, alexithymic individuals are thought to “1) have a marked difficulty in expressing feelings in words and 2) not have fantasies appropriate to (or expressive of) feelings, their thought content being dominated by details of events in their external environment” (Nemiah, 1977). However, in more recent times, alexithymia is often simply referred to as ‘emotion-blindness’, and as a ‘difficulty in identifying and understanding one’s own emotions and the emotions of others’ (Mateos, 1993). Although the aetiology of alexithymia was recognised prior to Sifneos 1973, it was not until it was given its title that it was recognized as a phenomenon in its own right. Alexithymia is a concept that has emerged from the field of health research, in particular somatoform disorders (for review see Koch et al. (2015)). Evidence suggests that alexithymia measures can be used as an indicator of mental ill health (Hartwig et al 2014) and has been noted to accurately assess general psychological distress disorder and not just alexithymia (Leising et al 2009). In addition to this, Dongues and Suslow (2017) revealed that there are deficits in the automatic processing of emotional stimuli in alexithymia at a behavioural *and* neurobiological level. Hence, alexithymia is a concept that has been established in several different disciplines and has weaved a narrative through medical, psychiatric,

psychosomatic and cognitive research domains (Zackheim, 2007). Because of this, much of the early research conducted on alexithymia approached the concept from a medical perspective, in particular in the context of psychosomatic medicine. Over the past 20 years, the assessment of alexithymia has expanded into several wider areas of research, especially clinical and experimental psychology.

The first demographic assessment of Alexithymia was in Finland by Salminen (1999). Which has been reported as being stable across time (Hirola et al 2017). Alexithymia is thought to be found in around 10% of the typical population (Salminen, Saarijarvi, Aarela, Toikka, & Kauhanen, 1999). It has been suggested that 50% of people with an ASD diagnosis have alexithymia (Berthoz & Hill, 2005). Despite this, there is limited experimental evidence about the contribution of Alexithymia to FER ability in individuals with ASD. Research has found that alexithymia is significantly higher in individuals who have difficulty attributing mental states to others (but not ASD *per se*; Moriguchi et al. (2006)). Alexithymia was first assessed in relation to emotion perception in individuals with ASD by E. Hill, Berthoz, and Frith (2004), who concluded that individuals with ASD had significantly higher alexithymia levels compared to TD, individuals assessed using the Toronto Alexithymia Scale (TAS; Bagby, Parker, and Taylor (1994)).

Cook, Brewer, Shah, and Bird (2013) found that FER ability could be predicted by alexithymia score (using the TAS) over and above ASD severity, concluding that alexithymia, not ASD, is the source of FER difficulties. Further, Bird and Cook (2013) asserted that FER difficulty is not necessarily a symptom of ASD *per se*, but, since FER difficulty can be found in the typical population as alexithymia, the sub-group of individuals with ASD who have FER difficulty would be better defined as individuals with ASD and *co-occurring* alexithymia. This extends previous research that found that individuals with ASD are more likely to have alexithymia (Fitzgerald & Bellgrove, 2006). However, Nishimura et al. (2009) caution that the subjective difficulties with describing and identifying emotions are associated with empathetic and linguistic abilities; therefore using a self-report questionnaire (such as the TAS) to assess alexithymia requires particularly careful delivery and administration. Although Nishimura et al. (2009) do not identify ASD specifically in this cautionary note, it is apparent that the issues described would be applicable to individuals with ASD.

Alexithymia, as assessed using the TAS-20 has been evaluated in many contexts. For example, scores of alexithymia have been subject to meta-analysis in the field of

depression (Li, B., Guo, & Zhang, 2015) post traumatic stress disorder (Frewen, Dozois, Neufeld, & Lanius, 2008) and criminal offenders (D. Jolliffe & Farrington, 2004). However, the efficacy of the TAS-20 as an alexithymia measure has not been assessed on this scale in individuals with ASD. There are also intrinsic issues with self-report and measuring alexithymia. In other words, the ability to accurately reflect on one's own feelings and behaviours (as is necessary in order to complete the TAS-20), may be hampered by having the very traits that alexithymia is testing (i.e. issue with 'describing feelings and emotions', 'identifying feelings and emotions' and 'externally-oriented thinking'). Therefore, the field of ASD moves forward tentatively with the measurement of alexithymia, and more research is needed in order to provide reliable evidence that alexithymia can be measured effectively in individuals with ASD.

The contribution of alexithymia to better understand emotion perception in ASD is gaining traction (Ketelaars, In't Velt, Mol, Swaab, & van Rijn, 2016; Livingston & Livingston, 2016; P. Shah, Hall, Catmur, & Bird, 2016), however, further evidence is needed to ratify the notion that FER may be a co-occurring condition to ASD in the form of alexithymia.

#### **1.4.4 Experimental methods used to assess FER in individuals with ASD**

The most enduring conceptualization of human perception of facial expressions of emotion comes from Ekman and Friesen (1978) who argue that there is a universal 'language' evoked by facial expressions of emotion that can be boiled down into six basic components: happy, surprise, fear, disgust, anger and sadness (Ekman, 1992). There are, however, several accounts of facial expression perception that do not agree with this theory, e.g. (Cohen, 2005; Jack et al., 2014; Jack et al., 2016). Within ASD research, the majority of emotion research is rooted in the theory proposed by Ekman and Friesen (1978) and hence, though this thesis is not directly testing the validity of Ekman and Friesen's theory, the research methodologies described in the experimental chapters stem from the conceptualization of emotion based on Ekman and Friesen's theory.

Measuring the perception of emotions in individuals with ASD has, since the seminal research in this area, proven to be difficult e.g. Hobson (1991). These difficulties are often due to intrinsic factors related to ASD research (as discussed in section 1.1), such as limitations regarding methodological techniques when designing research studies that are appropriate to individuals with ASD that have complex needs or lower-than-average cognitive functioning.

There has been a demand within the literature for a development of more advanced, novel techniques in order to satisfy questions regarding inconclusive findings in the FER in ASD literature (Harms et al., 2010; Uljarevic & Hamilton, 2013), with a particular demand for rigorous, psychophysical methods to be utilised (Dakin & Frith, 2005; Simmons et al., 2009). The majority of methods employed to assess FER in individuals with ASD use static photograph images of actors evoking high intensity (i.e. full-expression) emotions, most notably using the battery of images from P. Ekman and W. V. Friesen (1976), schematic drawings, or cartoon representations of facial expressions. Some recent studies have taken a different approach. For example, a study by Eack, Mazefsky, and Minshew (2015) look at the precise *confusions* between facial expressions of emotion in ASD (although the stimuli were still static photographs), reporting that adults with ASD have most difficulty differentiating neutral expressions from emotional expressions at different intensity levels. Another study that pushed the boundaries of basic FER methodology was conducted by Evers, Steyaert, Noens, and Wagemans (2015). They assessed facial expression perception in ASD using stimuli that were not only high quality *and* dynamic, but that also varied in intensity (here they had ‘medium’ which was a morph from 0% emotion to 50% of the full emotion and ‘high’ which was a morph from 0% emotion to 100% of the emotion). They also controlled for response bias, finding that there was a significant difference in emotion perception accuracy between children with ASD and children who were typically developed. Overall, advances in methodologies used in the field of FER in ASD research have largely been made by introducing either: 1) a timing component or 2) an intensity component. What remains difficult in this area of research is the trade-off between ‘life-like’ (and, hence, more ecologically valid) facial expressions and control over stimulus features. For example, videos of actors displaying facial expressions of emotion are very life-like insofar as they have high ecological validity for everyday life experiences; however, videos allow for very little manipulation of the stimuli and do not allow for sensitivity of measurement (participants can be measured only on a single response). Methods presented in this thesis introduce a novel technique for measuring FER (Yu, Garrod, & Schyns, 2012) and apply this to measuring FER in individuals with ASD.

Despite a call for more rigorous psychophysical methods to be applied to FER ASD research (Dakin & Frith, 2005), few studies have employed these methods. There are intrinsic aspects of ASD research that limit the scope of these types of methods with a participant sample of individuals with ASD (such as level of ASD symptoms; discussed in section 1.2). Spezio, Adolphs, Hurley, and Piven (2007b) used a novel psychophysical

method called ‘bubbles’ (Gosselin & Schyns, 2001) to reveal information about eye movement patterns when participants with ASD look at faces. Adolphs, Spezio, Parlier, and Piven (2008) utilised a similar method to assess FER strategies in parents of children with ASD. Therefore, it is evident that novel techniques are required for better understanding of the complex findings currently presented in the FER ASD literature and although advances have been made, this area of research would benefit from further research utilizing novel methodologies.

#### **1.4.5 Summary of FER research**

Thus far, the research reviewed in section 1.4 highlights that there are conflicting evidence-based arguments as to whether individuals with ASD have atypical FER. Employing a multi-dimensional assessment of FER (e.g. looking at intensity and dynamics of realistic facial expressions) has provided a greater depth of evidence and greater detail about the atypicalities in FER in individuals with ASD. However, it is also apparent that a complex web of factors may impede this, such as heterogeneity of presentation of ASD, ‘severity’ of ASD symptoms and alexithymia.

### **1.5 Role of the ‘Autism Spectrum Quotient’ in FER ASD research**

Research has explored the scope of ASD heterogeneity, suggesting that ASD can also be considered in terms of a continuum between autism and typicality that extends beyond the boundaries of clinical diagnosis. In other words, individuals who do not have a clinical diagnosis of ASD may exhibit traits of autism-like cognitive style and behaviour. The idea of ASD as a continuum arose firstly from Wing (1988) and was later extended by S. Baron-Cohen (1997). To test the idea of an autism-trait continuum, the Autism-spectrum Quotient (often abbreviated to ‘AQ’) was developed by S. Baron-Cohen, Wheelwright, Skinner, et al. (2001). This is a measure that can be applied to any adult with average or above average IQ (i.e. no learning disability). The AQ is a fifty-item self-report questionnaire that probes information about the level of autism-like traits an individual has. Responses are given in terms of a four-point Likert scale (‘definitely agree’, ‘slightly agree’, ‘slightly disagree’ or ‘definitely disagree’). The questions target five subscales, which are thought to be characteristic of an ASD diagnosis: willingness of adequate social interaction and interest, degree of repetition, imagination, degree of empathy and attention to detail. S. Baron-Cohen, Wheelwright, Skinner, et al. (2001) claimed that, at the time the

study was conducted, 80% of people with an ASD diagnosis had a score of between 32 and 50, whereas in a control group (typically developing) only 2% had a score this high.

Some research has indicated comparable results between high AQ scorers and those with diagnosed ASD. Reed et al (2011), for example, found comparable results on the Navon letters task (Navon, 1977) and Grinter et al. (2009) reported that participants with high AQ scores performed similarly to individuals with ASD on an embedded figures task. Robertson and Simmons (2013) found that increased levels of sensory sensitivity correlated with increased AQ scores. In the context of facial expression perception, Poljac, Poljac, and Wagemans (2013) showed that FER accuracy is reduced in individuals with high AQ as opposed to individuals with low AQ scores. S. Baron-Cohen, Wheelwright, Skinner, et al. (2001) argue that the AQ is a valuable instrument for ‘rapidly quantifying’ where any given individual is situated from ‘autism to normality.’ Whilst there are other methods for quantifying autism-like trait levels such as the Social Responsiveness Scale (Constantino & Todd, 2003) and there have been strong criticisms of the use of AQ score as a proxy diagnosis (Gregory & Plaisted-Grant, 2013), it remains a commonly used and well-validated instrument for use with typical populations (Ruzich, Allison, Chakrabarti, et al., 2015; Ruzich, Allison, Smith, et al., 2015).

## 1.6 Theoretical context

Several theories have been established in an attempt to explain the cognitive mechanisms that underlie the ASD phenotype (Pellicano, 2011). The most prominent cognitive models of ASD are Theory of Mind (S. Baron-Cohen, Leslie, & Frith, 1985), Executive Dysfunction (U. Frith, 1972) and Weak Central Coherence (Happe & Frith, 2006) as reviewed in Pellicano (2011) and Rajendran and Mitchell (2007). Additionally, ASD has been conceptualised using a ‘fractionated’ approach (Brunsdon and Happe, 2013) as well as the ‘Enactive Mind’ theory (Klin 2003), the Bayesian theory (Pellicano & Burr, 2012) and the ‘Social Motivation’ theory (Chevallier, Kohls, Troiani, Brodtkin, & Schultz, 2012). Neural theories have also been presented such as the Amygdala theory (Schultz, 2005) and Neural Noise theory (Simmons, et al 2009). In relation to the research presented in this thesis, two cognitive theories are particularly relevant to FER: 1) Theory of Mind (Baron-Cohen, Leslie, & Frith, 1985) 2) Weak Central Coherence (Happe & Frith, 2006).

Theory of Mind (ToM) is the ability to attribute mental states to oneself and to others (sometimes termed ‘mentalizing’ (Uta Frith, 2003)) and was first established in Dennett

(1978). Applied in an experimental format to children with ASD, S. Baron-Cohen et al. (1985) found that children with ASD were significantly less able in a task that involved attributing a mental state to another person (the ‘False Belief Task’) compared to TD children. Subsequent to this, many research studies have provided evidence to support this finding (see Tager-Flusberg, 2007 or Baron-Cohen et al 2000 for a review of this topic). However, although this theory is useful in explaining some aspects of ASD characteristics, it does not provide an exhaustive answer to the cognitive aspects of ASD (Pellicano, 2011). In terms of FER, ToM is informative because accuracy in social information processing requires an element of ‘reading’ another person’s mind. It could be argued that skills in ToM may be directly related to skills in FER, for example, individuals with ASD who have particularly good ToM skills compared to other individuals with ASD may be particularly good at FER. However, does FER really require effective skills in ‘metaling’?. It could be argued that FER is a skill that can be improved via learning visual cues (e.g. (Golan et al., 2010) and hence FER difficulties may be a social-specific as opposed to facial-specific (Weigelt, Koldewyn, & Kanwisher, 2013).

Central Coherence theory (CC), initially described as ‘Weak Central Coherence’ (WCC), refers to a specific cognitive ‘style’ used by individuals with ASD which favours thinking about things in the smallest possible parts as opposed to seeing the ‘bigger picture’ (Uta Frith, 1989; Happe & Frith, 2006). This theory was particularly useful in helping explain findings from visuo-spatial tasks in which children with ASD, were found to perform *better* than TD children in tasks which involved searching for embedded shapes within images (T. Jolliffe & Baron-Cohen, 1997; A. Shah & Frith, 1983). Hence, the ‘weak’ aspect of WCC was used to explain visual processing in ASD as an impairment (or ‘weakness’) in global processing of visual information (i.e. the ‘bigger picture’) which, in turn, causes a greater focus on smaller parts of the visual scene, and hence results in better performance in visual search tasks (Uta Frith, 1989; Happe & Frith, 2006; Pellicano, Gibson, Maybery, & Durkin, 2003; Pellicano, Gibson, Maybery, Durkin, & Badcock, 2005). However, the idea that superior processing of small parts of a visual scene is due to a ‘weakness’ in processing the bigger picture has been challenged in favour of the idea that superior processing of small parts exists alongside intact processing of the whole scene, sometimes referred to as ‘Enhanced Perceptual Functioning (EPF; (Mottron, Burack, Iarocci, Belleville, and Enns (2003); (Mottron, Dawson, Soulieres, Hubert, & Burack, 2006)). Hence, the term ‘Central Coherence’ has been adopted as a more fitting term to describe this theory (Chan & Naumer, 2014). Central Coherence theory suggests that it would be difficult for individuals with ASD to group ‘local’ features into meaningful

wholes. Consequently, it might be expected that judgements involving complex visual stimuli, like faces, may be compromised due to the necessity of grouping local features across space and time (Campatelli, Federico, Apicella, Sicca, & Muratori, 2013; Weigelt et al., 2012). In contrast to ToM, CC theory helps account for the visual aspect of FER, i.e. CC would explain FER improvements due to learning compensatory strategies via visual cues. CC does not, however, provide an all-encompassing account of FER because it does not explain convergent literature of FER ability in ASD. In other words, if the CC theory were correct, all individuals with ASD would have a clear FER impairment, even with static photographs of facial expressions, which is not the case (Harms et al., 2010). Hence, both of these prominent theories of cognition in ASD are useful in understanding FER in this population, however, neither provide a comprehensive account. The implications of the results presented in this thesis for theoretical approaches understanding ASD will be discussed further in Chapter 5.

## **1.7 Aims of the thesis**

The aim of this thesis is to assess the extent to which FER is atypical in individuals with ASD as compared to TD individuals. This will be assessed in adults, children and in the typical population (measured using autism-like traits). A secondary aim of this thesis is to assess the contribution of alexithymia to FER. The research presented in this thesis is, primarily, exploratory in nature, however some more specific hypotheses are presented at the beginning of each experimental chapter.

## **1.8 Overview of chapters**

In Chapter 2, a novel, psychophysics measure of FER is introduced. Then, results from a research experiment using this novel technique in adults with ASD is presented. The purpose of this study is to investigate the subtleties of FER that may have been overlooked in previous ASD research. Namely, this study assesses: individual differences in FER (as allowed by this particular method), the specific pattern of facial muscle components (action units) utilised from the stimuli, the intensity of stimuli favoured by individuals with ASD and the relative timing of the action units (AUs). The impact that these factors have on FER ability between ASD and TD groups is then discussed.

Chapter 3 presents a novel technique, similar to that presented in Chapter 2, now applied to *children* with ASD in order to ascertain FER abilities from participants at a younger age as

compared to age and IQ matched controls. The aim of the study presented in this chapter is to model FER in children with ASD in detail, by assessing the use of AUs, intensity and timing of facial expression stimuli that were utilised to categorise emotions. The data presented in chapter 3 are critiqued in relation to current literature surrounding FER in ASD, as well as compared to results reported in Chapter 2.

The aim of Chapter 4 is to assess FER in relation to alexithymia and autism-like personality traits in the typical population. This adds to the current debate in the literature that FER may be better defined in terms of co-occurring alexithymia (Bird & Cook, 2013). This also contributes to research assessing the extent to which FER atypicalities extend *beyond* clinically diagnosed ASD, into autism-like personality traits in the typical population.

Chapter 5 brings the aforementioned chapters together, in an evaluation of the results in the context of previous research and theoretical conceptualizations of FER in ASD. The implications of these findings are then reviewed, followed by a discussion of the limitations of the research presented in this thesis and projections for future research.

## Chapter 2 Facial Expression Recognition in Adults with ASD

### 2.1 Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder that is characterised by difficulties in social interaction and social communication (*DSM-5*, 2013). Because of these particular characteristics, researchers have been interested in how individuals with ASD perceive facial expressions of emotion. This interest is evident in the large number of research papers published on the topic, spanning over thirty years (Harms et al., 2010). Hence, because facial expressions are deemed intrinsic to effective and meaningful social communication, researchers have long hypothesised that difficulty reading and understanding another person's facial expression may be integral to the social difficulties experienced by people with ASD. Alas, despite the logical and somewhat 'common sense' nature of this hypothesis, research evidence indicates an impairment in FER in individuals with ASD in some instances, and no impairment in FER others (Harms et al., 2010; Jemel et al., 2006; Simmons et al., 2009). Hence, this hypothesis may be a much more complicated story than first assumed. See section 1.4 for a full review of the FER literature in ASD.

Researchers have suggested that, among other factors, limitations in the methods used to measure FER are likely to contribute to the confusing results in the ASD FER literature (Uljarevic & Hamilton, 2013). For example, many studies use static (photograph) images of facial expressions to measure FER in an ASD population. While static facial expression images serve well to assess a) the ability to recognise an emotion at its peak 'amplitude' (i.e. peak expression of the emotion) and b) the ability to recognise an emotion without the interference of timing, they do not allow for measurement beyond these boundaries. Having said this, there are discernible reasons for using these methodologies. Firstly, it is very difficult to assess 'real life' facial expression recognition in a quantitative manner and secondly, there are methodological difficulties which are incidental to conducting research with individuals with ASD (as discussed in Section 1.2). There is evidence, however, that measuring FER using dynamic stimuli (as apposed to unmodified, static images) reveals more pronounced FER difficulty in participants with ASD (Harms et al., 2010). For example, Sato et al. (2013) reported that individuals with ASD found slow presentations of the 6 basic emotions to be 'natural', whereas typically developed individuals found faster presentation of emotion stimuli to be 'natural'. From this, they suggested that the dynamics

of facial expressions are important for accurate FER in ‘real life’ and hence, the lack of dynamic stimuli in the literature is partly to blame for the heterogeneous findings. Law Smith et al. (2010) found that when facial expression content was manipulated using a series of different morphs (i.e. stimuli were displayed from 0% to 20% emotion; 0% to 30% emotion and so on until 0% to 100% emotion), participants with ASD were less accurate at labelling the emotions than TD individuals. This suggests that the emotional intensity of a facial expression may be an important predictor of FER ability in participants with ASD. Law Smith et al. (2010) did, however, find this difference only in surprise, disgust and anger, out of the 6 basic emotions. Enticott et al. (2014), on the other hand, argue against the idea that FER difficulties would become more distinct in individuals with ASD when presented with dynamic stimuli. Instead, they reported that accurate recognition of some emotions improved when the stimuli were dynamic as compared to static. Hence, the present study aims to address the contribution of dynamics, as well as the basic muscle components of the face, to FER ability in individuals with ASD. In addition to this, the contribution of factors such as amplitude and timing are assessed in order to ascertain if these components help to better explain FER in adults with ASD.

Researchers have also suggested that poorer FER ability in individuals with ASD can be explained by a particular difficulty with the eye region of the face (Simmons et al., 2009). For example, studies have suggested that individual with ASD find it difficult to accurately categorise facial expressions when the eye region only is presented (S. Baron-Cohen, Wheelwright, Hill, et al., 2001; S. Baron-Cohen et al., 1997). However, a study by Leung, Ordqvist, Falkmer, Parsons, and Falkmer (2013) found that adults with ASD used the eye information just as much as the typical adults when asked to work out the correct facial expression based on stimuli that depicted fragments of the face. Hence, there is still confusion about whether perception of specific parts of the face influences FER ability in individuals with ASD and so the present study assesses the relative contribution of the eye and mouth regions.

In addition to methodological limitations, the heterogeneity of research findings in FER in ASD research has also been explained by the inconsistency of individual participant variables. For example, variables which have been indicated as contributing to FER ability include levels of autism-severity (B. T. Williams & Gray, 2013), levels of autism-like traits (Poljac et al., 2013) and alexithymia, a condition of ‘emotion-blindness’, which affects typically developing individuals as well as those with ASD (Bird & Cook, 2013; Cook et al., 2013). Therefore this study sought to be mindful of individual differences within ASD

and TD groups when assessing FER ability by examining data not only at a group level but also at an individual level.

This study utilises dynamic stimuli that can be manipulated on several levels: specific muscles of the face, amplitude and timing. Because of this, a rich tapestry of data is formed for each individual, providing a complex and comprehensive portrait of the action units (AUs; i.e. the muscle movements on the face), the amplitude and the timing components that are meaningful to each participant, represented in an individual ‘model’. Hence, the major aim of this study was to apply this novel technique in order to assess FER ability in adults with ASD. Although this study is therefore somewhat exploratory in nature, the following hypotheses can be drawn from the literature: 1. The pattern of AUs in the ASD models will be atypical compared to the TD group. 2. The amplitude of the AUs will be higher in the ASD group 3. The timing of the AUs will be atypical in the ASD group and 4. The ASD group will use information from the mouth area of the face in favour of information from the eyes.

## **2.2 Method**

### **2.2.1 Instruments for characterising participants**

#### **2.2.1.1 Autism Spectrum Quotient (AQ)**

The Autism Spectrum Quotient (AQ; S. Baron-Cohen, Wheelwright, Skinner, et al. (2001)) is a 50-item, self-report questionnaire that probes information about the level of autism-like traits an individual has. Questions in the AQ target five subscales, which are thought to be characteristic of an ASD diagnosis: 1) social skills 2) attention switching 3) attention to detail 4) communication and 5) imagination. Responses are given in terms of a four-point Likert scale (‘definitely agree’, ‘slightly agree’, ‘slightly disagree’ or ‘definitely disagree’) to statements which target the above 5 ASD characteristics. For example, statements include: “I enjoy social chit-chat.” (social skills; reverse score), “I usually notice car number plates or similar strings of information” (attention to detail; regular score), “When I’m reading a story, I can easily imagine what the characters might look like.” (imagination, reverse score). A response of ‘definitely agree’ or ‘slightly agree’ was scored as a ‘1’ and ‘slightly disagree’ or ‘definitely disagree’ as a ‘0’, and the opposite for reverse scored items. A total score on the AQ consists of a sum of all 50 items (i.e. with some questions reversed scored), hence, higher scores indicate higher levels of autism-like

traits. This follows the typical scoring technique for this questionnaire (S. Baron-Cohen, Wheelwright, Skinner, et al., 2001).

### **2.2.1.2 The twenty-item Toronto Alexithymia Scale (TAS-20)**

The 20-item Toronto Alexithymia Scale (TAS-20; (Bagby, Parker, et al., 1994; Bagby, Taylor, & Parker, 1994)) is a self-report questionnaire that investigates an individual's ability to understand their own, and other people's emotions. Having alexithymia is sometimes referred to colloquially as having 'emotion-blindness'. The TAS-20 has 3 subscales: 'describing feelings and emotions', 'identifying feelings and emotions' and 'externally-oriented thinking'. Items are statements regarding the above 3 subscales, for example "I am able to describe my feelings easily" (describing feelings and emotions, reverse score), "When I am upset, I don't know if I am sad, frightened, or angry" (identifying feelings and emotions, regular score), "Looking for hidden meanings in movies or plays distracts from their enjoyment." (externally-oriented thinking, regular score). Participants responded to the statements using a 5-point Likert scale: 1 = 'Strongly Disagree', 2 = 'Disagree', 3 = 'Neither Disagree or Agree', 4 = 'Agree' and 5 = 'Strongly Agree'. For each participant all 20 items are summed (note, 5 items are reverse scored in which their responses are scored as follows: 'Strongly Disagree' = 5, 'Disagree' = 4, 'Agree' = 2 and 'Strongly Agree' = 1). The sum of the items indicates the participant's total TAS score, and therefore, higher TAS scores indicate higher levels of alexithymia.

### **2.2.2 Participants**

Eleven adults with ASD and eleven typically developed adults participated in the study. Participants were recruited through the California Institute of Technology (Caltech) and the University of Glasgow. Eleven adults with ASD and seven TD adults were recruited at Caltech: the remainder of the TD adults (n = 4) were recruited at the University of Glasgow (see limitation section for discussion of this recruitment technique). Typically developed participants were matched on age, verbal IQ (VIQ), performance IQ (PIQ) and full-scale IQ (FSIQ) at a group level (see section **Error! Reference source not found.**). All participants in the study were educated to a high school or greater level. Participants with ASD underwent ADOS-II (Module IV) and MINI (Mini-International Neuropsychiatric Interview) administration from a trained practitioner at California Institute of Technology. Only individuals who met diagnostic threshold in ADOS-II and MINI were admitted to the experiment. Participants were also administered the WASI (Wechsler Abbreviated Scale of Intelligence), AQ (Autism Quotient Test; S. Baron-Cohen,

Wheelwright, Skinner, et al. (2001) and TAS-20 (Toronto Alexithymia Scale- 20; Bagby, Taylor, et al. (1994). In order to take part in the study all participants were also required to have normal or corrected-to-normal vision. This was self-reported at the stage of signing up for the experiment. Participants were compensated for their participation at the rate of £6 per hour in UK and \$20 per hour in USA. These payment levels reflect the standard participant payment for their respective institutions. All participants gave written, informed consent to participate and the study met ethical approval at both the University of Glasgow and the California Institute of Technology.

### **2.2.2.1 Demographics and descriptive Information**

Demographic and descriptive information for all participants can be found in Table 2-1, Table 2-2 and Table 2-3. Independent samples t-tests indicated no significant difference between the two groups in age ( $t(20) = 0.37, p = 0.71$ ), verbal IQ (VIQ;  $t(20) = 0.44, p = 0.67$ ), performance IQ (PIQ;  $t(20) = 0.75, p = 0.46$ ) and full-scale IQ (FSIQ;  $t(20) = 0.98, p = 0.34$ ). The ASD group was found to be significantly different from the TD group on AQ score ( $t(20) = 6.25, p < 0.01$ ) and also on TAS-20 score ( $t(20) = 2.97, p < 0.01$ ). Due to tight restrictions on matching IQ and age, difficulties were met when trying to meet the high number of male ASD participants. Hence, slight differences are found in gender ratio between groups. Gender ratio was, however, not significantly different between the groups  $\chi^2(1, N = 22) = 0.78, p = 0.62$ .

Table 2-1 Demographics of ASD and TD participants (n = 22) \*p < 0.01

	ASD (n = 11)			TD (n = 11)			
	Count	Percentage		Count	Percentage		
<b>Gender</b>	8 male 3 female	73% 27%		6 male 5 female	55% 45%		$\chi^2 (1, N = 22) = 0.78, p = 0.62$
	Mean	SD	Range	Mean	SD	Range	T-test
<b>Age (years: months)</b>	31.5	12.4	20-60	33.4	10.4	20 - 53	t(20) = 0.37, p = 0.71
<b>Verbal IQ</b>	102.6	22.7	50-131	106.0	11.5	85-119	t (20) = 0.44, p = 0.67
<b>Performance IQ</b>	110.9	11.3	99-128	114.4	10.4	102-133	t(20) = 0.75, p = 0.46
<b>Full Scale IQ</b>	108.5	13.7	91-126	111.5	8.1	97-122	t(20) = 0.98, p = 0.34
<b>AQ</b>	29.0	7	17-38	13.3	4.5	4 - 19	t (20) = 6.25, p < 0.01
<b>TAS-20</b>	55	8.2	44 - 63	45	6.5	36 - 56	t(20) = 2.97, p < 0.01

Note, the AQ is a measure of autism-like traits where higher scores indicate higher levels of traits. Scores on the AQ can be between 0 and 50, where scores over 32 have been suggested to be a useful cut-off point, at which a clinical diagnosis of ASD is likely (S. Baron-Cohen, Wheelwright, Skinner, et al., 2001). The TAS is scored between 20 and 100, where low scores indicate low levels of alexithymia and high score indicate high levels of alexithymia. According to Bagby, et al., (1994) a score on the TAS-20 that is equal to or less than 51 indicates ‘no alexithymia’, 52 to 60 indicates ‘possible alexithymia’ and scores equal to or greater than 61 indicates ‘alexithymia’.

**Table 2-2 Autism Diagnostic Observation Schedule (2<sup>nd</sup> Edition) Score for ASD Participants**

	<b>Module</b>	<b>ADOS A</b>	<b>ADOS B</b>	<b>ADOS C</b>	<b>ADOS D</b>	<b>ADOS Total</b>
ASD 1	4	3	6	0	1	14
ASD 2	4	2	6	1	4	17
ASD 3	4	4	11	2	0	21
ASD 4	4	2	7	0	2	15
ASD 5	4	5	11	0	1	21
ASD 6	4	4	9	1	0	18
ASD 7	4	6	11	1	0	22
ASD 8	4	7	13	1	3	28
ASD 9	4	7	14	1	3	29
ASD 10	4	5	10	2	0	21
ASD 11	4	3	7	1	0	15

Note, here that ‘module 4’ of the ADOS-2 is specifically designed for adults or adolescents that have fluent speech. ADOS A refers to scores on ‘Communication’ subset, ADOS B ‘Reciprocal Social Interaction’ subset, ADOS C ‘Imagination and Creativity’ and ADOS D ‘Stereotyped Behaviours and Restricted Interests’. All eleven participants who were assessed on the ADOS-2 met cut-off criteria for an autism spectrum classification.

**Table 2-3 Autism Spectrum Quotient Score (AQ) and Toronto Alexithymia Scale Score (TAS)**

	<b>AQ score</b>	<b>TAS score</b>	<b>TAS Cut-Off</b>
ASD 1	29	52	Possible Alexithymia
ASD 2	20	70	Alexithymia
ASD 3	31	60	Possible Alexithymia
ASD 4	37	63	Alexithymia
ASD 5	17	46	No Alexithymia
ASD 6	34	58	Possible Alexithymia
ASD 7	30	57	Possible Alexithymia
ASD 8	21	46	No Alexithymia
ASD 9	28	48	No Alexithymia
ASD 10	38	44	No Alexithymia
ASD 11	34	55	Possible Alexithymia
<hr/>			
TD 1	14	56	Possible Alexithymia
TD 2	11	55	Possible Alexithymia
TD 3	9	50	No Alexithymia
TD 4	16	48	No Alexithymia
TD 5	18	45	No Alexithymia
TD 6	14	44	No Alexithymia
TD 7	10	38	No Alexithymia
TD 8	18	42	No Alexithymia
TD 9	19	41	No Alexithymia
TD 10	13	41	No Alexithymia
TD 11	4	36	No Alexithymia

### **2.2.3 Stimulus generation**

Stimuli used in this study were developed by Yu, Garrod and Schyns (2012) following several steps:

#### **2.2.3.1 Facial Action Unit Coding System (FACS)**

Firstly, four actors who were trained in the Facial Action Unit Coding System (FACS; Ekman and Friesen (1978)) were recruited to pose as ‘actors’ for the development of the stimuli. FACS is a taxonomy of the muscle groups found in the face that create facial movements. These muscle groups are defined as ‘action units’ (AUs). For example, AU 12 codes for the muscle group ‘zygomaticus major’, which is defined as ‘lip corner puller’ in

FACS and appears as a smile. Likewise, AU 2L codes for activation of left ‘frontalis, pars lateralis’, which is defined as ‘left outer brow raiser’ in FACS and appears as the lifting of the left eyebrow. FACS consists of approximately 40 AUs along with several non-facial movements such as eye and head movements. Actors trained in FACS movements had undergone training to enable them to move each AU on their face (as specified in FACS) independently. The four actors were recorded making each AU movement in a 4D stereo imaging system (DI4D Facial Motion Capture System; Dimensional Imaging Limited, Glasgow) located in the Institute of Neuroscience at University of Glasgow. The 4D stereo imaging system captured videos of actors making facial movements in a 3-dimensional space (the term ‘4D’ here indicates a 3D image across time). Hence, each AU created on the actor’s face was recorded as a motion 3D picture. This resulted in four templates of all FACS AUs in 4D format. These templates were averaged to form one ‘basic’ template of the FACS AUs, thus providing a model for all the muscle groups that can be active in the face for any given facial expression of emotion. Stimuli used in this study were developed by the Schyns laboratory (Yu et al., 2012) and adapted for use in the current study.

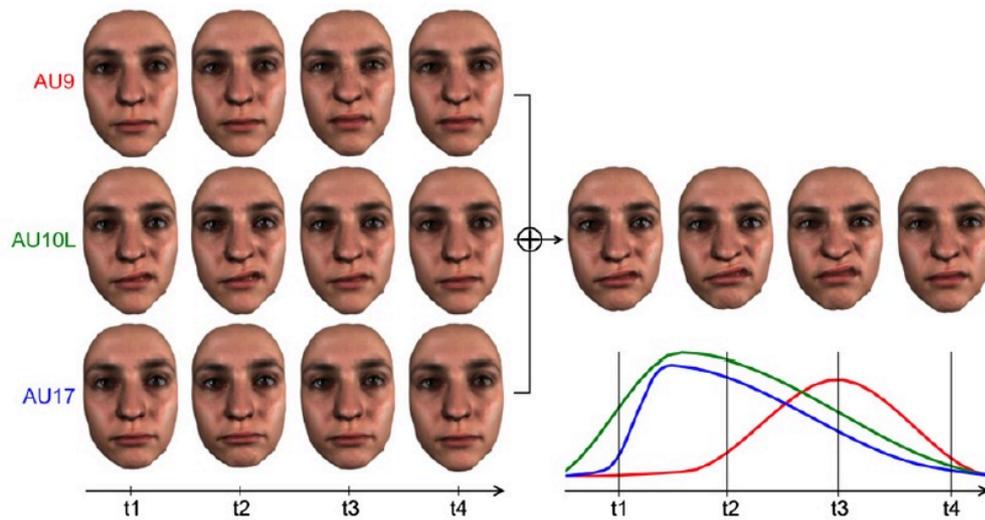
### **2.2.3.2 Synthesis of facial expression stimuli**

The basic template produced from the actor data (described above) was then applied to other faces captured in the 4D stereo imaging system. The actor template provided a 4D map of AUs in the face, consequently any face that was captured in the 4D stereo imaging system could be manipulated using this map. For example, if an individual has their face captured in the 4D stereo imaging system, although the resulting image is of their *neutral* face, the image can be manipulated to produce any of the FACS AUs and any combination of these AUs. In addition to this, the timing and intensity of the combination of AUs that are produced in the image can also be varied across seven parameters: action unit, peak amplitude, peak latency, onset latency, offset latency, acceleration, deceleration. Therefore, this stimulus generation technique is unparalleled in the control that can be exerted over the resultant stimulus.

### **2.2.3.3 Generating random facial expressions**

For the purposes of this experiment random facial expression stimuli were used. Figure 2–1 illustrates the generation of a single stimulus, where a combination of AUs (in this example AU 9, AU 10L and AU 17), each with their own timing and intensity trajectory, were combined to create a single, ‘random’, facial expression. A dynamic figure of how the stimuli are generated can be found in Appendix 1. The resulting facial expression that

is created is essentially ‘random’ in that it combines an arbitrary set of AUs and timing/intensity parameters, however a ‘Generative Face Grammar’ (GFG) was applied in order to avoid producing facial expressions that do not ‘make sense’. The GFG is an established concept (Edelman, 2003) but was applied to the generation of these stimuli by Yu et al. (2012) and serves to ensure that the stimuli that were generated were constrained in such a way that they produced only physiologically plausible facial expressions. In the present study, 2400 different, random facial expressions, produced across 25 male and 25 female western Caucasian facial identities, were used. These same stimuli have been used in previous experiments such as Jack, Garrod, Yu, Caldara, and Schyns (2012), Jack, Caldara, et al. (2012a), Jack et al. (2014), Gill, Garrod, Jack, and Schyns (2014), Richoz, Jack, Garrod, Schyns, and Caldara (2015).



**Figure 2–1. Random facial expression synthesis. Figure taken from Yu, Garrod and Schyns (2012) p153.**

## 2.2.4 Procedure

A total of 52,800 stimuli were categorised into emotion (happy, surprise, fear, disgust, anger, sadness, other) and intensity (very strong, strong, medium, weak, very weak) by 22 participants (11 ASD; 11TD). A step-by- step run through of the experiment is described below:

Step 1. Participants were briefed about the experiment and asked to complete a consent form. They logged into the experiment with help from the lab assistant and viewed the following instructions. Please note, these instructions were given in addition to an information sheet and verbal direction from the experimenter.

1. Watch the facial animation – it will only be played once.
2. Select the emotion you think the facial animation represents- choices are happy, surprise, fear, disgust, anger, sadness or other.

PLEASE NOTE: It is important that you are confident with your answer: If you think the face contains no emotion or you cannot accurately describe the facial animation using the available options, please click ‘other’.

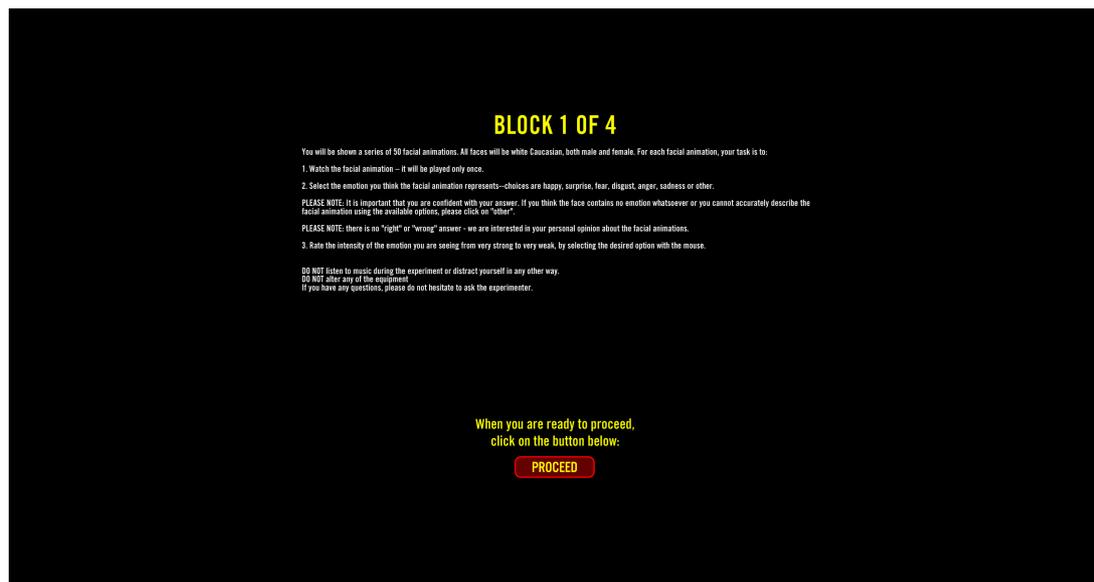
PLEASE NOTE: there is no ‘right’ or ‘wrong’ answer- we are interested in your personal opinion about the facial animation

3. Rate the intensity of the emotion you are seeing from very strong to very weak by selecting the desired option with the mouse.

DO NOT listen to music during the experiment or distract yourself in any other way.

DO NOT alter the equipment

If you have any questions, please do not hesitate to ask the experimenter.



**Figure 2–2. Online instructions at the beginning of the experiment. Please note these instructions were given in addition to an information sheet and verbal direction from the experimenter.**

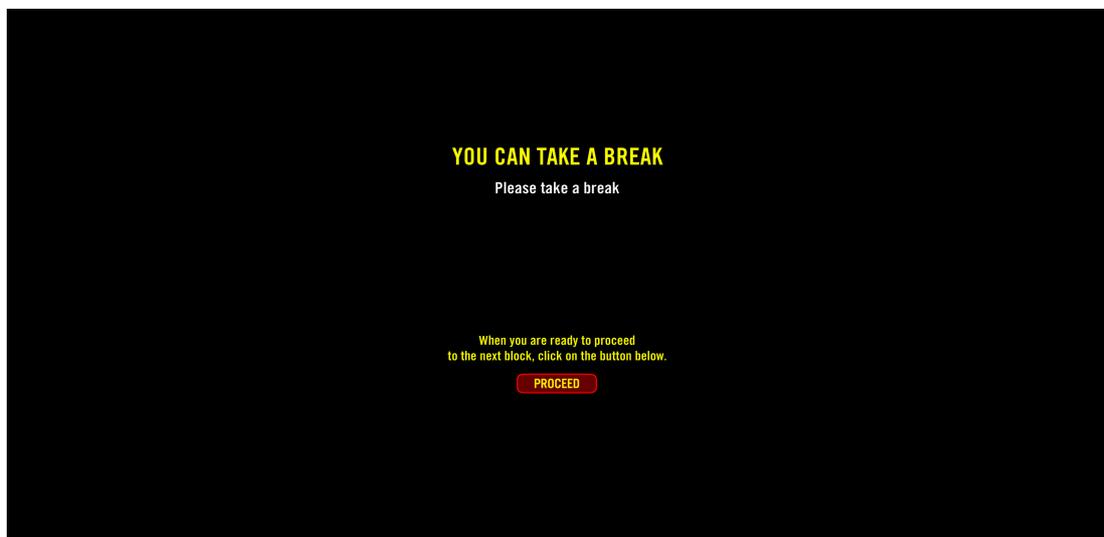
Step 2. Participants viewed 2400 randomised stimuli each, separated over 4-6 different sessions. Stimuli were presented on a computer monitor at a viewing distance of 60cm. Participants were given instructions to sit still during testing. Although no head restraint or chin-rest was used, participants were instructed not to move their head during testing. Each stimulus video was displayed in the central viewing field for 1.25 seconds and remained on screen until the observer responded. Response options appeared on the right hand side of the screen after the stimulus video finished playing. After 50 trials participants were offered a break and were reminded of the task instructions upon beginning the next set of stimuli. Participants completed the first session (around 600 trials) in the lab and the rest of the sessions were completed at home. Participants were given full instructions on how to conduct the experiment at home. This included set up instructions (including strict instructions for viewing the experiment at home, as described above) and contact information for the experimenters i.e. Caltech lab assistant (Catherine Holcomb) and University of Glasgow researcher (Kirsty Ainsworth). The experimenters were on-hand via email to help the participant if they encountered any issues. All participants reported that they followed the instructions accurately and had encountered no problems with the experiment. The following screenshots depict a visualization of this step:



**Figure 2–3. Snapshot of dynamic stimulus as viewed by the participants, which was displayed in the central viewing field for 1.25 seconds (snapshot taken at the middle of the time course). Participants viewed the stimuli 60cm away from the screen.**



**Figure 2–4. Snapshot of the categorization options as viewed by the participant. The stimulus (neutral) stayed on the screen until the participant selected their option.**



**Figure 2–5. Snapshot of the break instructions as viewed by the participant after periods of 50 stimuli.**

Step 3. On completion of the experiment participants were invited back into the lab where they were debriefed and paid for their time (£6/\$20 per hour). Participants were given an opportunity to ask any questions about the experiment and they were thanked for their time.

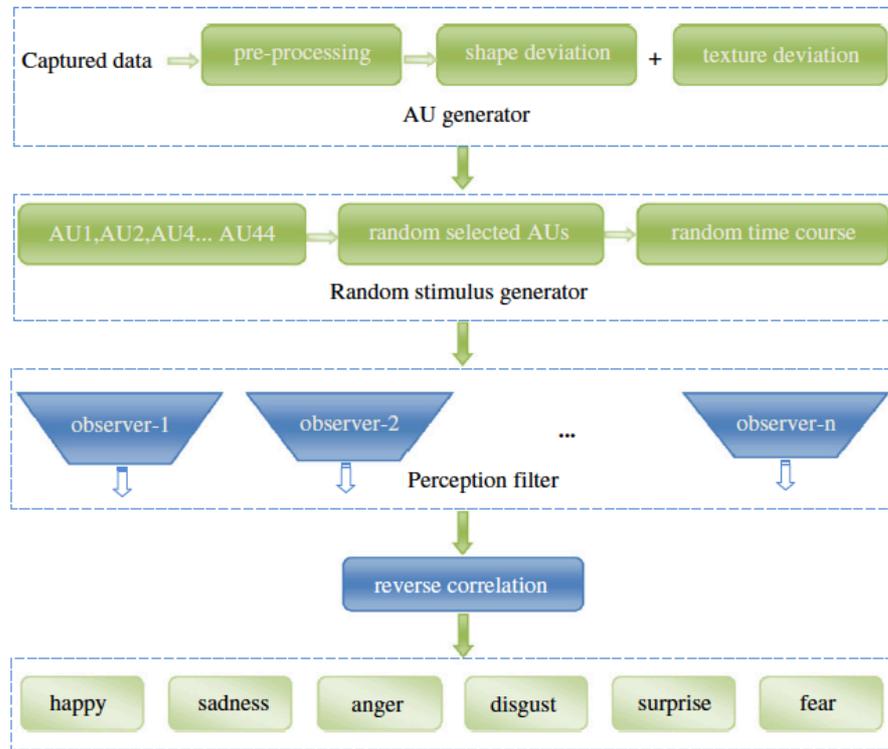
### 2.2.5 Analysis

Reverse correlation analysis is a technique that reveals a precise ‘model’ of the information used to categorise the stimuli i.e. the signals within a stimulus which are considered most indicative of fitting a given categorization label. This technique was developed by

Ahumada and Lovell (1971) who investigated the acoustic properties of auditory stimuli by asking participants to categorise a 500-Hz signal tone bursts as either present or absent. Reverse correlation has been applied to various areas of low-level vision such as motion perception (Borghuis et al., 2003), surface recognition (Gosselin, Bacon, & Mamassian, 2004) and depth perception (Neri, Parker, & Blakemore, 1999). The technique was first applied to face perception by Gosselin and Schyns (2001) using a stimulus classification task termed ‘bubbles’. In this task participants were presented with images of faces that were covered with a mask so that only partial areas of the image could be seen. The participants categorised two identities via thousands of images that showed only partial information. From this, reverse correlation analysis revealed which information was continually used (i.e. could be seen through the mask) to categorise each identity. Therefore, information that was useful for identity categorization was revealed for each participant. Reverse correlation has been applied to several subsequent vision tasks that have specifically assessed visual processing of facial expressions of emotion: Uddenberg and Shim (2015); Das et al. (2013); Ethier-Majcher, Joubert, and Gosselin (2013); Kontsevich and Tyler (2004) and facial expression perception in ASD (Adolphs et al., 2005; Adolphs et al., 2008; Blais, Roy, Fiset, Arguin, & Gosselin, 2012; Song, Kawabe, Hakoda, & Du, 2012; Spezio, Adolphs, Hurley, & Piven, 2007a; Spezio et al., 2007b). A study using identical stimuli and a similar paradigm to the present study can be found in Jack, Garrod, et al. (2012).

As depicted in Figure 2–6, the present study applied reverse correlation to the raw data. The raw data were the output from the categorization labels of emotion (happy, surprise, fear, disgust, anger, sadness, other) and intensity (very strong, strong, medium, weak, very weak) given to 2400 arbitrary stimuli per participant. A linear model (using linear multiple regression and Pearson’s correlation) was fitted to each AU parameter and to the intensity ratings for those trials. The end product is a template for each participant, which reflects the subjective diagnostic information used in their categorizations, termed hence forth as their personal ‘model’ of each emotion. For example, to gain a participant’s model for ‘happy’, a Pearson’s correlation would be performed between the AUs that were ‘on’ in the stimuli viewed by the participant, and the stimuli that were categorised by that participant as ‘happy’, over all 2400 trials. Then, a regression would be performed between the other AU parameters (peak amplitude, peak latency, onset latency, offset latency, acceleration and deceleration) and emotional intensity rating for the trials where the AU was ‘on’ and the emotion was categorised as ‘happy’. This would reveal the spatial (action

units) and temporal components (starting/ending points and trajectory) of the stimuli, which, by their activation, led to each emotional response category.

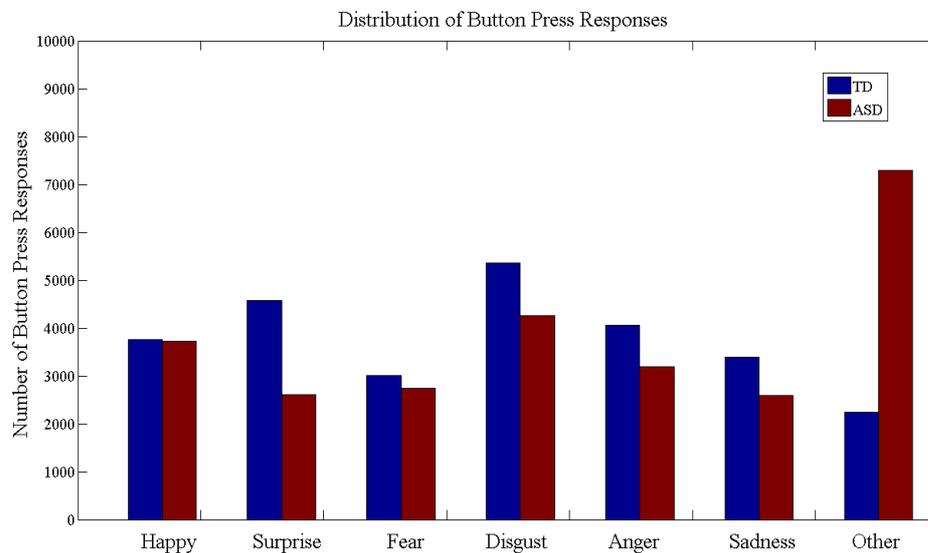


**Figure 2–6. Experimental steps from stimuli generation to reverse correlation analysis (Image source: Yu, Garrod and Schyns, 2012, p. 154)**

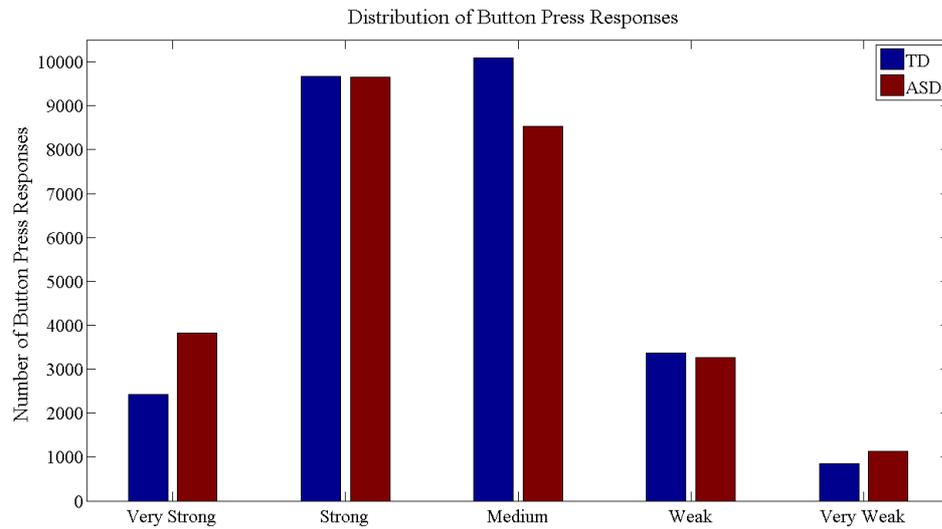
## 2.3 Results

### 2.3.1 Response patterns

In the first instance, the pattern of responses (i.e. button presses) for each group were assessed in order to establish an initial picture of group differences in the task. The distribution of responses was compared between ASD and TD groups to assess any indication of a preference for one particular emotion (see Figure 2–7) or for one particular intensity level (Figure 2–8). These figures indicate a preference in the ASD group to select ‘other’ while the TD group are more evenly distributed across all emotion categories. The ASD group also indicated more ‘very strong’ and ‘very weak’ responses, while the TD group indicated more ‘medium’ responses.



**Figure 2–7 Distribution of responses (emotion button presses) to stimuli (2400 trials per participant)**



**Figure 2–8 Distribution of responses (intensity button presses) to stimuli (2400 trials per participant)**

In order to further assess statistical differences in the response patterns outlined above, the ASD and TD group data were compared using independent samples t-tests. To assess differences in emotion response, data were collapsed over intensity, i.e. responses were summed for each emotion regardless of responding using the ‘very high’, ‘high’, medium’, ‘low’ or ‘very low’ response option (as depicted in Figure 2–7). Results revealed the ASD group responded ‘other’ significantly more compared to the TD group ( $t(20) = 3.3, p < 0.01$ ) and the TD group responded ‘surprise’ significantly more than the ASD group ( $t(20) = 3.77, p < 0.01$ ). However, no significant difference was found between groups for the remaining emotions: Happy ( $t(20) = 0.04, p = 0.97$ ), Fear ( $t(20) = 0.40, p = 0.69$ ), Disgust ( $t(20) = 1.31, p = 0.20$ ), Anger ( $t(20) = 1.20, p = 0.24$ ), Sadness ( $t(20) = 1.39, p = 0.18$ ). Similarly, a between-groups analysis was conducted for intensity only (i.e. collapsing the data across emotions and hence using the frequency of each intensity category only (as depicted in Figure 2–8). No significant differences were found between groups in the categorization of intensity: V Strong ( $t(20) = 1.81, p = 0.09$ ), Strong ( $t(20) = 0.02, p = 0.99$ ), Medium ( $t(20) = 1.35, p = 0.19$ ), Weak ( $t(20) = 0.12, p = 0.90$ ) and Very Weak ( $t(20) = 0.42, p = 0.67$ ).

To assess individual response patterns, and hence assess the dataset for ‘atypical’ response patterns, the distribution of button presses was examined per individual. As described in the methods procedure (section 2.2.4), each participant was required to categorise 2400 facial stimuli using seven possible emotion labels (‘happy’, ‘surprise’, ‘fear’, ‘disgust’, ‘anger’, ‘sadness’, ‘other’) and five possible intensity labels (‘very strong’, ‘strong’,

‘medium’, ‘weak’, ‘very weak’). Hence, figures depict the number of responses for each possible emotion response (from left to right: ‘happy’, ‘surprise’, ‘fear’, ‘disgust’, ‘anger’, ‘sadness’, ‘other’) and each possible intensity response (from left to right: ‘very strong’, ‘strong’, ‘medium’, ‘weak’, ‘very weak’). TD participants responded in an overall similar pattern, responding in a mostly even distribution of emotion and intensity (see Figure 2–9, Figure 2–10 and Figure 2–11). ASD participants, however, responded slightly more atypically, with some individuals showing greater bias to selecting ‘other’ (see Figure 2–12, Figure 2–13 and Figure 2–14). Taken together, these data suggest that the ASD group (or at least a proportion of participants in this group) have responded to the stimuli in an atypical fashion. In addition to appearing to have the most atypical response patterns, ASD participants ‘8’ and ‘9’ also have the highest autism severity levels out of the ASD group (as shown in the ADOS scores in section 2.2.2.1).

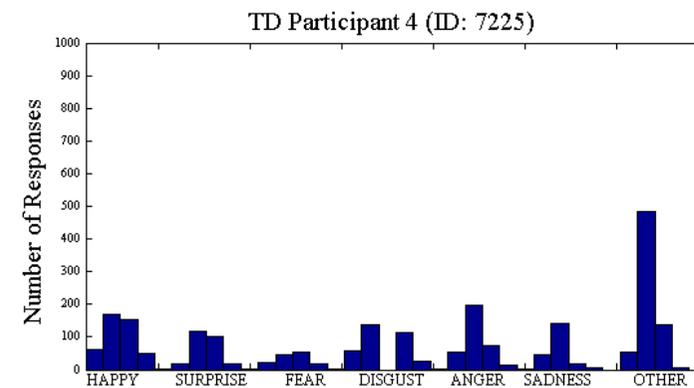
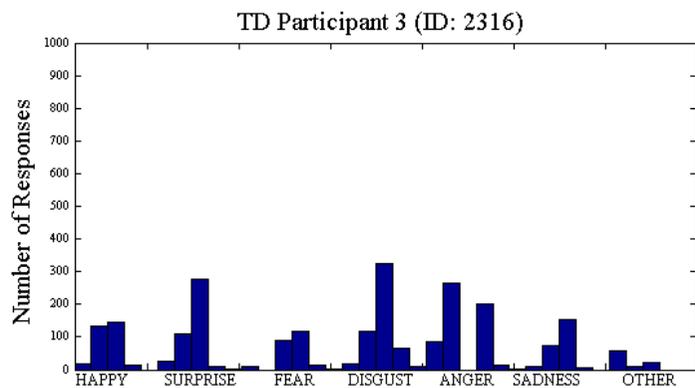
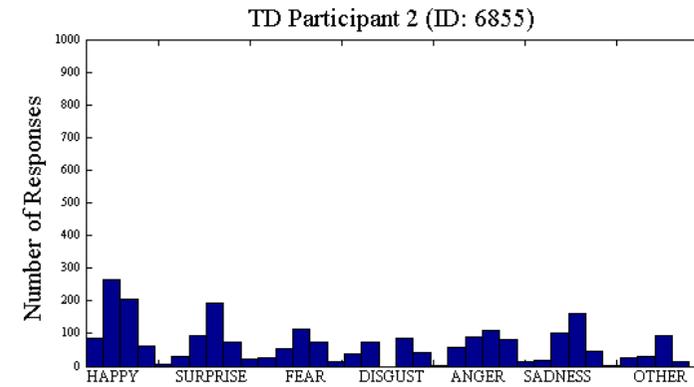
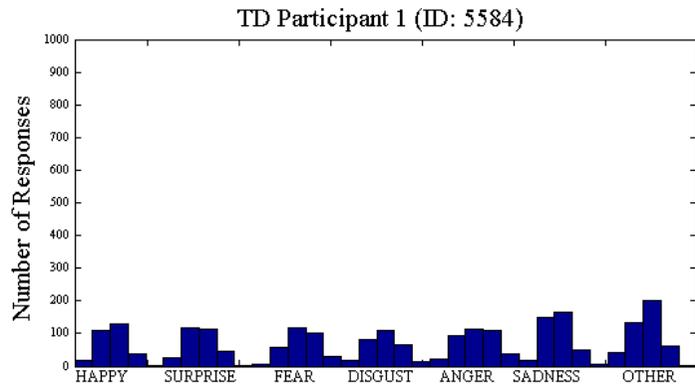


Figure 2-9 TD Group Part 1: Frequency of stimuli responses per TD participant for emotion and intensity (2400 trials per participant).



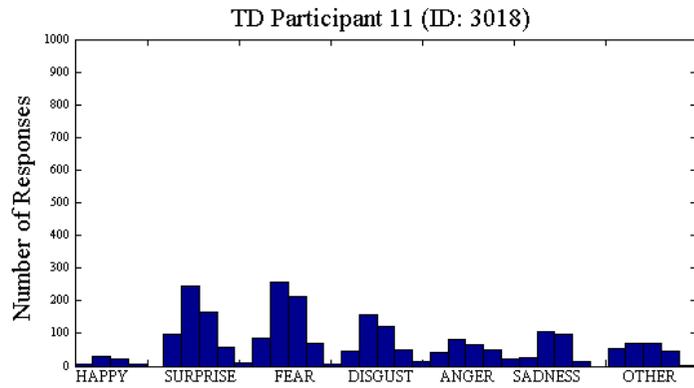
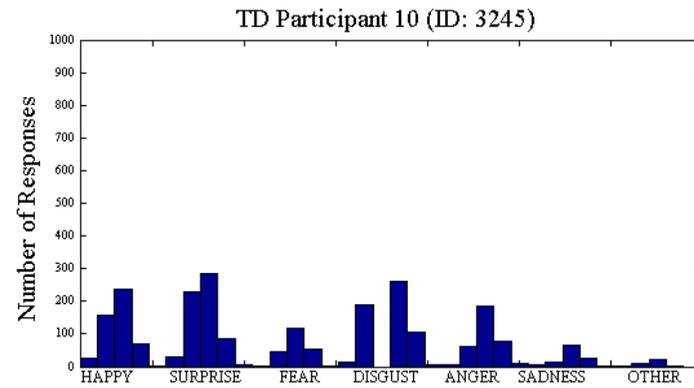
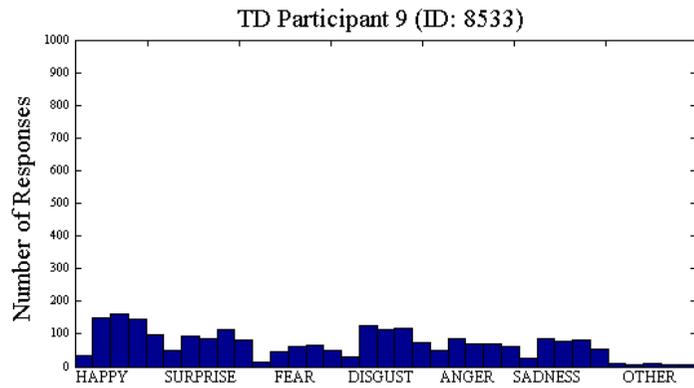


Figure 2–11 TD Group Part 3: Frequency of stimuli responses per TD participant for emotion and intensity (2400 trials per participant).

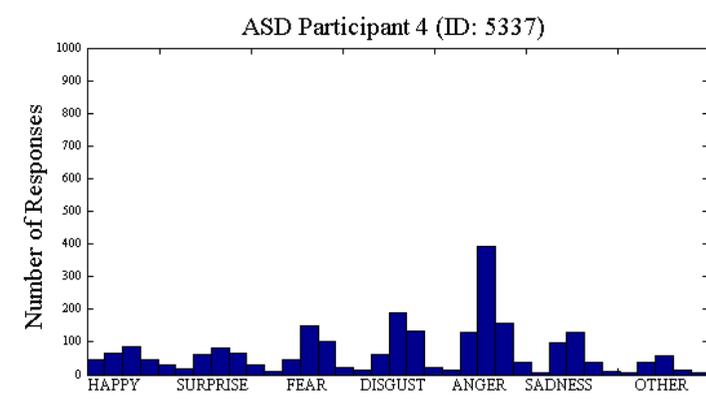
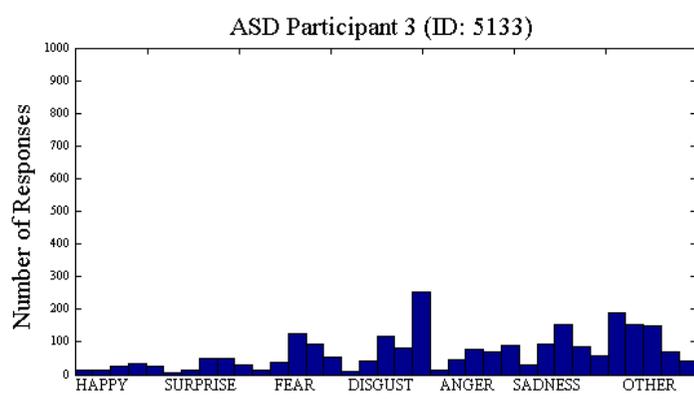
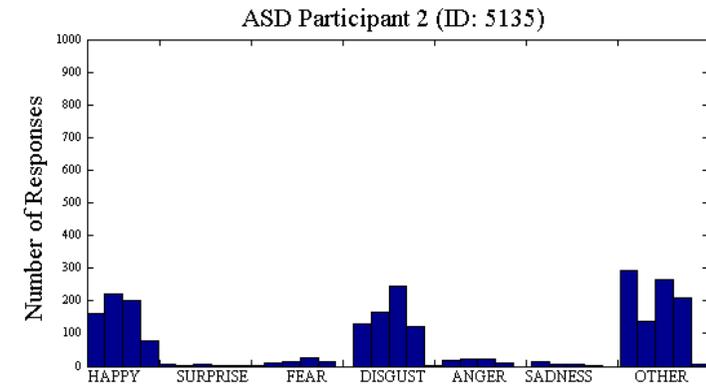
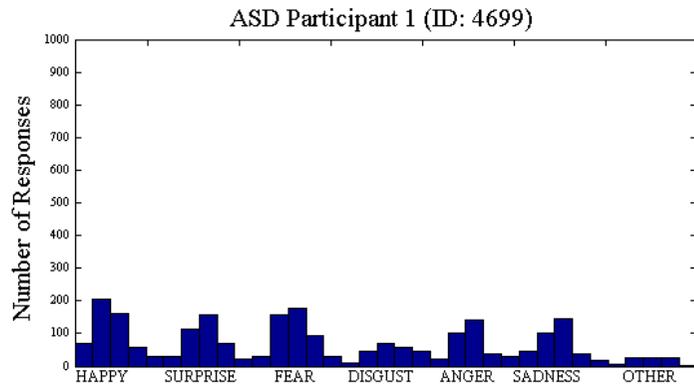


Figure 2-12 ASD Group Part 1: Frequency of stimuli responses per ASD participant for emotion and intensity (2400 trials per participant).

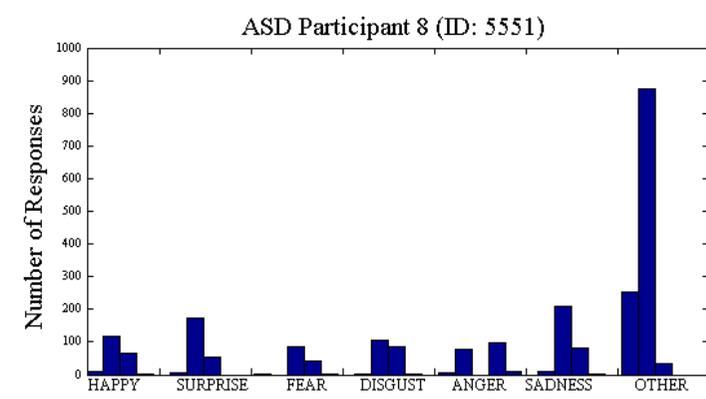
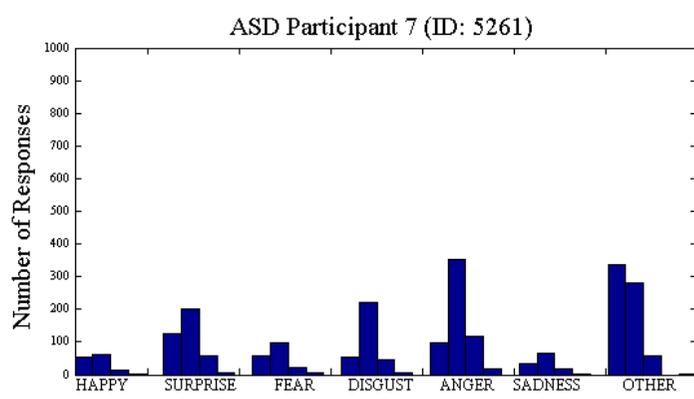
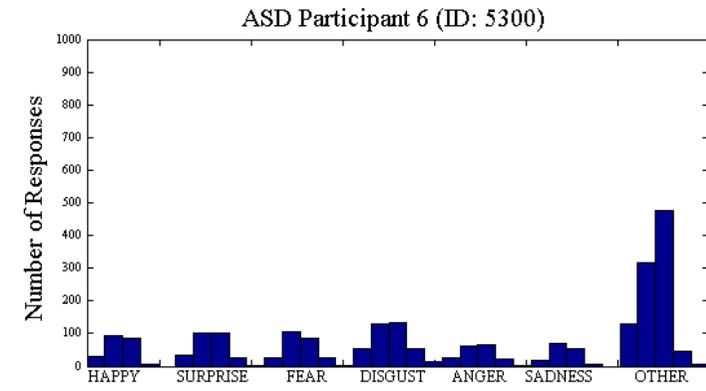
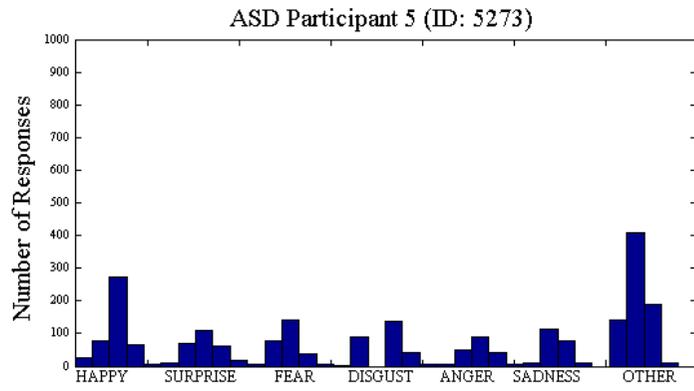


Figure 2–13 ASD Group Part 2: Frequency of stimuli responses per ASD participant for emotion and intensity (2400 trials per participant).

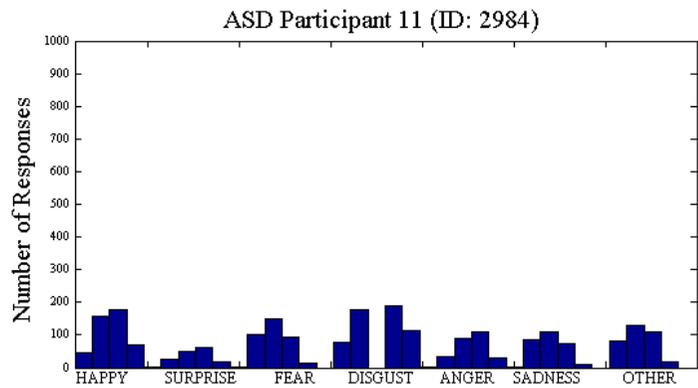
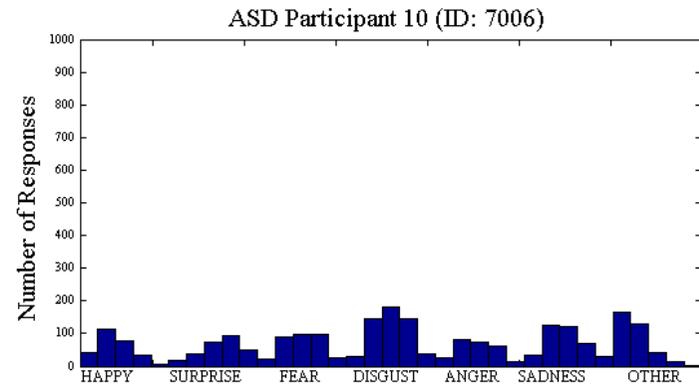
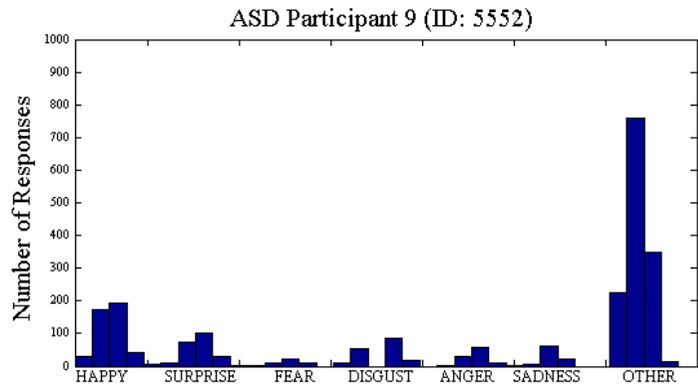


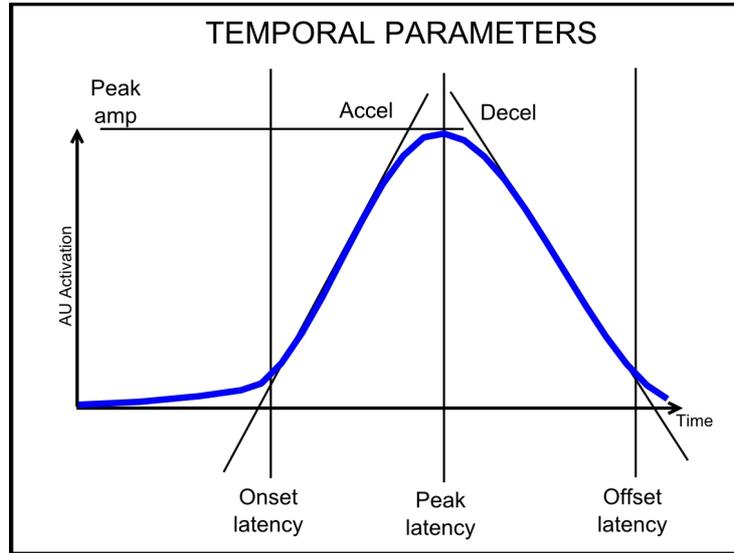
Figure 2-14 ASD Group Part 3: Frequency of stimuli responses per ASD participant for emotion and intensity (2400 trials per participant).

### **2.3.2 Reverse correlation: modelling individual differences**

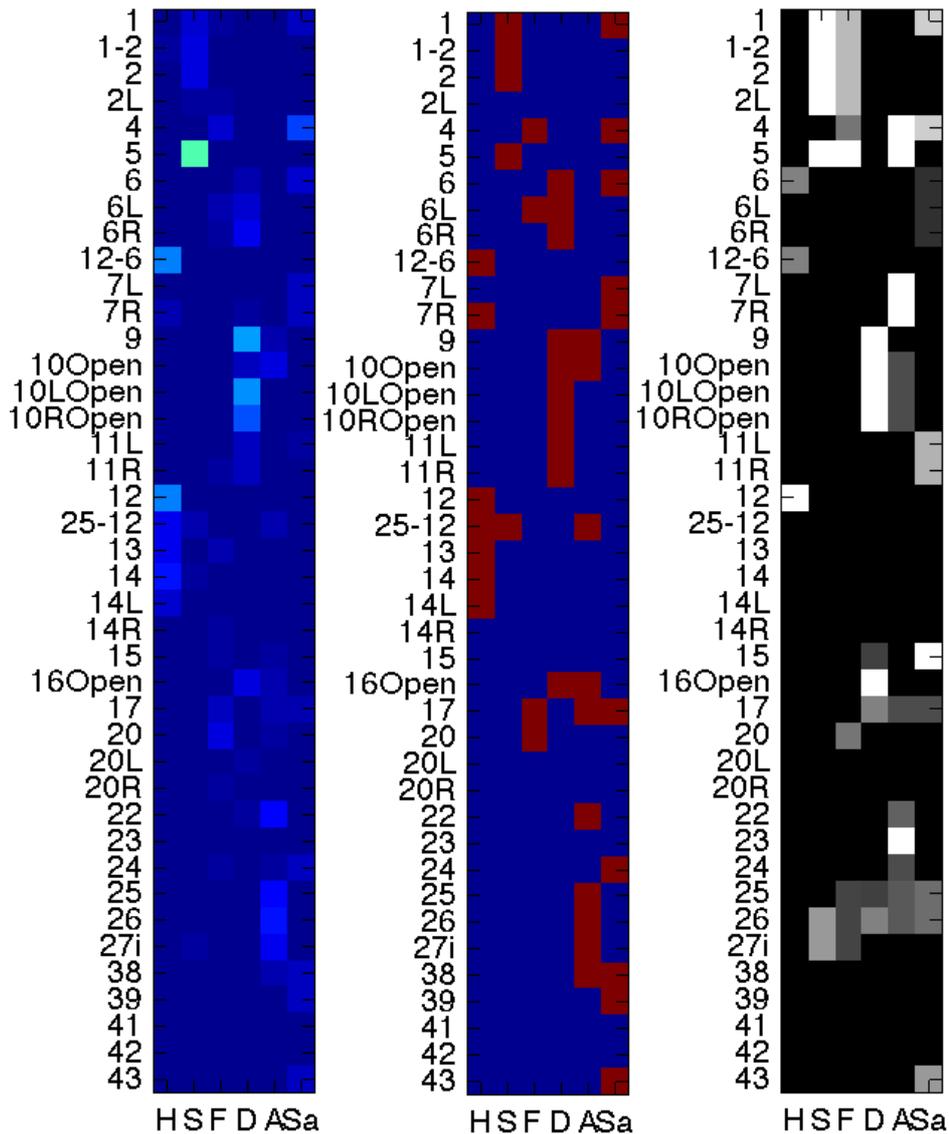
In order to obtain a detailed picture ('model') of emotion perception for each participant, the 2400 categorizations made by each participant were analysed using reverse correlation (as described in Section 2.2.5). Each model contains the AUs (Figure 2–15), and temporal parameters (Figure 2–16) that were associated with each participant's responses to a significant level ( $p \leq 0.05$ ) and a separate model was produced for each emotion. Below is an example of the AUs that emerged in the reverse correlation process for one participant. Figure 2–17 depicts, for one participant with ASD (participant 11 (ID 2984)), the AUs that were indicated by the reverse correlation analysis (lighter blue indicating higher correlation coefficients), the AU that reached a correlation to a significant level ( $p \leq 0.05$ ) and also the AUs which are regarded as 'core' AUs for each emotion expression, as described in Ekman and Friesen (1977).

Cell Number	AU Number	AU Name
1	AU1'	Inner Brow Raiser
2	'AU1-2'	Inner Brow Raiser - Outer Brow Raiser
3	'AU2'	Outer Brow Raiser
4	'AU2L'	Outer Brow Raiser Left
5	'AU4'	Brow Lowerer
6	'AU5'	Upper Lid Raiser
7	'AU6'	Cheek Raiser
8	'AU6L'	Cheek Raiser Left
9	'AU6R'	Cheek Raiser Right
10	'AU7L'	Eye Lid Tightener Left
11	'AU7R'	Eye Lid Tightener Right
12	'AU9'	Nose Wrinkler
13	AU10 Open'	Upper Lip Raiser Open Mouth
14	AU10L Open'	Upper Lip Raiser Open Mouth Left
15	AU10R Open'	Upper Lip Raiser Open Mouth Right
16	'AU11L'	Nasolabial Deepener Left
17	'AU11R'	Nasolabial Deepener Right
18	'AU12'	Lip Corner Puller
19	'AU25-12'	Lips Part - Lip Corner Puller
20	'AU12L'	Lip Corner Puller Left
21	'AU12R'	Lip Corner Puller Right
22	'AU13'	Cheek Puffer
23	'AU14'	Dimpler
24	'AU14L'	Dimpler Left
25	'AU14R'	Dimpler Right
26	'AU15'	Lip Corner Depressor
27	'AU16Open'	Lower Lip Depressor Lips Open
28	'AU17'	Chin Raiser
29	'AU20'	Lip Stretch
30	'AU20L'	Lip Stretch Left
31	'AU20R'	Lip Stretch Right
32	'AU22'	Lip Funneler
33	'AU23'	Lip Tightener
34	'AU24'	Lip Pressor
35	'AU25'	Lips Part
36	'AU26'	Jaw Drop
37	AU27'	Mouth Stretch
38	'AU38'	Nostril Dilator
39	'AU39'	Nostril Compressor
40	'AU43'	Eyes Closed
41	'AU7'	Eye Lid Tightener
42	'AU12-6'	Lip Corner Puller - Cheek Raiser

Figure 2–15 Action Unit numbers and their respective names

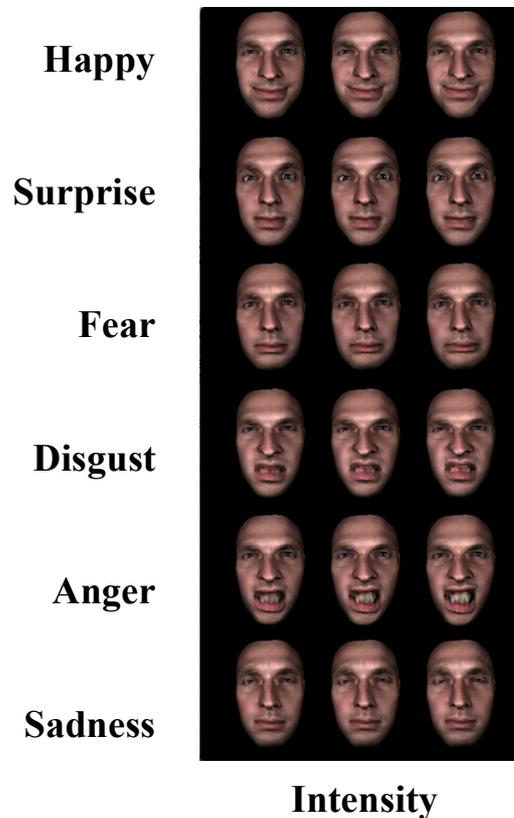


**Figure 2-16 Six temporal parameters associated with each action unit**



**Figure 2-17** From left to right: 1) Single participant data (ASD participant 11 (ID 2984)) of AUs that emerge from the data (lighter blue indicates higher correlation coefficient ( $\rho$  values), and hence lower  $p$  values) for all emotions (happy, surprise, fear, disgust, anger sadness) 2) This is the same as the image 1 but only the  $\rho$  values that are significant are displayed (i.e.  $p \leq 0.05$ ) in red and 3) Action Units indicated as being 'core' action units from Ekman and Friesen 1978.

Dynamic models for each participant can be viewed in video format (see Appendix 2). A snapshot at the middle time point (0.75 seconds) of one participant can be seen in Figure 2–18. This dynamic model allows the viewer a detailed view of the nuances of any one individual participant’s data.



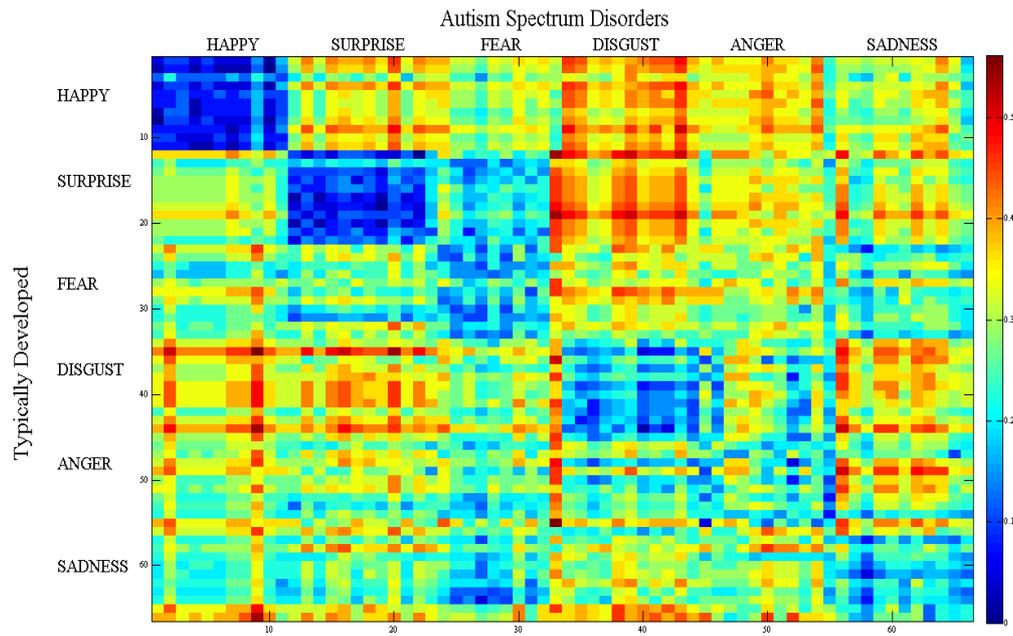
**Figure 2–18 Example: ASD participant 11 (ID 2884)’s model at the midpoint of the dynamic videos at 3 intersections of intensity.**

### 2.3.2.1 Analysis of models: action units

Group models, i.e. depicting the AUs and temporal dynamics as averaged for each group, can be seen in video format (see Appendix 3).

In order to assess the differences between ASD and TD participants, contingent to the specific AUs (i.e. AUs identified as contributing to a significant level), individual models (as described in section 2.3.2) were compared between groups. Firstly, the AUs that were present in each individual’s model were compared group-wise using Hamming distance (Hamming, 1950). Hamming distance is a measure of difference between binary units, which is applicable to the AU data presented here because each AU, out of 42, is either ‘on’ (1) or ‘off’ (0). Hence, a binary vector can be used to represent the AUs that reached significance for each model. For example, in the figure above (Figure 2–17) a depiction of the AUs that are ‘on’ for participant 11 can be seen. This figure could thus just as easily be

depicted as a combination of '0' values for (i.e. the blue boxes) and '1' values (i.e. for the red boxes), where the '1' values indicate each AU that was found to be 'on' to a significant level in the reverse correlation analysis. Hamming distance compares the 'distance' between two binary vectors by calculating the number of positions at which the vectors are different in its '1' or '0' content, divided by the total number of positions (i.e. the length of the vector). Hence, a hamming distance of '0.2' between two binary vectors would indicate a difference in 20% of the binary pattern. In the current study the AUs for each model were depicted in a vector of 42 '1's or '0's. Hamming distance was used to measure the difference between pairs of models. Hamming distance for 66 model pairs (i.e. 11 participants per group x 6 emotion categories), as shown in the dissimilarity matrix in Figure 2–19. Colder colours (e.g. dark blue) indicate high similarity between models (low hamming distance values) and hotter colours (e.g. dark red) indicate a larger difference between models (high hamming distance values). Hence, each small square indicates one individual model compared to another individual model and the larger squares indicate the overall similarity between the groups per emotion category. The dissimilarity matrix indicates that the two groups have similar AU composition in models for happy and surprise, relatively similar in models for disgust, and less similarity for the other 3 emotions (fear, anger and sadness). The figure also suggests that there are overlaps/confusions between surprise, fear and sadness.

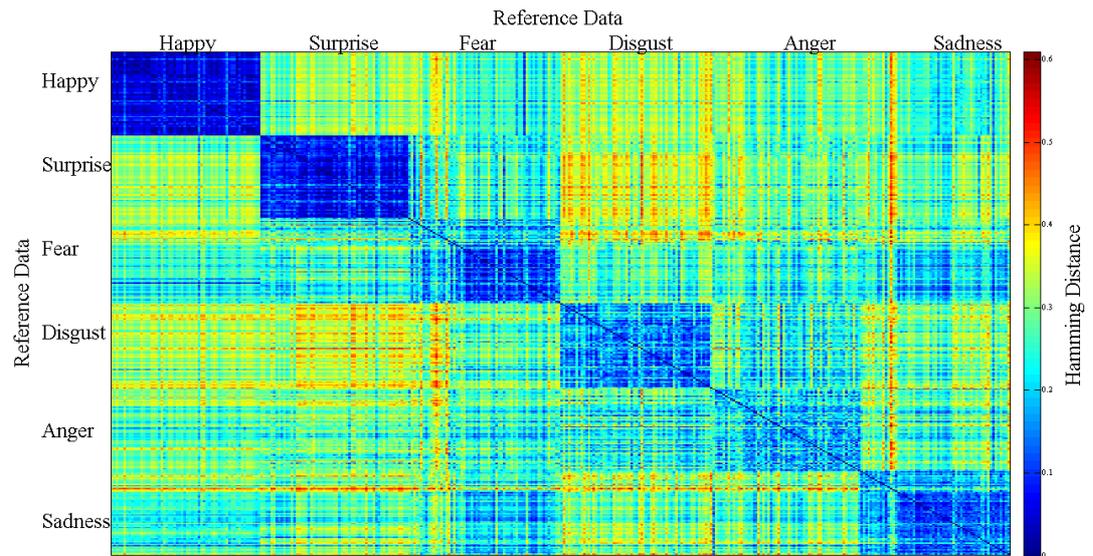


**Figure 2–19** Dissimilarity matrix of hamming distance (i.e. percentage of binary values that differ) between ASD ( $n = 11$ ) and TD ( $n = 11$ ) for all 6 emotions. Cool colours (e.g. dark blue) indicate low hamming distance values and hot colours (e.g. dark red) indicate high hamming distance values.

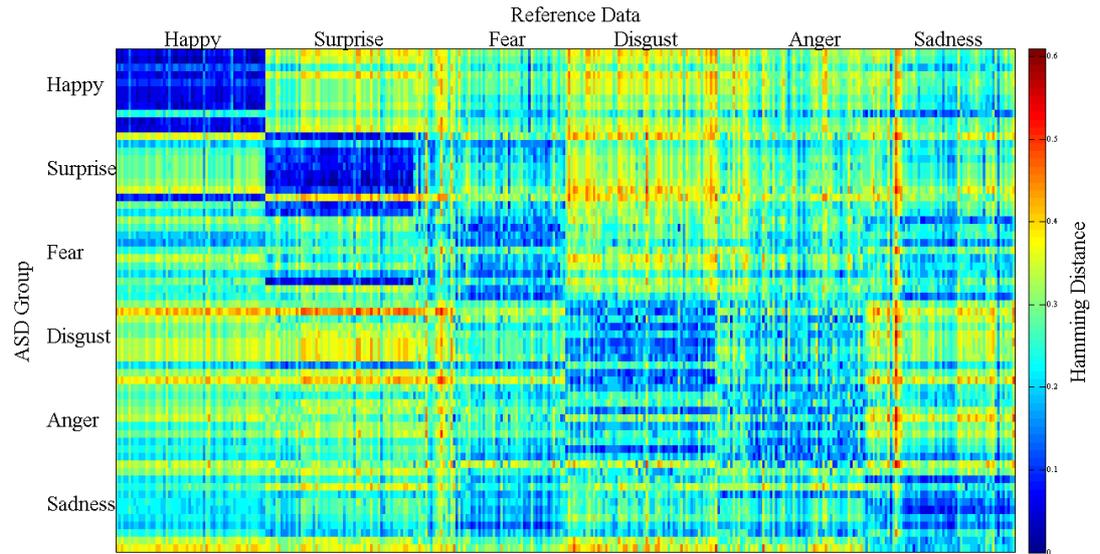
The above dissimilarity matrix suggests that group differences in AUs can be found between the ASD and TD groups for some emotions. However two questions remain unresolved: 1) Do the differences between ASD and TD groups meet statistical significance? 2) Are a number of ‘atypical’ individuals driving the group difference? (as suggested in Figure 2–9 and Figure 2–10).

Between group analyses were conducted on the AUs from each group’s models in order to assess whether the ASD and TD groups were statistically different from each other. In order to do this, a quantitative measurement of how ‘typical’ the AUs were in each model was sought. Therefore, a dataset of 60 models per emotion that depicted ‘typical’ AU usage were introduced. This dataset was derived from a previous study by Jack, Caldara, et al. (2012b) in which models of emotion perception were derived from 15 participants using the same method and reverse correlation analysis procedure as described above. As these participants were self-reportedly ‘typical’, (i.e. they reported no clinical diagnoses) and previous research has demonstrated their success in an all-round representation of each of the basic 6 emotions, this group of models was used as a reference group upon which to compare ASD and TD groups. This reference group was made up of 15 Western Caucasian

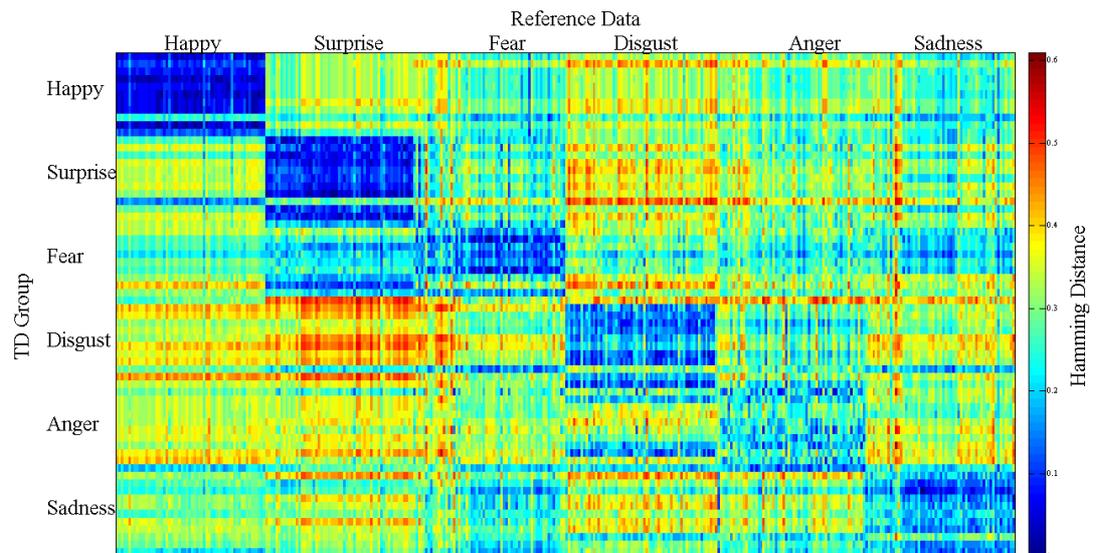
adults (6 men and 9 women) with a mean age of 21.3 years (SD = 1.2 years). For full details of the study please see (Jack, Caldara, et al., 2012b)). The ‘atypicality’ of AUs used by each group was, therefore, defined by its closeness to the reference group (by means of Hamming distance). Likewise, Hamming distance was measured for each TD individual’s model to each individual in the reference group. Therefore, distance values represent a measure of ‘typicality’ for both the ASD and the TD group, and hence controlling for possibilities of atypicality in the TD group.



**Figure 2–20 Reference data (n = 15): comparison of models within group**



**Figure 2–21** Hamming distance between AU composition of models for the ASD group compared to reference dataset



**Figure 2–22** Hamming distance between AU composition of models for the TD group compared to reference dataset

Figure 2–24 and Figure 2–26 suggest that, when compared to an independent reference dataset, the AUs of the ASD models are no more ‘atypical’ than the TD group. This finding is further ratified by a t-test, which compared the mean distance of each emotion in which no significant differences were found between ASD and TD groups in their mean distance values as compared to the reference group:

Happy:  $t(20) = 0.02$ ,  $p = 0.98$

Surprise:  $t(20) = 0.52$ ,  $p = 0.61$

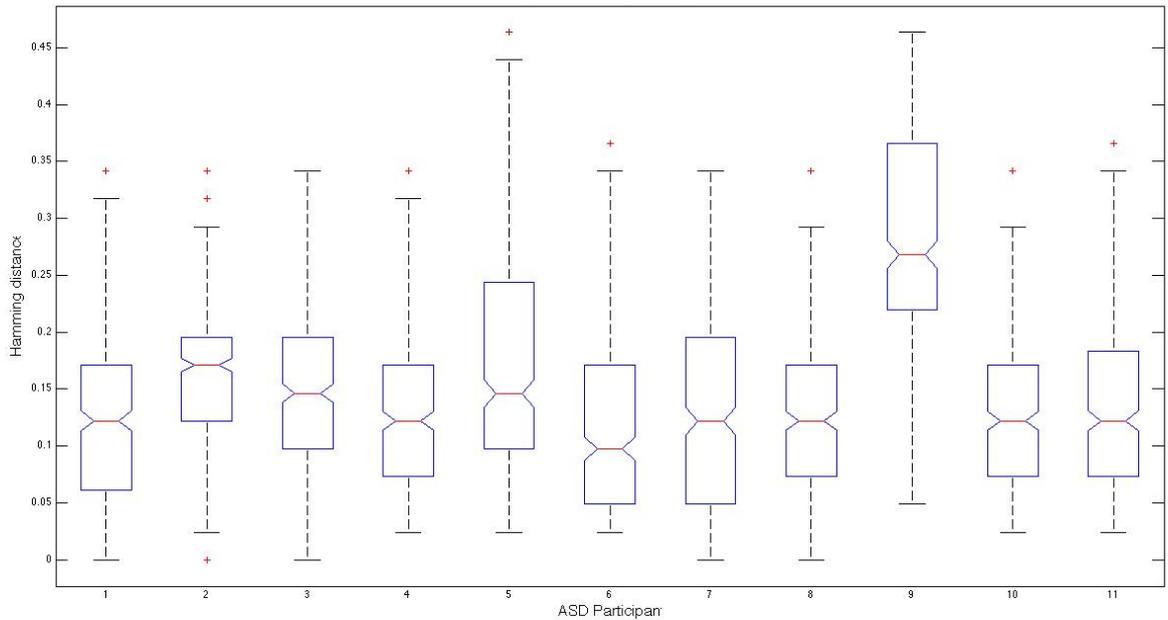
Fear:  $t(20) = -0.09$ ,  $p = 0.93$

Disgust:  $t(20) = 0.06$ ,  $p = 0.95$

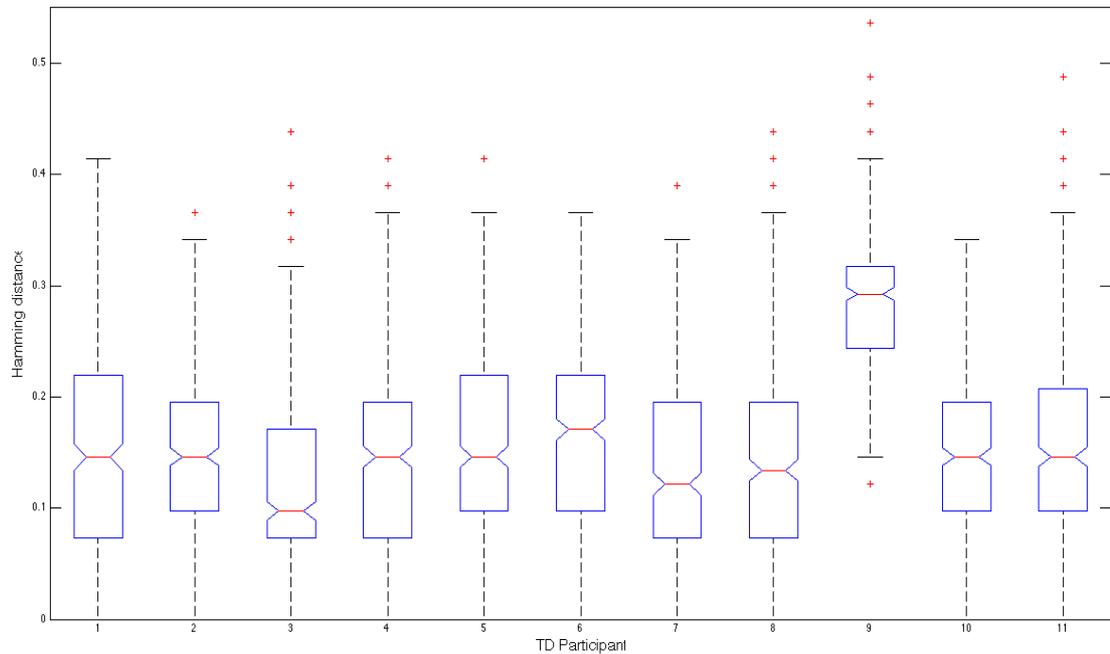
Anger:  $t(20) = 1.80$ ,  $p = 0.09$

Sadness:  $t(20) = 0.46$ ,  $p = 0.65$

It could be suggested that there may be some individuals who, when isolated, show themselves as being ‘atypical’ - a finding which may have previously been overlooked when analysing the data at group level. The mean, upper and lower quartiles, and the range of hamming distance values, for each individual, are displayed for the ASD group (Figure 2–23) and for the TD group (Figure 2–24). These figures suggest that both the ASD and TD groups contain ‘atypical’ individuals who do not follow suit with the rest of the group (e.g. participant number 9 in the ASD group and participant number 9 in the TD group). Again, ASD participant number ‘9’ also has the highest autism severity level out of the ASD group (as shown in the ADOS scores in section 2.2.2.1). Participant ‘9’ in the TD group has the highest level of autism-like traits out of the TD group (as shown in in the AQ scores in section 2.2.2.1).

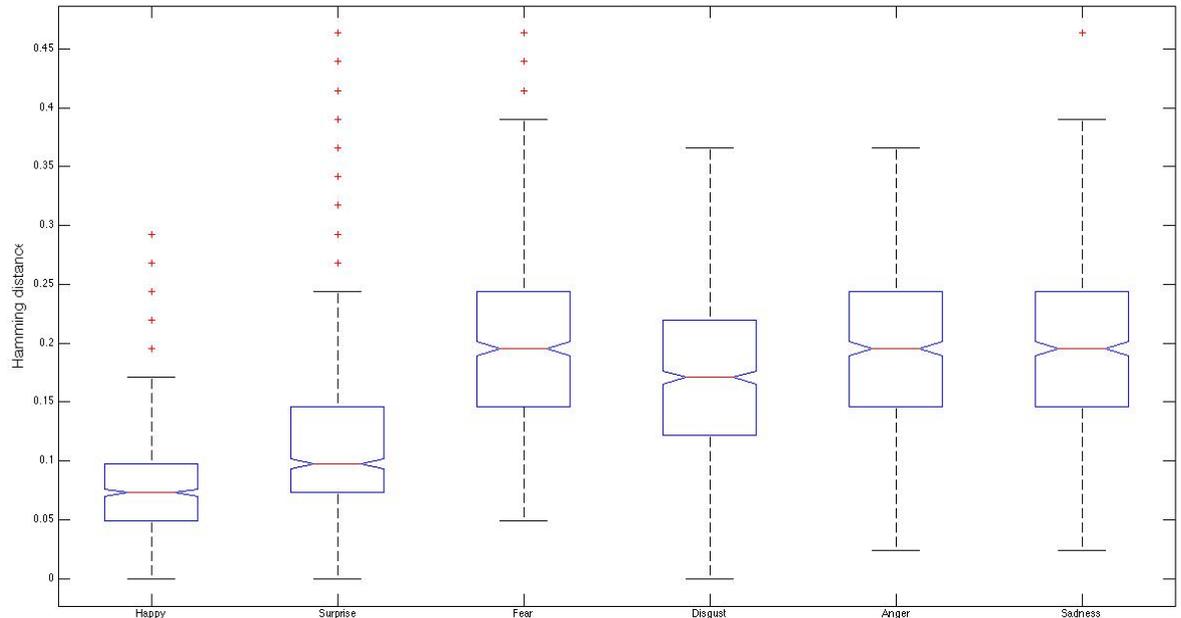


**Figure 2–23. ASD vs. template hamming distance per individual means and variance**

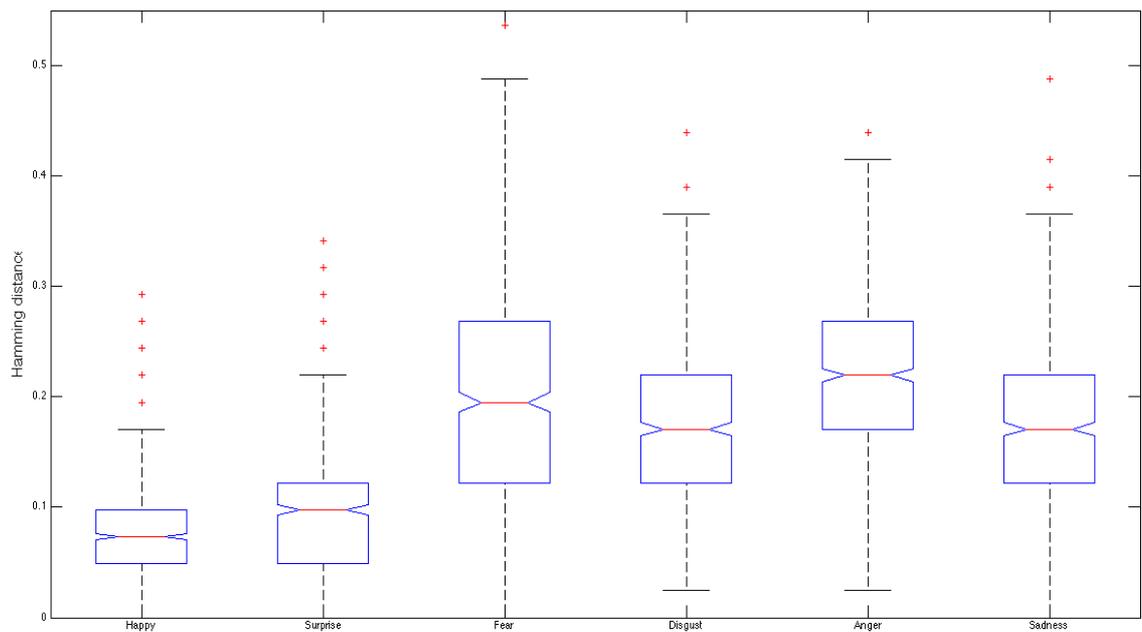


**Figure 2–24. TD vs. template hamming distance per individual means and variance**

Further to this, in order to assess within-group differences in ‘atypicality’, mean, upper and lower quartiles, and the range of hamming distance values for each emotion are displayed in Figure 2–25 and Figure 2–26. These figures indicate that both the ASD and TD group are consistently less like the reference data for fear, disgust, anger and sadness as suggested in the analyses described above. Hence, this suggests that both groups have a more disjointed categorization of the ‘negative’ emotions (fear, disgust, anger and sadness) compared to the emotions that are often described as the ‘happy’ emotions of the six (happy and surprise).



**Figure 2–25. ASD vs. template hamming per emotion distance means and variance**

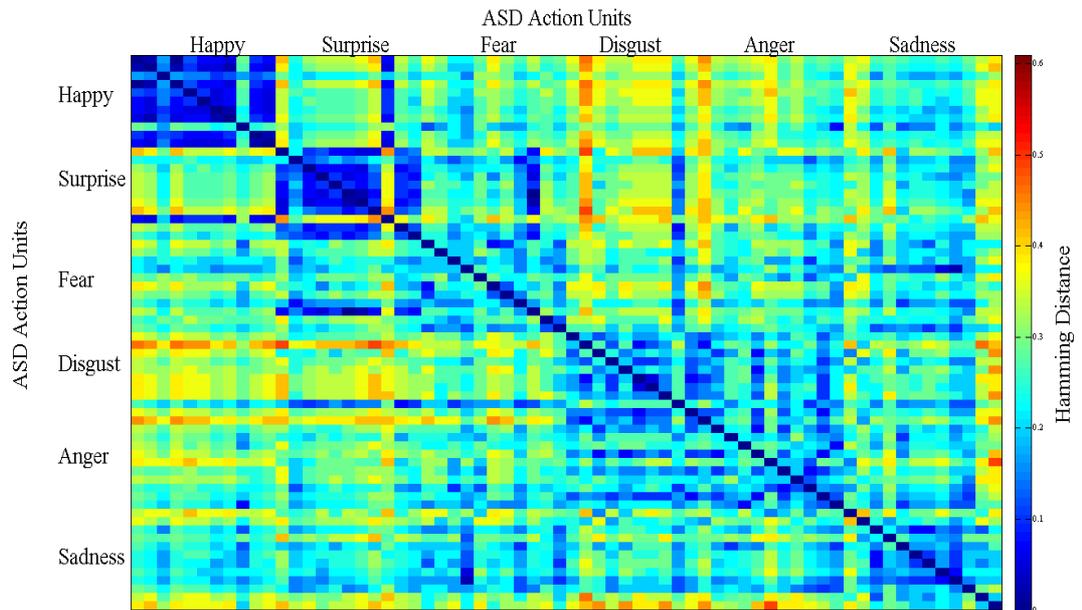


**Figure 2–26 TD vs. template hamming distance per emotion means and variance**

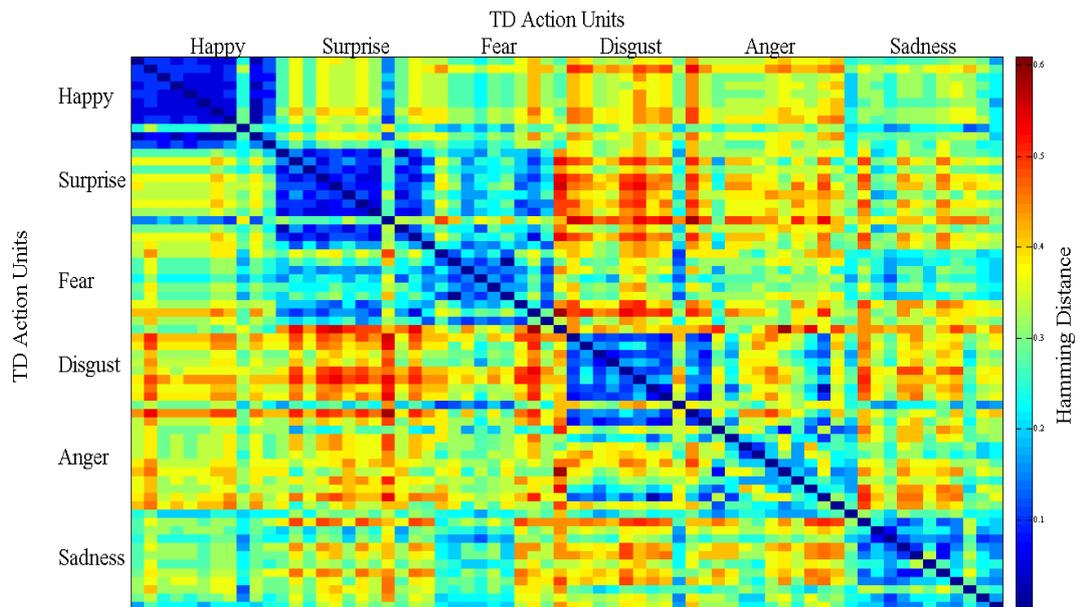
### 2.3.2.2 Within groups analysis of models: action units

As of yet, the data have not revealed whether the ASD group's models contained AU patterns that were consistent with each other. This analysis was considered important because heterogeneity within the ASD group was highlighted in the way in which some participants responded to the stimuli (i.e. atypical categorisations seen in Figure 2–9 and Figure 2–10). Hence, in addition to the analyses described thus far, within-group differences were also assessed by obtaining Hamming distance values for each model,

compared to each other model within one group. To do this, Hamming distance was measured for each ASD pair of models in the ASD group (66 comparisons: 11 participants x 6 emotions) and the same analysis was carried out for the TD group.



**Figure 2–27** Within-group comparison of AU composition of models using Hamming distance



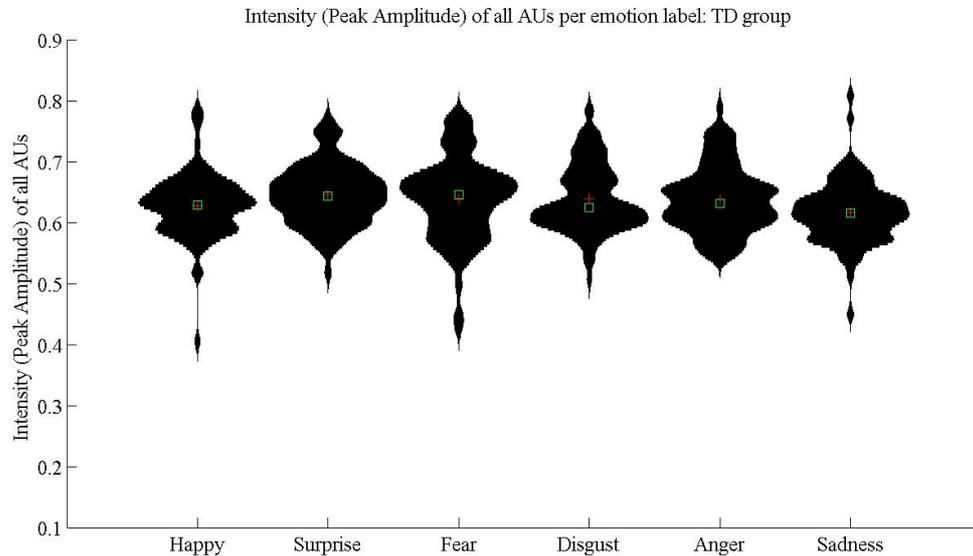
**Figure 2–28** Within-group comparison of AU composition of models using Hamming distance

The dissimilarity matrices in Figure 2–27 and Figure 2–28 suggest that the ASD group were more inconsistent with each other (more blue colour across the figure indicating less distance between the models of different emotions and hence more ‘confusion’ between emotions). In particular, the ASD group appear to confuse disgust and anger and do not appear to have distinct classifications for fear or sadness (indicated by a lack of defined blue box that would be expected along the middle diagonal line of the figure). The TD group, on the other hand, appear to be more consistent with each other, as indicated by the more distinct blue boxes across the diagonal and clear red colouration throughout the rest of the figure. However, it is apparent that the TD group are still relatively dissimilar for anger, which, in the above figure, is indicated by Hamming distance values in the middle of the scale (around 0.3) in contrast to very low hamming distance values for the other emotions (around 0.1). In order to ascertain if these data indicated that the TD group were significantly more consistent with each other, a two-way ANOVA was applied to the data. Results revealed that the variance in distance values could not be accounted for by group ( $F(1,120) = 0.28, p = 0.60$ ) however, the variance could be accounted for by emotion ( $F(5,120) = 16.33, p < 0.01$ ). Variance was not, however driven by an interaction between group and emotion ( $F(6,120) = 1.1, p = 0.36$ ).

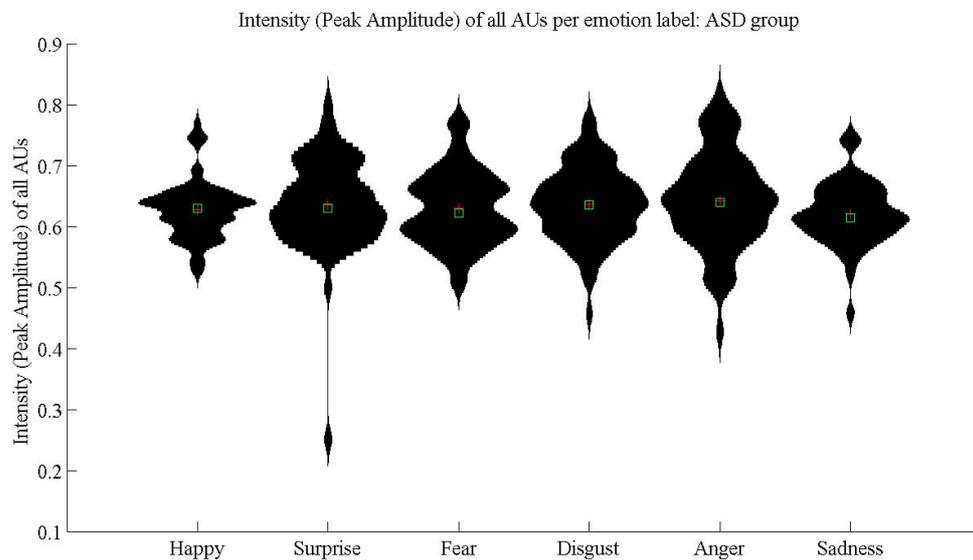
### 2.3.2.3 Analysis of models: amplitude

The data were then analysed for differences in ‘amplitude’ of the models (i.e. how strong an AU was depicted in the models). A measure of amplitude of the models was established by obtaining, for each AU that was ‘on’ (i.e. was present in the model to a significant level), the peak amplitude of the AU. Amplitude values were, therefore a value between 0 and 1, where 0 indicates ‘off’ and 1 indicates maximum amplitude. Figure 2–29 and Figure 2–30 provide a multiple distribution visualisation (violin plot) of the peak amplitude values for each group. For example, Figure 2–29 displays the mean (red cross), the median (green box) and the distribution (black opaque section) of values across the whole TD group. Figure 2–29 indicates that the mean amplitude values for the TD group are relatively similar across all six emotions. The distribution of values is also relatively equal across the emotions for the TD group. For the ASD group, Figure 2–30 suggests relatively equal mean intensity levels, however surprise displays a long tail relative to the other emotions, suggesting that a small subset of models within the ASD group had low intensity values for surprise. In general, the violin plots presented in Figure 2–29 and Figure 2–30 indicate that the ASD group has a relatively wider distribution of values compared to the TD group.

In order to assess the hypothesis that ASD models would have higher intensity values a 2 x 2 ANOVA was conducted between the groups. Two-way ANOVA revealed that the variance in amplitude of the models was not due to group ( $F(1, 125) = 0.35, p = 0.55$ ), nor emotion ( $F(5, 125) = 0.85, p = 0.51$ ).



**Figure 2–29 Multiple distribution plot denoting mean (red cross), median (green box) and distribution of intensity values of TD models**



**Figure 2–30 Multiple distribution plot denoting mean (red cross), median (green box) and distribution of intensity values of ASD models**

As it has been suggested that group differences between ASD and TD participants may be found between the eye and mouth regions of the models, a further analysis was conducted,

isolating AUs associated with the eye and mouth only. The specific AUs that were deemed ‘eye AUs’ and ‘mouth AUs’ were as follows:

Eye AUs (Total = 10):

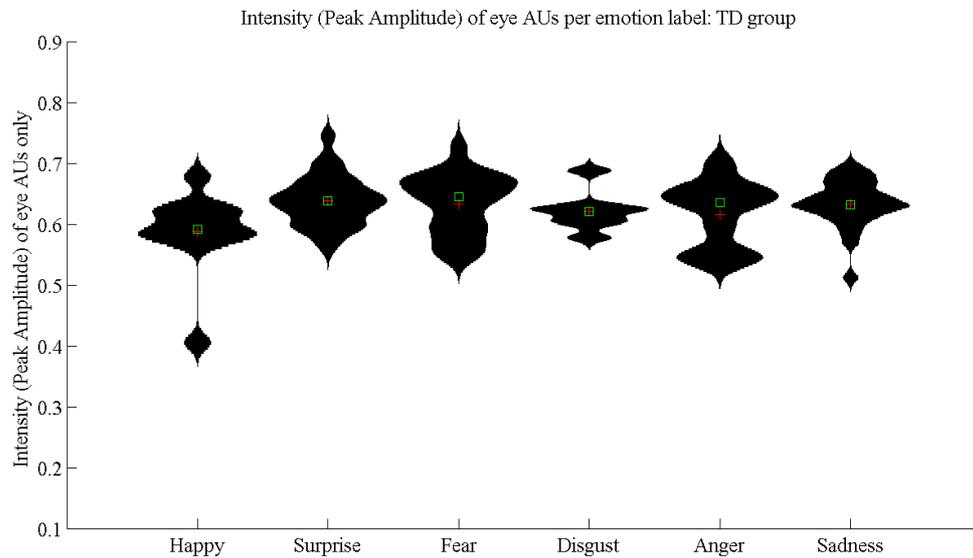
Inner Brow Raiser  
 Inner and Outer Brow Raiser  
 Outer Brow Raiser  
 Brow Lowerer  
 Upper Lid Raiser  
 Eye Lid Tightener Left  
 Eye Lid Tightener right  
 Eye Lid Drop  
 Eye Slit  
 Eyes Closed

Mouth AUs (Total = 25):

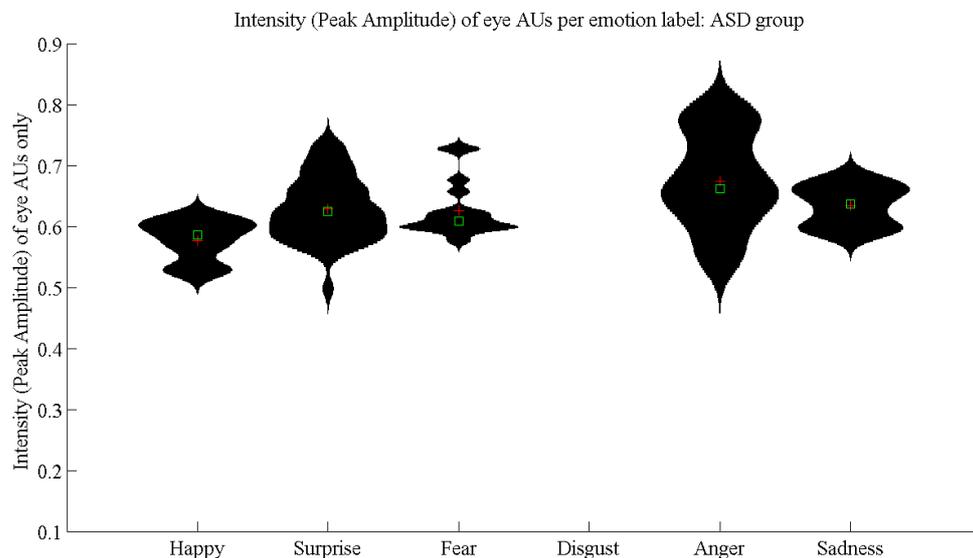
Cheek Raiser  
 Cheek Raiser Left  
 Cheek Raiser Right  
 Lip Corner Puller and Cheek Raiser  
 Upper Lip Raiser Open  
 Upper Lip Raiser Open Left  
 Upper Lip Raiser Open Right  
 Lip Corner Puller  
 Lip Corner Puller -Lips Part  
 Cheek Puffer  
 Dimpler  
 Dimpler Left  
 Dimpler Right  
 Lip Corner Depressor  
 Lower Lip Depressor Lips Open  
 Chin Raiser  
 Lip Stretch  
 Lip Stretch Left  
 Lip Stretch Right  
 Lip Funneler  
 Lip Tightener  
 Lip Pressor  
 Lips Part  
 Jaw Drop  
 Mouth Stretch

First, multiple distribution plots were created in order to visualize the variance of the intensity values across participant group and emotions for the eye AUs. Figure 2–31 displays the distribution of intensity values within the TD group for each emotion. The figure indicates that the TD group was relatively consistent across emotions. Happy had the lowest mean intensity value (red cross) however a small cluster of low values within this emotion may have influenced this. Similarly, small clusters are found for disgust, however the higher intensity cluster is relatively small. Figure 2–32 indicates that, despite

relatively similar means, the ASD group were slightly more variable across emotions compared to the TD group. For example, for the emotion anger, the ASD group appear to have a wide distribution of intensity scores. Note that for disgust, no intensity scores are shown because no eye AUs were ‘on’ to a significant level in the eye region in disgust for the ASD group.



**Figure 2–31 Intensity values of TD models for eye AUs only**

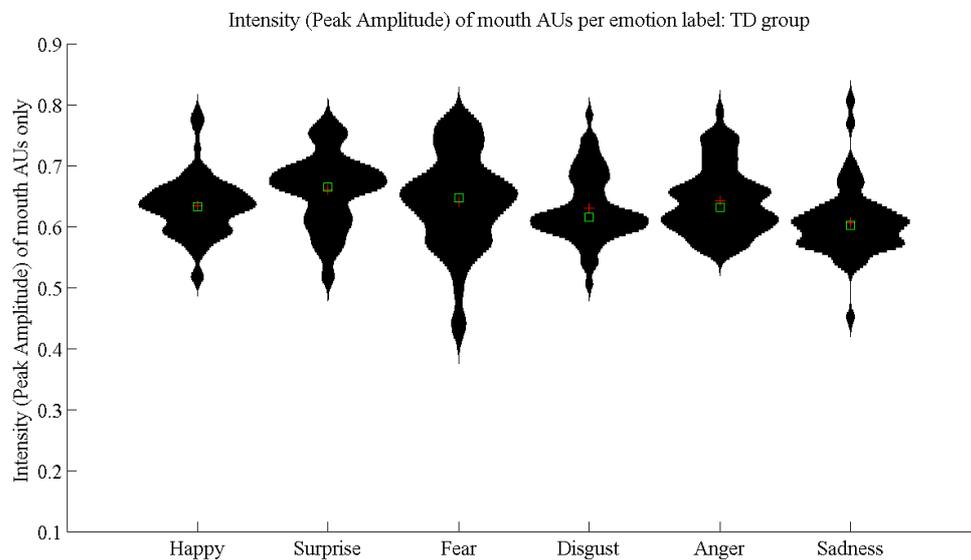


**Figure 2–32 Intensity values of ASD models for eye AUs only**

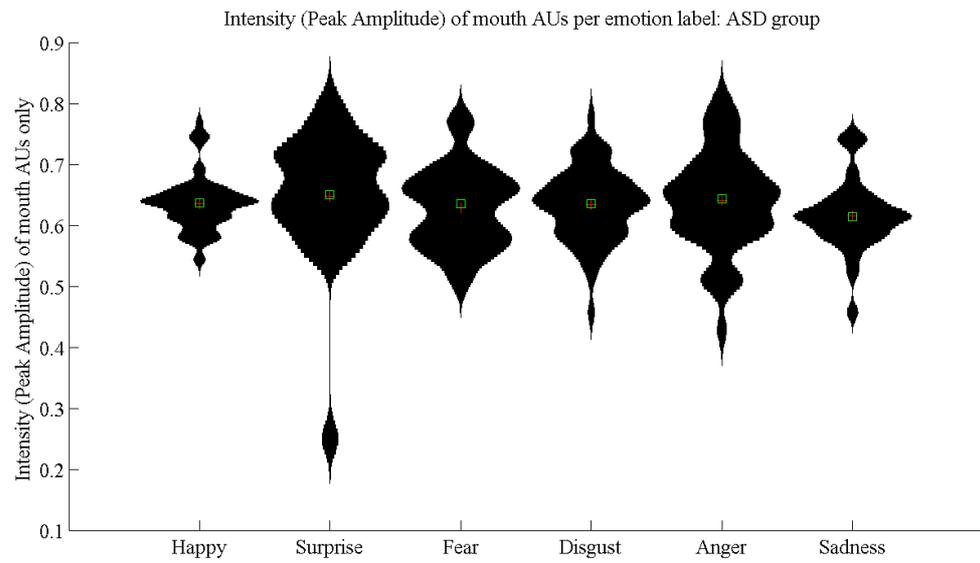
A 2-way ANOVA was performed on the peak amplitude of the eye AUs. This ANOVA allowed for unbalanced data because some AUs were not ‘on’ for the eye region (for example in disgust for the ASD group). As expected from the visualizations show in the

violin plots, results revealed no significant difference for the main effect of group  $F(1, 88) = 0.44$ ,  $p = 0.51$ , and no significant difference for the main effect of emotion  $F(5, 88) = 1.46$ ,  $p = 0.21$ .

The mouth area was assessed using the same analysis steps as the previous analysis on eye AUs. Distribution plots (Figure 2–33 and Figure 2–34) display the mean (red cross), median (green square) and distribution (black section) of the mouth AUs for each group's models. Overall, the means of each group are relatively similar, slight variations in distribution of values can be seen, for example, the ASD group appear to have slightly larger distribution of values across all emotions except happy. Also, a somewhat wider distribution of intensity values can be seen for surprise in the ASD group as compared to the TD group. Again, an unbalanced 2-way ANOVA was applied to these data in order to account for unequal vector lengths due to any AUs that were not 'on'. Results revealed that there was no significant difference for the main effect of group  $F(1, 122) = 0.05$ ,  $p = 0.83$  and there was no significant difference for the main effect of emotion  $F(5, 122) = 0.46$ ,  $p = 0.80$ .



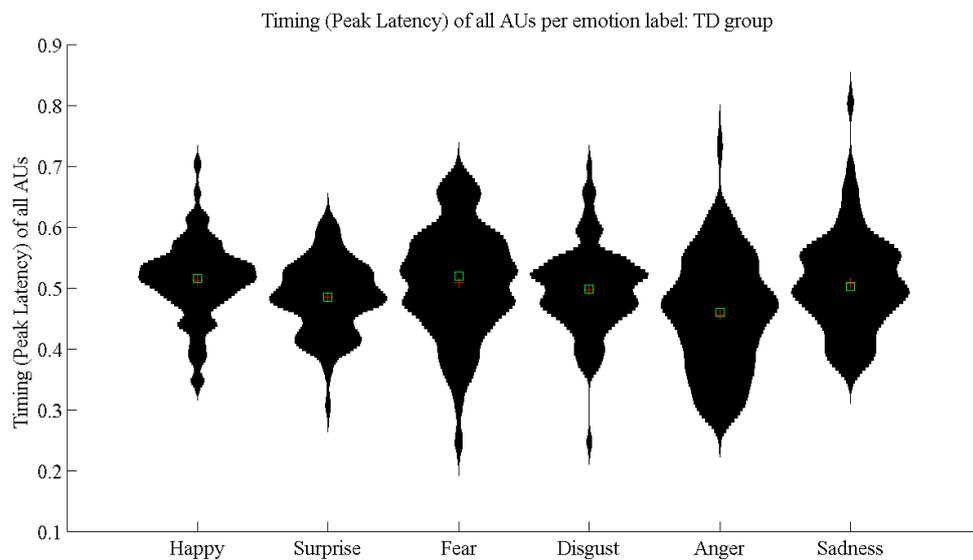
**Figure 2–33 Intensity values of TD models for mouth AUs only**



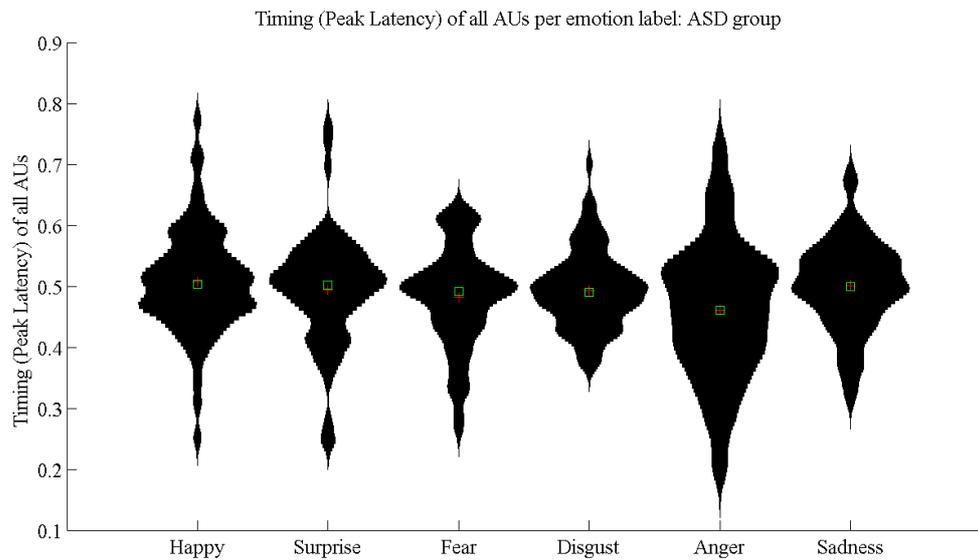
**Figure 2–34 Intensity values of ASD models for mouth AUs only**

### 2.3.2.4 Analysis of models: temporal

In order to assess differences in the temporal aspect of the models (i.e. the point at which the AUs were active across the 1.25 second model), ‘peak latency’ was assessed. Peak latency refers to the time point at which the ‘peak’ of the facial movement was displayed. This was firstly assessed using visualization of the data via violin plots. Below, the violin plots for all TD AUs indicate relatively similar mean values across all emotions. This is reflected also in the ASD group, who also remain consistently around the 0.5 mark. The distribution of values is not wildly different between groups, although there are nuanced differences in specific emotions, for example, for anger the ASD group continues further into the very early time frames while the TD group does not. A 2-way ANOVA revealed that there was no significant difference for group ( $F(1,125) = 0.29, p = 0.59$ ) or for emotion ( $F(5,125) = 0.38, p = 0.86$ ).

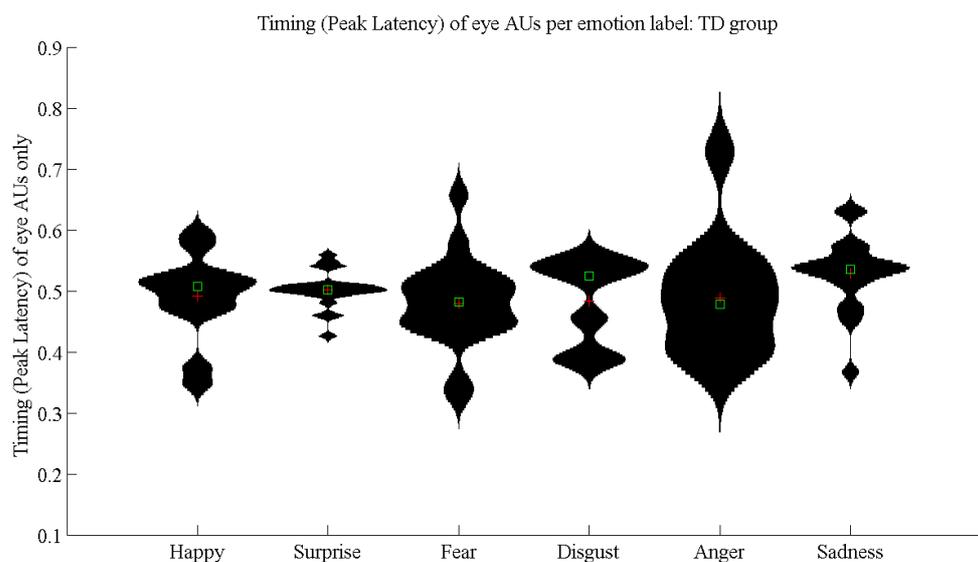


**Figure 2–35 Peak temporal values for TD models**

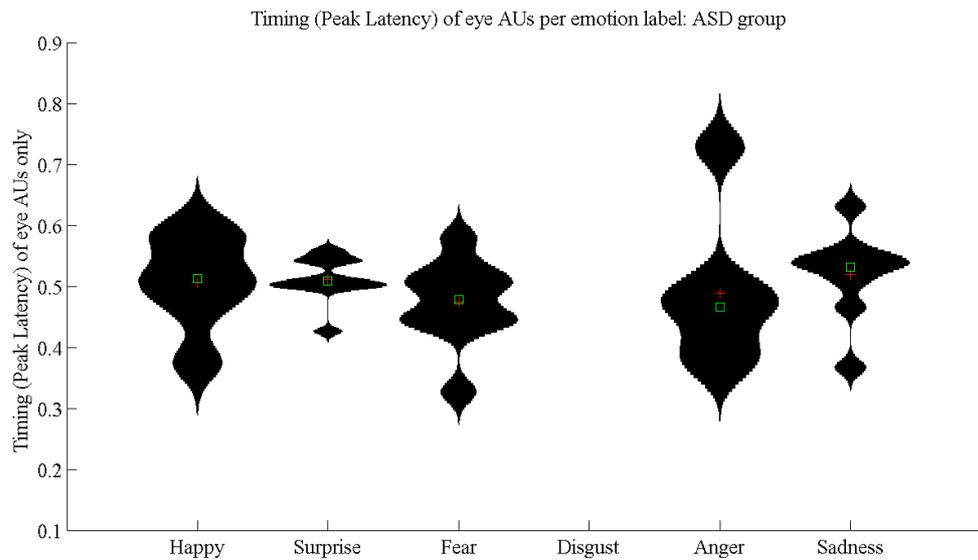


**Figure 2–36 Peak temporal values for ASD models**

In order to assess whether differences between groups or emotions may have been region specific, an independent assessment of the eye region and the mouth region was conducted (see Section 2.3.2.3 for break down of eye AUs and mouth AUs). Figure 2–37 and Figure 2–38 reveal that both TD and ASD groups varied across emotions. For example, both groups had a relatively narrow distribution of timing values for surprise (around the midpoint of 0.5), while having a relatively wide distribution of timing values for anger. Statistical analysis of these data indicated, however that there was no significant difference between groups ( $F(1, 88) = 0.04, p = 0.85$ ) or emotion ( $F(5, 88) = 0.63, p = 0.68$ ).

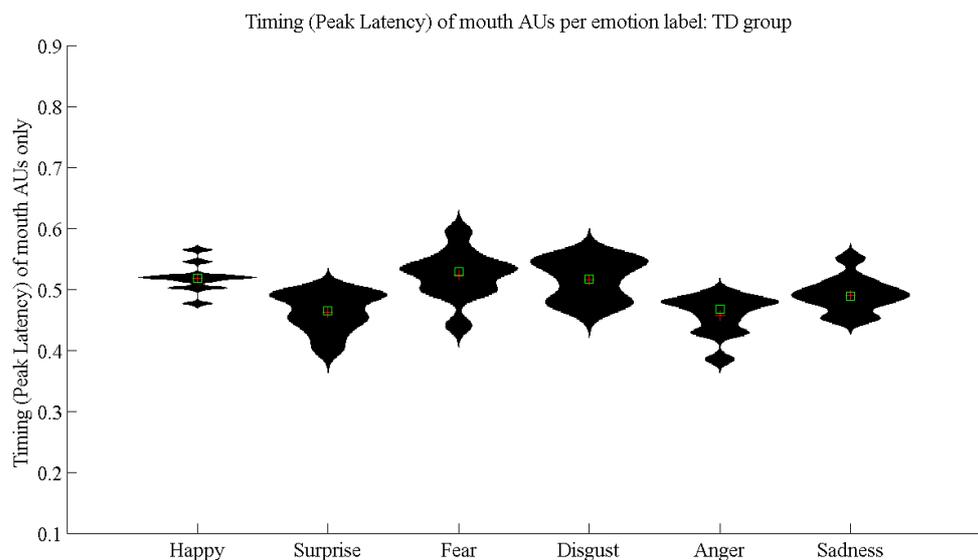


**Figure 2–37 Peak temporal values for TD models: eye region only**

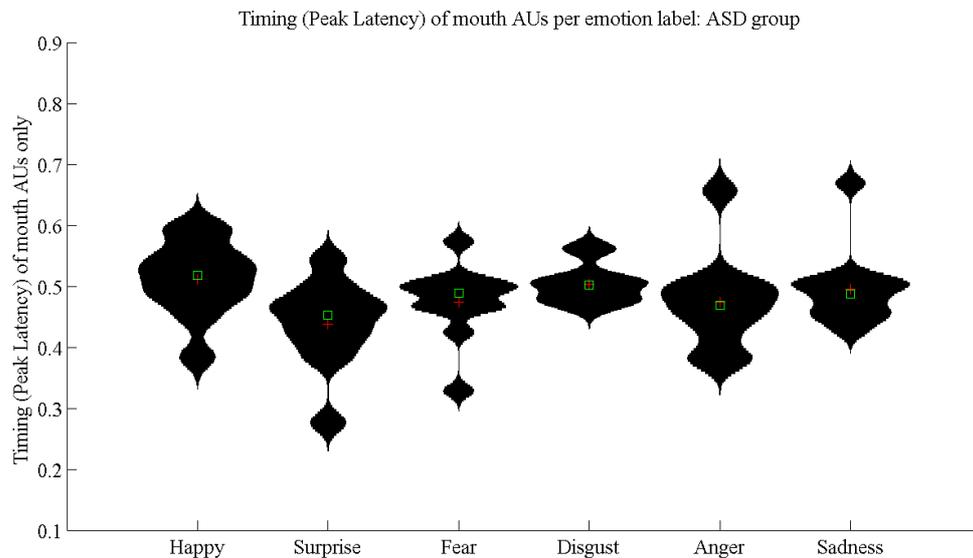


**Figure 2–38 Peak temporal values for ASD models: eye region only**

Subsequent to this, the timing data were assessed specifically for the mouth region of the models. Figure 2–39 and Figure 2–40 indicate that the ASD group had somewhat wider distribution of timing values as compared to the TD group, especially for happy, surprise, anger and sadness. Results of a 2-way ANOVA revealed that this group difference was, however, not statistically significant ( $F(1, 122) = 1.68, p = 0.19$ ). The main effect of emotion was also found to be non significant  $F(5, 122) = 0.76, p = 0.76$ .



**Figure 2–39 Peak temporal values for TD models: mouth region only**



**Figure 2–40 Peak temporal values for ASD models: mouth region only**

## 2.4 Discussion

The aim of this study was to provide an in-depth assessment of facial expression perception in individuals with ASD, as compared to TD individuals. This study builds upon previous research in two key ways: 1. It provides a novel approach to assess FER in ASD, offering a comprehensive portrait of facial expression perception per individual and 2. It explores several hypotheses drawn from the literature based on one rich dataset. The methods used in this study allow for a depth of measurement on facial expression perception that has not previously been attempted in ASD research. This study extends current methodologies, firstly by employing stimuli that are not only dynamic but also completely controllable across several parameters (i.e. the AUs are controlled as well as amplitude of the stimuli and the timing of the movements) and secondly, by applying a reverse correlation technique to this dataset, a representation of emotion perception based on responses to these stimuli is established. Four hypotheses were proposed: 1. The pattern of AUs in the ASD models will be atypical compared to the TD group. 2. The amplitude of the AUs will be higher in the ASD group 3. The timing of the AUs will be atypical in the ASD group 4. The ASD group will use information from the mouth area of the face in favour of information from the eyes. Each of these hypotheses will be now discussed in turn, as well as additional important information that emerged from the dataset.

Firstly, results revealed that the ASD group responded to the stimuli using the label ‘other’ significantly more times than the TD group. Interventions for emotion recognition in children with ASD (as reviewed by Kuou and Egel, 2016) often use a heuristic approach to learning emotions (e.g. ‘happy’ is when the corners of the mouth are upturned). Therefore, it could be argued that the ASD group are more inclined to select ‘other’ because they require the stimulus to be very like the heuristic basic emotions in order to commit to selecting an emotion label. Curiously, the TD group responded to the stimuli using the label ‘surprise’ significantly more times than the ASD group. This may be due to the connotations associated with surprise, i.e. perhaps the TD group are drawn to the ‘surprise’ label when shown random facial expressions because surprise can be composed of several different combinations of AUs as opposed to, for example, ‘happy’ which would require a more specific set of AUs (e.g. an upturned mouth and relaxed eye area).

The hypothesis that ‘the pattern of AUs in the ASD models will be atypical compared to the TD group’ was not supported in the results, which revealed that ASD and TD individuals are similar to each other in the AUs that are present in their models. Hence, this suggests that categorizing emotions based on facial features alone does not elicit obvious group differences between the ASD and TD groups. This is consistent with Harms et al. (2010) and Simmons et al. (2009) who suggest that group differences between individuals with ASD and TD individuals are harder to detect in static (i.e. AU information only) facial expression images as compared to dynamic ones. However, these data suggest that despite relative equality between the groups, some individuals were atypical compared to others, which appeared to be driven by higher autism severity in the ASD group and higher AQ scores in the TD group. This provides evidence to support assertions that FER ability is directly linked to severity of autism (Lerner et al., 2013; Uono et al., 2013; B. T. Williams & Gray, 2013). However, because these particular individuals scored low on Alexithymia, it does not support the hypothesis that FER is a separate construct to ASD itself as suggested by Cook et al. (2013). It is also notable that, despite a link between AQ level and FER ability in the TD group (and thus supporting assertions that FER can be predicted by levels of autism-like traits (Poljac et al., 2013)), no link was found for autism quotient score in the ASD group (i.e. those that scored high in autism severity did not score themselves high on the AQ). Facial expression recognition and its relationship with Alexithymia and AQ are explored further, in a larger population sample, in Chapter 4.

It was hypothesised that the amplitude of the AUs would be higher in the ASD group, however the data presented from this study do not support this. Conversely, these data

revealed that the TD group displayed greater amplitude in their models, i.e. stronger expressions. This is surprising given that previous research has suggested individuals with ASD may require a stronger facial expression in order to recognise it, or at least that they are better at recognising emotions that have a higher amplitude (e.g. Law Smith et al. (2010)). The results of this study also did not indicate any temporal bias for individuals with ASD as was suggested in the original hypotheses. Hence, this finding supports the first set of results reported in the current study that individuals with ASD are not significantly more ‘atypical’ in FER compared to TD individuals. These findings go somewhat against previous research that has suggested that adults with ASD have a slower comprehension of facial expressions compared to TD adults (Enticott et al., 2014; Sachse et al., 2014). In particular, since these data are reported in the context of information from the other aspects of FER (i.e. AUs), the current results also go against previous research that has suggested that reaction time is slower, despite intact FER accuracy (Leung et al., 2013). Because there is only a small amount of research that has looked specifically at the temporal domain of FER in ASD, these results warrant further research on the temporal characteristics of FER in individuals with ASD.

In addition to this, no significant differences were found between the ASD and TD group in their use of eye and mouth region of the face, which counters the original hypothesis that the ASD group will use information from the mouth area of the face in favour of information from the eyes. This provides evidence against claims that individuals with ASD utilise information from the mouth area over and above the eye area (Spezio et al., 2007a). However, in the current study information from both the eyes and mouth were presented in conjunction with each other (i.e. the whole face was presented) and so future research would aim to assess perception from the eye-only stimuli and mouth-only stimuli in order to ratify this finding.

### **2.4.1 Limitations**

Despite having very rich individual data, the study is limited in its overall sample size ( $n = 22$ ). The present study required a large time commitment from the participant (6 hours in total spread over several sessions) and therefore the strengths of this study lie in the richness of the data collected per individual but a drawback of this is the difficulty in obtaining large numbers of participants. In addition to this, the present study may be limited due to the fact that a proportion of the participants were recruited in the UK, while the rest were recruited in the U.S. The reasons for using this method of recruitment were

due to complications regarding the international collaboration involved in this project. Specifically, resources for recruitment of typically developed adults were only partially available at Caltech, and hence, in order to manage this issue, a proportion of the typically developed adults recruited for the present study were recruited in the UK.

In terms of general limitations, it could also be argued that, as adults, individuals in the ASD group may have learnt compensatory strategies to decode facial expression of emotion using a 'rules-based' approach (Walsh et al., 2014). Hence, these data do not rule out the idea that, in an experimental setting, adults with ASD may perform well yet impairments are still experienced in 'real life' emotion perception.

The task used in the present study adopted a 'seven – alternative forced choice task' (7AFC) which allowed participants to select 'other' if the six basic emotions were not deemed suitable to label the GFG stimulus presented. This was built into the method of the study as a mechanism to avoid participants randomly assigning an emotion label to a stimulus that was perceived as arbitrary (or as an emotion not within the 6 basic labels given). Thus, by having the option of 'other', the amount of noise introduced into the 6 emotion categorizations was limited. However, there are drawbacks to using this method. For example, it transpired that the ASD group selected a greater number of 'other' responses as compared to the TD group. It is possible that the greater number of 'other' responses selected by the ASD group may have had an effect on the outcomes of the study. However, it is important to note that due to the large number of trials (26,400 trials per group) any influence from an increased 'other' response from the ASD group is likely to be marginal.

Another possible limitation to this study is the reporting of results from selected individuals. In this study, an in depth analysis was conducted, assessing both group differences and individual differences. Individual differences were highlighted for two reasons: 1. To display the benefits of the novel technique presented in this study i.e. the level of individual detail presented has not previously been achieved in the ASD FER field and 2. because individual data are rarely reported in this field despite calls for more transparency regarding group heterogeneity in ASD (Harms et al., 2010; Uljarevic & Hamilton, 2013).

### **2.4.2 Conclusion**

In conclusion, this study provides a comprehensive, individual representation of the key components associated with the 6 basic emotion categories (happy, surprise, fear, disgust, anger and sadness). These rich data allow for an assessment of the AUs, amplitude and timing components that are most important in emotion categorization. Results revealed that, despite a comprehensive, fine-grained assessment of FER, adults with ASD are not significantly different from TD adults. These themes are further explored in Chapter 3.

## Chapter 3 Facial Expression Recognition in Children with ASD

### 3.1 Introduction

For typical children, facial expression recognition begins in early infancy and develops rapidly throughout the first year of life (Grossmann, 2010). Although some evidence of facial expression perception has been reported in neonates (i.e. < 4 weeks old, Farroni et al. (2007)), there is a general consensus in the research literature that infants begin to recognise facial expressions around 7 months of age (Caron et al., 1982; Nelson, 1987; WalkerAndrews, 1997). Compared to typical children, evidence has suggested that children with Autism Spectrum Disorder (ASD) have difficulty with face perception (Campatelli et al., 2013), and in particular, facial expression recognition (FER) (e.g. Davies et al., 1994; Gross, 2004; Hobson, 1986a; B. T. Williams & Gray, 2013b; see section 1.4 for further discussion of this). How children with ASD perceive facial expressions has been intensively researched for over 30 years (Harms et al., 2010). As discussed in Chapter 1, this is somewhat unsurprising given that ASD is characterised by impairments in social interaction and communication (*DSM-5*, 2013) and that the face is arguably the most important cue to an individual's thoughts and feelings (and hence crucial for effective social interaction and communication). Despite over thirty years of research dedicated to facial expression recognition (FER) in children with ASD, conclusions from this field of research still paint a confusing picture (Collin, Bindra, Raju, Gillberg, & Minnis, 2013; Harms et al., 2010; Simmons et al., 2009). The volume of conflicting results is, consequently, also paired with criticism that non-significant findings may be less likely to be reported in this field (Nuske, Vivanti, & Dissanayake, 2013). This has resulted in an atmosphere of confusion and an appetite for a better explanation of FER in individuals with ASD.

It has been argued that, as individuals with ASD grow older, and hence are exposed to support, training and interventions in relation to their diagnosis, compensatory mechanisms can be learned in order to improve FER ability (Walsh et al., 2014). Still, there is a large volume of conflicting findings in research on FER, specifically in children with ASD, with some papers indicating a significant 'impairment' in FER in this population while others indicate no impairment relative to TD individuals. Evidence of impairment includes: (Ashwin, Baron-Cohen, Wheelwright, O'Riordan, & Bullmore, 2007; Balia et al., 2014; Braverman et al., 1989; Capps et al., 1992; Celani et al., 1999; Davies et al., 1994; Domes,

Kumbier, Heinrichs, & Herpertz, 2014; Eack et al., 2015; Evers et al., 2015; Feldman et al., 1993; Golan, Sinai-Gavrilov, & Baron-Cohen, 2015; Gross, 2004; Grossman, Klin, Carter, & Volkmar, 2000; Hobson, 1986a; Lindner & Rosen, 2006; Loukusa, Makinen, Kuusikko-Gauffin, Ebeling, & Moilanen, 2014; Ozonoff et al., 1990; Peterson et al., 2015; Rutherford, Walsh, & Creighton, 2015; Sachse et al., 2014; Tantam et al., 1989). Evidence of no impairment includes: (Bormannkischkel et al., 1995; Castelli, 2005; Evers et al., 2014; Fein et al., 1992; Gepner et al., 2001; Lacroix, Guidetti, Roe, & Reilly, 2014; Loveland et al., 1997; Prior et al., 1990; Robel et al., 2004; Tell & Davidson, 2015).

One of the limiting factors in FER research, particularly within the ASD field, is the limited scope of the methods used (Uljarevic & Hamilton, 2013). Despite a movement towards more advanced methods, very few studies within the FER field have moved beyond using stimuli of basic static photos, avatars or actors to assess FER ability in ASD (Simmons et al., 2009). In addition to this, several studies suggest intensity (i.e. the ‘amplitude’ of the expression) is a key, yet understudied, variable in FER (Law Smith et al., 2010; Mazefsky & Oswald, 2007; G. L. Wallace et al., 2011). Despite the likelihood of the intensity of facial expressions having a contributory influence on FER ability, very few studies to date have aimed to, or have had the methodological capacity to, directly measure the contribution of facial expression intensity. In addition to this, timing of emotional expression has been indicated as influencing FER in children with ASD. For example, Tardif, Laine, Rodriguez, and Gepner (2007) found that FER was significantly enhanced in children with ASD when the facial expression was slowed down. The effects of timing on FER are, however, not often examined specifically in FER in ASD research.

The aim of this study was therefore, to utilise a novel technique to assess FER in children with ASD. In particular, nuances in FER that may be found between children with ASD and typical children that might not otherwise be picked up by different methods. It was hypothesised that children with ASD would respond in an inconsistent/atypical manner and so their responses would be more ‘atypical’ than the TD children. In addition to this, it was hypothesised that children with ASD would, in comparison to TD individuals, require the stimuli to be more ‘obvious’ (i.e. greater intensity of facial components) in order to categorise the stimulus as the target emotion. It was also hypothesised that for children with ASD they would require the stimuli to have greater intensity in order to categorise them as the target emotion. Lastly, it was hypothesised that children with ASD would display a lag in timing of FER. These hypotheses were drawn despite finding divergent

results in adults with ASD in section 2.3, because it was argued that children with ASD may show more pronounced FER difficulties compared to adults.

## **3.2 Methods**

### **3.2.1 Participants**

Data were collected from 16 children with ASD and 56 typically developing (TD) children aged between six and fourteen years old. An outline of these two sample groups is given in Table 3-1. All children were chaperoned throughout the study, meaning that they had an adult experimenter or assistant with them at all times to encourage the highest level of concentration possible from the child. All caregivers provided informed, written consent and all children gave informed, verbal assent before testing began. Ethical approval was granted from the University of Glasgow and the University of Victoria before the study commenced.

#### **3.2.1.1 Children with Autism Spectrum Disorder**

Children with ASD were recruited from the University of Victoria, Centre for Autism Research Technology and Education (CARTE) and MOSAIC Learning Society (a school for children with ASD) in Victoria, British Columbia. Twenty participants were recruited to take part in the experiment. Of these, two participants failed to turn up and another two participants, despite coming to the session, did not take part in the experiment. Sixteen participants completed the experiment. Of these, thirteen were boys and three were girls with a mean age of eleven years and two months. This sample had a higher ratio of boys than girls (4 boys to 1 girl). The reasons for this were twofold: 1. There were a greater number of boys with ASD who were willing to take part in research at CARTE and MOSAIC 2. It is largely recognised that, in the general population, there are more boys diagnosed with ASD than girls (Brugha et al. (2011); Kim et al. (2011)), with the majority of estimates across studies suggesting a 4:1 ratio of boys to girls (Fombonne, 2003, 2009). A debate about the exact gender ratio (as discussed in section 1.1) in the ASD population is, however, still ongoing (Kirkovski, Enticott, and Fitzgerald (2013); Mattila et al. (2011)). The inclusion criteria required children to be aged between six and fourteen years old, have normal or corrected-to-normal vision and have a diagnosis of autism or Asperger's syndrome based on standard ICD-10 (World Health Organization, 1993) criteria. All children were of western Caucasian ethnicity. Caregivers of ASD participants were required to bring evidence of the child's diagnosis to be viewed by the experimenter

before testing commenced, hence, although a re-test for ASD was not performed on-site by the experimenter, all of the children met the standard criteria for a clinical diagnosis of ASD.

### **3.2.1.2 Typically developing children**

Children who were typically developing (TD) were used as a control comparison group. Eighty-four children were recruited through the University of Glasgow and the Glasgow Science Centre (a popular, hands-on science museum in Glasgow). Caregivers completed the Autism Quotient-10 questionnaire (AQ-10; Allison, Auyeung, and Baron-Cohen (2012)), which is a shortened version of the Autism Spectrum Quotient (AQ; S. Baron-Cohen, Wheelwright, Skinner, et al. (2001) and often used in clinical settings as a referral tool for children suspected of having ASD (Allison et al., 2012). We used this to ensure that children in the TD group had low levels of autistic traits. IQ was also measured and using the Kaufman Brief Intelligence Test II (Kaufman & Kaufman, 2004). Of the total eighty-four children, five children were excluded from the analysis because their AQ score was higher than 5. From the remaining seventy-nine children, a sample of children who reflected the ASD group on age, gender, verbal IQ (VIQ), performance IQ (PIQ) and ethnicity (Western Caucasian) were selected. Of these, forty-five were boys and eleven were girls in order to match the gender ratio found in the ASD group as described above. Selections were made in order to make the ASD and TD group as similar as possible in terms of age, gender and IQ, but all experimental data analysis was conducted after the selection of participants had been made. The typically developing group had a mean age of ten years and two months. No significant differences were found between the two groups on age ( $t(70) = 1.31, p = 0.19$ ), verbal IQ ( $t(70) = 0.41, p = 0.68$ ), performance IQ ( $t(70) = 0.61, p = 0.54$ ) or full scale IQ ( $t(70) = 0.46, p = 0.63$ ). All TD children had an AQ-10 score of below 4 (i.e. below average) and all caregivers reported their children to have normal or corrected-to-normal vision.

**Table 3-1 Demographic information from children with ASD and TD children**

ASD (n = 16)				TD (n = 56)			
	Count	%		Count	%		
<b>Gender</b>	13 m	81%		46 m	81%		
	3 f	19%		11 f	19 %		
	Mean	SD	Range	Mean	SD	Range	T-test
<b>Age (yrs: mnths)</b>	11:08	01:08	07:07-14:02	11:00	02:11	6:02 - 14:07	$t(70) = 1.32, p = 0.19$
<b>VIQ</b>	103.1	23.4	60 - 141	105	13.4	72 - 136	$t(70) = 0.41, p = 0.68$
<b>PIQ</b>	103.5	13.9	73 - 124	106	13.5	84 - 131	$t(70) = 0.61, p = 0.54$
<b>FS IQ</b>	104.4	19.1	73 - 135	106	14.1	83 - 131	$t(70) = 0.46, p = 0.62$

### 3.2.2 Experimental set-up

The aim of the study was to use a novel, sensitive measure to assess the perception of facial expressions of emotion in children with ASD and TD children. To do this, a procedure similar to that described in Chapter 2 (section 2.2) was implemented, however, the procedure was altered to make the experiment more suitable for children by reducing the 7AFC task to a 2AFC task. Children were, therefore, asked only about two emotions: Happy and Angry. These emotions were chosen in order to tap into both positive and negative types of emotional expression. In addition to this, the task was made more ‘child-friendly’ by lowering the number of trials completed per participant (and therefore lowering the overall task length) and introducing cartoons and encouragement messages between blocks of trials to facilitate attention to the task. The experiment was set up in a ‘fun-day’ style environment consisting of games, activities and educational play, which the children took part in in addition to the experiment. In Victoria, this was located in the Centre for Autism Research, Technology and Education at the University of Victoria and was advertised as ‘Face Lab’. At Face Lab each child was paired with a ‘buddy’ for the day. Buddies consisted of CARTE’s director (Professor James Tanaka), CARTE’s Lab Coordinator (Bonnie Heptonstall), Kirsty Ainsworth (the author) and twelve third-year psychology student volunteers who had knowledge about ASD or had previous experience working with children with ASD. Before each Face Lab session, buddies were fully briefed

about the experimental task and the format of Face Lab. On the day, they were provided with a clipboard which consisted of a detailed outline of the child they would be working with, including any particular likes, dislikes or behavioral triggers. The clipboard also included a 'visual schedule' of the day that showed the child their exact time slot for each Face Lab activity: Face Game (the experiment), Recess (play room activities such as drawing, jigsaws, origami) and Puzzle Games (the Kaufman Brief Intelligence Test- II (Kbit-2)). Children could stamp their visual schedule after each session using stickers. In order to incentivise attention in the experimental task, children could earn tokens for each block of trials of Face Game they completed and were rewarded with prizes at the end of the task in exchange for their tokens. The buddy's role was to help the child with the experimental task, to oversee that data were collected in a consistent way, and to take care of the child to ensure they enjoyed their experience at Face Lab. Face Lab lasted for two hours and many of the children came back a second or third time at later dates. In Victoria, all children were rewarded with a \$10 book voucher for taking part.

In Glasgow the experimental set-up was developed to be as closely similar to Face Lab as possible. 'Brain Lab' was developed as a fun day for children to attend which involved games, activities and educational play. Children took part in Brain Lab at the Glasgow Science Centre on an ad-hoc basis so a specific visual schedule could not be provided. However, the sequence of tasks was very similar to CARTE insofar as children completed a combination of Face Game (the experimental task), activities (hands-on educational brain workshop) and puzzles (the Kbit-2). In Glasgow, sixteen volunteers were recruited to help out as 'buddies', most of whom were psychology undergraduates and postgraduates from the University of Glasgow, with an interest in working with children. In order to encourage attention to the experimental task, children earned tokens for each block of trials of Face Game they completed and were rewarded with prizes at the end of the task in exchange for their tokens. The experiment was also mobile in Glasgow, meaning that for children who were interested in taking part but could not make it to the Glasgow Science Centre, the experimenter could bring the task to the children in their own homes. When data were collected in the home of the child, instead of earning tokens for prizes, £10 cash was awarded to the child at the end of the experiment.

Due to constraints met with accessing participant samples, the experimenter had a relatively short time with typically developed children as compared to the time available with children with ASD. Due to these constraints, the ASD group completed a total of 19,000 trials and the TD group completed a total of 21,200 trials (see Table 3-2),

however, there were a larger number of participants in the TD group. In order to address this, data were aggregated for each group and data were assessed only using aggregated group-wise analyses, as further described in section 3.2.5. The limitations imposed on the results due to these circumstances are discussed in section 3.4.1.

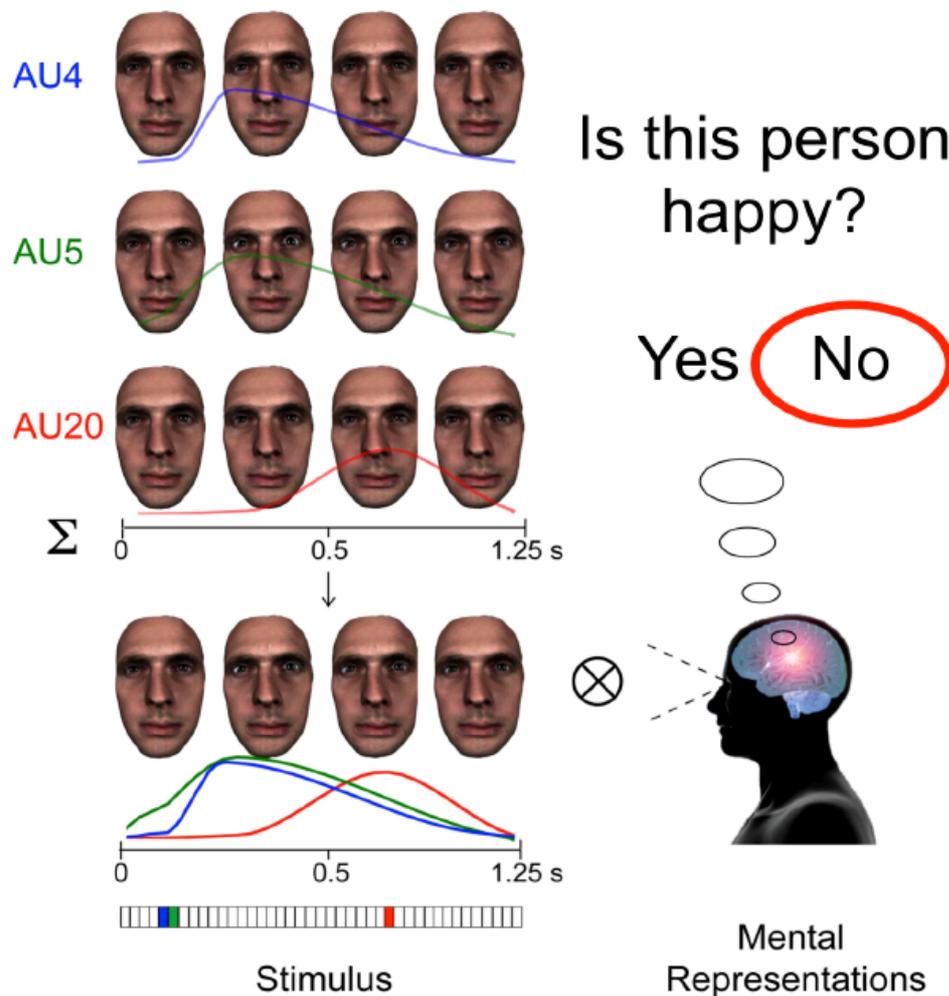
**Table 3-2 Trial Numbers of Aggregate Group Data**

	ASD	TD	Total
Happy	7,800	9,400	<b>17,200</b>
Angry	11,200	12,200	<b>23,400</b>
<b>Total</b>	<b>19,000</b>	<b>21,600</b>	<b>40,600</b>

### 3.2.3 Stimuli

The stimuli were produced using the Generative Face Grammar (GFG) developed by Yu and colleagues (Yu et al. (2012)). The GFG technique was used to create 2400 dynamic facial expression stimuli, which displayed a facial expression that moved in ‘random’ ways (i.e. like a ‘digital puppet’ that was being moved in different places on the face, at different times and with varying levels of intensity). The dynamic stimuli were constructed from a randomised selection of 42 possible facial muscle groups (called “action units” AUs; Ekman & Friesen, 1978) and 6 temporal parameters (onset latency, acceleration, peak latency, peak amplitude, deceleration, offset latency). For each stimulus generation the GFG selected a number of AUs from a binomial distribution (median = 3), for example in Figure 3–1 the GFG has selected AU 17 (Chin Raiser), AU 10 L (Upper Lip Raiser Left) and AU 9 (Nose Wrinkler). Then, for each AU of the stimulus, the GFG randomly assigns a value for the six temporal parameters: These values were taken from a uniform distribution (coloured lines on the figure). On each stimulus generation the GFG randomises its selection of AUs and temporal parameters but, importantly, the algorithm constrains the pattern of selected AUs and parameters so that the combination is one that is physically possible by the human face. For example it would be near impossible to express a downturned mouth (AU 15 Lip Corner Decompressor) with a lifting of the cheeks (AU 6 Cheek Raiser). Similarly, there is a point at which if the intensity value of an AU is too high it becomes non-human like (Katsyri, Forger, Makarainen, & Takala, 2015; Yu et al., 2012). The GFG, therefore, constrained the selection of AUs and temporal parameters to create a random, but physiologically plausible, facial expression stimulus that had no real

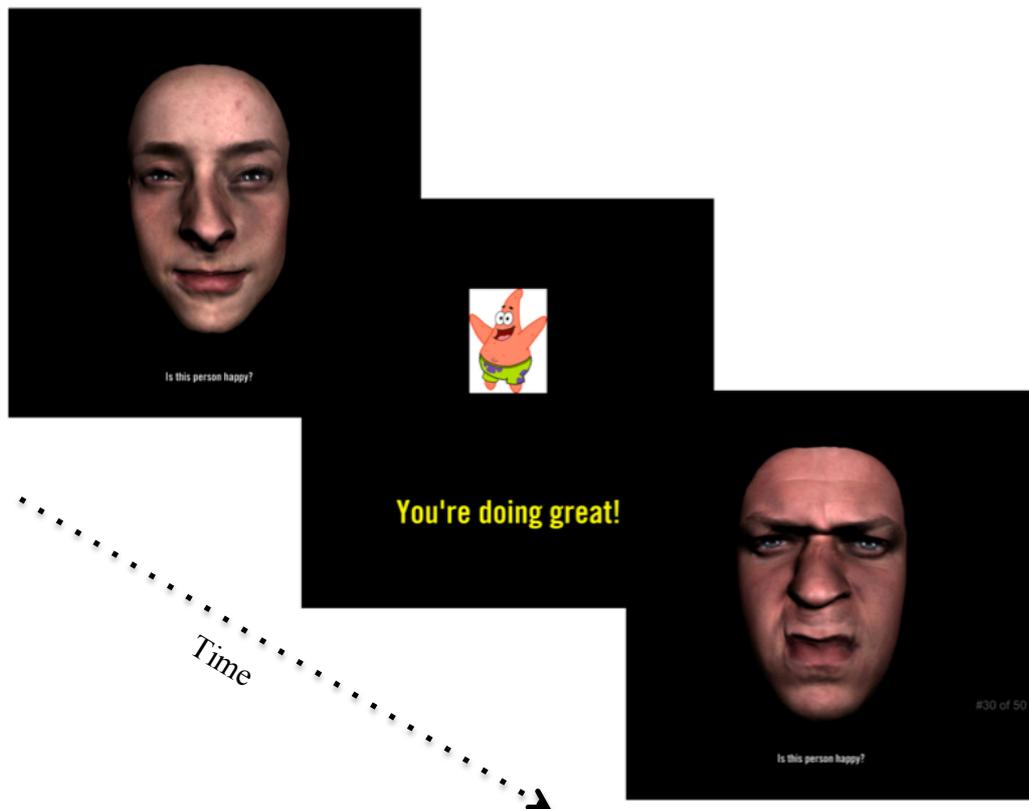
meaning/affect associated with it (i.e. each stimulus was a ‘random’ facial expression). This study used 2400 individual, dynamic stimuli generated by the GFG. Because each stimulus displayed random combinations of AUs and temporal dynamics, a stimulus will be referred to henceforth as a ‘random facial stimulus’ (note here the adjective ‘random’ is descriptive of the random nature in which the features and timing of the facial expression move). The 2400 stimuli used in this experiment were identical to those used in Chapter 2. The simplification of the study required that only two emotions could be assessed (because to the large number of trials needed). The study sought to assess both positive and negative emotions, therefore happy was chosen to assess positive emotion FER and anger was chosen to assess negative emotion FER.



**Figure 3–1 Stimulus Generation using the Generative Face Grammar (Yu et al. 2012). For details see text (section 3.2.3).**

### 3.2.4 Procedure

After taking the IQ test (K-Bit 2) the children were introduced to the experiment. The task of each participant was to categorise each facial expression stimulus using computer keys labeled ‘yes’ or ‘no’ in response to the question displayed below the stimulus (in the happy condition this was ‘is this person happy?’, and in the angry condition this was ‘is this person angry?’). The order of stimuli and order of condition was randomised across participants. Before the task commenced, the experimenter or assistant working with the child explained that it was important to look at the way the face moves and not what the person looks like. All efforts were made to ensure that the child focused on the dynamic expression that was produced by the stimulus and not the resting expression of the face, or the identity of the person. Stimuli were interleaved with cartoon images displaying motivational messages (e.g. ‘good job!’; ‘keep it up!’) in order to encourage attention to the task. Stimuli were displayed on a computer monitor 60 cm from the child’s face, as measured by a distance meter (a 60cm measure of string), which was attached to the base of the computer monitor. Children touched the end of the string to their nose before testing commenced and were asked to remain in the same position throughout the experiment. Each facial expression stimulus lasted 1.25 seconds, starting and ending in a neutral expression. The stimulus remained on the screen until the child responded.



**Figure 3–2 Example of two trials within the experimental task. Random facial expressions lasted 1.25 seconds (displayed here as a snapshot at 0.75 seconds). The facial expression trials were interleaved with cartoon and motivational messages.**

### 3.2.5 Analysis

The original set of 2400 stimuli (as described in Section 3.2.3) were, when in numerical format, a matrix of 3 dimensions: size  $A \times T \times S$  where  $A$  indicates the AUs that are active (a binary indicator for each AU: 1 to 42),  $T$  indicates time (1.25 seconds is split into 30 even frames) and  $S$  is 2400 stimuli. The values in the matrix are amplitude values i.e. values from 0 to 1 representing the intensity of the stimulus. During the experiment, stimuli are categorised into datasets of ‘yes’ or ‘no’ for happy condition (‘Is this person happy?’) and also for the angry condition (‘Is this person angry?’). Therefore, there were four matrices of raw data, each containing the set of stimuli for which participants selected ‘yes’. One dataset for the ASD group happy condition (‘Is this person happy?’), one for the ASD group angry condition (‘Is this person angry?’), one for the TD group happy condition (‘Is this person happy?’), and one for the TD group angry condition (‘Is this person angry?’). The input to the analysis was therefore four separate datasets of the size  $A \times T \times N$  where  $A \times T$  is a stimulus in numerical format (as above) and  $N$  is the number of stimuli. The number of stimuli was different for each dataset because a different number of stimuli were found to fit a ‘yes’ response for each group/condition (further details given in section 3.2.3). The aim of the analysis was to reveal the patterns in AU activations and intensity values that were consistently present in the stimulus when participants responded ‘yes’. In other words, what aspects of the stimuli were important in order for a ‘yes’ response to occur? Because of the nature of the stimuli, non-negative matrix factorization (NMF) was identified as an ideal analysis tool to answer this question, as detailed below.

Prior to applying NMF, the raw data (‘yes’, ‘no’ responses) and their associated stimuli (AU number and intensity values across time) were refined into a smaller number of data points, consisting only of the commonly agreed categorizations. This step was necessary in order to identify the stimulus-response pairs that were not only representative of an individual but also representative of the group as a whole. A group-wise mean average response was calculated for each stimulus where ‘yes’ response = 1 and ‘no’ response = 2 (i.e. for each stimulus the sum of the responses divided by the number of participants who viewed that stimulus). A threshold was selected so that the stimuli that had an average response reaching  $<1.2$  (i.e. on average close to a 1 (‘yes’) response) were inputted into a dataset of ‘selected’ stimuli. Stimuli that had an average response  $>1.8$  (i.e. on average

close to a 2 ('no') response) were inputted into a dataset of 'rejected' stimuli. Only the 'selected' stimuli datasets were used in the data analysis.

### 3.2.5.1 Non-negative Matrix Factorization (NMF)

Non-negative matrix factorization (NMF) is a dimensionality reduction technique, argued to be particularly useful for analysing datasets of faces, in comparison to older techniques of dimensionality reduction such as Principle Components Analysis (PCA) and Vector Quantization (Lee & Seung, 1999). NMF uses a parts-based algorithm (i.e. determining the data structure via individual components) as opposed to a holistic pattern (determining the data structure using a group of many different parts combined together). This enables the user to gain information about the *components* of the face, which is ideal for this particular dataset. When applied to a dataset of face stimuli, NMF is essentially attempting to answer the question 'what are the key ingredients of the stimuli that received a 'yes' response?'. In order to reveal the 'key ingredients' of the dataset, NMF reduces the dimensions of the dataset to a linear combination of bases and weights. In other words, NMF approximates dataset  $V$  into two smaller datasets  $W$  and  $H$ , where  $V \approx W * H$ .  $W$  is a  $m \times r$  matrix of 'bases' where  $m$  is equal to the rows in  $V$ .  $H$  is a  $n \times r$  matrix of 'weights' where  $n$  is equal to the columns in  $V$ .

Equation 1 - Non-Negative Matrix Factorization

$$V_{i\mu} \approx (WH)_{i\mu} = \sum_{a=1}^r W_{ia} H_{a\mu}$$

The value  $r$  is set by the user and provides boundaries for NMF. Therefore, NMF uses  $W$  and  $H$  to provide a simpler, dimensionally-reduced version of the original matrix  $V$ .  $W*H$  can also be thought of as a 'compressed' version of  $V$  (Lee & Seung, 1999) i.e. by weighting the bases and adding them together NMF provides an approximation of  $V$ . In order to calculate the closest approximation of  $V$ , NMF applies gradient descent (Snyman, 2005): minimizing the squared distance of  $V - W * H$  until a tolerance limit is met.

In this study the decomposition output from NMF details the AUs and intensity values that complement each other. For example, if a decomposition revealed AU 6 (Cheek Raiser) and AU 12 (Lip Corner Puller) for the TD happy condition it indicates that when a

stimulus displayed ‘Cheek Raiser’ and ‘Lip Corner Puller’ together, this *combination* of AUs was important to selecting ‘yes’ to ‘Is this person happy?’. The value ‘r’ (set by the user) specifies the number of decompositions NMF will attempt to approximate the dataset with (this is similar to cluster analysis when the user specifies the number of clusters the dataset is desired to be broken into). Here, r was specified as 3 in order to reveal a snapshot of the underlying structure of the dataset whilst gaining information on the pertinent combinations of AUs and timing parameters. In other words, each of the 3 decompositions contains a representation of the ‘key ingredients’ present in the stimuli which elicited a ‘yes’ response. NMF output for this dataset is an ‘A x T x 3’ matrix containing intensity values of each AU (A) across 30 timeframes (T) for each decomposition (of which there are 3). NMF was, therefore, deemed a particularly well suited analysis tool for this dataset and has recently been demonstrated to perform well on data similar to these (e.g. (Delis et al., 2016)).

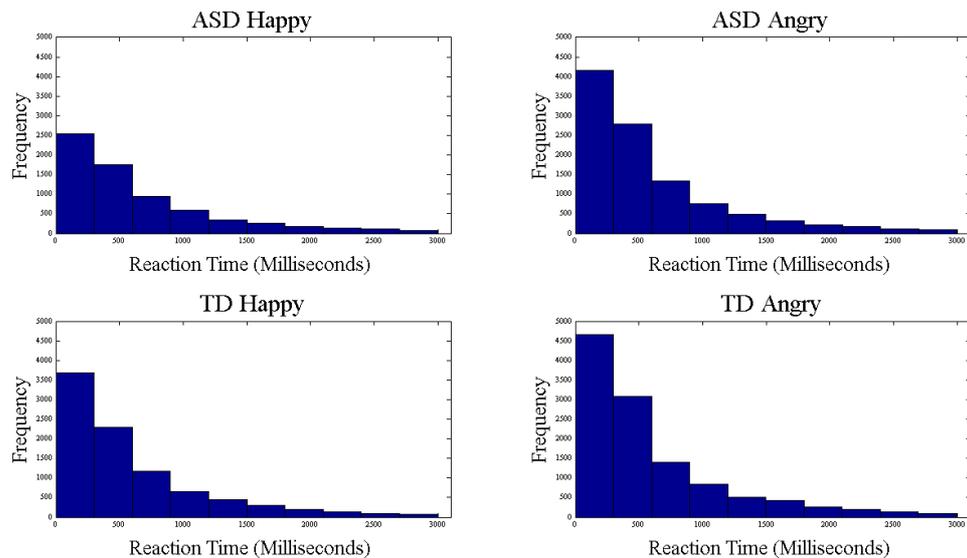
### **3.3 Results**

#### **3.3.1 Response characteristics**

##### **3.3.1.1 Reaction time**

Since task performance on facial expression recognition tasks can be influenced by reaction time (Harms et al., 2010), an initial analysis comparing reaction time between ASD and TD groups was conducted. It was expected that children with ASD might take more time over their judgements of the stimuli due to previous research findings that children with ASD are slower than TD children on FER tasks (Leung et al., 2013). Although this task was not primarily a reaction time (RT) task (i.e. children were asked to ‘answer quickly’ but were not told their RTs were being recorded), RT was recorded for each trial, starting as soon as the stimulus finished and ending as soon as the participant pressed a valid response key (‘yes’ or ‘no’). In order to control for outliers, RTs that were greater than 3000ms were eliminated. 3000ms was deemed an appropriate cut-off point as it was apparent that reactions after 3000ms were due to the participant being distracted, and not reflective of a true response (based on reports from the helpers who were paired with each child during experimentation). This cut-off was implemented following established guidelines from Whelan (2008), who claims that when RT are positively skewed (as they are in this case) trimming RTs from a point relative to the standard deviation is not appropriate due to the biased distribution (i.e. the mean of all datasets does not reflect the positive skew). The median was also not an appropriate measure of central

tendency for this dataset because there were a different number of trials in each group (Whelan, 2008). Therefore, the absolute cut-off point for RTs being  $>100\text{ms}$  and  $< 3000\text{ms}$  was deemed appropriate for this sample. Figure 3–3 displays the distribution of reaction times (in milliseconds) after the aforementioned trimming procedure was conducted. The figure suggests that the distribution of trimmed RTs is very similar across all groups and conditions, despite there being varying numbers of trials in each.



**Figure 3–3 Reaction time trimmed ( $<3000\text{ms}$ ) for all trials**

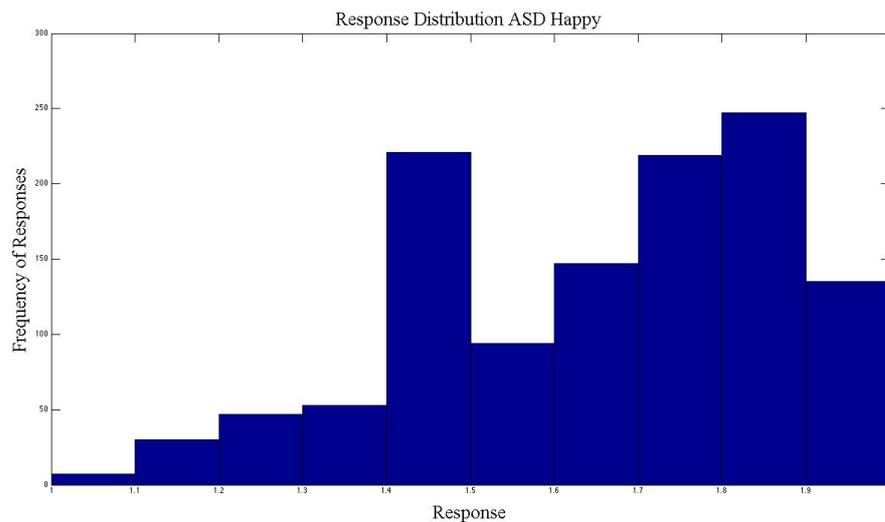
### 3.3.1.2 Response distribution: groups

In order to obtain a better understanding of the way in which each group responded for each condition, descriptive information was obtained from the experimental responses. It was hypothesised that children with ASD will, in comparison to the TD group, require the stimuli to be more ‘obvious’ (i.e. greater intensity of facial components) in order to categorise them as the target emotion. An indicator of requiring more intense stimuli was the proportion of stimuli selected (i.e. responding ‘yes’) compared to stimuli rejected (i.e. responding ‘no’). Being sensitive only to very intense stimuli would likely result in the participant selecting rather few (i.e. only intense) stimuli as ‘yes’ and therefore rejecting a larger number of stimuli (i.e. answering ‘no’). To assess this, the percentage number of yes and no responses was computed for each group. As expected, the ASD group did answer ‘no’ more than they answered ‘yes’, however this was to a greater extent in the happy condition (33% yes, 67% no) than the angry condition (49% yes, 51% no). The TD group answered ‘no’ more often than ‘yes’ for both conditions (see Table 3-3).

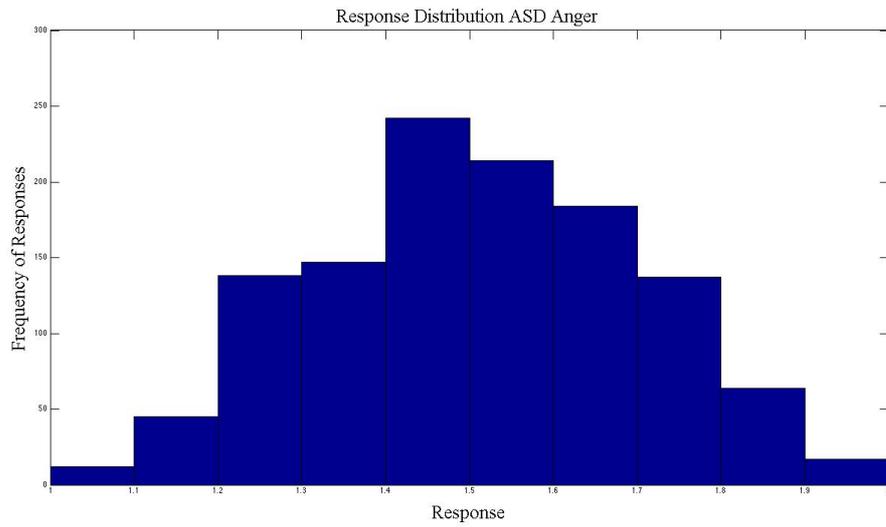
Table 3-3 Percentage of responses for each condition

	ASD		TD	
	Yes	No	Yes	No
Happy	33%	67%	36%	64%
Angry	49%	51%	39%	61%

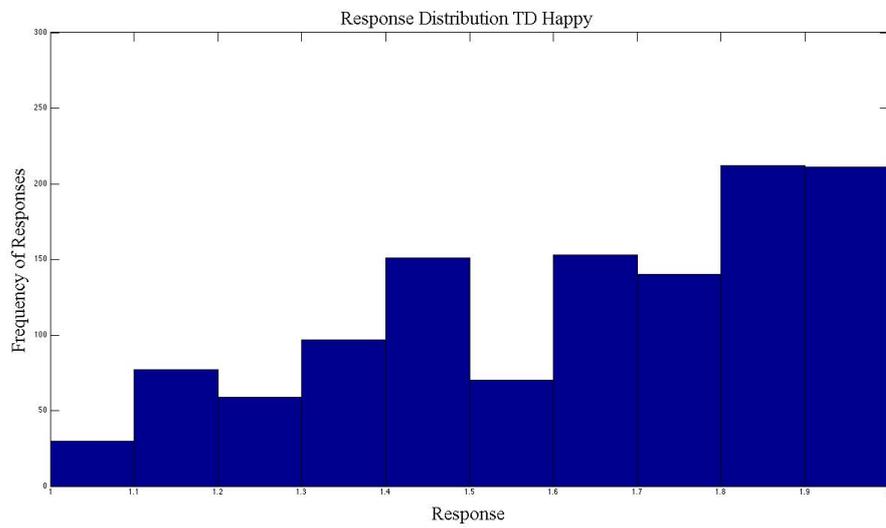
In order to deconstruct, and further assess, the response data displayed in Table 3-3, the distribution of average scores was computed and visualised in Figure 3-4, Figure 3-5, Figure 3-6 and Figure 3-7. Average scores were computed by taking the mean of the total responses to each stimulus: the sum of all participants' responses (sum of '1' = yes or '2' = no), divided by the number of participants that viewed that stimulus. Hence, a mean response score (a value between 1 and 2) was computed for each video stimulus. The frequency of all mean response scores were plotted to assess the distribution of responses. Plots that skew to the right indicate a group preference towards responding 'no' (i.e. close to '2' response) and plots that skew to the left indicate a group preference towards responding to 'yes' (i.e. closer to a '1' response). Figure 3-4 and Figure 3-5 suggest that average responses of the ASD group tended towards 'no' for the happy condition but reflected a normal distribution (not biased to 'yes' or 'no' responses) for the angry condition. Figure 3-6 and Figure 3-7 indicate that the TD group tended towards 'no' for the happy condition and for the angry condition. This supports the group/condition response pattern indicated in Table 3-3.



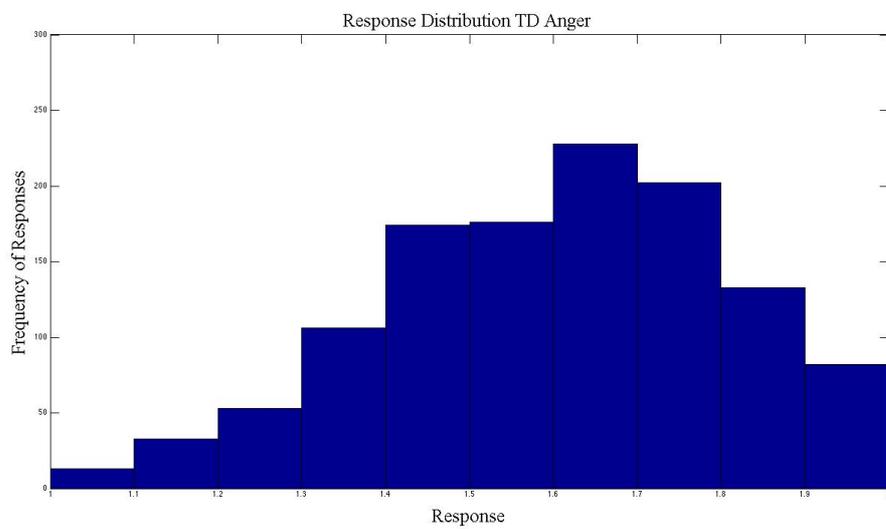
**Figure 3-4 Response distribution of children with ASD in the happy condition.**



**Figure 3–5 Response distribution of children with ASD in the angry condition.**



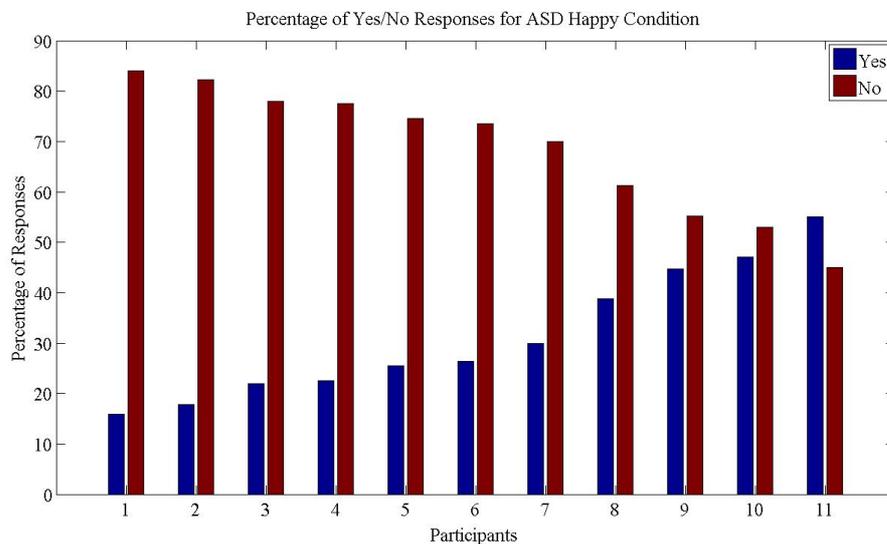
**Figure 3–6 Response distribution of TD group in the happy condition.**



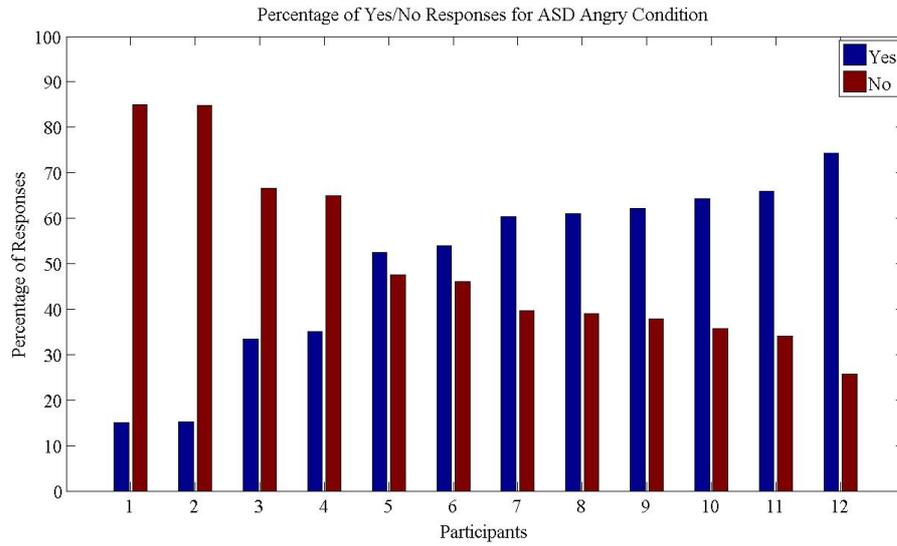
**Figure 3–7 Response distribution of TD group in the angry condition.**

### 3.3.1.3 Response distribution: individuals

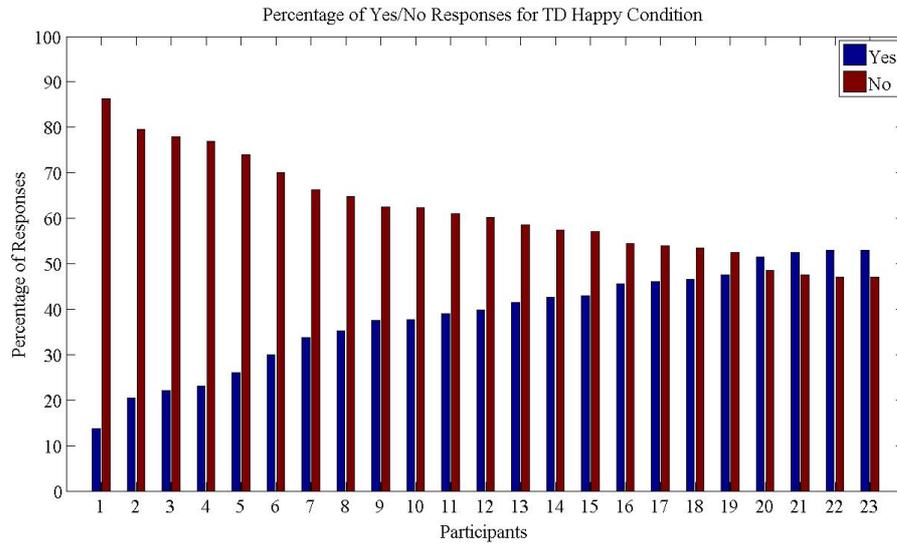
In order to assess whether specific individuals may have influenced results reported in section 3.3.1.2, the percentage of yes/no responses for each condition were assessed per individual. Figure 3–8, Figure 3–9, Figure 3–10 and Figure 3–11 indicate that individuals in both the ASD and TD group have similar patterns of responses for each condition. Therefore, there is no evidence to suggest that there is a specific ‘responding pattern’ characteristic only to the ASD group. Note, figures indicate  $n = 11$  in the happy condition and  $n = 12$  in the angry condition. This is due to slight variation in the number of trials completed by each participant; hence some participants completed only one of the conditions. We do find, however, that for both ASD and TD groups there are a proportion of people who are very ‘choosy’ (i.e. they say ‘yes’ very selectively) and a proportion of people who say ‘yes’ and ‘no’ to almost equivalent levels. It should also be noted that for both groups, there were a proportion of individuals willing to select many ‘yes’ responses in the angry condition despite this not being found for the happy condition. This suggests that a selection of ‘yes’ over ‘no’ response in the happy condition is harder to do, and may require more specific features from the random facial expressions, while angry can be attributed to a larger number of random facial expressions.



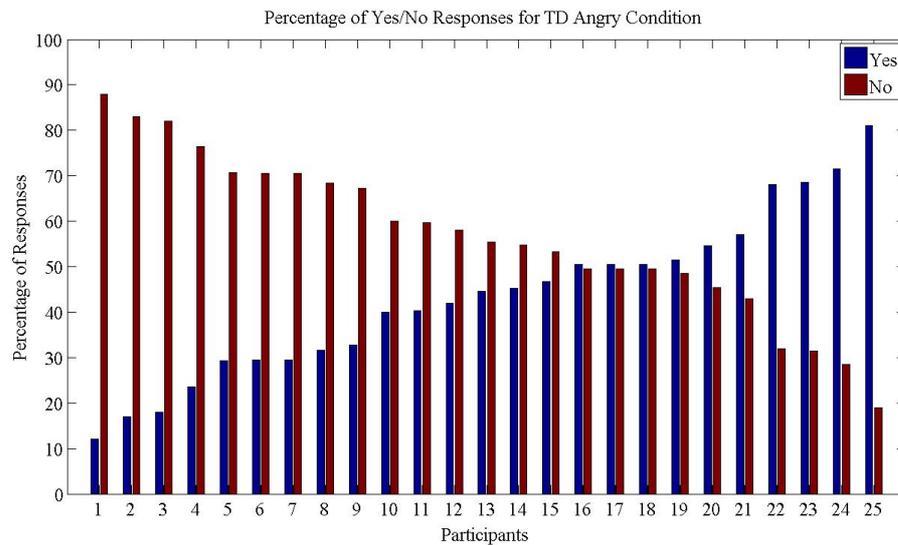
**Figure 3–8 Percentage of ‘yes’ and ‘no’ responses from ASD participants in the happy condition i.e. ‘Is this person happy?’**



**Figure 3–9 Percentage of ‘yes’ and ‘no’ responses from ASD participants in the angry condition i.e. ‘Is this person angry?’**



**Figure 3–10 Percentage of ‘yes’ and ‘no’ responses from TD participants in the happy condition i.e. ‘Is this person happy?’**



**Figure 3–11 Percentage of ‘yes’ and ‘no’ responses from TD participants in the angry condition i.e. ‘Is this person angry?’**

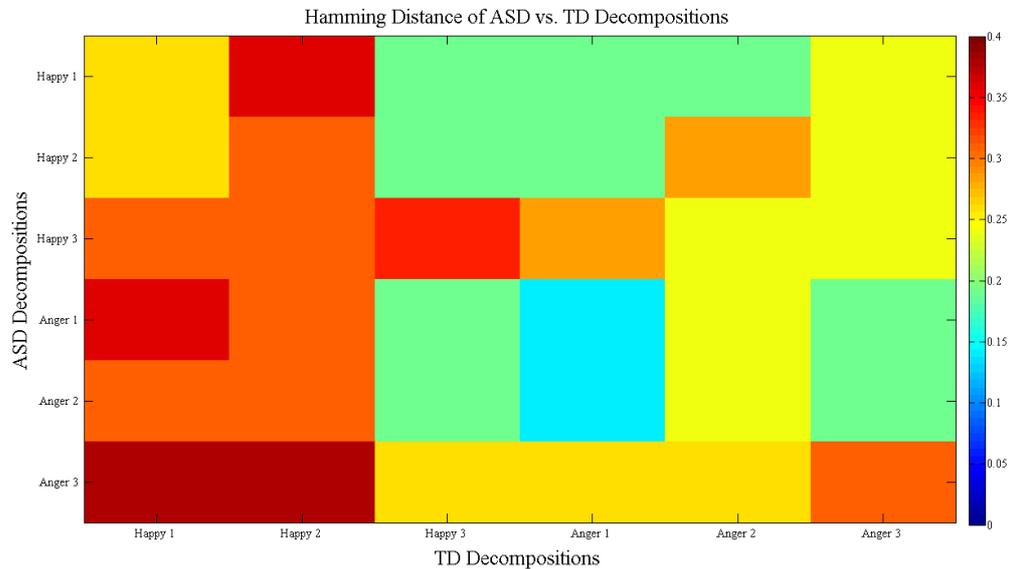
### 3.3.2 ASD and TD decompositions derived from NMF

As described in section 3.2.5, the raw data were analysed using a non-negative matrix factorization (NMF) technique. The result of this was a large matrix of AU, timing and intensity information for each group (ASD and TD), and for each emotion (happy and angry) decomposition. Hence, NMF output for this dataset was a ‘A x T x 3’ matrix containing intensity values of each AU (A), across 30 timeframes (T), for each decomposition (of which there are 3). Decompositions were rendered into a format that can be viewed in the context of the face (i.e. face avatars) and can be viewed in video format (see Appendix 4).

#### 3.3.2.1 Atypicality of AUs in the decompositions

It was hypothesised that children with ASD would respond in an inconsistent or atypical manner to the GFG random stimuli and so the underlying AUs that make up their ‘yes’ responses would be atypical in comparison to the TD children. In order to compare the AUs of the decompositions, hamming distance was applied to measure the distance between each binary vector. Figure 3-12 displays distance values in the form of a heat map where cool colours (i.e. dark blue) indicate distance values close to zero and hot colours (i.e. dark red) indicate higher distance values. Highest distance values were found between ASD angry and TD happy decompositions. Similarly, relatively high distance values were found between ASD happy and TD happy decompositions. The most similar decompositions in terms of AUs were TD angry and ASD angry. Hence,

these data suggest happy decompositions are least like each other while angry decompositions are most like each other. However, note that the maximum distance value between the two groups is 0.4 despite the maximum possible distance being 1.0, and so, it could be argued that although some differences can be found between the groups, these are relatively small.



**Figure 3–12 Hamming distance between ASD and TD decompositions.**

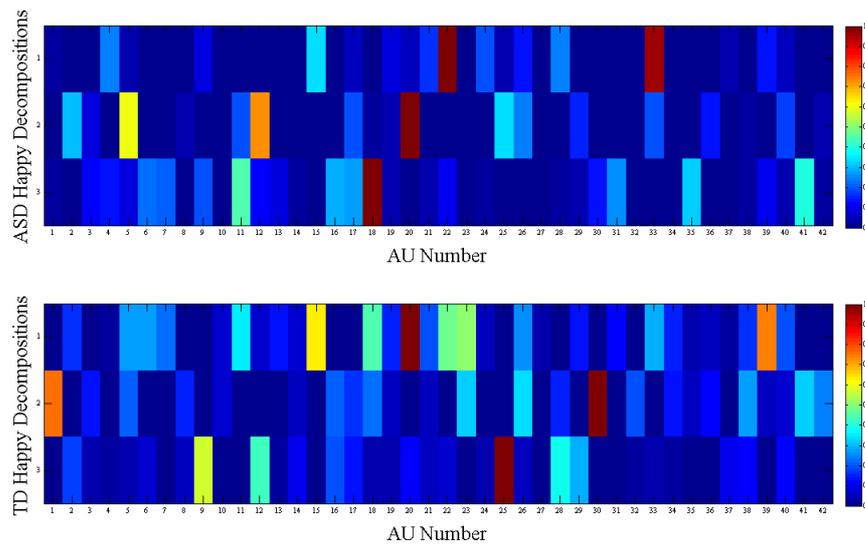
### 3.3.2.2 Amplitude of decompositions

Decompositions were also assessed for variance in amplitude values (i.e. the ‘intensity’ of the AUs). Intensity values of AUs were between 0 and 1, where 0 is no intensity (i.e. the AU is not active) and 1 is maximum intensity, which, controlled by the GFG (see Section 3.2.3), remained within the bounds of what is physiologically possible (meaning a maximum intensity of ‘1’ is still human-like). Figure 3–13 and Figure 3–14 indicate at least 1 AU reaches maximum intensity (red) in each decomposition across all conditions/groups. The figures also illustrate that there are fewer medium or high intensity (yellow or orange) AUs in the angry condition, compared to the happy condition (for both groups). The key in Table 3-4 can be used to find the associated AU number and AU name for the AU cell (i.e. numbers given along the x-axis in Figure 3-10 and Figure 3–14). AUs that reached intensities in the upper quartile ( $> 0.75$ ) in the happy condition were, for ASD, Dimpler, Lip Tightener, Nose Wrinkler, Lip Corner Puller and Lip Corner Puller Right and for the TD group, Inner Brow Raiser, Lip Corner Puller Left, Dimpler Right, Lip Stretch Left and Nostril Compressor. Ekman and Friesen (1978) suggest in their taxonomy of facial expression that, in a happy expression, the AUs Cheek Raiser and Lip Corner Puller

must be present. Therefore, two AUs that are particularly dissimilar to Ekman and Friesen's happy expression were found in the ASD group (Lip Tightener, Nose Wrinkler), while one that was particularly dissimilar to happy was found in the TD group (nostril compressor). These data suggest that, although the ASD group's more intensive AUs are not completely dissimilar from what would be expected to be seen in a happy expression (according to Ekman and Friesen (1978)), they are not as representative of this 'expected' happy when compared to the most intensive AUs found in the TD group. Hence it could be suggested from these data that there is more confusion in the ASD group compared to the TD group, but there is little evidence that the ASD group are completely impaired in their perception of the emotion happy.

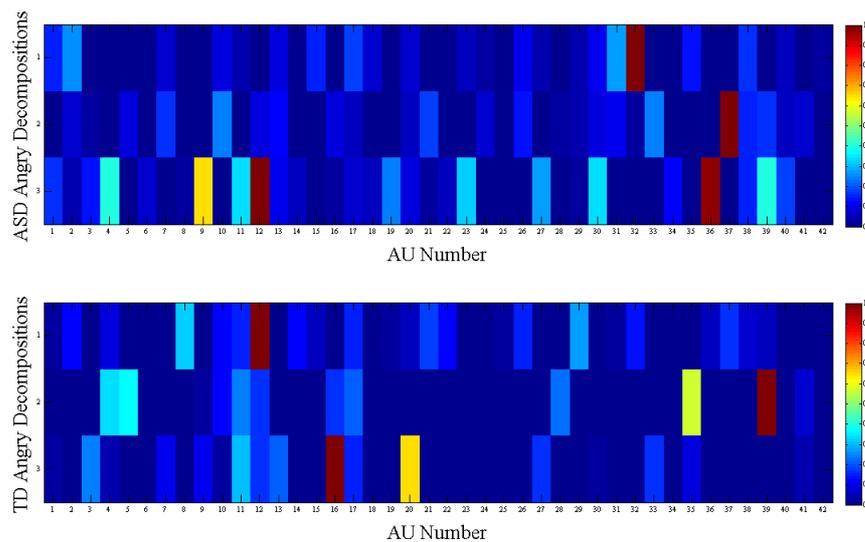
Table 3-4 Action Unit Number and Associated AU Name

Cell Number	AU Number	AU Name
1	AU1'	Inner Brow Raiser
2	'AU1-2'	Inner Brow Raiser - Outer Brow Raiser
3	'AU2'	Outer Brow Raiser
4	'AU2L'	Outer Brow Raiser Left
5	'AU4'	Brow Lowerer
6	'AU5'	Upper Lid Raiser
7	'AU6'	Cheek Raiser
8	'AU6L'	Cheek Raiser Left
9	'AU6R'	Cheek Raiser Right
10	'AU7L'	Eye Lid Tightener Left
11	'AU7R'	Eye Lid Tightener Right
12	'AU9'	Nose Wrinkler
13	AU10 Open'	Upper Lip Raiser Open Mouth
14	AU10L Open'	Upper Lip Raiser Open Mouth Left
15	AU10R Open'	Upper Lip Raiser Open Mouth Right
16	'AU11L'	Nasolabial Deepener Left
17	'AU11R'	Nasolabial Deepener Right
18	'AU12'	Lip Corner Puller
19	'AU25-12'	Lips Part - Lip Corner Puller
20	'AU12L'	Lip Corner Puller Left
21	'AU12R'	Lip Corner Puller Right
22	'AU13'	Cheek Puffer
23	'AU14'	Dimpler
24	'AU14L'	Dimpler Left
25	'AU14R'	Dimpler Right
26	'AU15'	Lip Corner Depressor
27	'AU16Open'	Lower Lip Depressor Lips Open
28	'AU17'	Chin Raiser
29	'AU20'	Lip Stretch
30	'AU20L'	Lip Stretch Left
31	'AU20R'	Lip Stretch Right
32	'AU22'	Lip Funneler
33	'AU23'	Lip Tightener
34	'AU24'	Lip Pressor
35	'AU25'	Lips Part
36	'AU26'	Jaw Drop
37	AU27'	Mouth Stretch
38	'AU38'	Nostril Dilator
39	'AU39'	Nostril Compressor
40	'AU43'	Eyes Closed
41	'AU7'	Eye Lid Tightener
42	'AU12-6'	Lip Corner Puller - Cheek Raiser

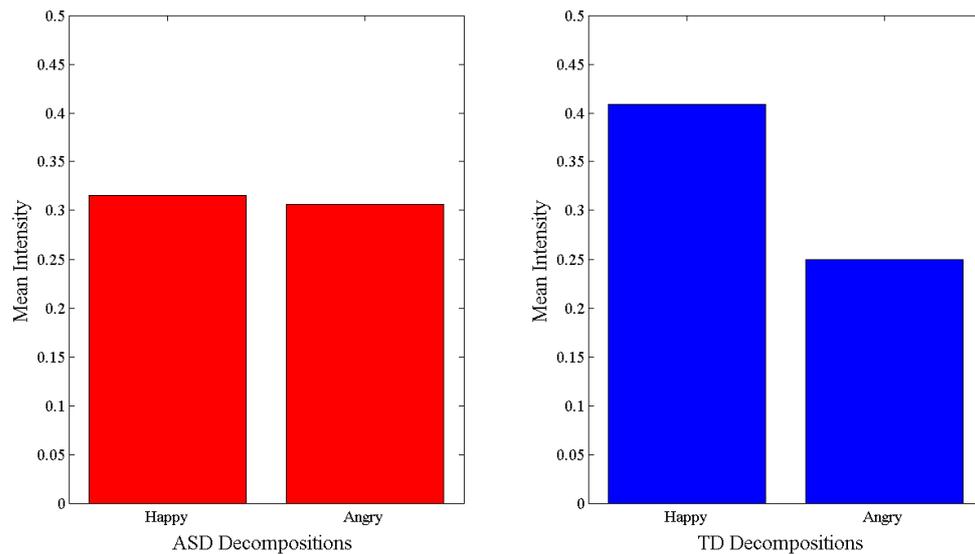


**Figure 3–13 Amplitude values (0:1) of ASD and TD Decompositions for the happy condition.**

Action units that reached threshold intensity in the upper quartile ( $> 0.75$ ) in the angry condition were, for the ASD group: Lip Funneler, Mouth Stretch, Nose Wrinkler and Jaw Drop, and for the TD group: Nose Wrinkler, Nostril Compressor and Nasolabial Deepener Left. Therefore, the AUs that emerged as having the highest intensity from *both* the ASD and TD group were dissimilar to the AUs expected for an angry expression (according to Ekman and Freisen, 1978 anger should contain ‘Brow Lowerer’, ‘Upper Lid Raiser’, ‘Eye Lid Raiser’ ‘Lip Tightener’). Therefore, the data suggest that the AUs that emerge from both the ASD and TD group are not representative of the typical components of anger.



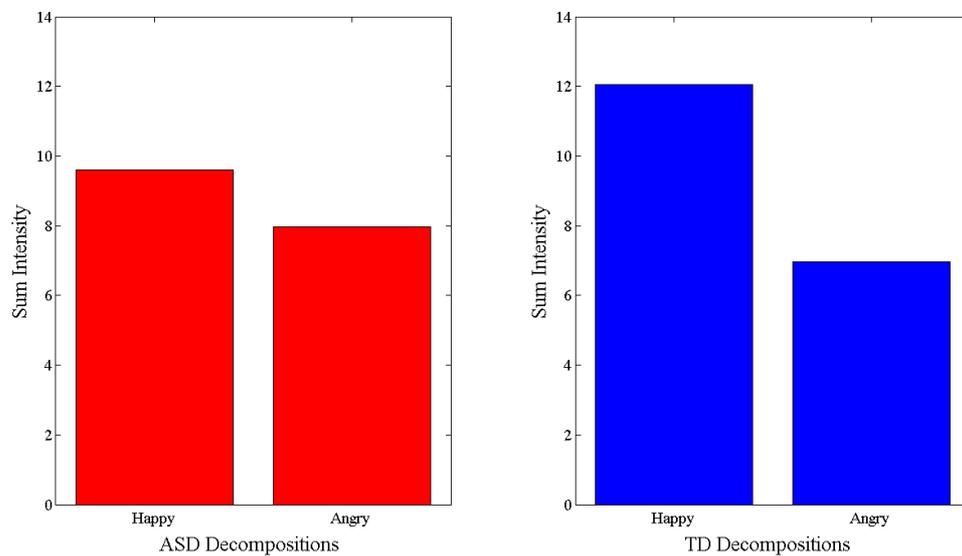
**Figure 3–14 Amplitude values (0:1) of ASD and TD Decompositions for the angry condition.**



**Figure 3–15 Mean amplitude (intensity) values for all decompositions for each group/condition**

Decomposition number was collapsed to provide one set of intensity values for each condition/group. Figure 3–15 suggests that the overall mean of these factors have slightly higher intensity in the happy condition, particularly for the TD group. A 2 x 2 ANOVA was applied to the collapsed intensity values (i.e. one vector of 42 values per emotion/group pair). There was, therefore, two factors being assessed: group (ASD or TD) and emotion (Happy or Angry), with 42 values of intensity for each group/emotion pairing. The results did not indicate the difference to be large enough to be significantly different from chance: no significant main effect of group ( $F(1, 164) = 3.3, p = 0.07, \eta_p^2 = 0.02$ ); no significant main effect of emotion ( $F(1, 164) = 0.15, p = 0.70, \eta_p^2 < 0.02$ ) and no significant interaction between group and emotion ( $F(1, 164) = 2.61, p = 0.11, \eta_p^2 = 0.02$ ).

It was recognised that even at relatively low intensities, the cumulative effect of many AUs being ‘on’ (i.e. active) in the decomposition could be considered perceptually more intense than only one or two AUs being ‘on’ at very high intensities. In order to assess intensity value while accounting for the greater weight of a large numbers of AUs, an additional analysis was conducted. The sum of the intensity values for all ‘on’ AUs was calculated, i.e. creating a single representative value of intensity for each decomposition. Therefore, this measure captures AU activation numbers and intensity values along a single continuum, where the lowest values reflect very few ‘on’ AUs at very low intensity and the highest values reflect a large number of ‘on’ AUs at high intensity. Note, values included in the analysis were  $>0.2$  because intensity lower than 0.2 were below detection threshold (Delis, Berret, Pozzo, & Panzeri, 2013).

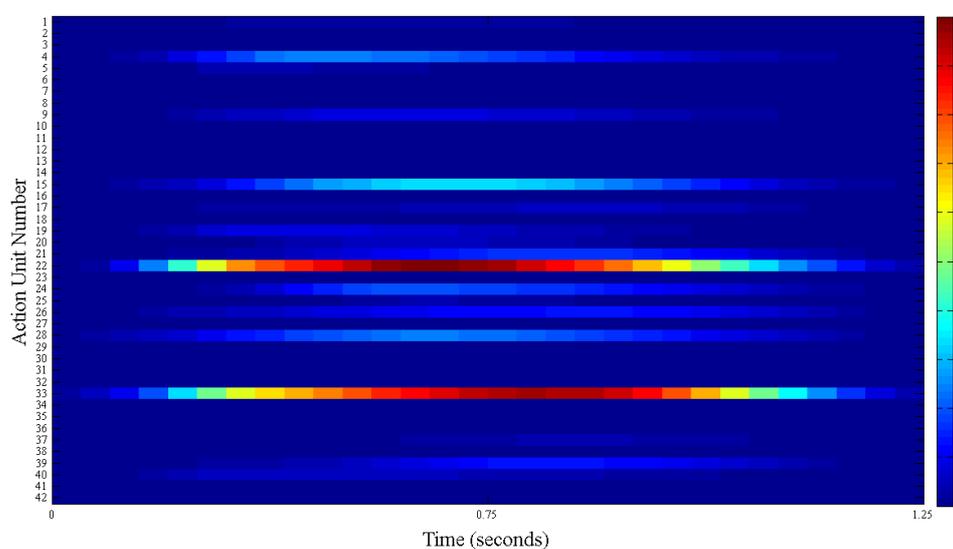


**Figure 3–16 Sum of the ‘on’ amplitude (intensity) values (i.e. values  $> 0.2$ ) for all decompositions**

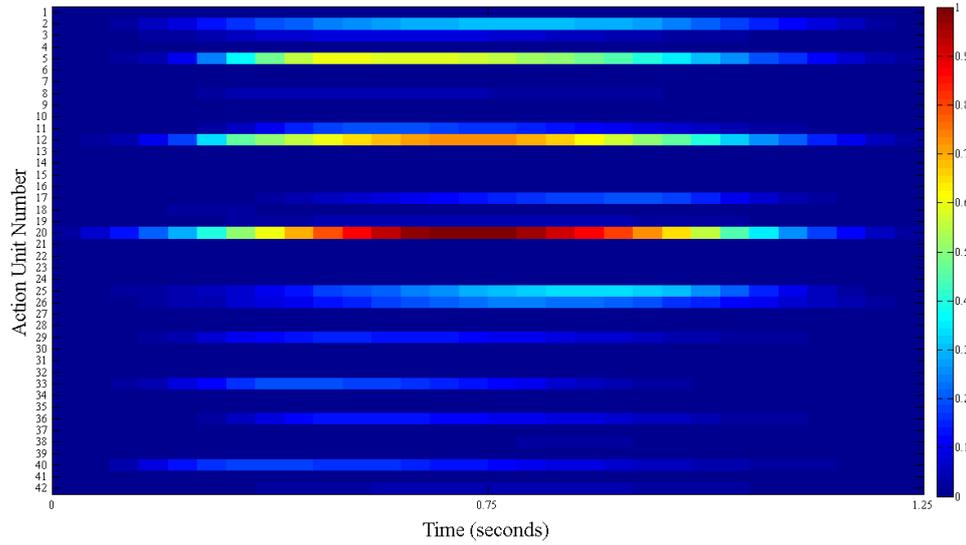
Summed intensity values in Figure 3–16 indicate similar patterns to the mean intensity values in Figure 3–15, with higher intensity values in the happy condition, particularly for the TD group. The data reveal that, visually, the intensity values seem to be slightly higher in the happy condition compared to the angry condition. This is particularly pronounced for the TD group. It can be concluded therefore, that there is no evidence to suggest the ASD group requires a more intense stimulus as indicated in the initial hypothesis, which is consistent with findings from adults with ASD in section 2.3.

### 3.3.2.3 Timing of decompositions

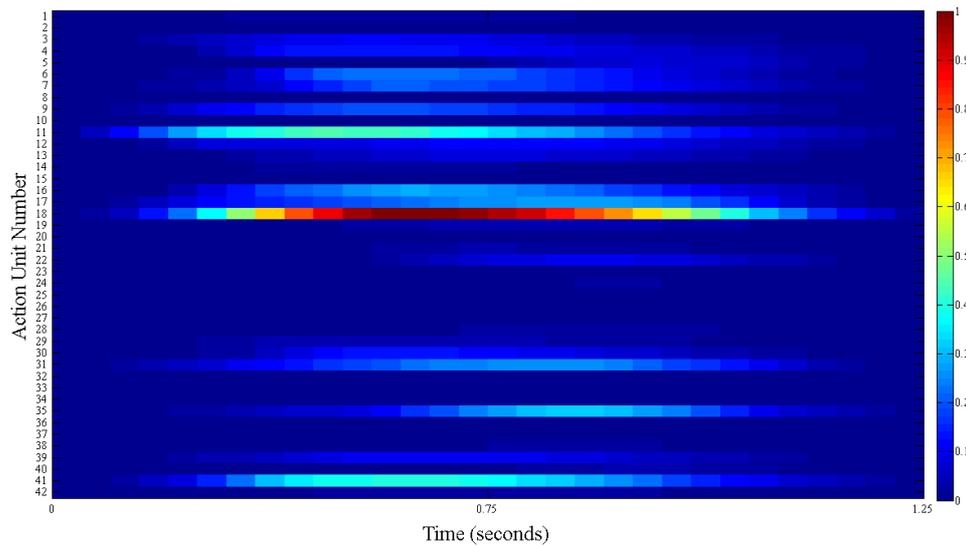
Analyses of timing components were also applied to these data. The timing (x-axis) and amplitude (as indicated by the right hand colour bar) of each AU are displayed for the ASD happy condition in Figure 3–17, Figure 3–18, and Figure 3–19 and for the TD happy condition in Figure 3–20, Figure 3–21, and Figure 3–22. These figures indicate that the ASD group use AUs from the mouth region at relatively central time points (around the mid point of 0.75 seconds) and the TD group use mouth and eye AUs, again around the midpoint of the expression latency.



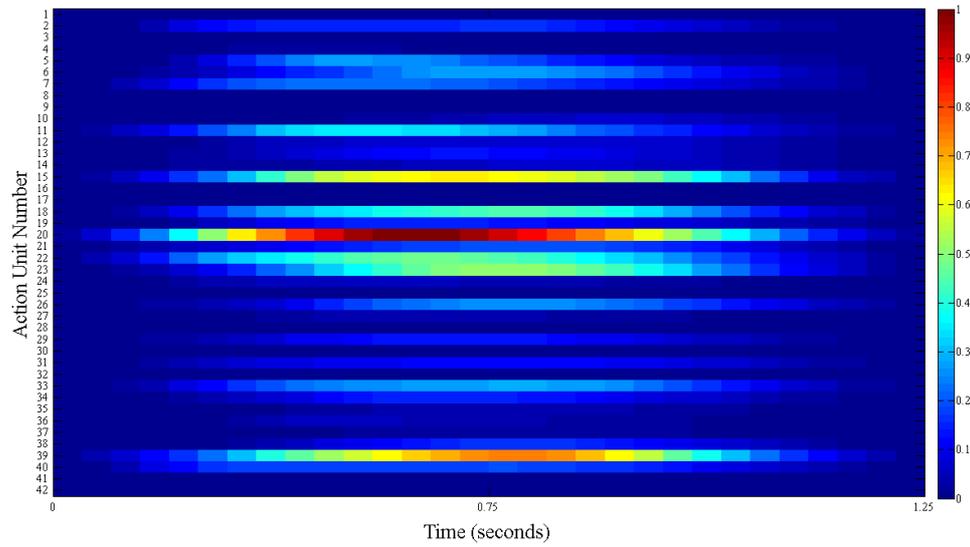
**Figure 3–17 ASD happy Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 22 (AU 13: ‘Cheek Puffer’) and AU number 33 (AU23: ‘Lip Tightener’).**



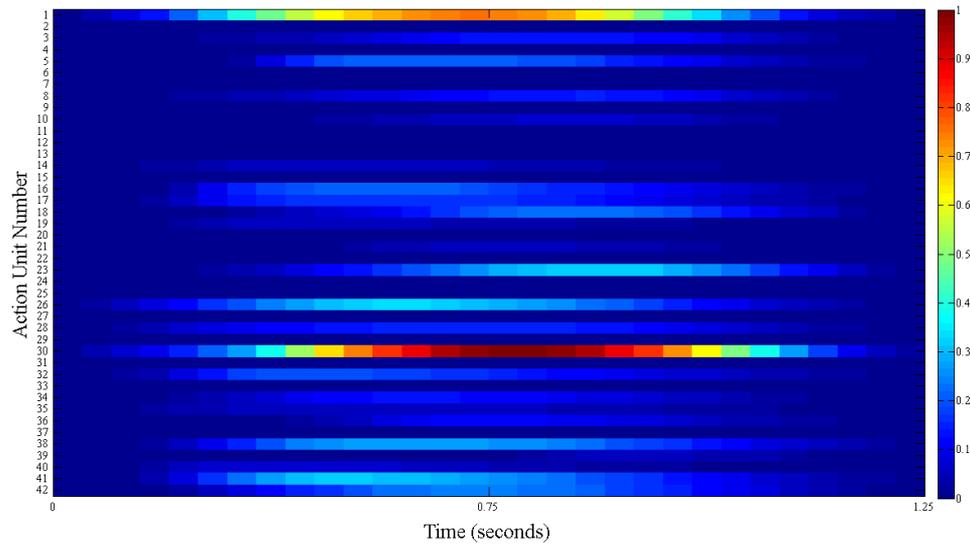
**Figure 3–18 ASD happy Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 20 (AU12L: ‘Lip Corner Puller Left’)**



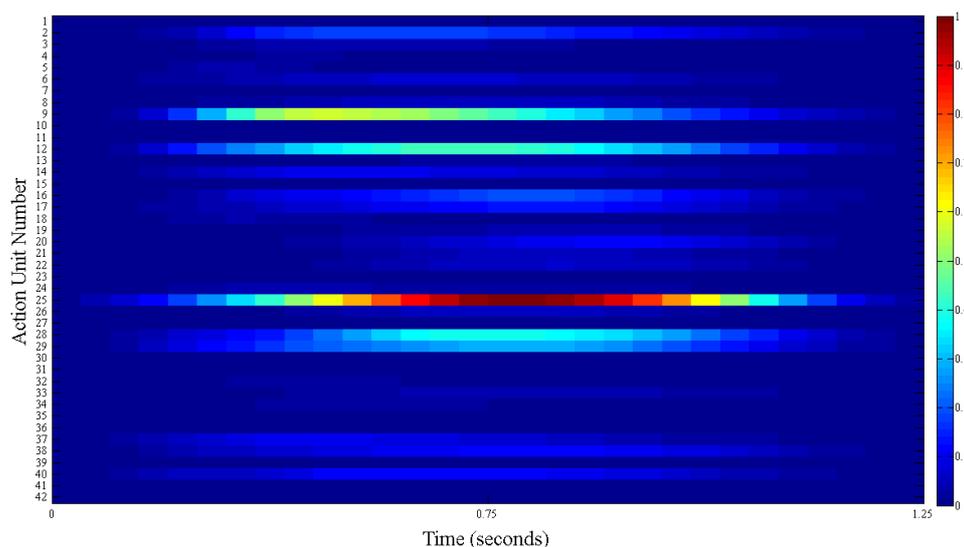
**Figure 3–19 ASD happy Decomposition 3: Activation of AUs across time and amplitude (intensity) Highest amplitude AU is AU number 18 (AU 12 Lip Corner Puller)**



**Figure 3–20 TD happy Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 20 (AU12L: ‘Lip Corner Puller Left’) and AU number 39 (AU 39 ‘Nostril Dialator’).**

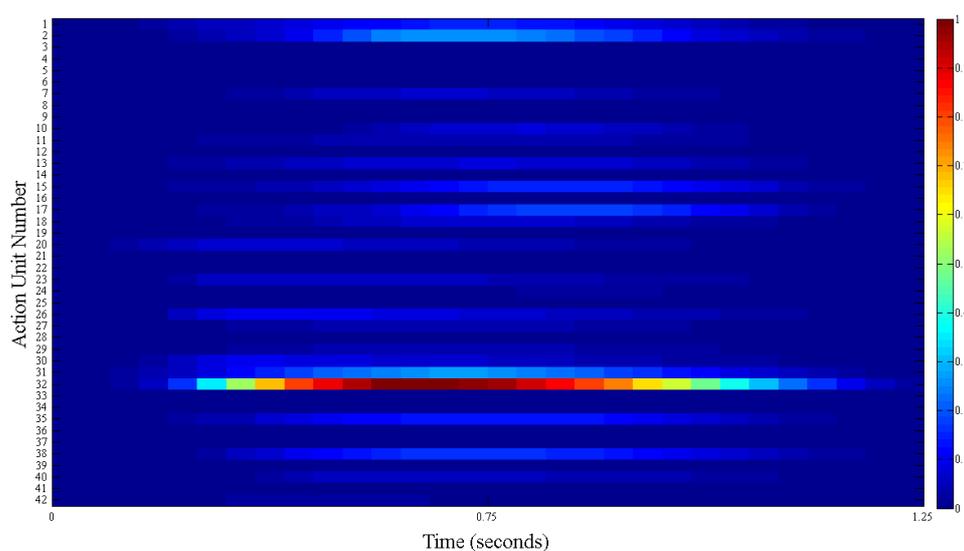


**Figure 3–21 TD happy Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 1 (AU1: ‘Inner Brow Raiser’) and AU number 30 (AU20L ‘Lip Stretch Left’).**

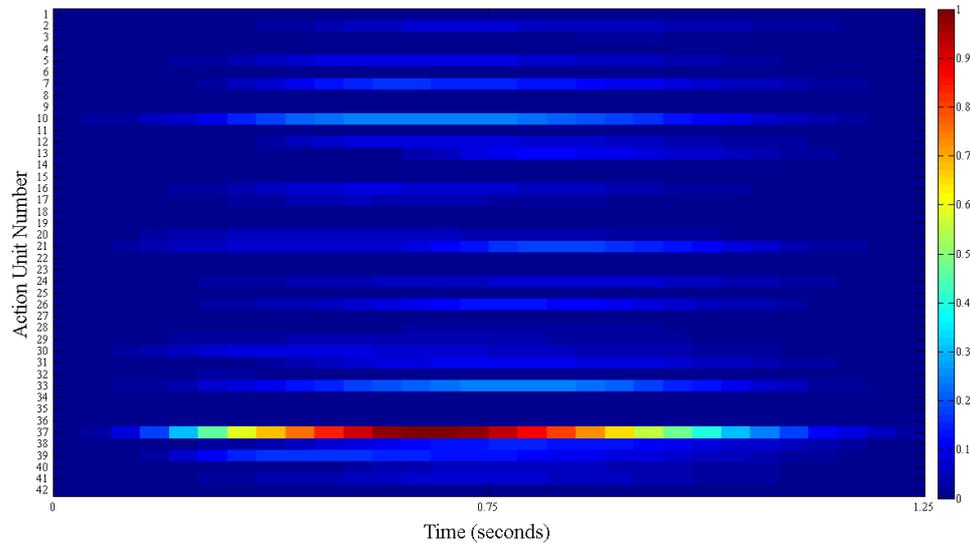


**Figure 3–22 TD happy Decomposition 3: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 25 (AU 14R ‘Dimpler Right’).**

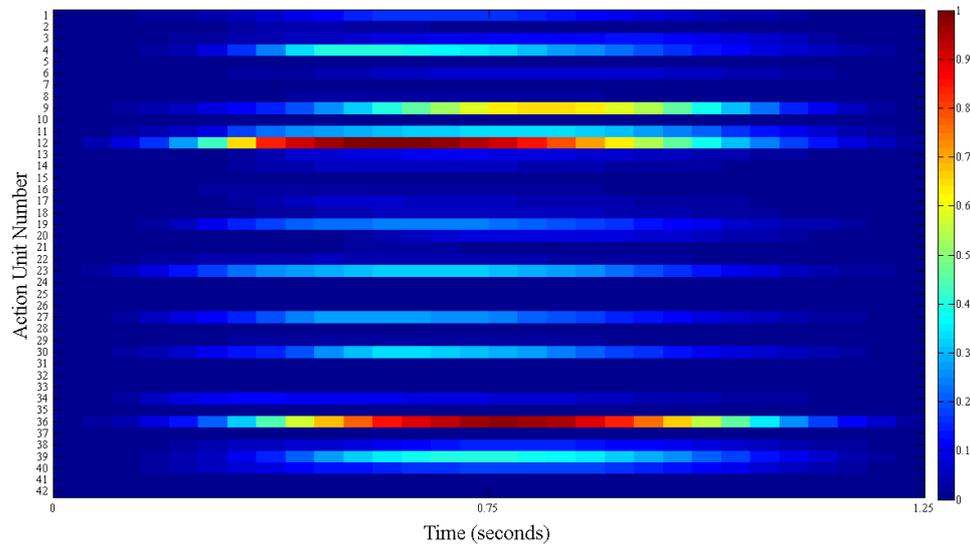
Similarly, the amplitude of the AU across time for the ASD angry condition can be found in Figure 3–23, Figure 3–24, and Figure 3–25 and for the TD angry condition in Figure 3–26, Figure 3–27 and Figure 3–28. The figures indicate slightly left of centre peak AUs in the ASD group (except for AU 36 in the third decomposition, which is slightly after the midpoint). The TD group show slightly left of centre peak amplitudes (except for AU 39 in the second decomposition which is directly on the midpoint). Hence, thus far, timing data of the decompositions indicate 1. no obvious timing differences between ASD and TD groups and 2. no evidence to suggest a ‘late’ timing bias as seen in the previous chapter.



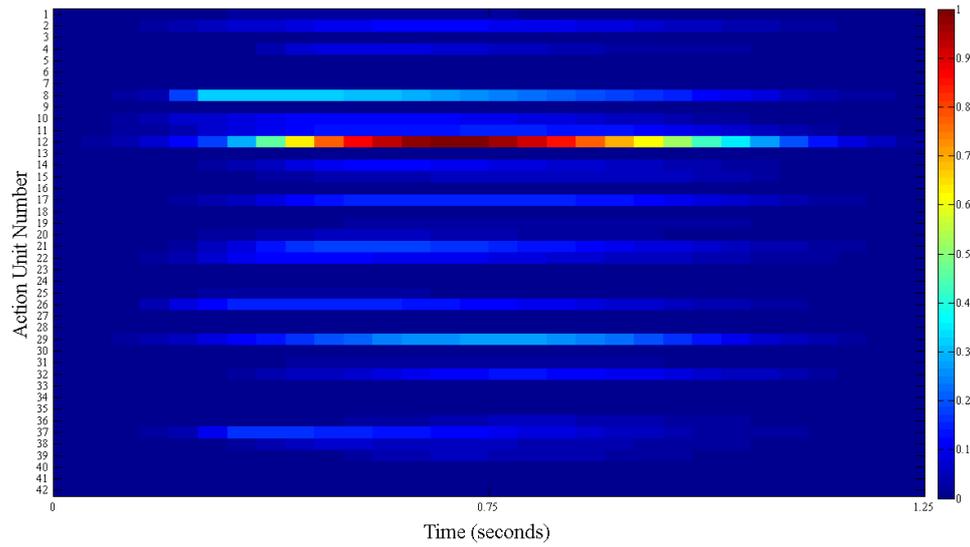
**Figure 3–23 ASD angry Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 32 (AU 22 ‘Lip Funneler’).**



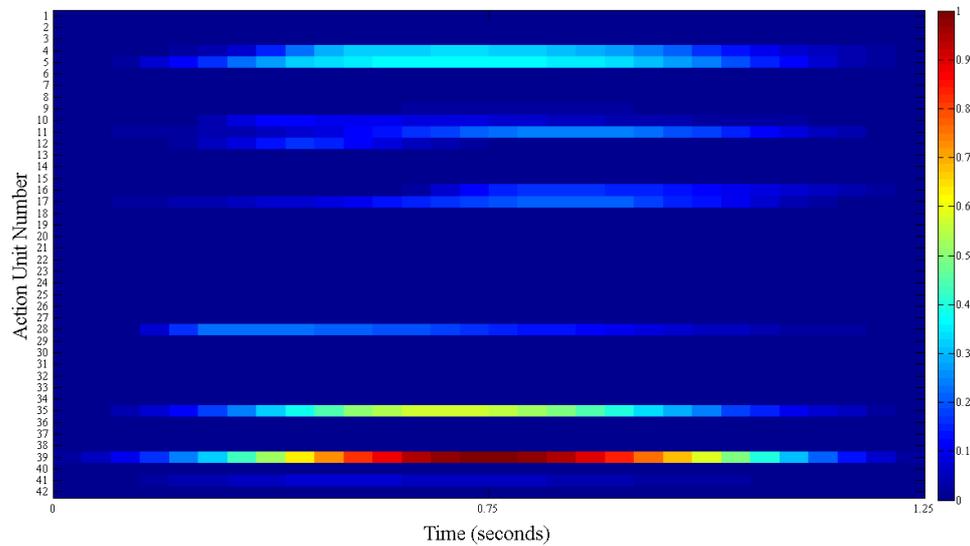
**Figure 3–24 ASD angry Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 37 (AU 27 ‘Mouth Stretch’).**



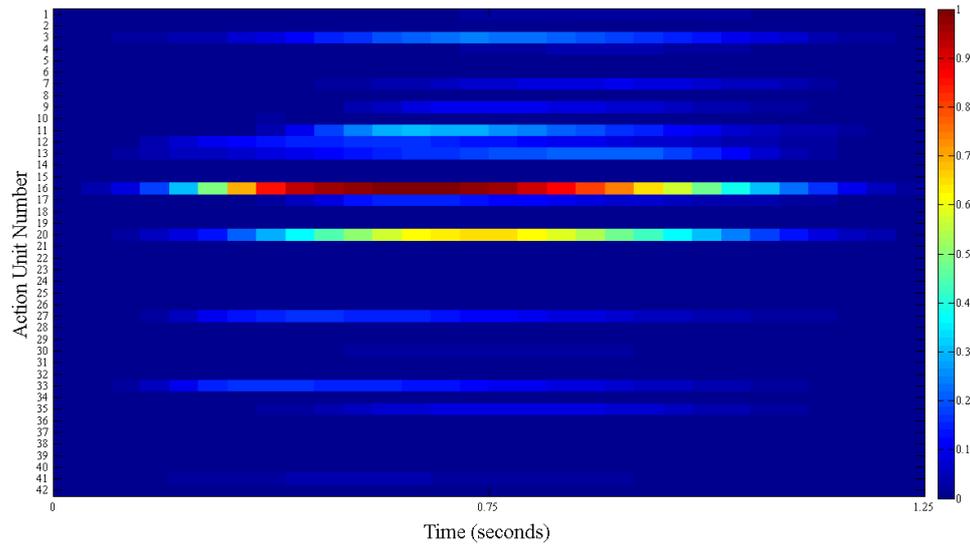
**Figure 3–25 ASD angry Decomposition 3: Activation of AUs across time and amplitude (intensity). Highest amplitude AUs are AU number 12 (AU 9 Nose Wrinkler) and AU number 36 (AU 26 ‘Jaw Drop’).**



**Figure 3–26 TD angry Decomposition 1: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 12 (AU 9 ‘Nose Wrinkler’).**

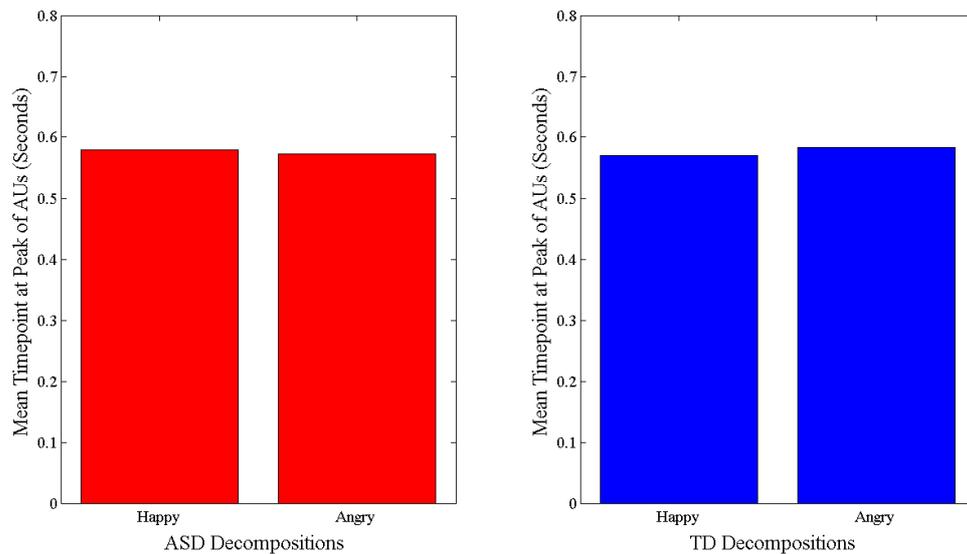


**Figure 3–27 TD angry Decomposition 2: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 39 (AU 39 ‘Nostril Compressor’).**



**Figure 3–28 TD angry Decomposition 3: Activation of AUs across time and amplitude (intensity). Highest amplitude AU is AU number 16 (AU 11L ‘Nasolabial Deepener Left’).**

The overall mean timing of each group, visualised in Figure 3–29, indicated that timing appeared to be similar between groups. However, a statistical analysis was conducted to establish whether the null hypothesis (i.e. the variance in timing of the decomposition was due to chance) could be rejected. A 2-way ANOVA assessed the contribution of group (ASD or TD) and of emotion (Happy or Angry) to the variance in timing (i.e. 42 points across 3 decompositions (126 values) for each group/emotion pair). Note, ‘peak time point’ is displayed along the x axis in where the red colour bar is found (see above figures).. As expected from the figure, no significant difference was indicated for group:  $F(1, 500) = 0.07$ ,  $p = 0.79$ ,  $\eta_p^2 < 0.00$ ; emotion:  $F(1, 500) = 0.01$ ,  $\eta_p^2 < 0.00$ ,  $p = 0.92$  or interaction:  $F(1, 500) = 0.61$ ,  $p = 0.43$ ,  $\eta_p^2 < 0.00$ .



**Figure 3–29 Timepoint (seconds) of AUs when taken together (mean average)**

### 3.4 Discussion

This study aimed to expand findings reported in Chapter 2 by exploring facial expression recognition in children with ASD. This study adopted a method similar to that reported in Chapter 2, which was adapted to be more suitable for use with children. Despite finding adults with ASD to be relatively equal to TD adults in FER (reported in Chapter 2), this study hypothesised that children with ASD may show more pronounced FER difficulties compared to adults. Thus, the first hypothesis of the current study was that children with ASD would respond in an atypical manner as compared to TD children. It was also hypothesised that children with ASD would require the stimuli to have greater intensity in order to categorise them as the target emotion. Lastly, it was hypothesised that children with ASD would display a lag in timing of FER.

The characteristics of the children’s responses (i.e. the proportion of ‘yes’ and ‘no’ responses, and reaction time) were evaluated in order to ascertain whether children with ASD had an atypical response pattern compared to TD children. In the present study, children with ASD had a slight response bias to ‘yes’ for the angry condition (i.e. most children with ASD pressed a higher proportion of ‘yes’ responses than ‘no’ responses). This suggests that children with ASD were less ‘conservative’ in their categorizations of the stimuli and so rejected less of the random facial expressions. This was, however, driven by a less conservative response in the angry condition only (i.e. being asked ‘is this person angry?’) as opposed to the happy condition (‘is his person happy?’). It is unclear from the

data why children with ASD responded with a greater proportion of ‘yes’ responses in the angry condition, while having a similar response pattern to the TD group in the happy condition. However, it could be argued that the children with ASD found it more difficult to recognise an angry expression from the random stimuli as opposed to a happy expression because of the greater complexity of affect associated with angry. Several studies have suggested that anger, as well as fear and other ‘negative emotions’, are particularly difficult for children with ASD (Ashwin et al., 2007; Humphreys et al., 2007; Lindner & Rosen, 2006). Due to the findings in Chapter 2 that adults with ASD selected a higher proportion of ‘other’ responses compared to the 6 basic emotion labels, it was expected that the children with ASD would select a higher proportion of ‘no’ responses. It could be argued that adaptations made to the experimental design in order to suit children (and hence the removal of ‘other’ as a response option) may have altered the way in which the children responded to the task (task limitations are discussed further in section 3.4.1). An evaluation of response characteristics found there to be no major differences in reaction time between children with ASD and TD children. This counters previous research suggesting that children with ASD may be slower in RT compared to TD children, despite having equivalent accuracy scores in FER tasks (Leung et al., 2013). Therefore, despite slightly different response patterns in the ASD group compared to TD group, the present study provides no evidence for a specific response speed characteristic to children with ASD. However, because this task was not a ‘speeded task’ (i.e. not RT was not a pre-defined dependent variable) further evidence, specifically aiming to measure RT in this task, may be required.

This study employed a novel experimental method (Generative Face Grammar), paired with non-negative matrix factorization to obtain subtle information about the principal components used to categorise happy and angry facial expressions. Despite recent advances in FER research, a study of this kind has never been applied to examine FER in individuals with ASD. This method revealed information about the subtleties of FER that cannot be obtained from static image stimuli or actor video stimuli. From the data, NMF revealed the principle AUs, amplitude (intensity) and timing associated with each group’s responses and these were then analysed for differences between conditions and between ASD and TD groups. It was hypothesised that children with ASD would respond in an atypical manner to the stimuli and so the underlying components (i.e. the decompositions) which made up their ‘yes’ responses would be more atypical than those of the TD children. AU differences were relatively small (hamming distance  $< 0.4$ ) between ASD and TD group. Overall, the AU comparison data suggested that happy decompositions were least

like each other while angry decompositions are most like each other however, these differences were relatively marginal. This result provides a somewhat different approach than the majority of FER ASD studies, that largely claim that FER is either ‘impaired’ or ‘typical’ in individuals with ASD (Harms et al., 2010; Simmons et al., 2009). Results also indicated that children with ASD did not need a more ‘intense’ version of a facial expression in order to categorise it as the target emotion, despite previous research suggesting stimulus intensity was a key factor in FER ability (Tell et al., 2014). This supports findings in Chapter 2 that intensity of the facial expression does not appear to be directly linked to FER ability in adults with ASD. Timing was also found to be a non-contributing factor to FER ability in this sample of children with ASD. These results do not support the original hypothesis that children with ASD would display a lag in timing of FER and do not support claims that timing components of FER may be particularly atypical in individuals with ASD (Lerner et al., 2013; Sato et al., 2013; Tardif et al., 2007). Taken together, these data indicate no strong FER differences between children with ASD and TD children, which is in disagreement with large proportion of the literature (Harms et al., 2010), but is in agreement with the data on adults with ASD preceding the present study (Chapter 2).

### **3.4.1 Limitations**

Although all efforts were made to ‘simplify’ the task in Chapter 2 to be more appropriate for children, the task was slightly more difficult than anticipated for some participants. A competitive aspect was incorporated into the task in an effort to avoid these issues, however, future research could aim to improve attentiveness of child participation by incorporating a game-like design into the task procedure from inception of the task design.

The ‘child-friendly’ alterations made to this experimental design may themselves impact the way in which comparisons can be made between the adult and child groups. For example, because this experimental design ‘primes’ the participant with the target emotion (for example, ‘Is this person angry?’) this may elicit a quicker response to the stimuli as opposed to the design used with adults (Chapter 2, section 2.2.4) which required the participant to choose from a choice of 7 emotion labels.

Unfortunately, time constraints were also a limiting factor in the present study because a longer time allocation was available with children with ASD as was available with TD children. To address this, data had to be aggregated and hence no within-group analyses

could be conducted. Results may, therefore, reflect a bias towards the ASD children because each individual in the ASD group completed a larger number of trials. Efforts were made to address this by aggregating the data and performing analyses that were suitable to this, however a more ideal situation would be equality between groups for sample size and trial number. Future versions of this study would aim to implement this step. The methodology of the present study could potentially be further improved by providing a practice session for participants. In addition to this, a more detailed analysis of the data may have been possible with the addition of AQ data (or a level of ASD severity) for the children with ASD and also possibly the inclusion of an alexithymia measure as was included in the previous chapter.

### **3.4.2 Conclusion**

This study aimed to provide an evidence base, unlike any methods previously used, in an attempt to better understand the complexity of FER in children with ASD. Despite employing a novel, sensitive technique to measure FER, results provided no evidence of a clear-cut FER atypicality in children with ASD. This conclusion builds upon evidence presented in Chapter 2, providing further indication that there are no discernible differences between individuals with ASD and TD individuals in FER ability. Therefore, this study demonstrates that even a sensitive measure of FER does not elucidate atypicalities of FER in the ASD population. This reflects the current tone of research in FER literature that more must be done to reconceptualise the notion that all individuals with ASD are impaired in facial expression recognition (Collin et al., 2013; Harms et al., 2010; Simmons et al., 2009; Uljarevic & Hamilton, 2013).

## Chapter 4 FER in Relation to Alexithymia and Autism-like Traits

### 4.1 Introduction

As discussed in Chapter 1, there is mixed evidence that individuals with ASD are impaired in perceiving facial expressions of emotion (Harms et al., 2010), despite emotion perception being a key component of a diagnosis of ASD (e.g. Autism Diagnostic Observation Schedule (Lord et al., 2000)). In fact, there is much evidence to suggest that we cannot clearly define a difference in facial expression perception in individuals with ASD compared to typically developing individuals (e.g. (Uljarevic & Hamilton, 2013)). The previous chapters presented in this thesis have provided evidence that adults and children with ASD were relatively similar to TD individuals in their FER abilities. This chapter therefore sought to expand the lens in order to take into account other, possibly unnoticed, variables that may influence FER in ASD. In particular, this chapter sought to assess the contribution of alexithymia because of the traction this argument has been gaining in the literature (discussed in further detail below). In addition, empathy, psychological wellbeing and demographic factors were assessed which are now also discussed in further detail.

Previous research has challenged the notion that emotion recognition impairment is integral to having an ASD diagnosis. Bird and Cook (2013) suggest the ASD population can be divided into two ‘sub-groups’: one with an emotion recognition impairment and an ASD diagnosis, and the other with no emotion recognition impairment but an ASD diagnosis. They suggest that, since a Facial Expression Recognition (FER) deficit can be found in the typical population, identified as ‘alexithymia’, the sub-group of individuals who have an FER deficit and an ASD diagnosis should instead consider the FER deficit as a co-occurring alexithymia condition. This theory provides a different perspective on why the FER ASD literature reports conflicting results.

In recent years, alexithymia (defined as a ‘relative constriction in emotional functioning, poverty of fantasy life, and inability to find appropriate words to describe emotions’ (Sifneos, 1973)) has been brought into the scientific discussion about FER abilities in ASD. Specifically, alexithymia has been presented as a possible explanation as to why some individuals with ASD have difficulty with FER (Bird & Cook, 2013). This is because alexithymia can be found in the typical population (Salminen et al., 1999) and is

significantly higher in individuals who have difficulty attributing mental states to others (Moriguchi et al. (2006)). It has been suggested that 50% of people with an ASD diagnosis have alexithymia (Berthoz & Hill, 2005). Hence, alexithymia has been explored as a comorbid construct to ASD and a possible reason as to why FER difficulties may be present in *some* individuals with ASD and not others (Bird & Cook, 2013). A wider discussion of the concept of alexithymia can be found in Section 1.3.3.3. The current study therefore sought to examine this contribution of alexithymia in further detail.

In addition to Cook et al. (2013), Poljac et al. (2013) also used AQ as a measure of autism-like traits, comparing emotion recognition skills between high and low AQ scorers. Poljac et al. (2013) reported a significant difference in both emotion accuracy (per cent correct) and emotion sensitivity (amount of emotion needed in order to provide a correct response) for anger, disgust and sadness. However, the differences, although found to be significant, were relatively small. For example, per cent correct results revealed, for anger, 70% accuracy for high AQ and 75% accuracy for low AQ scorers. In addition to this, they found both high and low AQ scorers to have a very low accuracy for fear (high AQ 40%; low AQ 45%). Also, for surprise, their results indicated that low AQ scorers were *less* accurate than high AQ scorers, which does not reflect their conclusion that “the ability to recognise emotions is highly related to the extent of autism-like traits” (Poljac et al. (2013), page 10.). There is little further evidence in the current literature upon which to compare the findings of Poljac et al. (2013), hence, the current study sought to explore the relationship between AQ score and emotion recognition further.

In addition to assessing the potential contribution of alexithymia to FER, this study also sought to assess the claim that the contribution of demographic factors is important, but often overlooked (Harms et al., 2010; Uljarevic & Hamilton, 2013). Uljaveric & Hamilton acknowledge the complex heterogeneity of the ASD phenotype in their meta-analysis of the FER in ASD literature, pointing to demographic factors as being a variable that is likely linked to individual differences in ASD symptoms (for example, age and IQ). Regardless of this, Uljaveric and Hamilton state that it is unfortunate that very few papers included in their meta-analysis report enough information about demographic variables in order to effectively measure their contribution FER ability. Uljaveric and Hamilton cite this as an interesting area of future work, advocating for more transparency of demographic factors and for their inclusion at analysis stage. Harms et al (2010) described demographic factors as accounting for a proportion of the discrepant findings in the literature in their review of FER ASD literature published before 2010. Harms et al

postulated that this has two guises: 1. ASD is heterogeneous and there is little solid evidence for specific demographic factors that may account for the large individual differences observed in the ASD phenotype 2. Despite this, demographic factors are not reported in great detail or included at analysis stage (as was also observed by Uljaveric and Hamilton (2013)). The present study sought to cast a wide net, measuring several different demographic factors and, in addition, all of the factors measured were included in the analysis stage. This was intended to highlight any potential contributory variables to FER that may have not otherwise been picked up. The demographic factors included in this study were location (i.e. whether the participants took part in Glasgow or in San Diego), age, gender, handedness, race, psychological history and education.

Empathy was also included in this study as a variable that may influence FER. The ability to ‘put oneself in another person’s shoes’, empathy has long been studied in relation to emotion perception and has been of particular interest to researchers in the field of ASD. The link between empathy and emotion perception has, in particular, attracted much research in the field of atypical psychology. For example, poor FER ability has been found to be associated with low levels of empathy in individuals with schizophrenia (Mandal, Pandey, & Prasad, 1998). A core characteristic of an ASD diagnosis is a difficulty understanding the mental states of others, or ‘mentalizing’ (S. Baron-Cohen et al., 1985). A diminished ability to accurately determine the mental state of others has been recognised in empirical studies with individuals with ASD: for example understanding another’s beliefs (S. Baron-Cohen et al., 1985; Leslie & Frith, 1988; Perner, Frith, Leslie, & Leekam, 1989), pretence (Leslie, 1987) and intentions (Phillips, Baron-Cohen, & Rutter, 1998). These seminal pieces of work culminated in one of the first cognitive theories of autism: Theory of Mind (C. Frith & Frith, 2005), as is discussed in Section 1.5. The ability to ‘mentalize’ is intrinsically linked to the ability to empathise with another person (Lawson, Baron-Cohen, & Wheelwright, 2004). Because of this, researchers have attributed poor mentalizing abilities in individuals with ASD to having less empathy for others. A conceptualization of this has been developed in the Extreme Male Brain (EMB) theory of autism (S. Baron-Cohen, 2002) which posits that a low-empathizing, high-systemizing brain (which is described as being ‘stereotypical’ of males) is presented in the ASD phenotype in an exaggerated form (also discussed further in Section 1.5). Empirical research has somewhat supported this claim, for example Sucksmith, Allison, Baron-Cohen, Chakrabarti, and Hoekstra (2013) found that both males and females with ASC had significantly lower empathy abilities when measured using the ‘Empathy Quotient’ self-report questionnaire (Simon Baron-Cohen & Wheelwright, 2004). Aaron, Benson, and Park (2015) assessed

empathy in typical college students who varied in their level of autism-like traits. Aaron et al concluded that an increase in AQ levels was significantly correlated with a decrease in empathic ability. Mazza et al. (2014) found that adolescents with ASD were comparable to TD adolescents in an empathizing task involving positive emotional valence (e.g. happy emotion) but were significantly worse than TD adolescents when the emotional valence was negative. The story of empathy in ASD is further complicated by research such as that of Schulte-Ruther et al. (2014) who demonstrated that the neural substrates of empathy may be influenced by age, insofar as adolescents with ASD performed poorly on an empathy behavioural task while adults with ASD performed relatively similarly to TD adults. In addition, they reported that in an fMRI scanner, adolescents with ASD showed less neural activation in the lateral pre-frontal cortex during the behavioural empathy task compared to TD adolescents, while adults with ASD were relatively similar to TD adults. There has been criticism from the autism community regarding the attribution of ASD with being ‘impaired’ in understanding empathy. Researchers have argued that individuals with ASD do not *lack* empathy but sensitivity to empathy is slightly less pronounced. For example children with autism were found to do well on an empathy task despite performing less well than typical children (Charman et al., 1997). Brewer and Murphy advocate that holding the view that individuals with ASD lack empathy can “distort the perception” of individuals with ASD and “possibly delay effective treatments” (Brewer & Murphy, 2016).

This study aimed to assess Facial Expression Recognition (FER) ability and its relationship to AQ scores (autism-like traits) and TAS scores (alexithymia). In addition to this, several other measures were taken such as Empathy, Psychological Diagnosis and Depression indicators, in an attempt to provide possible alternative predictors of poor FER. It was hypothesised that both autism-like traits and alexithymia would predict emotion discrimination ability. Moreover, alexithymia would account for a higher proportion of the variance than autism-like traits.

## **4.2 Methods**

### **4.2.1 Participants**

A total of 144 adults took part in this study, recruited from two locations: University of Glasgow (n = 54) and University California San Diego (UCSD; n = 90). Online and poster advertisements were used to recruit participants and all participants provided written,

informed consent prior to starting the study. Participation was voluntary and compensated by Psychology course credit or a small monetary payment (£9 at University of Glasgow or \$20 at UCSD). The study was carried out in accordance with ethical standards of the College of Science and Engineering ethics committee at Glasgow University and Institutional Review Board committee at UCSD. Approval was obtained from these respective committees before testing commenced. At the data analysis stage it was apparent that three participants displayed extreme outlier values (see Section 4.2.6) and so these participants were removed. Final data analysis was therefore conducted on 141 participants (Glasgow  $n = 54$ ; UCSD  $n = 87$ ).

## **4.2.2 Psychometric measures**

### **4.2.2.1 Autism Spectrum Quotient (AQ)**

The Autism Spectrum Quotient (AQ; S. Baron-Cohen, Wheelwright, Skinner, et al. (2001)) is a 50-item, self-report questionnaire that probes information about the level of autism-like traits an individual has. The questions target five subscales, which are thought to be characteristic of an ASD diagnosis: 1) social skills 2) attention switching 3) attention to detail 4) communication and 5) imagination. Responses are given in terms of a four-point Likert scale ('definitely agree', 'slightly agree', 'slightly disagree' or 'definitely disagree') to statements which target the above 5 ASD characteristics. For example, statements include: "I enjoy social chit-chat." (social skills; reverse score), "I usually notice car number plates or similar strings of information" (attention to detail; regular score), "When I'm reading a story, I can easily imagine what the characters might look like." (imagination, reverse score). A response of 'definitely agree' or 'slightly agree' was scored as a '1' and 'slightly disagree' or 'definitely disagree' as a '0', and the opposite for reverse scored items. A total score on the AQ consists of a sum of all 50 items (i.e. with some questions reversed scored), hence, higher scores indicate higher levels of autism-like traits. This follows typical scoring technique for this questionnaire (S. Baron-Cohen, Wheelwright, Skinner, et al., 2001).

### **4.2.2.2 The twenty-item Toronto Alexithymia Scale (TAS-20)**

The 20-item Toronto Alexithymia Scale (TAS-20; (Bagby, Parker, et al., 1994; Bagby, Taylor, et al., 1994)) is a self-report questionnaire that investigates an individual's ability to understand their own, and other people's emotions. Having alexithymia is sometimes referred to colloquially as having 'emotion-blindness'. The TAS-20 has 3 subscales:

‘describing feelings and emotions’, ‘identifying feelings and emotions’ and ‘externally-oriented thinking’. Items are statements regarding the above 3 subscales, for example “I am able to describe my feelings easily” (describing feelings and emotions, reverse score), “When I am upset, I don’t know if I am sad, frightened, or angry” (identifying feelings and emotions, regular score), “Looking for hidden meanings in movies or plays distracts from their enjoyment.” (externally-oriented thinking, regular score). Participants responded to the statements using a 5-point Likert scale: 1 = ‘Strongly Disagree’, 2 = ‘Disagree’, 3 = ‘Neither Disagree or Agree’, 4 = ‘Agree’ and 5 = ‘Strongly Agree’. For each participant all 20 items are summed (note, 5 items are reverse scored in which their responses are scored as follows: ‘Strongly Disagree’ = 5, ‘Disagree’ = 4, ‘Agree’ = 2 and ‘Strongly Agree’ = 1). The sum of the items indicates the participant’s total TAS score, and therefore, higher TAS scores indicate higher levels of alexithymia.

#### **4.2.2.3 Interpersonal Reactivity Index (IRI)**

The Interpersonal Reactivity Index (Davis, 1980) is a 28-item, self-report questionnaire that asks information about empathy, defined by Davis as “the reactions of one individual to the observed experiences of another” (Davis (1983), p 113). Participants indicate their response on the IRI by choosing one of five Likert scale choices from A (‘Does not describe me well’) to E (‘Describes me very well’). The IRI can be broken down into four sub-components: 1) perspective-taking, 2) fantasy, 3) empathetic concern and 4) personal distress. For example, “I sometimes find it difficult to see things from the "other guy's" point of view.” (perspective-taking, reverse scored), “I often have tender, concerned feelings for people less fortunate than me” (empathetic concern, regular score), “Being in a tense emotional situation scares me.” (personal distress, regular score). Following the preferred scoring technique (Davis, 1980), each statement was scored as follows: questions scored in the regular way are scored: A = 0, B = 1, C = 2, D = 3, E = 4 and for reverse-scored items the reverse score is given (i.e. A = 4, B = 3, C = 2, D = 1, E = 0). A total score on the IRI consists of a sum of all 28 items with higher values indicating higher levels of empathy.

#### **4.2.2.4 Personal information questionnaire**

In addition to the above psychometric measures, participants filled out questionnaires about 1) demographics and 2) medical history. Demographics included age, handedness, gender, ethnicity, race and level of education. Medical history questions asked whether the participant had had a history of a) mental health issues, b) a period of over two weeks of

unhappiness, c) a period of over two weeks of irritability, and d) autism in a sibling or family member.

### 4.2.3 Design

The present study used psychometric functions to plot data and draw useful information regarding participants' response to the stimuli. Previous chapters in this thesis utilized a Generative Face Grammar system paired with reverse correlation and non-negative matrix factorization techniques, however an alternative paradigm was deemed appropriate for the present study for the following reasons: 1. as AQ was being used as a measure of ASD, a large sample of participants was required (unsuitable because of the long length of time required to test participants using the GFG) and 2. Because the present study was in collaboration with a research group in California, the option to use the GFG was not available for use in the Californian lab. Instead, the current study drew from the psychophysics literature and from this built a paradigm to assess emotion perception in a quantitative manner. Previous research in this domain has utilized psychophysics methods to assess visual perception using difference thresholds and points of subjective equality (PSE) (Harris, Atkinson, Lee, Nithi, & Fowler, 2003; Pokorny, 1998; Tokita & Ishiguchi, 2010). In the field of emotion perception in ASD, Cook et al. (2013) utilized such techniques to assess discrimination between facial expressions. Cook et al. (2013) claimed that alexithymia, not autism, predicts poor recognition of emotional facial expressions. They assessed 64 adults with ASD and 42 TD adults on their ability to distinguish between two emotions (e.g. disgust/anger and surprise/fear) and two identities (Harold/Felix and Tracie/Maria). Stimuli were static images of one of 13 points along a continuum of the morph. Hence, the participant was presented with an image which was a hybrid of two emotions (e.g. disgust and anger) and two identities (e.g. Harold and Felix). In this example, participants were asked to indicate whether the emotion was 'disgust or anger' and 'Harold or Felix'. A measurement of how closely the participant was able to discriminate was taken by plotting each participant's responses to the stimuli then fitting a psychometric function to these data. Therefore higher slope values indicate better discrimination ability and is a standard measure used to report psychometric functions (Rose, 1988). Point of Subjective Equivalence (PSE) was also drawn from each psychometric function, which indicates the point at which the participant was equally likely to respond one way (e.g. disgust) and the other (e.g. anger). Again, the use of PSE is an established measure of perceptual bias using psychometric functions (Meese, 1995).

#### 4.2.4 Stimuli

Stimuli were generated from six base images of a female actor conveying happy, sad, disgust, anger, surprise and fear facial expressions obtained from the Karolinska Directed Emotional Faces database (Lundqvist, Flykt, & Öhman, 1998). These images made up the ‘endpoints’ for image morphs in 3 different emotional-pairs conditions: happy/sad, disgust/anger, fear/surprise (see Figure 4–2, Figure 4–3 and Figure 4–4).

##### Happy/Sad Condition ‘Endpoint’ Images



Base Image for ‘Happy’



Base Image for ‘Sad’

**Figure 4–1 Endpoint images used in stimuli generation for the happy/sad condition**

### Disgust/Anger Condition 'Endpoint' Images



Base Image for 'Disgust'



Base Image for 'Anger'

Figure 4-2 Endpoint images used in stimuli generation for the disgust/anger condition

### Surprise/Fear Condition 'Endpoint' Images



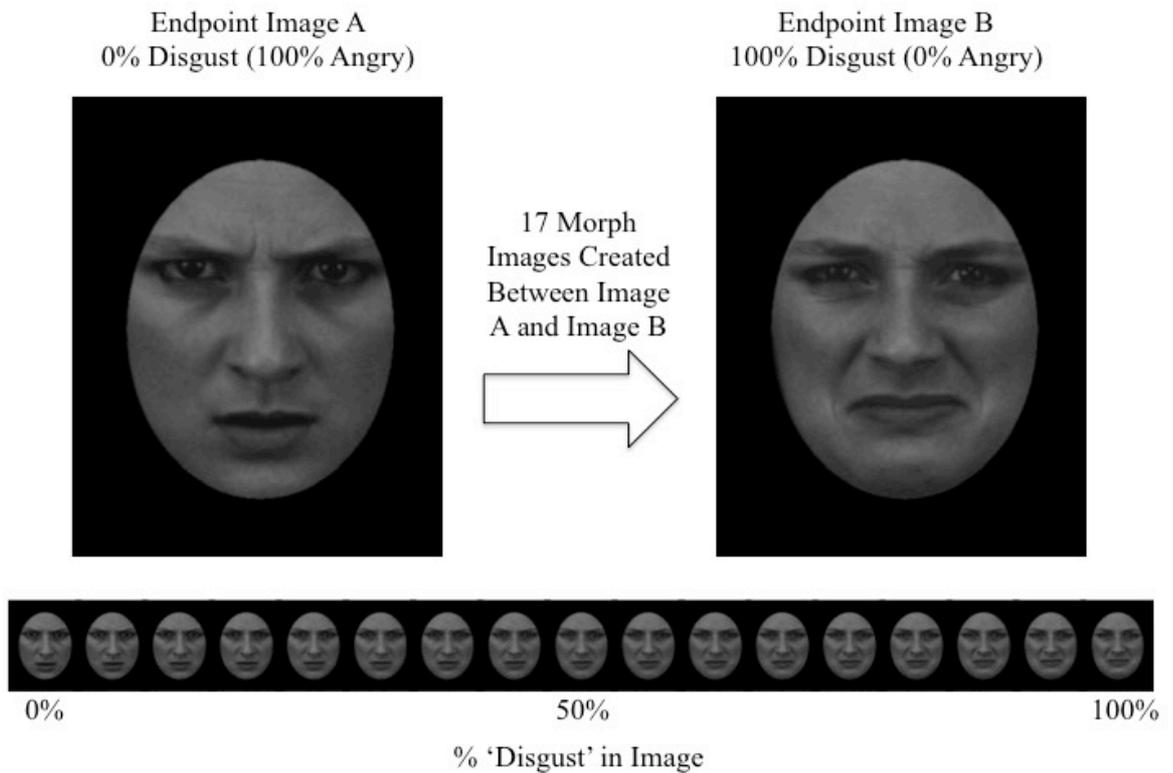
Base Image for 'Surprise'



Base Image for 'Fear'

Figure 4-3 Endpoint images used in stimulus generation for the surprise/fear condition

Using Fantamorph morphing software (Abrosoft, 2015) for each emotion-pair condition, 17 morphs were created between the endpoints, referring to the endpoints as 0% and 100%. For example, in the disgust/angry condition, 0% refers to ‘0% disgust’ (which is the angry endpoint face) and 100% refers to ‘100% disgust’ (which is the disgust endpoint face). A total of 17 morph steps were created: 0%, 10%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 90%, 100%. The morph steps are uneven because finer detail was required at the middle of the morph continuum.



**Figure 4–4 Stimulus Generation Example for disgust/angry Condition**

Before morphing the endpoints, images were fitted with an oval-shaped ‘mask’, which removed any external features of the images (e.g. ears, hair, background). Subsequent to this, the images were ‘normalised’ i.e. manipulated so that all stimuli had the same luminance and contrast. Note, however, only the area within the mask was normalised. The normalization process is typical of this kind of study (Pallett, Cohen, & Dobkins, 2014) and ensured that any performance differences observed between conditions could not be accounted for by differences in low-level visual spatial characteristics. The normalization process took the following steps:

1. All stimuli were converted to grey-scale using Adobe Photoshop (Adobe, 2015)
2. Images were converted back to their original luminance values. This was achieved by performing an 'inverse gamma' function on each of the endpoint images. An inverse gamma function is applied because, when compressed images (e.g. jpeg) are presented on a computer monitor, there is a non-linear relationship between gun values (i.e. proportion of white, red, blue and green) and luminance values (i.e. perceived brightness of the image). In order to have full control over the luminance values of all images, the inverse gamma function converts this non-linear relationship between gun values and luminance into a linear one. The luminance values were then averaged across all 6 'endpoint' images. This was achieved by firstly determining an overall mean luminance value. Then, for each 'endpoint' image, each pixel was multiplied by a value that made the mean luminance equal to the desired value. The mean luminance of all images was 15.17 cd/m<sup>2</sup>.
3. The root-mean-squared (rms) contrast of the stimuli was then manipulated for each baseline image. This was achieved by determining the rms contrast of the luminance-adjusted image, and then dividing each image-pixel value by a ratio (rms contrast of the luminance-adjusted image/desired rms contrast) in order to make the total rms contrast of the image the desired value.

Stimuli were presented on a gamma corrected monitor using Palamedes toolbox for Matlab (Palamedes Version 1.8.2, Prins 2009) at a angular size of 3.64 deg (width) 7.06 deg (height). Stimulus angular size was exactly the same in Glasgow and UCSD, as were all of the stimulus qualities described above.

#### **4.2.5 Procedure**

Before testing began, participants were seated in the lab where they were briefed about the experiment. All participants provided written, informed consent before testing began. Exact instructions for the experimental task were then given to the participant (both orally and in an information sheet), followed by a short period in which the participant was able to familiarise themselves with the response keys (on a computer keyboard in front of them) and the viewing distance of the task. Stimulus angular size was kept the same for Glasgow and for UCSD (stimuli width = 3.64 deg and stimuli length = 7.06 deg). Hence, viewing distance at Glasgow was set at 120cm and at UCSD it was set at 79cm, because at

Glasgow, stimuli were presented on a 32-inch (1920 x 1080 pixels) computer monitor with maximum luminance 112 cd/m<sup>2</sup>. At UCSD, stimuli were presented on an 18-inch (1280 x 800 pixels) computer monitor and max luminance was restricted to 112 cd/m<sup>2</sup> in order to match the maximum luminance in Glasgow.

In addition to oral instructions from the experimenter, participants were prompted with the following on-screen instructions before the experiment began:

Welcome to the Emotion Judgment Experiment! In this experiment, you will see faces and decide if they look more like one emotion or another. We will let you know before a block begins which emotions you will be judging. When you are ready to begin, press the SPACE bar.

After the participant pressed the spacebar a fixation-cross appeared at the centre of the screen. Participants were asked to look directly at the fixation-cross then, again, press the spacebar to initiate the first trial. At this point, a stimulus (see section 4.2.4) was presented in the centre of the screen for 400ms along with two response options on the left- and right-hand side of the stimulus. Response options were a '2-alternative forced choice' (2AFC) of 2 emotion labels. Emotions labels were 'happy' or 'sad' for the happy/sad condition 'disgust' or 'anger' for the disgust/anger condition and 'surprise' or 'fear' for the surprise/fear condition. Participants pressed the left or right key to indicate the emotion label (bright stickers were attached to the left key ('z') and the right key ('/')) and all other keys were disabled. Although the stimulus was presented for only 400ms, response options remained on the screen until the participant responded. No feedback was given and participants were not able to respond faster than 100ms – a message saying 'You responded too fast' appeared on the screen if this happened. If the participant did not respond before 3000ms a message appeared saying 'please make your response'.

The experiment consisted of three conditions: happy/sad, surprise/fear, disgust/anger. In total the experiment lasted around 90 minutes. Each condition was repeated with the emotion labels either side of the stimulus switched to the opposite side e.g. if the left emotion label read 'happy' and the right emotion label read 'sad' for the first block of the happy/sad condition, the second block would have left emotion label as 'sad' right emotion label as 'happy'. This measure was implemented in order to control for response bias (i.e. a participant favouring pressing one key over another). The experiment, therefore, comprised a total of 6 experimental blocks. Each block consisted of 170 trials (17 morph stimuli x 10

repetitions): 1020 trials in total. Before each block began, participants were prompted with on-screen instructions. For example the happy/sad condition would read:

In this round, you will be judging whether faces look more like happy or sad. You will begin with some practice trials to familiarise yourself with the emotions for this round.

Note, participants were prompted that the emotion labels would switch sides. For example, the second block of the happy/sad condition would read:

In this round, you will be judging whether faces look more like sad or happy. You will begin with some practice trials to familiarise yourself with the emotions for this round. PLEASE NOTE THAT THE KEYS ARE SWITCHED.

A practice session was given before each block in order to a. allow the participant to familiarise themselves with the stimuli and the keys and b. to ensure the participant was able to correctly label the six basic emotions (happy, sad, disgust, anger, sadness and fear). Participants were shown the 'endpoint' images (i.e. 100% and 0% emotion stimuli) and were required to correctly label these. For example, in the happy/sad condition participants were shown the 100% happy stimulus (100% happy face) and the 0% happy stimulus (100% sad face) and asked to label the stimuli as 'happy' or 'sad'. Participants were shown each stimulus 10 times and were required to obtain 100% correct in the practice round to move on to the real experiment.

After each block, participants were asked to take a break. During the breaks participants completed the psychometric and personal information questionnaires. Hence, the whole experiment was conducted in the following sequence (this was consistent for both Glasgow and UCSD):

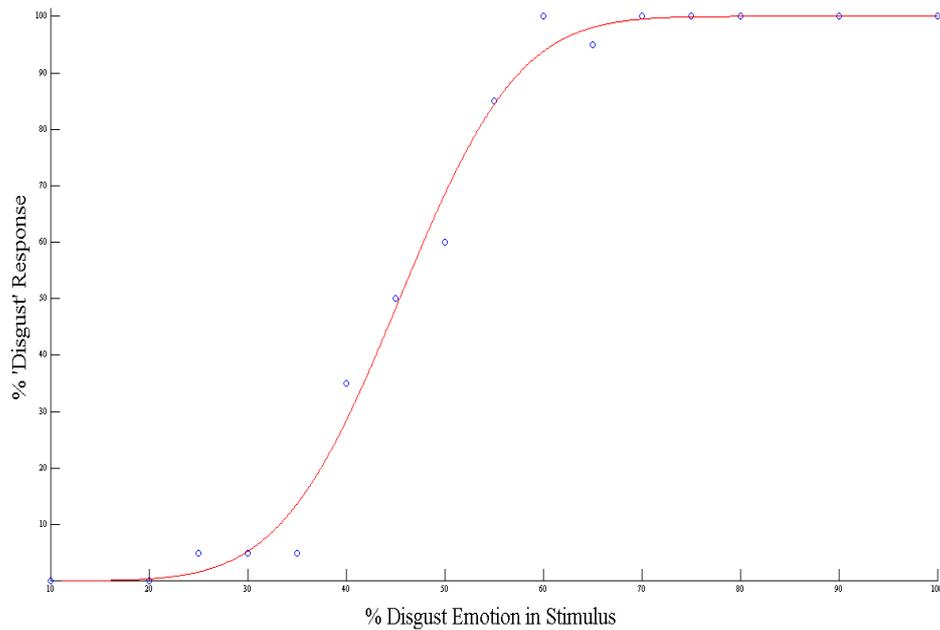
1. Introduction/ Consent form
2. Questionnaire 1: Demographic/Medical Questionnaire
3. Block 1 (randomised condition)
4. Questionnaire 2: IRI
5. Block 2 (randomised condition)

6. Own break (no questionnaire)
7. Block 3 (randomised condition)
8. Questionnaire 3: AQ
9. Block 4 (randomised condition)
10. Own break (no questionnaire)
11. Block 5 (randomised condition)
12. Questionnaire 4: TAS
13. Block 6 (randomised condition)

Participants were debriefed at the end of the experiment and compensated for their participation with a course credit or a small monetary payment (£9 at University of Glasgow or \$20 at UCSD). Each participant was offered the chance to ask questions or make comment at the end of the experiment as well as taking home an information sheet containing the experimenter's contact details in case needed at a later date. In total, the experiment lasted around 90 minutes for each participant.

## 4.2.6 Analysis

For each participant, data were obtained detailing the number of times they selected each emotion option in response to each of the 17 morph levels. In order to better characterise each participant's responses, a cumulative normal function was fitted to these data using Palamedes toolbox for Matlab (Palamedes Version 1.8.2, Prins 2009; see example in Figure 4–5).



**Figure 4–5 Example of one participant's data for the disgust/anger condition (blue dots) and the Cumulative Normal Function fitted to the data (red line)**

From the curve-fitting procedure two dependent variables were derived: 1) the slope of the cumulative normal curve (i.e. slope values closer 0 indicated poorer emotion discrimination ability) 2) the point of subjective equivalence (PSE) which is the point on the x-axis (morph level) that reflects 50 on the y axis (the point at which the participant was indicating equally 50% one emotion and 50% the other). PSE values that deviate from 50 indicate a perceptual bias, smaller values indicating a bias to happy, disgust or surprise and higher values indicating a perceptual bias towards sad, anger or fear. After the cumulative normal function was fit, outlier participants were detected and removed from the dataset. In order to be kept in the dataset, the function fit to each participant was required to achieve 15% or below on the y-axis at one end of the morph continuum and 85% or above on the y-axis at the other end of the morph continuum. It was recognised that if a participant curve did not meet 15% (i.e. the y intercept was  $> 15$ ) and if it did not meet

85% (asymptote of the curve < 85) then it was likely that the participant was responding at random or was not paying attention to the task. Therefore, if these criteria were not met (in any of the 3 conditions) the participant was removed from the dataset. These rejection criteria applied to 3 individuals, hence the final number of participants included in the analysis was 141.

## **4.3 Results**

### **4.3.1 Descriptive information**

Demographic information for all participants can be found in Table 4-1 and is displayed per location group (i.e. Glasgow and UCSD). Age was compared using an independent samples t-test (because these data were in scale format) and all other variables were compared using a Chi-Square test (i.e. because these data were in nominal format). After controlling for multiple comparisons using a Holm-Bonferroni correction (Holm, 1979), only race was found to be significantly different between Glasgow and UCSD:  $\chi^2 = (1, N = 144) = 29.33, p < 0.01$ . This indicated that UCSD had a significantly greater number of non-Western Caucasian participants (i.e. American Indian/Alaska native, Asian, Native Hawaiian or other Pacific Islander, Black or African American, White, Latin American or Unknown/Other) than Glasgow. Race was, therefore, included as a possible contributing factor to the below analyses, in order to assess if race had a significant contribution to the variance in the data.

**Table 4-1 Demographic Information**

	<b>Glasgow (n = 54)</b>		<b>UCSD (n = 87)</b>	
<b>Age</b>	Mean 23.04	(SD 4.63)	Mean 21.25	(SD 2.82)
<b>Gender</b>	31% M	69% F	29% M	71% F
<b>Handedness</b>	87% R	13% L	93% R	7% L
<b>Race</b>	70% White	30% Other*	31% White	69% Other*
<b>Education</b>	0% < High School	100% > High School	13% < High School	87% > High School
<b>Psych Diagnosis</b>	11% Yes	89% No	14% Yes	86% No
<b>Period of Unhappiness &gt; 2 weeks</b>	53% Yes	48% No	40% Yes	60% No
<b>Period Irritability &gt; 2 weeks</b>	28% Yes	72% No	18% Yes	82% No
<b>Relative with ASD</b>	4% Yes	96% No	5% Yes	95% No

\* Meets significance at  $p < 0.05$

**Descriptive statistics for psychometric measures were also compared between Glasgow and UCSD (see Table 4-2). Results revealed no significant difference between the groups on any of these measures. Dependent variables (i.e. Slope and PSE values) were also compared between Glasgow and UCSD (**

Table 4-3). Results revealed no significant difference between Glasgow and UCSD. Consequently, Glasgow and UCSD were included as one participant group in the analyses described henceforth. Location was, nevertheless, included as a possible contributing factor in the below analyses in order to remain watchful of any group differences in subsequent analyses.

**Table 4-2 Descriptive information about the psychometric measures**

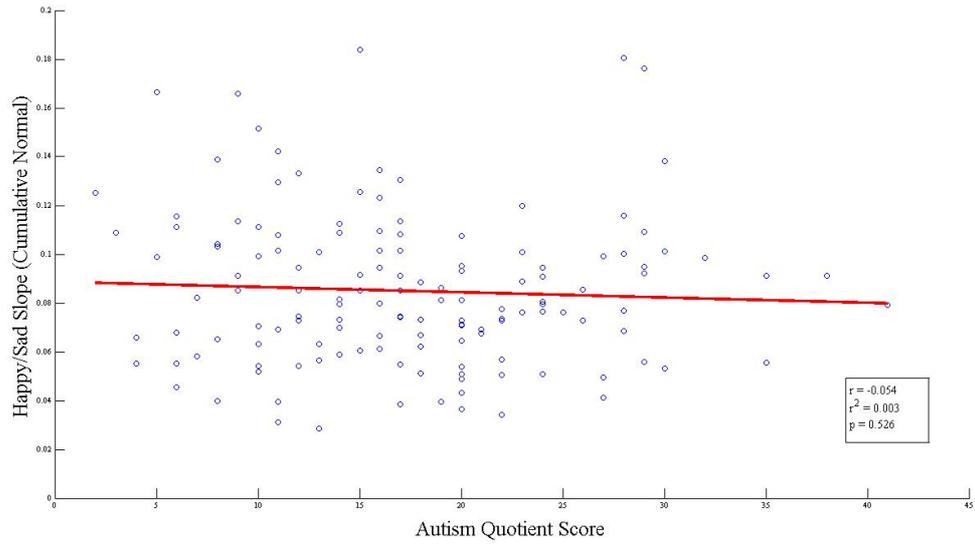
	Glasgow			UCSD		
	Mean	SD	Range	Mean	SD	Range
<b>AQ_Total</b>	18.28	9.13	2 - 41	16.50	6.84	4 - 35
<b>TAS_Total</b>	48.76	11.86	25 - 78	46.00	9.11	24 - 68
<b>IRI_Total</b>	68.04	11.07	37 - 90	69.25	9.56	50 - 92

**Table 4-3 Average scores for dependent variables for all participants, separated by location**

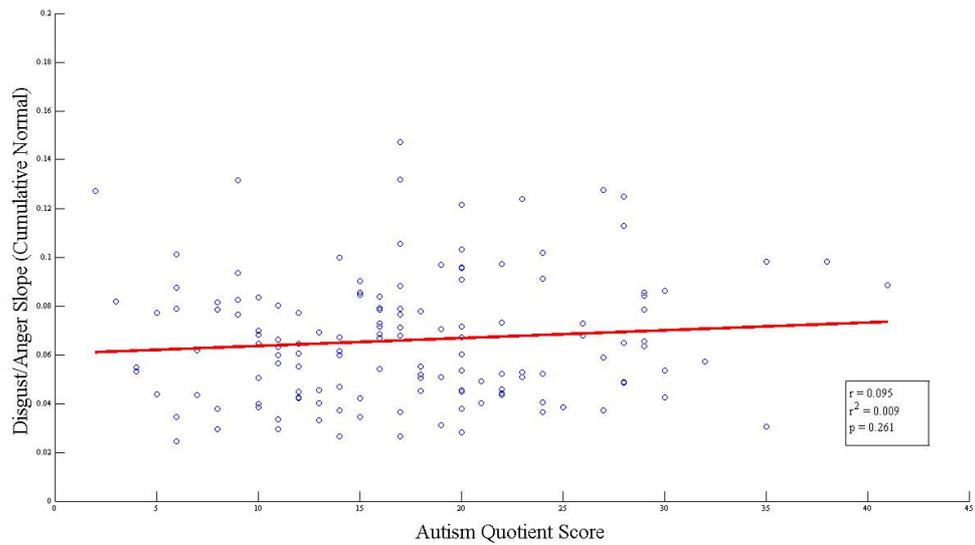
	Glasgow		UCSD	
	Mean	SD	Mean	SD
<b>Slope hap/sad</b>	0.09	0.03	0.08	0.03
<b>Slope dis/ang</b>	0.08	0.03	0.06	0.02
<b>Slope sur/fear</b>	0.06	0.02	0.06	0.02
<b>PSE hap/sad</b>	53.01	8.82	55.97	8.24
<b>PSE dis/ang</b>	54.16	6.41	54.43	6.71
<b>PSE sur/fear</b>	47.70	7.12	49.18	6.64

### 4.3.2 Pearson's correlation

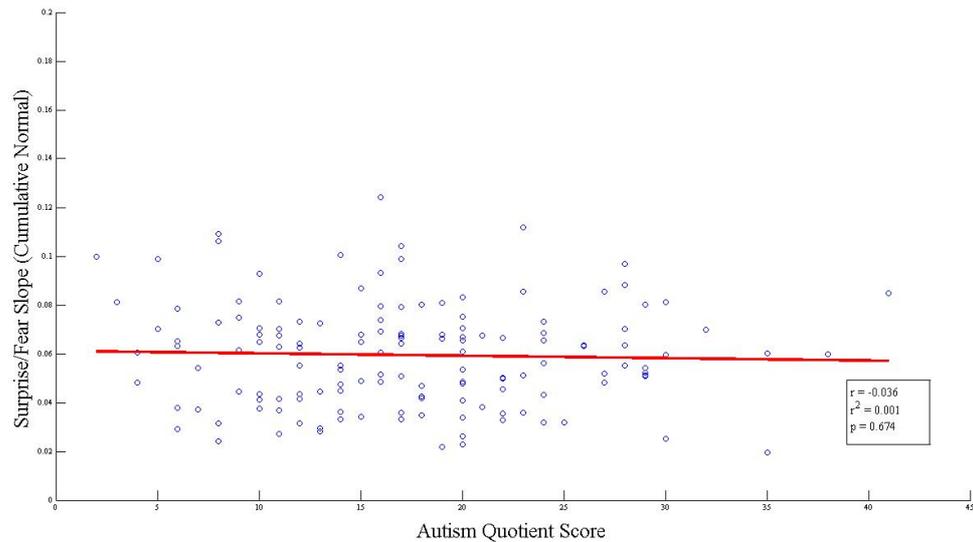
In the first instance the hypotheses that people with higher AQ scores would be poorer at emotion discrimination was assessed using a Pearson's correlation. This was performed between AQ and slope for all three conditions. Results revealed no significant correlation between AQ and slope for happy/sad ( $r = -0.05$ ,  $n = 141$ ,  $p = 0.53$ ; see Figure 4–6), disgust/anger ( $r = 0.1$ ,  $n = 141$ ,  $p = 0.26$ ; see Figure 4–7) or surprise/fear ( $r = -0.04$ ,  $n = 141$ ,  $p = 0.67$ ; see Figure 4–8).



**Figure 4–6 Pearson's correlation between slope values and AQ score (happy/sad condition)**

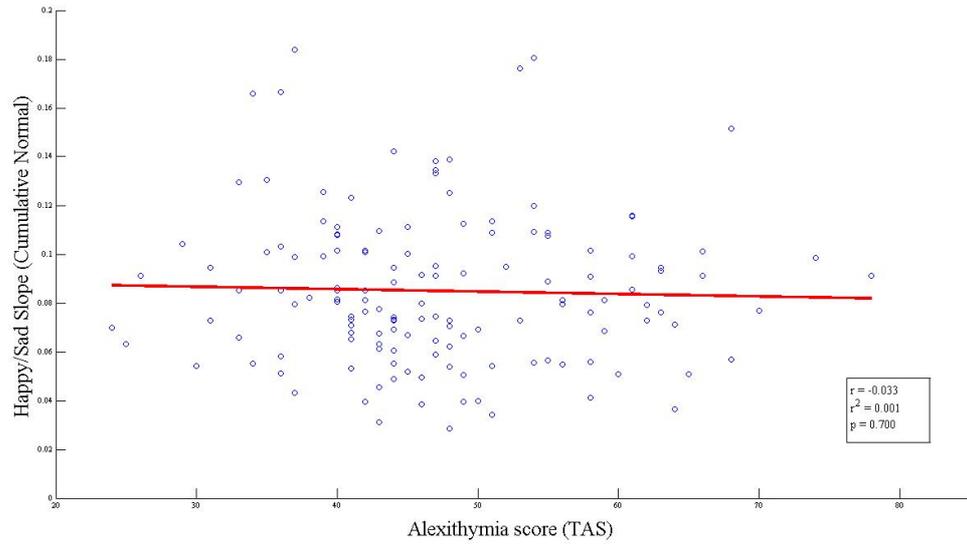


**Figure 4–7 Pearson's correlation between slope values and AQ score (disgust/anger condition)**

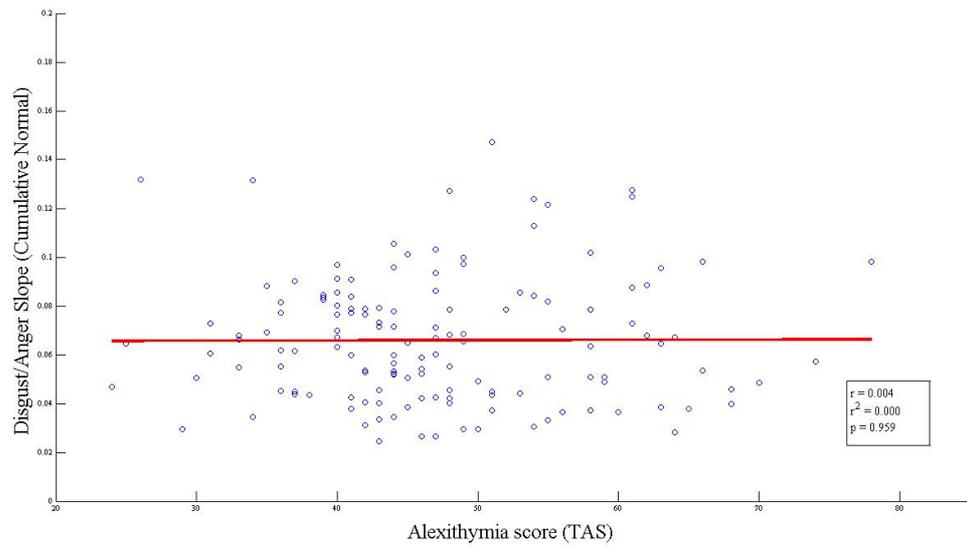


**Figure 4–8 Pearson's correlation between slope values and AQ score (surprise/fear condition)**

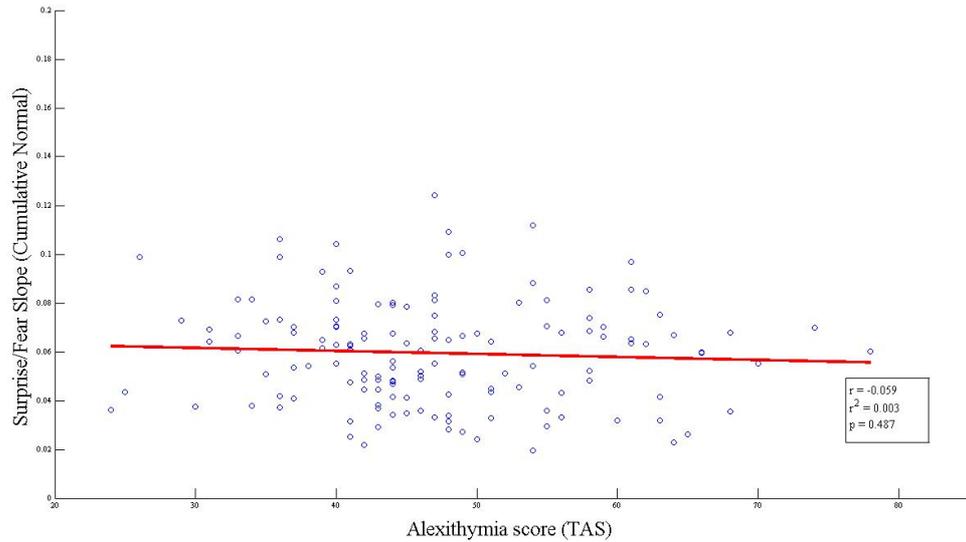
Since it was also hypothesised that, if AQ did not correlate with slope, alexithymia (TAS scores) would, independent samples t-tests were carried out between TAS score and slope values for each condition. No significant correlation was found between TAS and slope for happy/sad ( $r = -0.03$ ,  $n = 141$ ,  $p = 0.70$ ; see Figure 4–9), disgust/anger ( $r = 0.00$ ,  $n = 141$ ,  $p = 0.96$ ; see Figure 3-5) and surprise/fear ( $r = 0.06$ ,  $n = 141$ ,  $p = 0.49$ ; see Figure 4–11) conditions. Therefore, these initial analyses indicate that the null hypotheses are accepted because no correlation had been found between slope and AQ or TAS.



**Figure 4–9 Pearson's correlation between slope values and TAS score (happy/sad condition)**

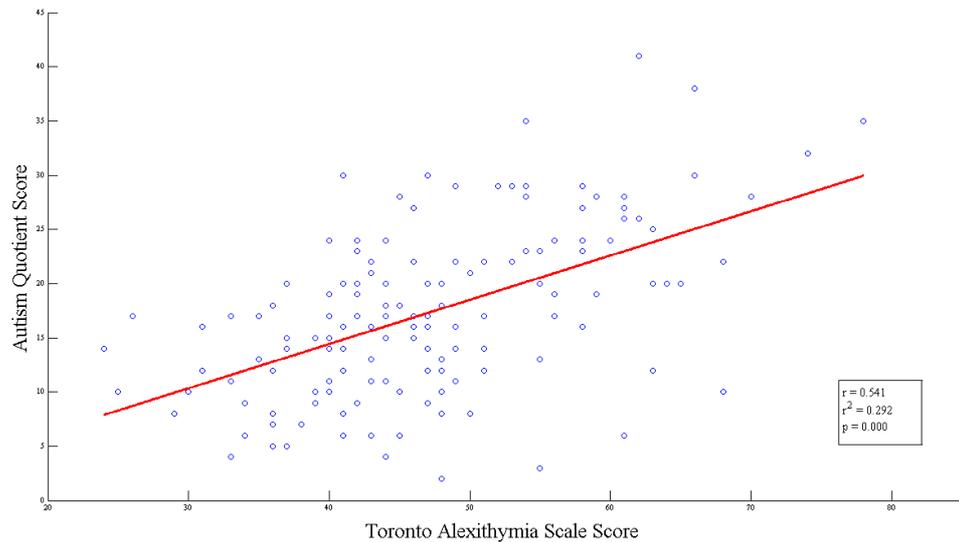


**Figure 4–10 Pearson's correlation between slope values and TAS score (disgust/anger condition)**



**Figure 4–11 Pearson's correlation between slope values and TAS score (surprise/fear condition)**

In order to further assess the relationship between AQ and TAS score, Pearson's correlation was performed between these variables. As shown in Figure 4–12, a significant positive correlation was found between AQ and TAS ( $r=0.541$ ,  $n = 141$ ,  $p < 0.01$ ), indicating that individuals with higher autism-like traits were more likely to have higher levels of alexithymia.



**Figure 4–12 Pearson's correlation between autism-like traits (AQ score) and alexithymia (TAS score)**

### 4.3.3 Multiple regression analysis

The above Pearson's correlation analysis indicated that neither AQ nor TAS scores predicted slope values. Therefore, Multiple Regression Analyses (MRAs) were performed in order to 1. further assess any nuanced contribution of AQ and TAS to the variance of slope and 2. assess the contribution of other variables that may better predict the variance in slope values. MRA was chosen as an ideal statistical measure due to the large number of possible contributing IVs. Because MRA requires, by its very nature, multiple IVs to be implemented into the model, it is susceptible to the unfavourable effects of multi-collinearity (i.e. two IVs being correlated with each other resulting in an artificially inflated prediction of the DV; Hair, Anderson, Tatham, and Black (1995)). The severity of multi-collinearity was then assessed by finding the variance inflation factor (VIF) via the equation:  $1/(1-R^2)$ . A VIF of 1 indicates no correlation and the most common cut off value for VIF is 10 (e.g., (Hair et al., 1995; Marquard.Dw, 1970; Neter, Wasserman, & Kutner, 1989)) However, more recent literature suggests a more conservative maximum VIF of 4 (e.g. Pan and Jackson (2008)). A cut off of point of  $<1$  and  $>4$  VIF was therefore implemented to this dataset. An MRA was conducted in order to assess the contribution of 13 possible IVs (including AQ and TAS) on the slope. Hence, the variables implemented in the multiple regression were as follows:

DV: Slope (derived from cumulative normal fit for each participant)

IVs:

- Location
- Age
- Gender
- Handedness
- Race
- Education
- Psych Diagnosis
- Period of Unhappiness > 2 weeks
- Period Irritability > 2 weeks
- Relative with ASD
- IRI Total
- TAS Total
- AQ Total

#### 4.3.3.1 Slope happy/sad condition

As a whole, the model did not account for a significant proportion of the variance in slope ( $R = 0.40$ ,  $R^2 = 0.16$ ,  $R^2$  Adjusted = 0.05;  $F_{(13,117)} = 1.50$ ,  $p = 0.13$ ). However, each IV was

assessed for its individual contribution to the model, calculated by dividing the unstandardised B value by its standard error. The B, beta and t values of all IVs were examined: significant t values indicated that the IV was making a significant contribution to the prediction of the DV. Age was found to be a significant contributor to the MRA model for happy/sad slope. Where  $B = 0.002$ ,  $\beta = 0.22$ ,  $t = 2.17$   $p = 0.03$ , i.e. increase in age indicated higher slope and therefore better emotion discrimination ability. For age, VIF was 1.246, meaning that age was not collinear to a significant level with any other IV when predicting slope. In addition to this, having a 'period of irritability > 2 weeks' was also found to be a significant individual contributor to the model (i.e.  $B = -0.02$ ,  $\beta = -0.21$ ,  $t = -2.06$ ,  $p = 0.04$ ,  $VIF = 1.33$ ). This provides an indication that emotion discrimination ability decreased as (the point biserial) group membership moved from 'no' to 'yes'. In other words, those who indicated they had a period of irritability for greater than 2 weeks were worse at emotion discrimination for happy/sad faces. Again, this IV was not found to be collinear to a significant level with any other IV when predicting slope, therefore risk of multi-collinearity was very low. These results suggest that, although this MRA model as a whole is non-significant, individual contributors may be age and having had a period of irritability. This analysis further verifies that neither AQ nor TAS predict emotion discrimination ability for happy/sad condition, and that age and a period of irritability may be better predictors of emotion discrimination.

In order to ratify the relative contribution of AQ and TAS. A second MRA was performed. Here, the IVs that were revealed to be most indicative of slope values were incorporated into the model, i.e. the variables incorporated into the MRA were as follows:

DV: Slope

IVs:

- age
- 'period of irritability > 2 weeks'

Results revealed that this model significantly accounted for the variance in slope ( $R = 0.22$ ,  $R^2 = 0.05$ ,  $R^2 \text{ Adjusted} = 0.03$ ;  $F_{(2,137)} = 3.35$ ,  $p = 0.03$ ). When AQ was added, it did not improve the model, but in fact reduced the significance of the model ( $R = 0.22$ ,  $R^2 = 0.05$ ,  $R^2 \text{ Adjusted} = 0.03$ ;  $F_{(3,137)} = 2.22$ ,  $p = 0.09$ ). AQ was removed from the model and this time TAS was added. Adding TAS to the model did not improve the prediction of slope, instead reducing the significance of the model ( $R = 0.22$ ,  $R^2 = 0.05$ ,  $R^2 \text{ Adjusted} = 0.03$ ;  $F_{(3,137)} = 2.24$ ,  $p = 0.08$ ). It is, therefore, apparent from the data that neither AQ score nor

TAS score predict slope values for the happy/sad condition and hence, an increase in autism-like traits or alexithymia is not associated with poorer emotion discrimination for happy and sad faces.

#### **4.3.3.2 Slope disgust/anger condition**

As above, the same MRA process was applied to slope values from the disgust/anger condition. Firstly slope values were compared to all 13 IVs. Second, only the IVs that were found to have a significant, or approach a significant, individual contribution to the model were incorporated in the MRA and thirdly AQ and TAS were applied in separate steps to the model in order to assess any possible contribution from them. The first analysis revealed that the model provided a significant prediction of slope ( $R = 0.48$ ,  $R^2 = 0.23$ , Adjusted  $R^2 = 0.13$ ;  $F_{(13,117)} = 2.40$ ,  $p < 0.01$ ). IVs were assessed for their individual contribution to the model, revealing that age was a significant predictor of slope ( $B = 0.002$ ,  $\beta = 0.24$ ,  $t = 2.45$   $p = 0.02$ ) (increase in age reflected better emotion discrimination) and period of unhappiness  $> 2$  weeks approached significance: ( $B = 0.01$ ,  $\beta = 0.20$ ,  $t = 1.99$   $p = 0.05$ ) (indicating 'yes' to having a period of unhappiness for greater than 2 weeks reflected better emotion discrimination). In order to assess the contribution of AQ and TAS to the model, each was added to a new model, which only implemented age and period of unhappiness. When age and period of unhappiness were entered into the MRA as a model, they significantly predicted slope ( $R = 0.35$ ,  $R^2 = 0.12$ , Adjusted  $R^2 = 0.110$ ;  $F_{(2,136)} = 9.281$ ,  $p < 0.01$ ). When AQ was added to this model it did not improve the fit of the model ( $R = 0.35$ ,  $R^2 = 0.12$ , Adjusted  $R^2 = 0.11$ ;  $F_{(3,136)} = 6.32$ ,  $p < 0.01$ ) with the contribution of AQ being:  $B = 0.00$ ,  $\beta = 0.05$ ,  $t = 0.55$ ,  $p = 0.58$ ). AQ was then removed from the model, and TAS was added. Results revealed that TAS did not improve the fit of the model ( $R = 0.35$ ,  $R^2 = 0.12$ , Adjusted  $R^2 = 0.10$ ;  $F_{(3,136)} = 6.32$ ,  $p < 0.01$ ; contribution of TAS:  $B = 0.00$ ,  $\beta = 0.00$ ,  $t = 0.00$ ,  $p = 0.99$ ). These findings provide further evidence that neither AQ nor TAS show any indication of predicting slope values.

#### **4.3.3.3 Slope surprise/fear condition**

As above, three separate multiple regression analyses were conducted in order to further explore the contribution of AQ and TAS to slope and also to unpick possible contributors, which may have otherwise been missed. As before, surprise/fear slope values were firstly compared to all 13 IVs. Second, only the IVs that were found to have a significant, or close to significant, individual contribution to the model were implemented in the MRA.

Thirdly, AQ and TAS were added to the MRA in separate steps in order to assess any possible contribution from them. The first analysis revealed that the model provided a significant prediction of slope ( $R = 0.46$ ,  $R^2 = 0.21$ , Adjusted  $R^2 = 0.11$ ;  $F_{13,117} = 2.12$   $p = 0.02$ ). Age was found to be a significant individual contributor ( $B = 0.002$ ,  $\beta = 0.34$ ,  $t = 3.48$ ,  $p < 0.01$ ;  $VIF = 1.25$ ), meaning that as age increased so did slope and therefore emotion discrimination was better for those who were older. No other IVs were indicated as being significant individual contributors to the MRA model. However, the next most significant predictor was Period of Unhappiness > 2 weeks which approached significance ( $B = 0.008$ ,  $\beta = 0.18$ ,  $t = 1.79$ ,  $p = 0.08$ ;  $VIF = 1.35$ ). This suggested that individuals who indicated that they had had a period of unhappiness for greater than 2 weeks may have been slightly *better* at emotion discrimination. Overall, these results suggest that neither AQ nor TAS predicted emotion discrimination for surprise and fearful faces. In addition to this, MRA results suggested that being older or having had a period of unhappiness may predict emotion discrimination ability.

For all conditions, neither AQ nor TAS predicted slope values. On the contrary, age was found to be a significant predictor of slope (as age increased so did emotion discrimination ability for all conditions). Also, having had a period of irritability of greater than 2 weeks may inhibit emotion discrimination ability (in the happy/sad condition only). In addition to this the data moved towards, but did not reach significance level  $p = 0.05$ , for period of unhappiness for greater than 2 weeks, giving a slight indication that emotion discrimination ability may be better if an individual reports a period of unhappiness.

#### **4.3.4 Slope: highest and lowest 10%**

It could be asserted that, by randomly sampling from the typical population, the aforementioned results are biased due to a large number of ‘intermediate’ individuals who have neither high nor low AQ scores. In order to assess the difference between those that had the very lowest and those that had the very highest AQ scores a t-test was performed on the top and bottom 10% of participants. Thus, the slope values of the lowest 10% AQ scorers ( $n = 14$ , mean AQ = 5.36, std. error = 0.44) and the highest 10% AQ scorers ( $n = 14$ , mean AQ = 31.71, std. error = 1.07) were compared using an independent samples t-test. No significant differences were found for happy/sad, disgust/anger or surprise/fear conditions. These results indicate that no significant difference was found in emotion discrimination ability for participants with very high autism-like traits compared to very low autism-like traits.

As highlighted above, it was hypothesised that despite a non-significant result between AQ score and slope, alexithymia (TAS score) may, none-the-less contribute in its own right. Therefore, the above analysis was repeated for TAS scores. Hence, the slope values of the lowest 10% TAS scorers ( $n = 14$ , mean TAS = 30.93, std. error = 0.99) and the highest 10% TAS scorers (mean TAS = 66.71, std. error = 1.23) were compared using an independent samples t-test. No significant differences were found for happy/sad, disgust/anger or surprise/fear conditions. These results indicate that relatively high alexithymic individuals did not significantly differ from relatively low alexithymic individuals in their ability to discriminate facial expressions of emotion.

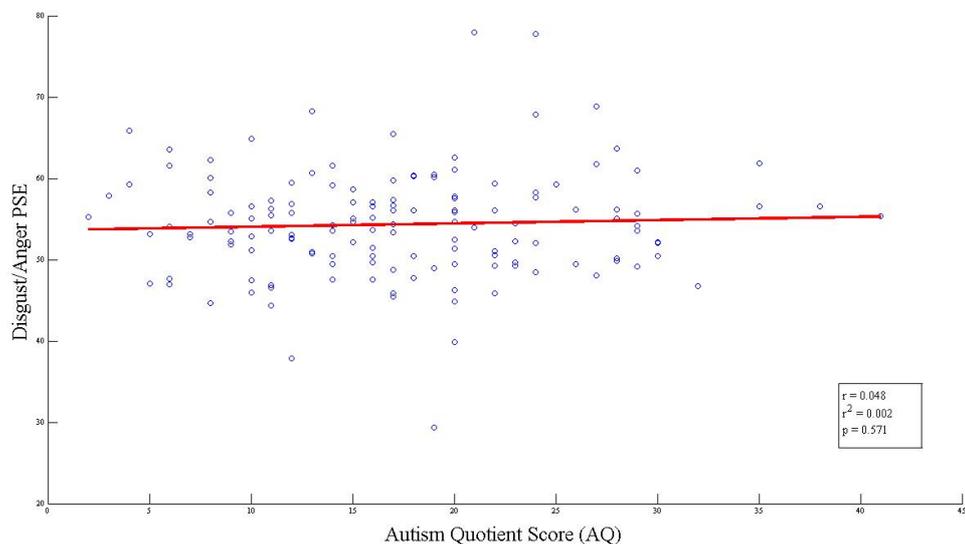
#### 4.4 Point of Subjective Equivalence (PSE)

Point of subjective equivalence (PSE) is the point at which the participant was equally likely to select either emotion label option in response to the stimulus. For example, in the happy/sad condition PSE value is the morph value (i.e. % of happy content in the face stimulus) when the participants selected happy 50% of the time. PSE is, therefore a measure of how biased the participant was to either of the emotion labels. For example, if the participant selected 'happy' 50% of the time and 'sad' 50% of the time when the stimulus displayed a 30% happy/70% sad morph then the individual would be biased towards happy emotion. i.e. although the stimulus was predominantly sad looking, the participant was equally likely to report it as happy than sad.

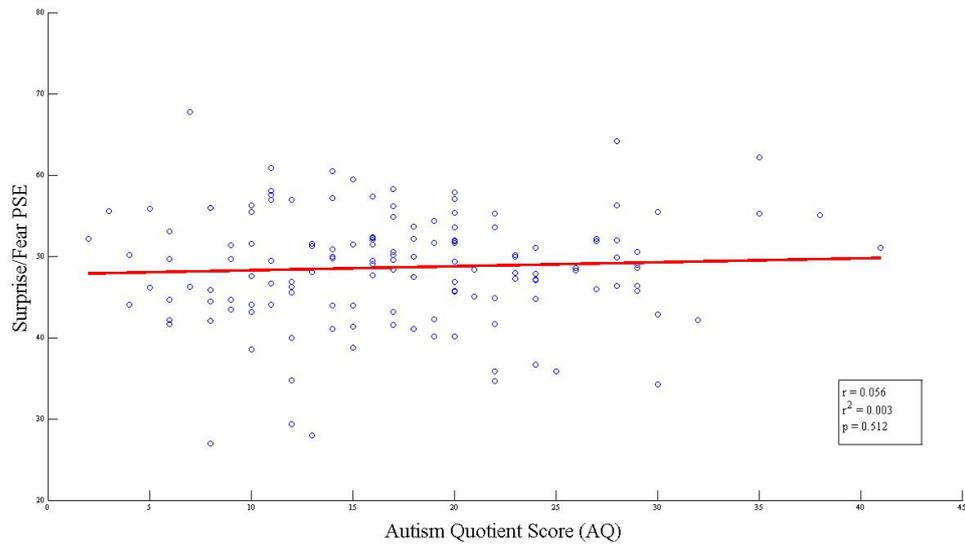
In addition to looking at discrimination (slope), the current study also measured perceptual bias (as measured using PSE). PSE values were compared using an independent samples t-test: between the highest 10% and lowest 10% AQ scores and the highest 10% and lowest 10% TAS scores. Therefore, for the lowest 10% AQ scorers ( $n = 14$ , mean = 5.36, std. error = 0.44) and the highest 10% AQ scorers ( $n = 14$ , mean = 31.71, std. error = 1.07), a t-test was performed for PSE values of all three conditions. A significant difference was found between high and low AQ scorers for the happy/sad condition  $t(26) = -2.16$ ,  $p = 0.04$ , where for the high AQ scorers the mean PSE was 58.25 (SD: 7.89, St. error of the mean: 2.11) and for low AQ scorers was 51.37 (SD: 8.98, St. error of the mean: 2.40). This suggested that the high AQ group was biased towards sad, meaning that they needed a higher proportion of 'happy' content in the face in order to classify it as 'happy', however it is noted that even for the low AQ scorers PSE is greater than 50, i.e. slightly biased towards sad. Therefore although a slight bias towards sad is found in low scorers this is significantly more pronounced in high AQ scorers.

A t-test was also performed on the lowest 10% TAS scorers ( $n = 14$ , mean = 30.93, std. error = 0.99) and the highest 10% TAS scorers ( $n = 14$ , mean = 66.71, std. error = 1.23) and, again, a significant difference was found for the happy/sad condition, with high TAS scorers having a bias towards happy  $t(26) = -2.51$ ,  $p = 0.02$ , no significant difference was found for disgust/anger or surprise/fear. For the happy/sad condition, high TAS scorers had a mean PSE of 58.45 (SD: 10.10, St error of the mean: 2.70) and low TAS scorers had a mean PSE of 48.99 (SD: 9.81, St. error of the mean: 2.62), hence individuals who were more alexithymic were significantly more biased towards the sad emotion.

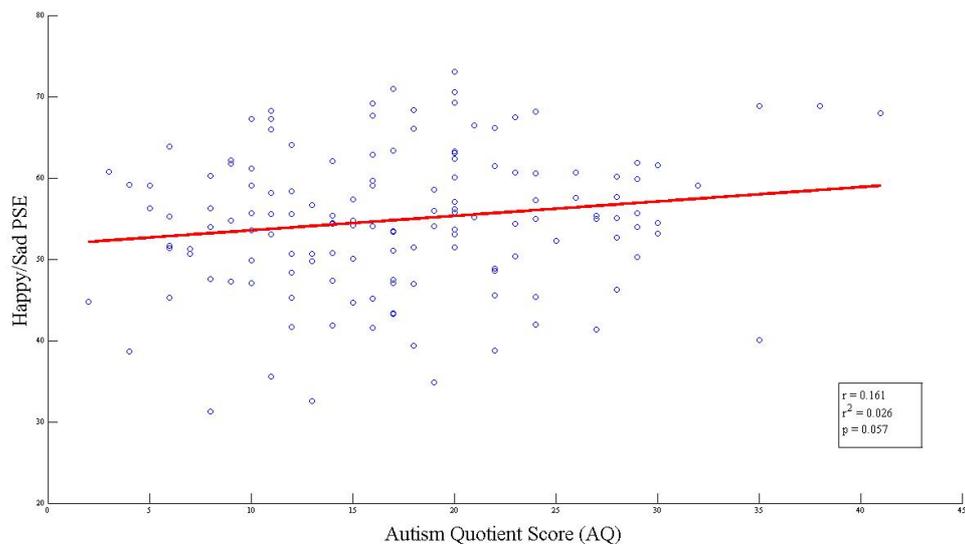
In order to assess the potential contribution of PSE further, a correlation of PSE with AQ was performed. A significant correlation was not found for the disgust/anger condition ( $r = 0.048$ ,  $n = 141$ ,  $p = 0.57$ ; see Figure 4–13) or the surprise/fear condition ( $r = 0.06$ ,  $n = 141$ ,  $p = 0.51$ ; see Figure 4–14), suggesting that AQ score did not influence being biased to one emotion or the other. However, for the happy/sad condition the correlation between AQ and PSE values approached significance ( $r = 0.16$ ,  $n = 141$ ,  $p = 0.06$ ; see Figure 4–15) indicating that having a higher AQ score may slightly bias the individual towards selecting ‘sad’ more than ‘happy’.



**Figure 4–13 Pearson's correlation between PSE values and AQ score (disgust/anger condition)**

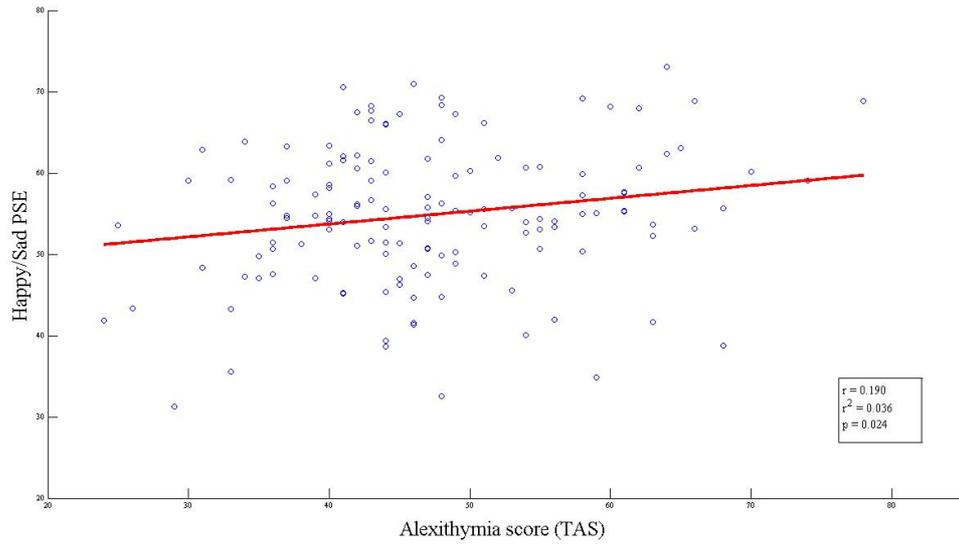


**Figure 4–14 Pearson's correlation between PSE values and AQ score (surprise/fear condition)**

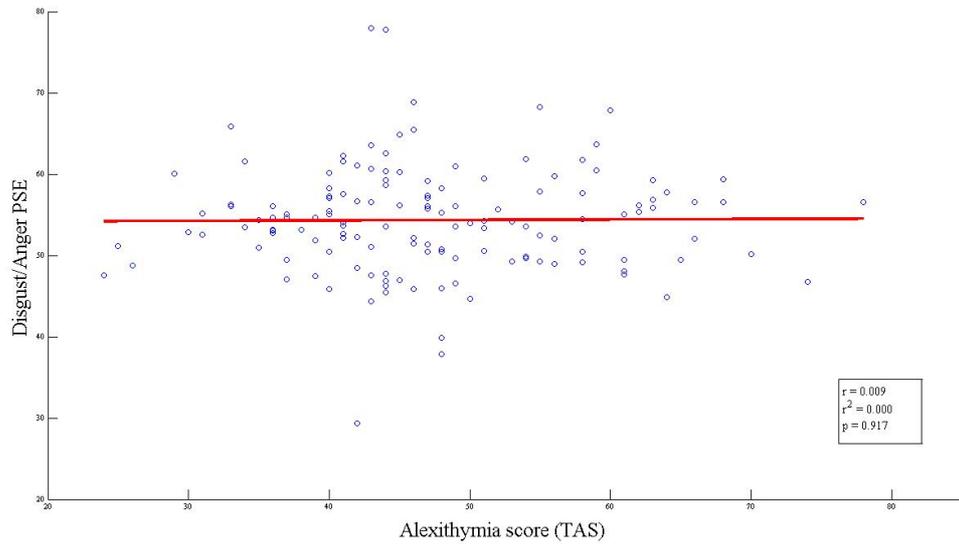


**Figure 4–15 Pearson's correlation between PSE values and AQ score (happy/sad condition)**

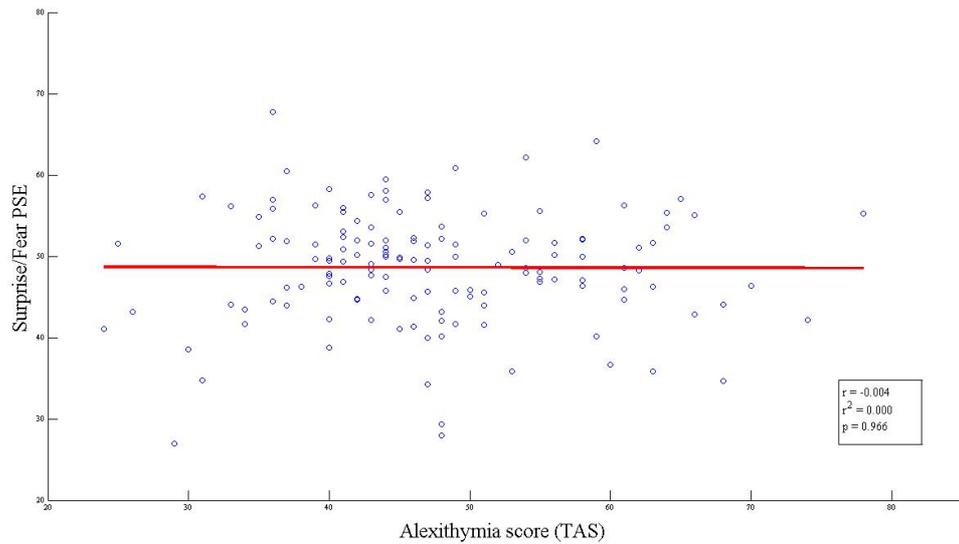
In order to assess the potential relationship between alexithymia and perceptual bias, a correlation between PSE and TAS was performed. A significant correlation was found between TAS and PSE for the happy/sad condition ( $r = 0.19$ ,  $n = 141$ ,  $p = 0.02$ ; see Figure 4–16). No significant correlation was found between TAS and PSE for the disgust/anger condition ( $r = 0.01$ ,  $n = 141$ ,  $p = 0.92$ ; see Figure 4–17) or for the surprise/fear condition ( $r = -0.01$ ,  $n = 141$ ,  $p = 0.97$ ; see Figure 4–18). Hence, these analyses suggest that participants who were more alexithymic were biased to the emotion ‘sad’ in the happy/sad condition but not for the disgust/anger or surprise/fear conditions.



**Figure 4–16 Pearson's correlation between PSE values and TAS score (happy/sad condition)**



**Figure 4–17 Pearson's correlation between PSE values and TAS score (disgust/anger condition)**



**Figure 4–18 Pearson's correlation between PSE values and TAS score (surprise/fear condition)**

## 4.5 Discussion

The current study assessed facial expression recognition ability and its relationship with autism-like traits and alexithymia in an adult population. The current study employed a procedure that carefully equated stimuli for luminance, contrast and visual angle, ensuring that group differences were not driven by the low-level visual characteristics of the stimuli. By fitting a cumulative normal function to each participant's dataset, a measure of emotion discrimination ability (slope) as well as perceptual bias (PSE) was obtained. These measures reflected an ability to discriminate between the emotions of happy and sad; disgust and anger; surprise and fear.

No relationship between level of autism-like traits and emotion discrimination ability was found in the current study. This is surprising given that there is much evidence to suggest that individuals with ASD have difficulty with emotion recognition (Harms et al., 2010), with some studies even suggesting particular confusions between disgust and anger, and fear and surprise (Jones et al., 2011). Poljac et al (2012) assessed emotion recognition in high and low AQ scorers and reported a significant difference between these groups for anger, disgust and sadness. Results from the current study are not consistent with this, instead finding no significant difference between high and low scorers in their emotion discrimination ability. Cook et al. (2013), on the other hand, found that autism severity in adults with ASD, as measured by AQ score, did not correlate with emotion discrimination

ability, a result that was reflected in the current study. However, Cook et al. (2013) also reported that, alexithymia, as measured by TAS score, correlated with emotion discrimination ability in the same sample of individuals, concluding that alexithymia, not AQ, predicted poor emotion discrimination. On the contrary, the current study found no significant correlation between alexithymia and emotion discrimination. Further, multiple regression analyses ratified the finding that neither AQ nor alexithymia score predicted emotion discrimination ability. Cook et al reported that TAS contributed to emotion discrimination ability *over and above* that of AQ because, when entered into a multiple regression analysis, AQ failed to significantly improve the model whereas TAS was successful in improving the model. The current study similarly assessed the relative contribution of AQ and TAS in improving the model of emotion discrimination ability, however neither AQ nor TAS significantly improved the model. Hence, the current study provides evidence against the proposal from Cook et al. (2013) that ‘alexithymia, not autism, predicts poor recognition of emotional facial expressions’.

The interplay between emotion perception, ASD and alexithymia is a growing area of research, with several studies in support of the notion that alexithymia, not autism, predicts emotion difficulties in ASD, e.g. (Bird & Cook, 2013; Bird, Press, & Richardson, 2011; P. Shah et al., 2016). The idea that alexithymia, not autism itself, predicts poor FER is compelling given that current literature in this area report conflicting findings. In other words, it is apparent that the research field is looking for an explanation as to why these conflicting results have arisen, and so, a co-morbid condition of alexithymia, present in some individuals with ASD but not others, would fit nicely as the missing piece to this puzzle. However, if this is true, is it unclear why alexithymia did not play a part in the participants’ ability to discriminate emotions in the current study. If alexithymia was the true reason for some individuals with ASD being impaired in FER and others not, then an obvious correlation between TAS score and FER ability would be expected. This study did not find that alexithymia score predicted FER ability and so it is apparent that more empirical evidence is needed in order to test this theory.

This study also aimed to explore the assertions that extraneous variables are often not accounted for in FER ASD research (Harms et al., 2010). A consistent predictor of emotion discrimination ability, across all three conditions, was age. It could be argued that participants who were older may have been more likely to pay close attention to the task, or perhaps took more time to consider their responses to each stimulus, but since reaction time was not measured this cannot be objectively assessed. Period of irritability also

emerged as a significant predictor of discrimination ability, but only for happy/sad faces, suggesting that individuals who have a tendency towards feeling irritable find it more difficult to discriminate between happy and sad faces. Period of unhappiness approached significance as a predictor of emotion discrimination ability (slope) for disgust/anger and surprise/fear. This suggests that individuals who have a tendency towards feelings of unhappiness may be more sensitive to subtle differences in these emotions. Despite 'period of unhappiness' emerging as a contributing independent variable to slope, a diagnosis of depression did not. It could be suggested that individuals who reported having a period of unhappiness may be more sensitive to the emotional affect of negative emotions, as has been found in neurotic vs. extrovert individuals (Larsen & Ketelaar, 1991). However, as these results only approached a significance level, more data would be needed in order to substantiate this suggestion.

In addition to the above findings, the present study also indicated that a perceptual bias towards sad faces in the happy/sad condition was found for individuals with high levels of autism-like traits. Similarly, a perceptual bias towards sad faces was found for participants who had greater alexithymia scores. This additional insight into perceptual bias adds to the aforementioned results because, although neither AQ nor alexithymia predicted differences in emotion discrimination, an atypical difficulty in perception of happy and sad faces emerged for those who had high AQ or high alexithymia scores. For example, Poljac et al. (2013) found high AQ scorers to be significantly worse than low AQ scorers for sad emotions, however, no measure of bias was assessed, hence results may have been indicative of a bias to sad as opposed to being 'better' at emotion detection. The present study was also careful to control for response bias, by requiring participants to switch response keys for half of the trials, which was not assessed in Poljac et al. (2013). Taken together, these results suggest that emotion discrimination ability is not predicted by autism-like traits or by alexithymia. However, nuances in perception of happy and sad faces indicate that those with high levels of autism-like traits and high levels of alexithymia may be biased towards sad emotions.

Overall, results from the present study suggest that neither AQ nor alexithymia predict facial expression discrimination. However, individuals who have high AQ traits (and therefore more likely to have ASD) cannot be considered equivalent to low AQ individuals because a perceptual bias towards sad was found for those with high levels of autism-like traits or alexithymia.

There are limitations to the current study which may have impacted the results described above, for example, overall, slope values were generally relatively low (all falling below 0.2) suggesting that the experimental task may have been somewhat difficult. The task itself required participants to discriminate between emotions after viewing them for a short period of time and therefore, some individuals may have found this challenging. It could be possible that, if floor effects were found, participants may have relied upon visual aspects of the stimulus and not the emotional affect itself. In comparison to similar studies (e.g. Poljac et al. (2013)), this study did not assess ‘accuracy’, so it could be suggested that the more complex design of this study may also have contributed to floor effects. Future studies would need to explore whether the present result could be replicated with a simpler design. Additionally, the present study could be improved by administering a more extensive measure of depression, for example the patient health questionnaire (PHQ-9; Kroenke, Spitzer, and Williams (2001)). In the present study, time constraints did not allow for the administration of established questionnaires on depression and so the general mood questions (e.g. ‘have you had a period of sadness for longer than two weeks?’) were used instead. These measures were considered as placeholders, with aims to assess depression measures in further detail in future studies. It could also be suggested that because the current study sampled from the ‘typical’ population (i.e. no participant had a known diagnosis of ASD), levels of autism may not have been high enough to facilitate difficulties in emotion perception. Future studies would also aim to incorporate a population of individuals with a diagnosis of ASD.

The current study explored the notion that alexithymia, not AQ may predict FER. Consequently, results indicated that neither alexithymia *nor* AQ score predict FER, based on evidence from an emotion discrimination task. Moreover, age was found to significantly predict emotion discrimination and so future research would benefit from considering covariates that may have otherwise been ignored. Overall, these data assert that neither AQ nor TAS predict emotion discrimination ability in a sample of non-ASD adults and hence expanding current research on autism like traits in the typical population as well as the relative contribution of alexithymia as a predictor of FER ability.

## **Chapter 5 General Discussion**

### **5.1 Overview of general discussion**

This thesis aimed to address the confusion in the literature surrounding facial expression recognition in individuals with ASD. To do this, a novel method of measuring FER was applied to a sample of adults (Chapter 2) and children with ASD (Chapter 3). Additionally, a sensitive measure of emotion discrimination was applied to typical adults, from which the contribution of autism-like traits and alexithymia to FER ability was explored (Chapter 4). The main findings of the thesis will be summarised regarding FER as measured by facial AUs, amplitude, latency and response patterns in adults with ASD and in children with ASD. Next, the contribution of alexithymia and autism-like traits to FER in the typical population will be discussed. Then, the contribution of these results to the field of facial expression processing in ASD will be evaluated and will be related to the theoretical accounts of FER in ASD. This will be followed by a discussion of the implications of these findings. General limitations of the research will be examined and, lastly, future directions of research will be suggested.

#### **5.1.1 Summary of findings**

##### **5.1.1.1 Atypicality of AUs**

Chapter 2 and Chapter 3 use a novel psychophysics technique (Generative Face Grammar (Yu et al., 2012)) to assess the recognition of facial expressions of emotion. By using this technique, a model, depicting the specific AUs that were used for FER, was drawn from a participant's response to the GFG stimuli. Hence, a rich dataset was provided, depicting the specific AUs that were meaningful for each emotion categorization (for adults: happy, surprise, fear, disgust, anger and sadness; for children: happy and anger). It was expected that the AUs used by individuals with ASD (children and adults) would differ (or be 'atypical') compared to TD individuals.

The results described in section 2.3.2 indicated that, as a group, adults with ASD were not significantly different from TD adults when both groups were compared to a reference dataset. This goes against our original hypothesis that adults with ASD would have atypical perception of AUs and also provides evidence against a basic FER 'impairment' in adults with ASD that has been presented in many research papers (S. Baron-Cohen et al., 1997; Clark, Winkielman, & McIntosh, 2008; Humphreys et al., 2007; Kennedy &

Adolphs, 2012; Philip et al., 2010). Further analysis suggested that specific individuals were different from the other members of their group, and hence suggest that: a) individuals with ASD vary in FER ability and b) reporting between-group differences alone may mask nuanced heterogeneity of FER ability across individuals with ASD. It was also found, in Chapter 2 that those individuals who were most ‘atypical’ had either the highest ADOS scores (in the ASD group) or the highest AQ scores (in the TD group). These results are tentative given that only two participants followed this trend, however, these results are meaningful given that previous research has found FER ability to be negatively correlated with measures of autism ‘severity’ (B. T. Williams & Gray, 2013). As described later in the text, further research would aim to ratify this relationship between ADOS/AQ scores and FER ability.

### **5.1.1.2 Amplitude**

Previous research has suggested that stimuli with greater ‘intensity’ (i.e. amplitude of the facial expression) elicit better accuracy in FER tasks compared to less intense stimuli (Law Smith et al., 2010). The results presented in Chapter 2 and Chapter 3 tested this in adults and children with ASD and revealed no connection between FER ability and intensity when the models were assessed as a whole. When the amplitude of the AUs were assessed separately for the eye and the mouth region, it emerged that the TD adults had significantly higher amplitude values for the eye region *and* the mouth region compared to adults with ASD. This provides evidence against claims that individuals with ASD require more intensity in a facial expression in order to accurately categorise (Law Smith et al., 2010) and also, somewhat surprisingly, provides evidence against the assertion that individuals with ASD utilise information from the mouth area over and above the eye area (Joseph & Tanaka, 2003; Spezio et al., 2007a).

### **5.1.1.3 Latency**

In Chapter 2 and Chapter 3 no significant differences in FER relative to timing were found for children or adults with ASD compared to TD individuals. This counters previous research that has suggested that timing is an important, and often overlooked, component of FER that influences the ability to accurately recognise emotions (Lerner et al., 2013; Philip et al., 2010; Sato et al., 2013). Lerner et al., (2013) argued for a ‘delayed social information processing speed’, indicating that individuals with ASD required additional time in order to process facial expressions and accurately recognise their emotion. The data presented in this thesis do not support this finding, instead suggesting equivalence in

this skill between ASD and TD groups for both adults and children. It could be suggested, therefore, that including a dynamic element to the stimuli (and hence making them more life-like), FER becomes somewhat *easier*. These sentiments resonate with the results discussed in several papers such as (Gepner et al., 2001; Loveland et al., 1997) and perhaps support Enticott et al. (2014), who argued that including a dynamic aspect to facial expression stimuli improves FER performance in individuals with ASD.

#### 5.1.1.4 Response patterns

Previous research such as Leung et al. (2013) found children with ASD to be slower in RT compared to TD children when categorizing facial expressions of emotion. It was expected, therefore, that in Chapter 3, children with ASD would have been slower in comparison to TD children. However, as described in section 3.3.1.1, the data indicated no RT bias in children with ASD as compared to TD children. These findings were, therefore, somewhat surprising. It could be argued that children with ASD are equivalent to TD children in their RT, despite some children responding in a relatively ‘atypical’ manner. In addition to this, it should be noted that the very nature of the task used in Chapter 3 might have influenced RTs between groups. For example, the task (unlike Leung et al. (2013)) required a ‘yes’ or ‘no’ response. Because of this, it could be argued that regardless of FER ability, reaction time would be quicker because participants did not have the increased cognitive demand of *choosing* between several emotion label options. In certain types of visual tasks, children with ASD have been found to be *faster* in RT compared to TD children (Plaisted, O’Riordan, & Baron-Cohen, 1998). Hence, it could also be argued that experimental designs that require less cognitive demand, in terms of decision-making, may override RT differences. Therefore, the findings reported here (that children with ASD were equivalent to TD children in RT to facial expression categorization) are, perhaps, not as surprising as first suggested.

Chapter 3 also indicated that children with ASD were more bias to the selection of ‘yes’ in their judgment of whether stimuli were ‘angry’ or not, compared to the TD group. TD children were more likely to select ‘no’ to random-looking stimuli when asked ‘is this person angry?’ However, both groups were equally ‘conservative’ in their judgment of ‘happy’. For adults (see Chapter 2), the task was slightly more taxing because instead of ‘yes’/‘no’ responses, participants responded from a selection of seven choices: ‘happy’, ‘surprise’, ‘fear’, ‘disgust’, ‘anger’, ‘sadness’ or ‘other’. For these participants, results indicated that the ASD group were more likely to select ‘other’ out of the seven options,

while the TD group were more likely to select ‘surprise’. This could indicate firstly that children with ASD may have greater difficulty with the emotion anger than typically developing children. It could also suggest that FER ability improves with age because while *children* with ASD appeared to be much less conservative to selecting the stimuli as ‘angry’, *adults* with ASD were significantly more conservative than TD adults for all emotions, favouring to choose ‘other’ instead of the six basic emotion labels.

#### **5.1.1.5 The contribution of age group, AQ and alexithymia**

Chapter 2 indicated that, although autism ‘severity’ was linked with FER ability, alexithymia scores were not. In an investigation of a larger sample of participants (Chapter 4), no evidence was found for a correlation between alexithymia score and FER ability. This provides evidence against the theory that alexithymia predicts FER ability over and above ASD, as suggested by Bird and Cook (2013). In addition to this, no correlation was found between AQ score and FER ability in this sample, however AQ score *was* linked to FER ability in typical individuals in Chapter 2. It could be argued that the tasks used in Chapter 2 and Chapter 3 were more demanding (i.e. required greater cognitive effort and FER skill), and, hence, may have produced AQ-dependent, FER atypicalities more so than in Chapter 4.

In addition to this, individual differences within groups were found to be informative about FER in ASD. In section 2.3.2 results indicated that, although no significant differences were presented *group-wise*, individual differences were found *within* groups. Hence, the particular ‘atypical’ individuals would not have been picked up in a group-wise analysis. The evidence given in this thesis for individual differences in adults with ASD suggests that assessing group differences alone may overshadow individual differences in some cases and, therefore, the proportion of studies that have shown a significant difference between ASD and TD groups in FER ability may be overestimated (Nuske et al., 2013; Uljarevic & Hamilton, 2013). It is also striking that the individuals who were found to be the most ‘atypical’ had the highest ADOS scores, hence suggesting that FER ability may be directly related to autism severity and thus supporting the notion that FER ability is negatively correlated with autism severity (Lerner et al., 2013; Uono et al., 2013; B. T. Williams & Gray, 2013). Also, it could be argued that some individuals may be at different stages in ‘social training’ (e.g. FER interventions), and so may be more skilled in compensatory strategies than others, depending on the resources that have been available to them (Rutherford & McIntosh, 2007; Walsh et al., 2014)

### 5.1.2 Condensed summary of findings

In a sample of adults with ASD, results indicated relatively equivalent FER abilities compared to TD adults. No differences were found in the timing of the models, suggesting that adults with ASD do not have a delayed social information processing speed. FER was also assessed in children with ASD and TD children using a similar, but simplified technique. Here, relatively equivalent results were also found between ASD and TD groups. Hence, both of these studies do not support theories of an ‘impairment’ in basic emotion recognition. These data also do not support the idea that facial expressions of greater ‘intensity’ are easier for people with ASD to recognise. AQ was found to be linked to FER in the adult TD group (Chapter 2), however alexithymia was not. The contributions of AQ and alexithymia were explored further in a facial expression discrimination task with TD adults (Chapter 4), which revealed that alexithymia did not predict FER ability more so than AQ.

## 5.2 Contributions to the field

As reviewed in Chapter 1, FER has been studied extensively in individuals with ASD. However, there are many conflicting results in the literature, which have been attributed to several factors, including methodological limitations. The studies reported in this thesis (Chapter 2 and Chapter 3) employ a novel technique, allowing for simultaneous measurement of the identity, activation level and relative timing of the facial action units contributing to FER in an individual with ASD, a method that has never previously been applied to the field of ASD. Results obtained with this technique revealed that individuals with ASD and TD individuals (both adults and children) were relatively similar in terms of the AUs and intensity that were associated with their recognition of facial expressions, however timing was atypical in the adult ASD group and, in particular, timing of FER was late in comparison to TD individuals. In addition, these methods provided an insight into individual differences within groups, suggesting that particular individuals struggled with the FER task more than others (which may have been overlooked if only group data were assessed). Hence, the studies presented in this thesis also provide a novel approach to assessing FER in individuals with ASD because FER can be modelled at the *individual* level. Additionally, this thesis assessed FER from a ‘modern’ perspective by assessing the contribution of AQ and alexithymia (Chapter 4), revealing that neither of these is consistently predictive of FER abilities.

### 5.3 Theoretical perspectives

Theory of Mind (ToM), first applied to ASD research by S. Baron-Cohen et al. (1985), proposes that the characteristics of ASD could be explained via an underlying cognitive impairment in the ability to attribute mental states to oneself and to others. The research presented in this thesis can be related to ToM insofar as accurate FER requires, by its very nature, proficient skills in attributing mental states to others. The results presented in this thesis suggest that adults with ASD are somewhat similar in FER as compared to TD individuals. This suggests that adults with ASD may have intact ToM, as they are able to attribute appropriate emotional affect from facial expression based on *some* components of the face. However it is difficult to separate recognition of facial expressions (in the perceptual sense) from the understanding of emotional affect (in the cognitive sense). This is also reflected in Chapter 3 and Chapter 4 and so ToM is only able to explain the results presented in this thesis to a partial extent. The Central Coherence (CC) theory also provides a useful framework to conceptualise the results presented in this thesis. CC posits that individuals with ASD have superior abilities to TD individuals in processing the ‘local’ aspects (i.e. small parts) of a visual scene. Results presented in this thesis indicated that adults and children with ASD were relatively typical in terms of the AUs, amplitude and timing that were present in their models. It could be argued that individuals with ASD may have intact visual perception of the face, but have difficulty with deriving social meaning from such stimuli (as suggested by Weigelt et al. (2013)) which would be consistent with CC theory.

### 5.4 Implications of these findings

Studies found in the FER in ASD literature have generally been split between two camps based on their conclusions: 1) individuals with ASD have a core difficulty with FER and 2) individuals with ASD have no difficulty with FER (Harms et al., 2010; Simmons et al., 2009; Uljarevic & Hamilton, 2013). The research presented in this thesis presented a range of abilities in FER, which vary across individuals with ASD due to a variety of different factors, including age and autism severity. However, this thesis argues that, overall, individuals with ASD are relatively typical in FER. It could be argued that each individual holds a unique internal representation of FER, observable through the methods used in this Chapter 2 and Chapter 3. The data suggest that each individual’s mental representation of an emotion is likely to vary, in differing degrees, to the mental representations experienced by TD individuals. In addition to this, it was apparent from the research presented in this

thesis that children with ASD are also relatively typical in their FER ability. Further, research presented in this thesis provided no evidence to support the idea that alexithymia predicts FER ability better than ASD or AQ score. This provides evidence against the notion that FER difficulty, instead of being a core characteristic of ASD, can be explained as co-morbid alexithymia.

In terms of implications for clinical practice, it was initially thought that the technique presented in Chapter 2 could be applied as a therapeutic tool for the improvement of FER skills. This was based on the notion that internal representations of facial expressions could be modelled at an individual level, and hence, these data could be used in a training setting to reinforce particular FER difficulties experienced by an individual. It was expected, therefore that an individually tailored intervention experience could be produced from this technique. However, it is not clear from the research presented in this thesis if this technique would be appropriate as an intervention for FER in individuals with ASD. Results have suggested that an individual model *can* be obtained for individuals with ASD, however the procedures specific to this method have proved somewhat challenging for individuals with ASD (e.g. difficulties in maintaining attention in children with ASD). Hence, this technique has implications for improving individually tailored interventions but is slightly hindered in terms of a clinical practicality due to the time and concentration commitments required by each individual.

## 5.5 General limitations

A tricky aspect of FER research is obtaining a balance between stimuli that are ‘life-like’ while retaining a high level of control over the presentation of the stimuli and the sensitivity of the measure. The stimuli used in Chapter 2 and Chapter 3 allow for a very high level of stimulus control as well as high sensitivity of measurement. However this requires a slight trade-off in terms of the ‘life-like’ quality of the stimuli, of which real, face-to-face interaction would be best. The stimuli are limited insofar as they lack eye-movements. Research has suggested that children with ASD label facial expression stimuli as more intense when the stimuli have directed (i.e. looking straight ahead) eye gaze (Tell et al., 2014). Hence, fixed gaze of the stimuli may have influenced some of the results presented in this thesis. The stimuli used in these chapters are currently being updated to include movements of the eyes and more realistic representations of the teeth, and so future research with these stimuli would involve an even higher level of ecological validity.

The studies presented in this text were in collaboration with Universities in North America. Because of this, a proportion of participants from each study were recruited in North American countries. The advantages of conducting research in this collaborative way include: access to facilities, access to clinical populations (i.e. children and adults with ASD) and research expertise. However, the drawbacks of using this method are that potential cultural differences between UK and North American participants cannot be completely accounted for. Efforts were made to display detailed demographic information about the participant groups, however the inclusion of both UK and North American participants is, naturally, a limitation to the studies presented in this thesis.

The studies presented in this thesis also recruited only individuals with average, or above average IQ. It was assumed that individuals on the spectrum with low IQ, complex needs or learning disabilities would not have been able to complete the tasks. Consequently, the data presented in this thesis do not reflect the ASD population as a whole. Due to the cognitive demands of the tasks presented in this thesis, adaptations would have been needed to be made in order to assess FER in individuals with ASD who had below average IQ. Future research would benefit from recruiting a wider population of individuals across the autism spectrum.

## **5.6 Future directions**

As highlighted in this thesis, few studies have assessed FER in ASD using novel techniques. This thesis presented the first application of the Generative Face Grammar (Yu et al., 2012) to individuals with ASD and is the first study in this field to simultaneously measure the contribution of AUs, timing and intensity to FER. Future research could build upon this by producing a replication of the study presented in Chapter 2. This would be useful in order to reveal any *consistent* atypicalities across individuals with ASD, if such consistencies exist. Also, further research would aim to ratify the relationship between ADOS/AQ scores and FER ability.

In addition to this, future research could build upon the findings reported in Chapter 4 by assessing the contribution of alexithymia and AQ score to the recognition of *complex* emotions. This would shed light on whether individuals who are high in alexithymic traits or autism-like traits have difficulty with FER in a more complicated facial recognition task.

This thesis was somewhat exploratory in nature, in part assessing the effectiveness of a novel technique and its ability to be applied practically to a clinical population. The research presented in this thesis may provide a platform for similar techniques to be applied to different groups of clinical populations, for example, as applied in a case study of an individual with prosopagnosia (Richoz et al., 2015).

## **5.7 Conclusions**

This thesis aimed to assess FER in individuals with ASD and its relationship to AQ and alexithymia. The research presented in this thesis revealed that children and adults with ASD were relatively typical in their FER ability. Hence, these data provide little evidence in support of a universal ‘impairment’ of FER. The research presented in this thesis also provide evidence against the notion that alexithymia predicts FER over and above ASD symptoms and revealed little evidence of a link between levels of autism-like traits and FER ability in the typical population. Hence, this thesis adds to current scientific knowledge by employing a new and innovative method to assess FER in individuals with ASD. From this, a rich tapestry of data has indicated that there are no obvious differences between ASD and TD individuals in their FER abilities.

## Appendices

### **Appendix 1: Videos of stimuli generation**

A dynamic figure of how the stimuli were generated from Yu et al. (2012) can be viewed or downloaded using the following link:

<https://owncloud.gla.ac.uk/cloud/s/I1SVFfPuGqWCSNZ>

### **Appendix 2: Videos of reverse correlation models for each participant**

Dynamic models for each participant can be displayed in video format, which can be viewed online or downloaded using the following link:

<https://owncloud.gla.ac.uk/cloud/s/AYiUgnoh32qxIDy>

### **Appendix 3: Videos of group models**

Group models, i.e. depicting the AUs and temporal dynamics as averaged for each group, can be displayed in video format, which can be viewed or downloaded using the following link:

<https://owncloud.gla.ac.uk/cloud/s/pv3lKH1AVxwHIEb>

### **Appendix 4: Videos of NMF decompositions**

Decompositions, rendered into a video format can be viewed or downloaded using the following link:

<https://owncloud.gla.ac.uk/cloud/s/17KGhEwbrIomN8i>

## References

- Aaron, R. V., Benson, T. L., & Park, S. (2015). Investigating the role of alexithymia on the empathic deficits found in schizotypy and autism spectrum traits. *Personality and Individual Differences, 77*, 215-220.
- Abdi, H., & Williams, L. J. (2010). Tukey's honestly significant difference (HSD) test. *Encyclopedia of Research Design* (pp. 1-5). Thousand Oaks, CA: Sage.
- Abrosoft. (2015). Abrosoft Fantamorph (Version 5).
- Adobe. (2015). Adobe Photoshop Elements.
- Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*(7021), 68-72.
- Adolphs, R., Spezio, M. L., Parlier, M., & Piven, J. (2008). Distinct face-processing strategies in parents of autistic children. *Current Biology, 18*(14), 1090-1093.
- Ahumada, A., & Lovell, J. (1971). Stimulus Features in Signal Detection. *Journal of the Acoustical Society of America, 49*(6), 1751-&.
- Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "Red Flags" for autism screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist for Autism in toddlers in 1,000 cases and 3,000 controls [corrected]. *J Am Acad Child Adolesc Psychiatry, 51*(2), 202-212 e207.
- Ashwin, C., Baron-Cohen, S., Wheelwright, S., O'Riordan, M., & Bullmore, E. T. (2007). Differential activation of the amygdala and the 'social brain' during fearful face-processing in Asperger Syndrome. *Neuropsychologia, 45*(1), 2-14.
- Ashwin, C., Chapman, E., Colle, L., & Baron-Cohen, S. (2006). Impaired recognition of negative basic emotions in autism: A test of the amygdala theory. *Soc Neurosci, 1*(3-4), 349-363.
- Back, E., & Jordan, T. R. (2014). Revealing variations in perception of mental states from dynamic facial expressions: a cautionary note. *Plos One, 9*(1), e84395.
- Bagby, R. M., Parker, J. D. A., & Taylor, G. J. (1994). The 20-Item Toronto-Alexithymia-Scale .1. Item Selection and Cross-Validation of the Factor Structure. *Journal of Psychosomatic Research, 38*(1), 23-32.
- Bagby, R. M., Taylor, G. J., & Parker, J. D. A. (1994). The 20-Item Toronto-Alexithymia-Scale .2. Convergent, Discriminant, and Concurrent Validity. *Journal of Psychosomatic Research, 38*(1), 33-40.
- Balia, C., Placini, F., Liori, A., Peddis, C., Carucci, S., & Zuddas, A. (2014). Face and emotion recognition in children with ADHD and autism spectrum disorders: effects of methylphenidate. *European Neuropsychopharmacology, 24*, S725-S726.
- Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. : MIT press.
- Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends Cogn Sci, 6*(6), 248-254.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*(1), 37-46.
- Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J Autism Dev Disord, 34*(2), 163-175.
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry, 42*(2), 241-251.

- Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger Syndrome. *Visual Cognition*, 4(3), 311-331.
- 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. *J Autism Dev Disord*, 31(1), 5-17.
- Baroncohen, S., Spitz, A., & Cross, P. (1993). Do Children with Autism Recognize Surprise - a Research Note. *Cognition & Emotion*, 7(6), 507-516.
- Batty, M., & Taylor, M. J. (2006). The development of emotional face processing during childhood. *Developmental Science*, 9(2), 207-220.
- 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.
- Bird, G., & Cook, R. (2013). Mixed emotions: the contribution of alexithymia to the emotional symptoms of autism. *Transl Psychiatry*, 3.
- Bird, G., Press, C., & Richardson, D. C. (2011). The role of alexithymia in reduced eye-fixation in Autism Spectrum Conditions. *J Autism Dev Disord*, 41(11), 1556-1564.
- Blais, C., Roy, C., Fiset, D., Arguin, M., & Gosselin, F. (2012). The eyes are not the window to basic emotions. *Neuropsychologia*, 50(12), 2830-2838.
- 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.
- Borghuis, B. G., Perge, J. A., Vajda, I., van Wezel, R. J. A., van de Grind, W. A., & Lankheet, M. J. M. (2003). The motion reverse correlation (MRC) method: A linear systems approach in the motion domain. *Journal of Neuroscience Methods*, 123(2), 153-166.
- Bormannkischkel, C., Vilsmeier, M., & Baude, B. (1995). The Development of Emotional Concepts in Autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 36(7), 1243-1259.
- Braverman, M., Fein, D., Lucci, D., & Waterhouse, L. (1989). Affect Comprehension in Children with Pervasive Developmental Disorders. *J Autism Dev Disord*, 19(2), 301-316.
- Brewer, R., & Murphy, J. (2016). People with autism can read emotions, feel empathy. *Scientific American*, 12.
- Brugha, T. S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., . . . Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Arch Gen Psychiatry*, 68(5), 459-465.
- Buitelaar, J. K., Van der Wees, M., Swaab-Barneveld, H., & Van der Gaag, R. J. (1999). Theory of mind and emotion-recognition functioning in autistic spectrum disorders and in psychiatric control and normal children. *Development and Psychopathology*, 11(1), 39-58.
- Campatelli, G., Federico, R. R., Apicella, F., Sicca, F., & Muratori, F. (2013). Face processing in children with ASD: Literature review. *Res Autism Spectr Disord*, 7(3), 444-454.
- Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of Simple and Complex Emotions in Nonretarded-Children with Autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 33(7), 1169-1182.
- Caron, R. F., Caron, A. J., & Myers, R. S. (1982). Abstraction of Invariant Face Expressions in Infancy. *Child Development*, 53(4), 1008-1015.
- Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism*, 9(4), 428-449.

- Cave, S. F. (2008). The History of Vaccinations in the Light of the Autism Epidemic. *Alternative Therapies in Health and Medicine*, 14(6), 54-57.
- Celani, G., Battacchi, M. W., & Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *J Autism Dev Disord*, 29(1), 57-66.
- Chan, J. S., & Naumer, M. J. (2014). Explaining autism spectrum disorders: central coherence vs. predictive coding theories. *J Neurophysiol*, 112(11), 2669-2671.
- Charman, T., Swettenham, J., BaronCohen, S., Cox, A., Baird, G., & Drew, A. (1997). Infants with autism: An investigation of empathy, pretend play, joint attention, and imitation. *Developmental Psychology*, 33(5), 781-789.
- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends Cogn Sci*, 16(4), 231-239.
- 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.
- Cohen, M. A. (2005). Against basic emotions, and toward a comprehensive theory. *Journal of Mind and Behavior*, 26(4), 229-253.
- Collin, L., Bindra, J., Raju, M., Gillberg, C., & Minnis, H. (2013). Facial emotion recognition in child psychiatry: A systematic review. *Res Dev Disabil*, 34(5), 1505-1520.
- Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Arch Gen Psychiatry*, 60(5), 524-530.
- 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.
- Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis*: Russell Sage Foundation.
- 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.
- Crawford, H., Moss, J., Anderson, G. M., Oliver, C., & McCleery, J. P. (2015). Implicit Discrimination of Basic Facial Expressions of Positive/Negative Emotion in Fragile X Syndrome and Autism Spectrum Disorder. *Am J Intellect Dev Disabil*, 120(4), 328-345.
- Dakin, S., & Frith, U. (2005). Vagaries of visual perception in autism. *Neuron*, 48(3), 497-507.
- Das, P., Coulston, C. M., Bargh, D. M., Tanious, M., Phan, K. L., Calhoun, V. D., & Malhi, G. S. (2013). Neural Antecedents of Emotional Disorders: A Functional Magnetic Resonance Imaging Study of Subsyndromal Emotional Symptoms in Adolescent Girls. *Biological Psychiatry*, 74(4), 265-272.
- Davies, S., Bishop, D., Manstead, A. S. R., & Tantam, D. (1994). Face Perception in Children with Autism and Aspergers Syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 35(6), 1033-1057.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalogue of Selected Documents in Psychology*, 10(85).
- Davis, M. H. (1983). Measuring Individual-Differences in Empathy - Evidence for a Multidimensional Approach. *Journal of Personality and Social Psychology*, 44(1), 113-126.
- De Jong, M. C., van Engeland, H., & Kemner, C. (2008). Attentional effects of gaze shifts are influenced by emotion and spatial frequency, but not in autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, 47(4), 443-454.

- Delis, I., Berret, B., Pozzo, T., & Panzeri, S. (2013). Quantitative evaluation of muscle synergy models: a single-trial task decoding approach. *Frontiers in Computational Neuroscience*, 7.
- Delis, I., Chen, C., Jack, R. E., Garrod, O. G. B., Panzeri, S., & Schyns, P. G. (2016). Space-by-time manifold representation of dynamic facial expressions for emotion categorization. *Journal of Vision*, 16(8).
- Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences*, 1(4), 568-570.
- Diagnostic and statistical manual of mental disorders: DSM-5*. (2013). Washington, D.C: American Psychiatric Association.
- Domes, G., Kumbier, E., Heinrichs, M., & Herpertz, S. C. (2014). Oxytocin Promotes Facial Emotion Recognition and Amygdala Reactivity in Adults with Asperger Syndrome. *Neuropsychopharmacology*, 39(3), 698-706.
- 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.
- Edelman, S. (2003). Generative grammar with a human face? *Behavioral and Brain Sciences*, 26(6), 675-+.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, 6(3-4), 169-200.
- Ekman, P., & Friesen, W. V. (1976). Measuring Facial Movement. *Environmental Psychology and Nonverbal Behavior*, 1(1), 56-75.
- Ekman, P., & Friesen, W. V. (1976). Pictures of Facial Affect. San Francisco, CA:: Human Interaction Laboratory, University of California Medical Center.
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA.
- Enticott, P. G., Kennedy, H. A., Johnston, P. J., Rinehart, N. J., Tonge, B. J., Taffe, J. R., & Fitzgerald, P. B. (2014). Emotion recognition of static and dynamic faces in autism spectrum disorder. *Cogn Emot*, 28(6), 1110-1118.
- Ethier-Majcher, C., Joubert, S., & Gosselin, F. (2013). Reverse correlating trustworthy faces in young and older adults. *Front Psychol*, 4.
- 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 Res Treat*, 2014, 345878.
- Evers, K., Steyaert, J., Noens, I., & Wagemans, J. (2015). Reduced recognition of dynamic facial emotional expressions and emotion-specific response bias in children with an autism spectrum disorder. *J Autism Dev Disord*, 45(6), 1774-1784.
- Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: a systematic literature review. *Eur Child Adolesc Psychiatry*, 22(6), 329-340.
- Farroni, T., Menon, E., Rigato, S., & Johnson, M. H. (2007). The perception of facial expressions in newborns. *European Journal of Developmental Psychology*, 4(1), 2-13.
- Fein, D., Lucci, D., Braverman, M., & Waterhouse, L. (1992). Comprehension of Affect in Context in Children with Pervasive Developmental Disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 33(7), 1157-1167.
- Feldman, R. S., Mcgee, G. G., Mann, L., & Strain, P. (1993). Nonverbal Affective Decoding Ability in Children with Autism and in Typical Preschoolers. *Journal of Early Intervention*, 17(4), 341-350.
- Fitzgerald, M., & Bellgrove, M. A. (2006). The overlap between alexithymia and Asperger's syndrome. *J Autism Dev Disord*, 36(4), 573-576.
- Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J Autism Dev Disord*, 33(4), 365-382.
- Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, 65(6), 591-598.

- Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., & Wallace, M. T. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Experimental Brain Research*, 203(2), 381-389.
- 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.
- Frewen, P. A., Dozois, D. J., Neufeld, R. W., & Lanius, R. A. (2008). Meta-analysis of alexithymia in posttraumatic stress disorder. *J Trauma Stress*, 21(2), 243-246.
- Frith, C., & Frith, U. (2005). Theory of mind. *Current biology : CB*, 15(17), R644-646.
- Frith, U. (1972). Cognitive Mechanisms in Autism - Experiments with Color and Tone Sequence Production. *J Autism Child Schizophr*, 2(2), 160-173.
- Frith, U. (1989). *Autism : explaining the enigma*. Oxford, UK ; Cambridge, MA, USA: Basil Blackwell.
- Frith, U. (2003). *Autism : explaining the enigma* (2nd ed.). Malden, MA: Blackwell Pub.
- Gepner, B., Deruelle, C., & Grynfeldt, S. (2001). Motion and emotion: a novel approach to the study of face processing by young autistic children. *J Autism Dev Disord*, 31(1), 37-45.
- 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.
- Gill, D., Garrod, O. G. B., Jack, R. E., & Schyns, P. G. (2014). Facial Movements Strategically Camouflage Involuntary Social Signals of Face Morphology. *Psychological Science*, 25(5), 1079-1086.
- Gillberg, C., Cederlund, M., Lamberg, K., & Zeijlon, L. (2006). Brief report: "The autism epidemic". The registered prevalence of autism in a swedish urban area. *J Autism Dev Disord*, 36(3), 429-435.
- Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., & Baron-Cohen, S. (2010). Enhancing Emotion Recognition in Children with Autism Spectrum Conditions: An Intervention Using Animated Vehicles with Real Emotional Faces. *J Autism Dev Disord*, 40(3), 269-279.
- Golan, O., Sinai-Gavrilov, Y., & Baron-Cohen, S. (2015). The Cambridge Mindreading Face-Voice Battery for Children (CAM-C): complex emotion recognition in children with and without autism spectrum conditions. *Molecular Autism*, 6, 22.
- Golarai, G., Grill-Spector, K., & Reiss, A. L. (2006). Autism and the development of face processing. *Clin Neurosci Res*, 6(3), 145-160.
- Gosselin, F., Bacon, B. A., & Mamassian, P. (2004). Internal surface representations approximated by reverse correlation. *Vision Res*, 44(21), 2515-2520.
- Gosselin, F., & Schyns, P. G. (2001). Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Res*, 41(17), 2261-2271.
- Gould, S. J. (1994). Faces are special. *Sciences-New York*, 34(1).
- Gregory, B. L., & Plaisted-Grant, K. C. (2013). The Autism-Spectrum Quotient and Visual Search: Shallow and Deep Autistic Endophenotypes. *J Autism Dev Disord*.
- 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. *J Autism Dev Disord*, 39(9), 1278-1290.
- Gross, T. F. (2004). The perception of four basic emotions in human and nonhuman faces by children with autism and other developmental disabilities. *Journal of Abnormal Child Psychology*, 32(5), 469-480.
- Gross, T. F. (2008). Recognition of immaturity and emotional expressions in blended faces by children with autism and other developmental disabilities. *J Autism Dev Disord*, 38(2), 297-311.

- 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.
- Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restor Neurol Neurosci*, 28(2), 219-236.
- Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate Data Analysis* (3rd ed.). New York: Macmillan. .
- Hamming, R. W. (1950). Error Detecting and Error Correcting Codes. *Bell System Technical Journal*, 29(2), 147-160.
- Happe, F., & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J Autism Dev Disord*, 36(1), 5-25.
- 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.
- Harris, J. P., Atkinson, E. A., Lee, A. C., Nithi, K., & Fowler, M. S. (2003). Hemisphere differences in the visual perception of size in left hemiParkinson's disease. *Neuropsychologia*, 41(7), 795-807.
- Head, A. M., McGillivray, J. A., & Stokes, M. A. (2014). Gender differences in emotionality and sociability in children with autism spectrum disorders. *Molecular Autism*, 5.
- Hill, E., Berthoz, S., & Frith, U. (2004). Brief report: cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. *J Autism Dev Disord*, 34(2), 229-235.
- Hill, E. L., & Berthoz, S. (2006). Response to "Letter to the editor: The overlap between Alexithymia and Asperger's syndrome", Fitzgerald and Bellgrove, *Journal of Autism and Developmental Disorders*, 36(4). *J Autism Dev Disord*, 36(8), 1143-1145.
- Hobson, R. P. (1986a). The Autistic Childs Appraisal of Expressions of Emotion. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 27(3), 321-342.
- Hobson, R. P. (1986b). The Autistic Childs Appraisal of Expressions of Emotion - a Further Study. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 27(5), 671-680.
- Hobson, R. P. (1991). Methodological Issues for Experiments on Autistic Individuals Perception and Understanding of Emotion. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 32(7), 1135-1158.
- Hobson, R. P., Ouston, J., & Lee, A. (1988). Whats in a Face - the Case of Autism. *British Journal of Psychology*, 79, 441-453.
- Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, 6(2), 65-70.
- 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.
- International Classification of Diseases 11th Revision. (2016). Retrieved 1st December 2016, from <http://www.who.int/classifications/icd/revision/en/>
- Jack, R. E., Caldara, R., & Schyns, P. G. (2012a). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *J Exp Psychol Gen*, 141(1), 19-25.
- Jack, R. E., Caldara, R., & Schyns, P. G. (2012b). Internal Representations Reveal Cultural Diversity in Expectations of Facial Expressions of Emotion. *Journal of Experimental Psychology-General*, 141(1), 19-25.
- Jack, R. E., Garrod, O. G. B., & Schyns, P. G. (2014). Dynamic Facial Expressions of Emotion Transmit an Evolving Hierarchy of Signals over Time. *Current Biology*, 24(2), 187-192.

- Jack, R. E., Garrod, O. G. B., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proc Natl Acad Sci U S A*, *109*(19), 7241-7244.
- 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.
- Jemel, B., Mottron, L., & Dawson, M. (2006). Impaired face processing in autism: fact or artifact? *J Autism Dev Disord*, *36*(1), 91-106.
- Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and violent behavior*, *9*(5), 441-476.
- Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *J Child Psychol Psychiatry*, *38*(5), 527-534.
- Jones, C. R., Pickles, A., Falcaro, M., Marsden, A. J., Happe, F., Scott, S. K., . . . Charman, T. (2011). A multimodal approach to emotion recognition ability in autism spectrum disorders. *J Child Psychol Psychiatry*, *52*(3), 275-285.
- Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *J Child Psychol Psychiatry*, *44*(4), 529-542.
- Katsyri, J., Forger, K., Makarainen, M., & Takala, T. (2015). A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Front Psychol*, *6*.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test*: John Wiley & Sons, Inc.
- Kennedy, D. P., & Adolphs, R. (2012). Perception of emotions from facial expressions in high-functioning adults with autism. *Neuropsychologia*, *50*(14), 3313-3319.
- 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.
- 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. *Res Autism Spectr Disord*, *21*, 51-60.
- Kim, Y. S., Leventhal, B. L., Koh, Y. J., Fombonne, E., Laska, E., Lim, E. C., . . . Grinker, R. R. (2011). Prevalence of autism spectrum disorders in a total population sample. *Am J Psychiatry*, *168*(9), 904-912.
- Kirkovski, M., Enticott, P. G., & Fitzgerald, P. B. (2013). A review of the role of female gender in autism spectrum disorders. *J Autism Dev Disord*, *43*(11), 2584-2603.
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch Gen Psychiatry*, *59*(9), 809-816.
- Koch, A. S., Kleiman, A., Wegener, I., Zur, B., Imbierowicz, K., Geiser, F., & Conrad, R. (2015). Factorial structure of the 20-item Toronto Alexithymia Scale in a large sample of somatoform patients. *Psychiatry Res*, *225*(3), 355-363.
- Kontsevich, L. L., & Tyler, C. W. (2004). What makes Mona Lisa smile? *Vision Res*, *44*(13), 1493-1498.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The Phq - 9. *Journal of general internal medicine*, *16*(9), 606-613.
- Lacroix, A., Guidetti, M., Roe, B., & Reilly, J. (2014). Facial emotion recognition in 4-to 8-year-olds with autism spectrum disorder: A developmental trajectory approach. *Res Autism Spectr Disord*, *8*(9), 1146-1154.
- Lacroix, A., Guidetti, M., Roge, B., & Reilly, J. (2009). Recognition of emotional and nonemotional facial expressions: A comparison between Williams syndrome and autism. *Res Dev Disabil*, *30*(5), 976-985.

- Larsen, R. J., & Ketelaar, T. (1991). Personality and Susceptibility to Positive and Negative Emotional States. *Journal of Personality and Social Psychology*, 61(1), 132-140.
- 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.
- Lawson, J., Baron-Cohen, S., & Wheelwright, S. (2004). Empathising and systemising in adults with and without Asperger Syndrome. *J Autism Dev Disord*, 34(3), 301-310.
- Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755), 788-791.
- Leonard, H. C., Annaz, D., Karmiloff-Smith, A., & Johnson, M. H. (2011). Brief Report: Developing Spatial Frequency Biases for Face Recognition in Autism and Williams Syndrome. *J Autism Dev Disord*, 41(7), 968-973.
- Lerner, M. D., McPartland, J. C., & Morris, J. P. (2013). Multimodal emotion processing in autism spectrum disorders: an event-related potential study. *Dev Cogn Neurosci*, 3, 11-21.
- Leslie, A. M. (1987). Pretense and Representation - the Origins of Theory of Mind. *Psychological Review*, 94(4), 412-426.
- Leslie, A. M., & Frith, U. (1988). Autistic Childrens Understanding of Seeing, Knowing and Believing. *British Journal of Developmental Psychology*, 6, 315-324.
- Leung, D., Ordqvist, A., Falkmer, T., Parsons, R., & Falkmer, M. (2013). Facial emotion recognition and visual search strategies of children with high functioning autism and Asperger syndrome. *Res Autism Spectr Disord*, 7(7), 833-844.
- Li, S., B., Z., Guo, Y., & Zhang, J. (2015). The association between alexithymia as assessed by the 20-item Toronto Alexithymia Scale and depression: A meta-analysis. *Psychiatry Res*, 227(1), 1-9.
- Lindner, J. L., & Rosen, L. A. (2006). Decoding of emotion through facial expression, prosody and verbal content in children and adolescents with Asperger's syndrome. *J Autism Dev Disord*, 36(6), 769-777.
- Liu, K. Y., King, M., & Bearman, P. S. (2010). Social Influence and the Autism Epidemic. *American Journal of Sociology*, 115(5), 1387-1434.
- Livingston, L. A., & Livingston, L. M. (2016). Commentary: Alexithymia, not autism, is associated with impaired interoception. *Front Psychol*, 7.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., . . . Rutter, M. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J Autism Dev Disord*, 30(3), 205-223.
- Loukusa, S., Makinen, L., Kuusikko-Gauffin, S., Ebeling, H., & Moilanen, I. (2014). Theory of mind and emotion recognition skills in children with specific language impairment, autism spectrum disorder and typical development: group differences and connection to knowledge of grammatical morphology, word-finding abilities and verbal working memory. *Int J Lang Commun Disord*, 49(4), 498-507.
- Loveland, K. A., TunaliKotoski, B., Chen, Y. R., Ortegon, J., Pearson, D. A., Brelsford, K. A., & Gibes, M. C. (1997). Emotion recognition in autism: Verbal and nonverbal information. *Development and Psychopathology*, 9(3), 579-593.
- Luce, P. A. (1986). A Computational Analysis of Uniqueness Points in Auditory Word Recognition. *Perception & Psychophysics*, 39(3), 155-158.
- Lundqvist, D., Flykt, A., & Öhman, A. (Producer). (1998). The Karolinska Directed Emotional Faces.
- Macdonald, H., Rutter, M., Howlin, P., Rios, P., Leconteur, A., Evered, C., & Folstein, S. (1989). Recognition and Expression of Emotional Cues by Autistic and Normal Adults. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 30(6), 865-877.

- Mandal, M. K., Pandey, R., & Prasad, A. B. (1998). Facial expressions of emotions and schizophrenia: A review. *Schizophr Bull*, 24(3), 399-412.
- Marquard, D. W. (1970). Generalized Inverses, Ridge Regression, Biased Linear Estimation, and Nonlinear Estimation. *Technometrics*, 12(3), 591-&.
- Mateos, J. L. A. (1993). Alexythymia - Concept and Measurement. *Archivos De Neurobiologia*, 56(2), 106-115.
- Mattila, M. L., Kielinen, M., Linna, S. L., Jussila, K., Ebeling, H., Bloigu, R., . . . Moilanen, I. (2011). Autism spectrum disorders according to DSM-IV-TR and comparison with DSM-5 draft criteria: an epidemiological study. *J Am Acad Child Adolesc Psychiatry*, 50(6), 583-592 e511.
- 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. *J Autism Dev Disord*, 37(6), 1086-1095.
- Mazza, M., Pino, M. C., Mariano, M., Tempesta, D., Ferrara, M., De Berardis, D., . . . Valenti, M. (2014). Affective and cognitive empathy in adolescents with autism spectrum disorder. *Frontiers in Human Neuroscience*, 8.
- McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, 126(3), 424-453.
- Meese, T. S. (1995). Using the Standard Staircase to Measure the Point of Subjective Equality - a Guide Based on Computer-Simulations. *Perception & Psychophysics*, 57(3), 267-281.
- Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., . . . Komaki, G. (2006). Impaired self-awareness and theory of mind: an fMRI study of mentalizing in alexithymia. *Neuroimage*, 32(3), 1472-1482.
- Mottron, L., Burack, J. A., Iarocci, G., Belleville, S., & Enns, J. T. (2003). Locally oriented perception with intact global processing among adolescents with high-functioning autism: evidence from multiple paradigms. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44(6), 904-913.
- Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *J Autism Dev Disord*, 36(1), 27-43.
- Navon, D. (1977). Forest before Trees - Precedence of Global Features in Visual-Perception. *Cognitive Psychology*, 9(3), 353-383.
- Nelson, C. A. (1987). The Recognition of Facial Expressions in the 1st 2 Years of Life - Mechanisms of Development. *Child Development*, 58(4), 889-909.
- Nemiah, J. C. (1977). Alexithymia. Theoretical considerations. *Psychotherapy and Psychosomatics*, 28(1-4), 199-206.
- Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, 401(6754), 695-698.
- Neter, J., Wasserman, W., & Kutner, M. H. (1989). *Applied Linear Regression Models*. Homewood, IL: Irwin.
- Nishimura, H., Komaki, G., Igarashi, T., Moriguchi, Y., Kajiwara, S., & Akasaka, T. (2009). Validity issues in the assessment of alexithymia related to the developmental stages of emotional cognition and language. *Biopsychosoc Med*, 3, 12.
- Nuske, H. J., Vivanti, G., & Dissanayake, C. (2013). Are emotion impairments unique to, universal, or specific in autism spectrum disorder? A comprehensive review. *Cognition & Emotion*, 27(6), 1042-1061.
- Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1990). Are There Emotion Perception Deficits in Young Autistic-Children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 31(3), 343-361.

- Pallett, P. M., Cohen, S. J., & Dobkins, K. R. (2014). Face and Object Discrimination in Autism, and Relationship to IQ and Age. *J Autism Dev Disord*, *44*(5), 1039-1054.
- Pan, Y., & Jackson, R. T. (2008). Ethnic difference in the relationship between acute inflammation and serum ferritin in US adult males. *Epidemiology and Infection*, *136*(3), 421-431.
- Pellicano, E. (2011). Psychological models of autism: an overview. In R. Iona & P. Rezaie (Eds.), *Researching the Autism Spectrum: Contemporary Perspectives* (pp. 219-265). UK: Cambridge University Press.
- Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends Cogn Sci*, *16*(10), 504-510.
- Pellicano, E., Gibson, L., Maybery, M., & Durkin, K. (2003). Elevated motion coherence thresholds in autism: Further evidence of 'weak' central coherence. *Australian Journal of Psychology*, *55*, 202-202.
- Pellicano, E., Gibson, L., Maybery, M., Durkin, K., & Badcock, D. R. (2005). Abnormal global processing along the dorsal visual pathway in autism: a possible mechanism for weak visuospatial coherence? *Neuropsychologia*, *43*(7), 1044-1053.
- Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *J Autism Dev Disord*, *32*(4), 249-261.
- Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the Autistic Child's Theory of Mind - Knowledge, Belief, and Communication. *Child Development*, *60*(3), 689-700.
- Peterson, C. C., Slaughter, V., & Brownell, C. (2015). Children with autism spectrum disorder are skilled at reading emotion body language. *J Exp Child Psychol*, *139*, 35-50.
- 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.
- Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. *British Journal of Developmental Psychology*, *16*, 337-348.
- Pierce, N. P., O'Reilly, M. F., Sorrells, A. M., Fragale, C. L., White, P. J., Aguilar, J. M., & Cole, H. A. (2014). Ethnicity reporting practices for empirical research in three autism-related journals. *J Autism Dev Disord*, *44*(7), 1507-1519.
- Piggot, J., Kwon, H., Mobbs, D., Blasey, C., Lotspeich, L., Menon, V., . . . Reiss, A. L. (2004). Emotional attribution in high-functioning individuals with autistic spectrum disorder: A functional imaging study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *43*(4), 473-480.
- Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *39*(5), 777-783.
- Pokorny, R. A. (1998). Using maximum-likelihood adaptive methods to estimate difference thresholds and points of subjective equality. *Behavior Research Methods Instruments & Computers*, *30*(4), 624-636.
- 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.
- Prior, M., Dahlstrom, B., & Squires, T. L. (1990). Autistic Children's Knowledge of Thinking and Feeling States in Other People. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *31*(4), 587-601.
- Rajendran, G., & Mitchell, P. (2007). Cognitive theories of autism. *Developmental Review*, *27*(2), 224-260.

- Richoz, A. R., Jack, R. E., Garrod, O. G. B., Schyns, P. G., & Caldara, R. (2015). Reconstructing dynamic mental models of facial expressions in prosopagnosia reveals distinct representations for identity and expression. *Cortex*, *65*, 50-64.
- Robel, L., Ennouri, K., Piana, H. N., Vaivre-Douret, L., Perier, A., Flament, M. F., & Mouren-Simoni, M. C. (2004). Discrimination of face identities and expressions in children with autism: Same or different? *Eur Child Adolesc Psychiatry*, *13*(4), 227-233.
- Robertson, A. E., & Simmons, D. R. (2013). The relationship between sensory sensitivity and autistic traits in the general population. *J Autism Dev Disord*, *43*(4), 775-784.
- Rose, D. (1988). Estimation of the Slope of the Psychometric Function. *Perception*, *17*(3), 396-396.
- 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.
- Rutherford, M. D., & McIntosh, D. N. (2007). Rules versus prototype matching: strategies of perception of emotional facial expressions in the autism spectrum. *J Autism Dev Disord*, *37*(2), 187-196.
- Rutherford, M. D., & Towns, A. M. (2008). Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *J Autism Dev Disord*, *38*(7), 1371-1381.
- Rutherford, M. D., Walsh, J., & Creighton, S. (2015). Emotion perception or social cognitive complexity: What drives face processing deficits in autism spectrum disorder? *J Vis*, *15*(12), 1223.
- Rutter, M. (1978). Diagnosis and Definition of Childhood Autism. *J Autism Child Schizophr*, *8*(2), 139-161.
- Ruzich, E., Allison, C., Chakrabarti, B., Smith, P., Musto, H., Ring, H., & Baron-Cohen, S. (2015). Sex and STEM Occupation Predict Autism-Spectrum Quotient (AQ) Scores in Half a Million People. *Plos One*, *10*(10), e0141229.
- 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*, 2.
- Sachse, M., Schlitt, S., Hainz, D., Ciaramidaro, A., Walter, H., Poustka, F., . . . Freitag, C. M. (2014). Facial emotion recognition in paranoid schizophrenia and autism spectrum disorder. *Schizophr Res*, *159*(2-3), 509-514.
- Salminen, J. K., Saarijarvi, S., Aarela, 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.
- Sato, W., Uono, S., & Toichi, M. (2013). Atypical recognition of dynamic changes in facial expressions in autism spectrum disorders. *Res Autism Spectr Disord*, *7*(7), 906-912.
- Schulte-Ruther, M., Greimel, E., Piefke, M., Kamp-Becker, I., Remschmidt, H., Fink, G. R., . . . Konrad, K. (2014). Age-dependent changes in the neural substrates of empathy in autism spectrum disorder. *Soc Cogn Affect Neurosci*, *9*(8), 1118-1126.
- Schultz, R. T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience*, *23*(2-3), 125-141.
- Shah, A., & Frith, U. (1983). An islet of ability in autistic children: a research note. *J Child Psychol Psychiatry*, *24*(4), 613-620.
- Shah, P., Hall, R., Catmur, C., & Bird, G. (2016). Alexithymia, not autism, is associated with impaired interoception. *Cortex*, *81*, 215-220.
- Sifneos, P. E. (1973). Prevalence of Alexithymic Characteristics in Psychosomatic Patients. *Psychotherapy and Psychosomatics*, *22*(2-6), 255-262.

- Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Res*, *49*(22), 2705-2739.
- Snyman, J. A. (2005). *Practical mathematical optimization : an introduction to basic optimization theory and classical and new gradient-based algorithms*. New York: Springer.
- Song, Y. N., Kawabe, T., Hakoda, Y., & Du, X. X. (2012). Do the eyes have it? Extraction of identity and positive expression from another's eyes in autism, probed using "Bubbles". *Brain & Development*, *34*(7), 584-590.
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007a). Abnormal use of facial information in high-functioning autism. *J Autism Dev Disord*, *37*(5), 929-939.
- Spezio, M. L., Adolphs, R., Hurley, R. S. E., & Piven, J. (2007b). Analysis of face gaze in autism using "Bubbles". *Neuropsychologia*, *45*(1), 144-151.
- Sucksmith, E., Allison, C., Baron-Cohen, S., Chakrabarti, B., & Hoekstra, R. A. (2013). Empathy and emotion recognition in people with autism, first-degree relatives, and controls. *Neuropsychologia*, *51*(1), 98-105.
- Tantam, D., Monaghan, L., Nicholson, H., & Stirling, J. (1989). Autistic Childrens Ability to Interpret Faces - a Research Note. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *30*(4), 623-630.
- Tardif, C., Laine, F., Rodriguez, M., & Gepner, B. (2007). Slowing down presentation of facial movements and vocal sounds enhances facial expression recognition and induces facial-vocal imitation in children with autism. *J Autism Dev Disord*, *37*(8), 1469-1484.
- Taylor, G. J. (1984). Alexithymia: concept, measurement, and implications for treatment. *Am J Psychiatry*, *141*(6), 725-732.
- Tell, D., & Davidson, D. (2015). Emotion recognition from congruent and incongruent emotional expressions and situational cues in children with autism spectrum disorder. *Autism*, *19*(3), 375-379.
- 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 Res Treat*, *2014*, 816137.
- Tokita, M., & Ishiguchi, A. (2010). Effects of element features on discrimination of relative numerosity: comparison of search symmetry and asymmetry pairs. *Psychological Research-Psychologische Forschung*, *74*(1), 99-109.
- Uddenberg, S., & Shim, W. M. (2015). Seeing the World Through Target-Tinted Glasses: Positive Mood Broadens Perceptual Tuning. *Emotion*, *15*(3), 319-328.
- Uljarevic, M., & Hamilton, A. (2013). Recognition of Emotions in Autism: A Formal Meta-Analysis. *J Autism Dev Disord*, *43*(7), 1517-1526.
- 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. *Res Autism Spectr Disord*, *7*(2), 361-368.
- Volkmar, F. R., & Reichow, B. (2013). Autism in DSM-5: progress and challenges. *Molecular Autism*, *4*(1), 13.
- WalkerAndrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, *121*(3), 437-456.
- 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. *J Autism Dev Disord*, *41*(11), 1475-1486.
- Wallace, S., Coleman, M., & Bailey, A. (2008). An investigation of basic facial expression recognition in autism spectrum disorders. *Cognition & Emotion*, *22*(7), 1353-1380.

- Walsh, J. A., Vida, M. D., & Rutherford, M. D. (2014). Strategies for perceiving facial expressions in adults with autism spectrum disorder. *J Autism Dev Disord*, *44*(5), 1018-1026.
- 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.
- Weigelt, S., Koldewyn, K., & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: A review of behavioral studies. *Neurosci Biobehav Rev*, *36*(3), 1060-1084.
- Weigelt, S., Koldewyn, K., & Kanwisher, N. (2013). Face recognition deficits in autism spectrum disorders are both domain specific and process specific. *Plos One*, *8*(9).
- Whelan, R. (2008). Effective analysis of reaction time data. *Psychological Record*, *58*(3), 475-482.
- Williams, B. T., & Gray, K. M. (2013). Are emotion recognition skills related to autism symptom severity in children with autism? *International Journal of Developmental Disabilities*, *59*(2), 118-133.
- Williams, D. (2010). Theory of own mind in autism: Evidence of a specific deficit in self-awareness? *Autism*, *14*(5), 474-494.
- Wing, L. (1988). The continuum of autistic characteristics *Diagnosis and assessment in autism* (pp. 91-110). US.: Springer.
- Wing, L. (1993). The definition and prevalence of autism: A review. *Eur Child Adolesc Psychiatry*, *2*(1), 61-74.
- Wingate, M., Kirby, R. S., Pettygrove, S., Cunniff, C., Schulz, E., Ghosh, T., . . . Monitoring, A. D. D. (2014). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010. *Mmwr Surveillance Summaries*, *63*(2).
- Yu, H., Garrod, O. G. B., & Schyns, P. G. (2012). Perception-driven facial expression synthesis. *Computers & Graphics-Uk*, *36*(3), 152-162.
- Zackheim, L. (2007). Alexithymia: The expanding realm of research. *Journal of Psychosomatic Research*, *63*(4), 345-347.