AUTISTIC TRAITS AND EVERYDAY SOCIAL BEHAVIOUR

A thesis submitted for the degree of Doctor of Philosophy

by

Leila Jameel

Department of Experimental Psychology

University College London

March 2016

I, Leila Jameel, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature… Date…

Leila Jameel

ABSTRACT

Autism Spectrum Disorder (ASD) is characterised by difficulties with social communication and interaction. A great deal of experimental work has examined the performance of people with ASD on social cognition tasks in laboratory settings, and a number of cognitive models have been postulated to account for observed differences in those with ASD. Meanwhile, clinical reports indicate that people with ASD face a range of difficulties in everyday functioning. However, very little experimental work has tried to elucidate how the postulated cognitive deficits in ASD might translate into difficulties in real-life-type settings, especially in adult populations.

A range of novel scenario-based tasks were developed for the present thesis which aimed to provide more sensitive tools than traditional social cognition tasks for identifying the nature and severity of impairments in everyday social functioning. These systematically examined different aspects of social performance, in particular pro-social behaviour, moral judgment and reasoning. The present thesis adopted a trait-based approach to investigate how high versus low levels of autistic traits influenced everyday social functioning. This is in line with the continuum conceptualisation of an autistic spectrum, whereby those with clinical levels of impairment (i.e. diagnosed with ASD) are thought to lie at the extreme end of a normal distribution of autistic traits.

Overall, two key findings emerged; firstly, people with high levels of autistic traits tended to be less behaviourally and emotionally responsive to others' needs. Secondly, people with high levels of autistic traits displayed relatively intact awareness of social and moral norms that underpin everyday situations, but their understanding of these appeared to be more limited. These findings are consistent with the conceptualisation of a continuum of trait severity, whereby those with high levels of autistic traits showed similar difficulties to those seen in people with ASD, although perhaps to a lesser extent. The body of work presented in this thesis has potential clinical implications for the assessment and management of adults with ASD.

ACKNOWLEDGEMENTS

I would like to thank my primary PhD supervisor, Professor Shelley Channon, for her invaluable guidance and support. Shelley has certainly gone 'above and beyond' in her supervision, devoting a huge amount of effort and time to both my research and my professional development. Shelley has helped me to engage my critical faculties, and has encouraged me to have an extremely rich and well-rounded academic training. She has also tried in vein to combat my grammatical afflictions! I am extremely fortunate to have Shelley as a mentor, and I look forward to a long and fruitful collaboration!

I would also like to thank my secondary PhD supervisor, Dr Sarah Crawford for her support. Her thoughtful feedback and clinical insights were extremely helpful.

A very special thanks to my fellow lab members, Giulia Bellesi and Kari Vyas, who I have really enjoyed working with. Our close collaboration has certainly enhanced my experience and improved the quality of my work. Thank you to Shelley for fostering a real 'team effort' mentality.

Thank you to the several thousand participants who took part in the screening and experimental phases of the studies included in this thesis. It was certainly an ambitious undertaking and I am grateful to all of them for their efforts and enthusiasm.

A big thank you also to Tom Hardwicke, a very special colleague and friend. Kari, Tom and I co-founded CRITUCL, which has been a tremendous success, and I have relished every discussion and debate we have had. I look forward to many, many more. I would also like to thank all my other UCL peers especially regular CRITUCL attendees, for their humour, passion and insights that have hugely enriched my experience.

Finally, I would like to thank all of my friends and family for their support and enthusiasm along the way. I have bored them to death talking, and often whinging about my PhD, and over the last few months when the going got tough, this has been a lifeline. In particular, thank you to my mum and sister, the best girls I could ever ask for. Andrew, thank you for putting up with me throughout this rollercoaster experience!

TABLE OF CONTENTS

Abstract 3
Acknowledgments 4
Table of contents 5
List of tables 13
List of figures 13

PART 1: LITERATURE REVIEW 15
CHAPTERS 1-4

Introduction to literature review 16

Chapter 1: Clinical Features of Autism Spectrum Disorder (ASD) and the autism continuum 17
1.1 Diagnostic criteria of ASD 18
1.2 Prevalence and prognosis of ASD 18
1.3 Measuring autistic traits in the general population 20
1.4 Chapter summary 21

Chapter 2: Aetiological models of ASD 22
2.1 Biological models of ASD 22
2.1.1 The genetics of ASD 22
2.1.2 The neurobiology of ASD 23
2.1.3 Prenatal, perinatal and neonatal risk factors in ASD 25
2.1.4 The role of oxytocin in ASD 26
2.2 Psychological models of ASD: Non-social deficits 28
2.2.1 Weak Central Coherence 28
2.2.2 Executive dysfunction 29
2.3 Psychological models of ASD: Social deficits 30
2.3.1 Impaired cognitive empathy 31
2.3.2 Preserved emotional empathy 32
2.3.3 Compensatory social knowledge 34
2.3.4 The 'Extreme Male Brain' hypothesis 35
2.4 A single psychological model of ASD? 37
2.5 Chapter summary 37

Chapter 3: Social skill training and pro-social functioning in ASD 39
3.1 Understanding everyday social functioning in ASD 39
3.2 Pro-social behaviour 40
3.3 Pro-social functioning in ASD 42
3.4 Brief review of social skill training in ASD 44
3.5 Supporting adults with ASD 45
3.6 Chapter summary 46
3.7 Introduction to Experiments A: Chapters 5-7 48

Chapter 4: Moral judgment and reasoning in ASD 49
4.1 The role of cognitive and emotional processes in morality 49
4.2 Exploring different aspects of morality in ASD 51
4.2.1 Normative versus descriptive morality 51

4.2.2 Black and white judgments of 'right and wrong' 52
4.2.3 Judgments of causality and intentionality 55
4.2.4 Judgments of deservingness 58
4.2.5 Coutnerfactual thinking 58
4.2.6 Moral dilemmas 60
4.3 Chapter summary 62
4.4 Introduction to Experiments B: Chapters 8-11 63

PART 2: EXPERIMENTS A 64
CHAPTERS 5-7

A note on experimental methods 65
Screening Phase 65
Autism-Spectrum Quotient 65
Selection of high and low AQ participants 65
Experimental Phase 66
Participants and procedure 66
Construction and nature of scenario-based tasks 66
Experimental pairings 67
Ethics 67

A note on statistical analysis 67
Assumptions of parametric tests 67
Significance levels 67

Chapter 5: Self- versus other- satisfaction for pro-social behaviour 68
5.1 Introduction 68
5.1.1 Hypotheses
5.1.1.1 Pro-social behaviour: generation versus selection 69
5.1.1.1 Self- versus other-satisfaction ratings 69
5.2 Methods 69
5.2.1 Screening phase 69
5.2.1.1 Screening participants and procedure 70
5.2.2 Experimental phase 70
5.2.2.1 Design 70
5.2.2.2 Experimental participants and procedure 70
5.2.3 The 'Above and Beyond' task 70
5.2.3.1 Scoring 71
5.2.3.1.1 Generation of pro-social responses 71
5.2.3.1.2 Selection of pro-social actions 72
5.2.3.1.3 Self- versus other-satisfaction ratings 72
5.3 Results 72
5.3.1 Data analysis 72
5.3.2 The 'Above and Beyond' task 73
5.3.2.1 Generation of pro-social responses 73
5.3.2.2 Selection of pro-social actions 73
5.3.2.3 Self- versus other-satisfaction ratings 73
5.4. Discussion 75
5.4.1 Summary of findings 75
5.4.2 Non-social accounts of the findings 76
5.4.3 Impaired cognitive empathy? 77

5.4.4 Impaired emotional empathy? 78
5.4.5 The role of social knowledge 79
5.4.6 Conclusion 80

Chapter 6: Examining how the clarity of the social expectation to help others influences pro-social behaviour 81
6.1 Introduction 81
6.1.1 Hypotheses 83
6.1.1.1 Pro-social behaviour and sympathy ratings 83
6.1.1.2 Understanding of social expectations: rationales versus ratings 83
6.2. Methods 83
6.2.1 Screening phase 83
6.2.1.1 Screening participants and procedure 83
6.2.2 Experimental phase 83
6.2.2.1 Design 83
6.2.2.2 Experimental participants and procedure 84
6.2.3 The 'Social Expectations' task 84
6.2.3.1 Scoring 85
6.2.3.1.1 Likelihood of helping ratings 85
6.2.3.1.2 Sympathy for character ratings 86
6.2.3.1.3 Strength of character expectation ratings 86
6.2.3.1.4 Verbal rationales: understanding of social expectations 87
6.3 Results 89
6.3.1 Data analysis 89
6.3.2 The 'Social Expectations' task 89
6.3.2.1 Likelihood of helping ratings 89
6.3.2.2 Sympathy for character ratings 89
6.3.2.3 Strength of character expectation ratings 89
6.3.3.4 Verbal rationales: understanding of social expectations 90
6.4 Discussion 92
6.4.1 Summary of findings 92
6.4.2 Can non-social and social models of ASD account for the findings? 93
6.4.3 The role of social knowledge 94
6.4.4 Conclusion 95

Chapter 7: Examining characters' justifications for requests of help on pro-social behaviour 97
7.1 Introduction 97
7.1.1 Hypotheses 98
7.1.1.1 The 'Social Favours' task 98
7.1.1.1.1 Pro-social behaviour 98
7.1.1.1.2 Acceptability of request ratings 98
7.1.1.1.3 Personal sacrifice ratings 98
7.1.1.2 Additional Measures 99
7.1.1.2.1 Executive functions 99
7.1.1.2.2 'Interpersonal Reactivity Index' (MIT) 99
7.1.1.2.3 'Mentalistic Interpretation Test' (IRI) 99
7.2. Methods 100
7.2.1 Screening phase 100
7.2.1.1 Screening participants and procedure 100
7.2.2 Experimental phase 100

7.2.2.1 Design 100
7.2.2.2 Experimental participants and procedure 100
7.2.3 The 'Social Favours' task 101
7.2.3.1 Scoring 102
7.2.3.1.1 Likelihood of helping ratings 102
7.2.3.1.2 Acceptability of request ratings 102
7.2.3.1.3 Personal sacrifice ratings 103
7.2.4 Additional measures 103
7.2.4.1 Executive functions 103
7.2.4.2 'IRI' 103
7.2.4.3 'MIT' 104
7.3 Results 107
7.3.1 Data analysis 107
7.3.2 The 'Social Favours' task 107
7.3.2.1 Likelihood of helping ratings 107
7.3.2.2 Acceptability of request ratings 107
7.3.2.3 Personal sacrifice ratings 108
7.3.3 Additional measures 109
7.3.3.1 Executive functions 109
7.3.3.2 'IRI' 109
7.3.3.3 'MIT' 110
7.3.4. Correlations between performance on the 'Social Favours' task and additional measures 111
7.4 Discussion 112
7.4.1 Summary of findings 112
7.4.2 Executive functioning in ASD 113
7.4.3 Cognitive empathy in ASD 114
7.4.4 Emotional empathy in ASD 116
7.4.5 Contributions of executive dysfunction to performance on the 'Social Favours' task 117
7.4.6 Contributions of empathy to performance on the 'Social Favours' task 118
7.4.7 Sensitivity of the 'Social Favours' task 119
7.4.8 Is the understanding of 'deservingness' intact in ASD? 122
7.4.9 Conclusion 123

PART 3: EXPERIMENTS B 124
CHAPTERS 8-11

Chapter 8: Judgments of deservingness for positive and negative outcomes 125
8.1 Introduction 125
8.1.1 Hypotheses 126
8.1.1.1 Character deservingness 126
8.1.1.2 Participant satisfaction 127
8.2 Methods 127
8.2.1 Screening phase 127
8.2.1.1 Screening participants and procedure 127
8.2.2 Experimental phase 127
8.2.2.1 Design 127
8.2.2.2 Experimental participants and procedure 127

8.2.3 The 'Deservingness' tasks 128
8.2.3.1 Development of the 'Deservingness' tasks 128
8.2.3.2 'Deservingness: Positive Outcomes' 129
8.2.3.3 'Deservingness: Negative Outcomes' 130
8.2.3.4 Administration of both tasks 131
8.2.3.5 Scoring of 'Deservingness: Positive Outcomes' 131
8.2.3.5.1 Character deservingness ratings 131
8.2.3.5.2 Participant satisfaction ratings 132
8.2.3.6 'Scoring of 'Deservingness: Negative Outcomes' 132

8.3 Results 132
8.3.1 Data analysis 132
8.3.2 'Deservingness: Positive Outcomes' 132
8.3.2.1 Character deservingness ratings 132
8.3.2.2 Participant satisfaction ratings 133
8.3.3 'Deservingness: Negative Outcomes' 133
8.3.3.1 Character deservingness ratings 133
8.3.3.2 Participant satisfaction ratings 133
8.4 Discussion 135
8.4.1 Summary of findings 135
8.4.2 The effect of congruence on judgments of deservingness 135
8.4.3 Why did the 'Deservingness' tasks fail to differentiate the high versus low AQ groups? 137
8.4.3.1 Characteristics of the high and low AQ participants 137
8.4.3.2 Sensitivity of the 'Deservingness' tasks 138
8.4.4 A preserved understanding of what is 'right and wrong' in ASD? 139
8.4.5 Conclusion 140

Chapter 9: Counterfactual thinking for self-versus other- and the experience of regret and blame 141
9.1 Introduction 141
9.1.1 Hypotheses 144
9.1.1.1 Counterfactual thinking fluency 144
9.1.1.2 Ratings of regret 144
9.1.13 Ratings of blame 145
9.2 Methods 145
9.2.1 Screening phase 145
9.2.1.1 Screening participants and procedure 145
9.2.2 Experimental phase 145
9.2.2.1 Design 145
9.2.2.2 Experimental participants and procedure 145
9.2.3 The 'Reflective Counterfactual Thinking' (RCFT) task 146
9.2.3.1 Scoring 150
9.2.3.1.1 Counterfactual thinking fluency 150
9.2.3.1.2 Ratings of regret 150
9.2.3.1.3 Ratings of blame 151
9.3 Results 151
9.3.1 Data analysis 151
9.3.2 The 'RCFT' task 151
9.3.2.1 Counterfactual thinking fluency 151
9.3.2.2 Ratings of regret 151

9.3.2.3 Ratings of blame 152
9.4 Discussion 154
9.4.1 Summary of findings 155
9.4.2 The role of congruence and agent identity 157
9.4.3 Counterfactual thinking, regret, blame and autistic traits 157
9.4.3.1 The role of executive functioning and imagination 157
9.4.3.2 The role of cognitive empathy 159
9.4.3.3 The role of emotional empathy 161
9.4.3.4 The role of social knowledge 162
9.4.4 The sensitivity of the 'RCFT' task 162
9.4.5 Conclusion 163

Chapter 10: Counterfactual judgments of others' mistakes and the experience of regret, blame and guilt 165
10.1 Introduction 165
10.1.1 Hypotheses 166
10.1.1.1 'Counterfactual Judgments' task 166
10.1.1.1.1 Ratings of regret and blame 166
10.1.1.1.2 Preferences and ratings of guilt for counterfactual alternatives 166
10.1.1.2 Additional measures 167
10.1.1.2.1 Executive functioning 167
10.1.1.2.2 'IRI' 167
10.1.1.2.3 'MIT' 167
10.2 Methods 167
10.2.1 Screening phase 167
10.2.1.1 Screening participants and procedure 167
10.2.2 Experimental phase 168
10.2.2.1 Design 168
10.2.2.2 Experimental participants and procedure 168
10.2.3 'Counterfactual Judgments' task 168
10.2.3.1 Administration 169
10.2.3.2 Scoring 169
10.2.3.2.1 Ratings of regret and blame 169
10.2.3.2.2 Preferences for counterfactual alternatives 170
10.2.3.2.3 Ratings of guilt for counterfactual alternatives 170
10.2.4 Additional measures 172
10.2.4.1 Executive functions 172
10.2.4.2 'IRI' 172
10.2.4.3 'MIT' 172
10.3 Results 173
10.3.1 Data analysis 173
10.3.2 'Counterfactual Judgments' task 173
10.3.2.1 Ratings of regret and blame 173
10.3.2.2 Preferences for counterfactual alternatives 173
10.3.2.3 Ratings of guilt for counterfactual alternatives 174
10.3.3 Additional measures 176
10.3.3.1 Executive functions 176
10.3.3.1.1 Colour-Word Interference test 176
10.3.3.1.2 Verbal Fluency test 176
10.3.3.2 'IRI' 176

10.3.3.3 'MIT' 177
10.3.4 Correlations between performance on the 'Counterfactual Judgments' task and additional measures 177
10.4 Discussion 179
10.4.1 Summary of findings from the 'Counterfactual Judgments' task 179
10.4.2 Performance on the 'Counterfactual Judgments' task 180
10.4.2.1 Judgments of regret and blame 180
10.4.2.2 Counterfactual preferences and guilt 181
10.4.2.2.1 The type of counterfactual thought 181
10.4.2.2.2 The target of the counterfactual thought 183
10.4.2.2.3 Ratings of counterfactual guilt 183
10.4.2.3 Counterfactual preferences and guilt in people with high versus low autistic traits 184
10.4.3 Limitations of the 'Counterfactual Judgments' task 185
10.4.4 Summary of findings from executive and empathy measures 186
10.4.5 Contribution of executive functioning and empathic processes 186
10.4.6 Conclusion 189

Chapter 11: Utilitarian decision-making for physical and social harm 190
11.1 Introduction 190
11.1.1 Hypotheses 193
11.1.2 The 'Utilitarian Judgments' task 193
11.1.1.1 Utilitarian/ non-utilitarian choices 193
11.1.1.2 Discomfort ratings 193
11.1.1.3 Verbal responses 193
11.1.3 Moral questionnaires 194
11.2 Methods 194
11.2.1 Screening phase 194
11.2.1.1 Screening participants and procedure 194
11.2.2 Experimental phase 194
11.2.2.1 Design 194
11.2.2.2 Experimental participants and procedure 194
11.2.3 The 'Utilitarian Judgments' task 195
11.2.3.1 Administration 196
11.2.3.2 Scoring 196
11.2.3.2.1 Scoring of utilitarian/ non-utilitarian choices 196
11.2.3.2.2 Scoring of discomfort ratings 199
11.2.3.2.3 Scoring of verbal responses 199
11.2.4 'Ethics Position Questionnaire' (EPQ) 201
11.2.4 'Moral Behaviour Inventory' (MBI) 201

11.3 Results 201
11.3.1 Data analysis 201
11.3.2 The 'Utilitarian Judgments' task 201
11.3.2.1 Utilitarian/ non-utilitarian choices 201
11.3.2.2 Discomfort ratings 202
11.3.2.3 Verbal responses 203
11.3.3 'EPQ' 207
11.3.4 'MBI' 207
11.4 Discussion 208
11.4.1 Summary of findings 208

11.4.1.1 Utilitarian/ non-utilitarian choices 208
11.4.1.2 Discomfort ratings 208
11.4.1.3 Verbal responses: considering the characters' and the agent's perspectives 209
11.4.1.4 Moral questionnaires 210
11.4.2 Performance on moral questionnaires 210
11.4.3 Performance on the 'Utilitarian Judgments' task 211
11.4.3.1 The influence of personal involvement and the type of harm involved for utilitarian decisions 211
11.4.3.2 Utilitarian decision-making and autistic traits 213
11.4.3.3 The possible contribution of executive and empathic processes 215
11.4.4 Strengths and weaknesses of the 'Utilitarian Judgments' task 216
11.4.5 Conclusion 218

PART 4: GENERAL DISCUSSION 219

Chapter 12: General Discussion 219
12.1 Overview 220
12.2 Novel contribution to the literature 220
12.3 Methodological contribution 222
12.3.1 The broader autistic phenotype 222
12.3.2 Balancing experimental control with ecological validity 221
12.3.3 Exploring the 'social' in social cognition 222
12.3.4 Extending the findings 224
12.3.5 Limitations of the methodological approach 225
12.3.5.1 Study design 225
12.3.5.2 A note on gender 227
12.3.5.3 Experimental materials 228
12.4 Theoretical contribution 229
12.4.1 Pro-social behaviour 229
12.4.1.1 Responsivity to others' needs in pro-social behaviour 230
12.4.1.2 Identifying and reasoning about the expectation of pro-social behaviour 231
12.4.2 Moral judgment 233
12.4.2.1 Emotional responses to moral judgment 234
12.4.2.2 Identifying and reasoning about moral rules 236
12.4.2.3 Counterfactual thinking in response to negative events 237
12.4.3 Understanding the influence of autistic traits on everyday social behaviour 238
12.4.3.1 Affective aspects of performance 239
12.4.3.2 Cognitive aspects of performance 240
12.5 Links to psychological models of ASD 241
12.5.1 Summary of findings from executive and empathic measures 242
12.5.2 The role of executive functioning and empathic processing in social performance 243
12.5.3 Alternative accounts 244
12.6 Clinical implications 245
12.6.1 Background and the relevance of the present work 245
12.6.2 Additional diagnostic tools 246
12.6.3 Transforming interventions 246

12.6.3.1 Using external cues to guide social performance 247
12.6.3.2 Targeting the discrepancy between identifying and reasoning about social rules 248
12.6.3.3 Enhancing social motivation 249
12.7 Concluding comments 251

REFERENCES 252

LIST OF TABLES

Table 1: Mean percentage scores and standard deviations for all measures for the 'Above and Beyond' task. Chapter 5.3 74

Table 2: Mean percentage scores and standard deviations for all measures for the 'Social Expectations' task. Chapter 6.3 91

Table 3: Mean percentage scores and standard deviations for all measures for the 'Social Favours' task. Chapter 7.3 108

Table 4: Mean scores and standard deviations for additional measures. Chapter 7.3 110

Table 5: Mean percentage scores and standard deviations for the 'Deservingness' tasks. Chapter 8.3 134

Table 6: Mean percentage scores and standard deviations for the 'RCFT' task. Chapter 9.3 153

Table 7: Mean percentage scores and standard deviations for the 'Counterfactual Judgments' task. Chapter 10.3 174

Table 8: Mean scores and standard deviations for additional measures. Chapter 10.3 197

Table 9: Mean percentage scores and standard deviations for scenarios from the 'Utilitarian Judgments' task. Chapter 11.3 206

Table 10: Mean percentage scores and standard deviations for the 'EPQ' and 'MBI'. Chapter 11.3 207

LIST OF FIGURES

Figure 1: False-belief paradigm example. Chapter 2.3.1 31

Figure 2: Examples scenario from the 'Above and Beyond' task. Chapter 5.2 71

Figure 3: Example scenario from the 'Social Expectations' task. Chapter 6.2 86

Figure 4: Scoring of example scenario from the 'Social Expectations' task. Chapter 6.2 88

Figure 5: Example scenario from the 'Social Favours' task. Chapter 7.2 102

Figure 6: Example items from the 'IRI'. Chapter 7.2 104

Figure 7: Example scenarios and answers from the 'MIT'. Chapter 7.2 106

Figure 8: Comparison of the 'Deservingness Tasks: Positive and Negative Outcomes'. Chapter 8.2 129

Figure 9: Example scenario from the 'Deservingness: Positive Outcomes'. Chapter 8.2 130

Figure 10: Example scenario from the 'Deservingness: Negative Outcomes'. Chapter 8.2 131

Figure 11: Design of the 'RCFT' task. Chapter 8.2 146

Figure 12: Example congruent scenario from the 'RCFT' task. Chapter 9.2 148

Figure 13: Example incongruent scenario from the 'RCFT' task. Chapter 9.2 149

Figure 14: Example scenario from the 'Counterfactual Judgments' task. Chapter 10.2 170

Figure 15: Example physical scenario from the 'Utilitarian Judgments' task. Chapter 11.2 197

Figure 16: Example social scenario from the 'Utilitarian Judgments' task. Chapter 11.2 198

Figure 17: Scoring of verbal responses from the 'Utilitarian Judgments' task. Chapter 11.2 200

PART 1: LITERATURE REVIEW

CHAPTERS 1-4

Chapter 1: Clinical features of autism spectrum disorder (ASD) and the autism continuum

Chapter 2: Aetological models of ASD

Chapter 3: Social skill training and pro-social functioning in ASD

Chapter 4: Moral judgment and reasoning in ASD

Introduction to literature review

Autism is a life-long condition that predominantly affects social interaction and communication, and is also associated with rigid or narrow interests and repetitive behaviours. Chapter 1 describes the core clinical features of autism and its prevalence, and updates to the diagnostic categorisation are reviewed. The re-conceptualisation of autism as a spectrum disorder and recent work examining autistic traits in the general population is also discussed. Whilst autism is a behaviourally defined disorder a substantial body of work has also been directed at identifying and examining possible biological substrates. Chapter 2 explores evidence for biological accounts of autism, including genetic candidates, studies exploring differences in anatomical and functional brain imaging, and putative biochemical associations. This Chapter also examines the most influential psychological accounts of both the non-social and social features of autism, and the relationship between these accounts.

These two Chapters provide a thorough overview of autism and a number of possible aetiological models. However, the main focus of this thesis is to explore how postulated deficits and differences associated with the autistic spectrum might translate into difficulties in everyday social behaviour. Thus, the importance of investigating the nature of everyday social impairments in ASD, and how this might inform social skill intervention programmes and support strategies are reviewed in Chapter 3. This Chapter also highlights the importance of pro-social behaviour and discusses how this might be impaired in ASD. This Chapter ends by introducing the first three experimental Chapters of this thesis that systematically examine pro-social behaviour in relation to the autism spectrum (Chapters 5, 6 and 7). Chapter 4 shifts the focus onto moral judgment reasoning, and reviews how aspects of this might be impaired in ASD and the evidence for this. At the end of this Chapter the rationale for the last four experimental Chapters is described, which examine different aspects of moral judgment and reasoning in relation to the autism spectrum (Chapters 8, 9, 10 and 11).

Chapter 1: Clinical features of autism spectrum disorder and the autism continuum

1.1 Diagnostic criteria of autism spectrum disorder

Autism is a neurodevelopmental disorder with associated social impairment, difficulties with communication, and narrow and repetitive interests and activities (Wing, 1981). The social deficits seen in autism, and its sub-types, consist of a lack of interest in others, a lack of eye contact, a preference for being alone, difficulty knowing how someone else thinks and feels, and how to react to their behaviour. The communication deficits include literal understanding of speech, language delay and echolalic speech. The non-social symptoms of autism include repetitive body movements such as hand flapping or spinning in circles, which are more strongly associated with more severe cases. They also include obsessional interests, highly repetitive behaviour, islets of intelligence, unusual memory and a need for sameness (Baron-Cohen, 2008).

The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) was published on the 18th May 2013. This update to the American Psychiatric Association's (APA) classification and diagnostic tool superseded DSM-IV, which was published in 2000. In DSM-IV there were four separate diagnoses: autism, Asperger's syndrome, childhood disintegrative disorder, or pervasive developmental disorder not otherwise specified (APA, 2013). A neurodevelopmental work-group from the National Institute of Mental Health, recommended that the DSM-5 subsumed the four different diagnoses into a single umbrella diagnosis of Autism Spectrum Disorder (ASD).

This shift was prompted by research that highlighted concerns regarding inter-rater reliability and the sensitivity/ specificity of the sub-types. Multi-site studies found that the separable diagnoses were not consistently applied across different clinicians (Klin, Land, Cicchetti & Volkmar, 2000; Lord et al., 2012), and studies assessing the impact of diagnosis for later life outcomes did not discriminate between the different sub-types (Howlin, 2000; Allik, Larsson & Smedje, 2006). The move towards a single category of ASD also reflects that autism and related disorders have increasingly been conceptualised as a continuum, with some individuals showing milder symptomatology and others more severe symptoms. The re-configuration of autism into a spectrum disorder therefore aims to allow clinicians to better account for variation in symptoms and behaviours from person to person.

Critics of the DSM-5 have argued that diagnostic sub-types remain clinically useful, and that rather than removing Asperger Syndrome, the diagnostic criteria should be refined to better discriminate between the different profiles, including subtler deficits in social impairment, and communication and language difficulties (Ghaziuddin, 2010). Furthermore, the omission of Asperger Syndrome from DSM-5 may be detrimental to individuals and their families who identify with the diagnosis (Kaland, 2011).

Further changes to the criteria were made regarding the grouping of the symptoms. In DSM-IV there were three clusters of symptoms: social reciprocity, communicative intent, and restricted and repetitive behaviours (APA, 2000). In DSM-5 these have been rearranged into two areas, with the removal of the division between social and communication features (APA, 2013). As with DSM-IV, impairment in both social and non-social areas is a requirement of diagnosis. To meet the criteria for the social communication/ interaction cluster (social features), the individual must display: i) problems in initiating, reciprocating and maintaining social and emotional interaction, ii) difficulty developing and maintaining relationships, and iii) difficulties understanding and using non-verbal communication. To meet criteria for the restricted and repetitive behaviour cluster (non-social features), the individual must display two of four problems, including i) stereotyped or repetitive speech or motor movement, ii) excessive adherence to routines and behaviour patterns, iii) restricted interests that are abnormal in intensity or focus, and iv) hyper/hypo reactivity to sensory information. As with DSM-IV these symptoms must be present from early childhood, although they may not become problematic or clear until later life when social demands increase, and they should not simply be present but also impair functioning. Symptom severity for each of the two areas is also now defined, based on the impact of the symptoms and the level of support required. Symptom severity may also reflect additional developmental problems such as intellectual disability and language impairment.

1.2 Prevalence and prognosis of ASD

Estimates put the prevalence rate of ASD at 1% in the UK (Baron-Cohen, Ashwin, Ashwin, Tavassoli, Chakrabarti, 2009). Other estimates put the worldwide prevalence below 1%, with figures converging on around 60-70/10,000 (Fombonne, 2009; Elsabbagh et al., 2012), suggesting that globally ASD is one of the most frequent childhood neurodevelopmental disorders (Fombonne, 2009). Despite the variability of estimates, a recent review found no evidence for variation in prevalence by geographic region, nor a strong impact of socio-economic factors (Elsabbagh et al., 2012). However, the study did note that at present there was very limited evidence available from low- and middle- income countries, which makes establishing an accurate picture of geographic patterns more challenging. There is also clear evidence to support the rising prevalence of ASD over time, which is thought to reflect a broadening of the diagnostic

concepts and an increased awareness in both general public and professionals (Fombonne, 2009; Elsabbagh et al., 2012).

Prior to publication of DSM-5, the new criteria were tested in real-life clinical settings, and the findings indicated that the new diagnostic classification should not significantly impact on the prevalence of the disorder (APA, 2013). However, the changes to the diagnostic criteria in DSM-5 have been heavily criticised by various stakeholders both prior to and post publication, partly on the grounds that the changes might significantly alter the composition of the autism spectrum (McPartland, Reichow & Volkmar, 2012). On one hand, the changes to the criteria regarding sensory disturbance might result in the inclusion of some individuals into the ASD population who may previously not have met criteria. Conversely, the re-groupings of the criteria, and additional requirement of the presence of two repetitive and restricted behaviours, may lead to the exclusion of some individuals who would have previously met the diagnostic criteria under DSM-IV. In either case, changes leading to alterations in the selection of those meeting the criteria for ASD may have serious ramifications. From a clinical perspective, expanding the diagnostic boundaries would stretch limited resources further, but narrowing them could result in vulnerable individuals losing eligibility for support. From a research perspective, if the ASD population composition changes dramatically, this could have implications for the comparability of different individuals diagnosed at different times, and of studies which use different diagnostic criteria (McPartland et al., 2012).

ASD is a lifelong neurodevelopmental condition for which there is no known cure, although there are a number of interventions that may improve functioning and quality of life. The difficulties associated with ASD increasingly hamper everyday functioning, as the complexities of social relationships and independent living become greater in adulthood (Beauchamp & Anderson, 2010). Reviews indicate that the prognosis of ASD for everyday living is generally poor; a minority of individuals live independently, few have social and intimate relationships, and education and employment levels are low, even when general intelligence is within the normal range (Gillberg, 1991; Howlin, 2000; Seltzer, Shattuck, Abbeduto & Greenberg, 2004; Sigman, Spence, Wing, 2006).

There is general agreement that ASD should be identified as early as possible in life to ensure appropriate support is provided, and to maximise any possible gains from treatment or intervention programmes (Fernell, Eriksson & Gillberg, 2013). However, because of the extremely heterogeneous, poorly understood aetiology of ASD, and the differing profiles of individuals on the spectrum, time of identification varies substantially from person to person. In particular, diagnosis may be later for those who are high-functioning, where subtle but pervasive difficulties may only become apparent or problematic relatively late in life, owing to the need to navigate increasingly complex environments and demands during transition into adulthood.

1.3 Measuring autistic traits in the general population

As highlighted in sections 1.a and 1.b, autism is now considered a spectrum disorder, and the distinctions previously made between different sub-types have been removed, which is a shift consistent with the broadening of the autism continuum. It is thought that autistic symptoms range from mild to severe across the population, and those with a diagnosis are at the extreme end of the continuum. This is supported by work indicating that there is no evidence of a bimodal distribution separating out clinical and non-clinical levels of impairments (Skuse, Mandy & Scourfield., 2005).

Based upon a continuum view of autism, Baron-Cohen, Wheelwright, Skinner, Martin & Clubley (2001) developed the 'Autism Spectrum Quotient' (AQ). This is a self-report questionnaire, consisting of 50 questions across five dimensions, designed to measure the level of autistic traits in the general population. The AQ was designed for high-functioning individuals with normal intelligence capabilities, and assesses the following core diagnostic traits of ASD: poor communication, poor social interaction, limited imagination, superior attention to detail and poor attention switching. The final questionnaire was developed using a clinical sample of those diagnosed with ASD, a control sample and a further control sample of students. Baron-Cohen et al. (2001) found that 80% of people with an ASD diagnosis scored over 32, whereas just 2% of controls did. The AQ has shown good reliability and good validity (for more details on reliability and validity please see page 65; Hoekstra, Bartels, Cath & Boomsma, 2008). All those scoring highly on the AQ (above 32) reported significant social interaction impairments throughout their lives with regard to peer relationships, bullying and social isolation. The mean AQ score for the control group did not differ significantly from that of the control student group, suggesting that a student sample is representative of the general population.

Measuring autistic traits in the population can be a valuable tool for examining the relationship between levels of traits and degree of impairment, and has been explored in a number of domains. Within the non-social domain, work has focused on elucidating differences in the broader phenotype, including visuospatial skills (Almedia, Dickson, Maybery, Badcock, Badcock, 2012; Bayliss & Kritikos, 2011; Grinter et al., 2009) and abnormal sensory responsivity (Horder, Wilson, Mendez & Murphy, 2013). Other studies have explored the link between cognitive processes and the social domain, including biological motion processing (van Boxtel & Lu, 2013), the identification of animate versus inanimate objects (Burnett & Jellema, 2013) and the learning of social information (Hudson et al., 2012). Further work has explored emotion processing and its role in social functioning (Cooper, Simpson, Till, Simmons & Puzzo, 2013; Gökçen, Petrides, Hudry, Frederickson, & Smillie, 2014; Poljac, Poljac & Wagemans, 2012). This body of work indicated a pattern of cognitive and emotional features in individuals with high autistic traits scores similar to, but perhaps less severe than, that typically seen in ASD.

1.4 Chapter summary

This first literature Chapter outlined the clinical features, epidemiology and prognosis of ASD. It also highlighted recent changes to the diagnostic criteria of ASD, and the reconceptualization of autism as a spectrum condition. The Chapter discussed evidence for the autism continuum, whereby autistic traits are thought to be present throughout the general population and those with a diagnosis of ASD to lie at the extreme end of the continuum. Measuring autistic traits in the general population and comparing individuals who lie at different points along the continuum for performance on tasks that tap different cognitive, emotional and social abilities has proved fruitful for broadened understanding of the autism spectrum. Despite burgeoning interest in this new approach, very little work has explored autistic traits in relation to social functioning, which is a central area of impairment in ASD; this thesis adopted a continuum approach to ASD, comparing university students with high versus low AQ scores on different facets of social functioning.

Chapter 2: Aetological models of ASD

2.1 Biological models of ASD

2.1.1 The genetics of ASD

ASD is thought to be influenced by strong but complex genetic factors (Cook, 1998). The strongest evidence for genetic factors comes from twin studies that show a high concordance of ASD in monozygotic twins (genetically identical) and a relatively small concordance in dizygotic twins (share 50% of their DNA as with any sibling pair). For instance, Bailey et al., (1995) found that 60% of monozygotic twins (N = 25) were concordant for ASD, compared with none of the dizygotic pairs (N = 20) studied. Additionally, 92% of monozygotic pairs were also found to be concordant for a broader spectrum of related cognitive or social abnormalities, versus just 10% of dizygotic pairs.

It has also been estimated that if one child in a family has ASD, the risk of a younger sibling also having ASD is 4.5%, which is a substantial elevated risk compared to the 1% or less prevalence rate for the general population (Jorde et al., 1991). Evidence from studies with first-degree relatives supports the notion that autistic traits are on a spectrum. Findings suggest that first-degree relatives have an increase in behavioural or cognitive features associated with ASD, albeit in lesser forms, when compared with the population prevalence (Losh et al., 2009). This has been called "the broader phenotype" and includes restrictive repetitive behaviours and sub-threshold deficits in social cognition, as well as language dysfunction (Warren et al., 2001; Constantino, 2011; Gamliel, Yirmiya, Jaffe, Manor, O., & Sigman, M, 2009)

Another approach is to explore the increased incidence of chromosomal and single gene disorders in ASD. Such disorders include Down's syndrome (Ghaziuddin, 1997) and Turners's syndrome (Skuse et al., 1997). However, it is often difficult to establish if the relationship with ASD is specific, since studies that exclude participants with significant intellectual disability tend to report lower rates of co-morbidity for chromosomal disorder (Cook, 1998). Several single gene disorders are also associated with ASD including phenylketonuria (Cohen, Young, Lowe, & Harcherik, 1980), tuberous sclerosis (Smalley, Tanguay, Smith, & Gutierrez, 1992; Hunt & Shepherd, 1993) and fragile X syndrome (Reiss & Freund, 1990). Whilst it is very unlikely that the majority of cases of ASD are due to single gene or chromosomal disorders, identifying associations between these disorders and ASD has helped to ascertain that these candidate genes are potential risk factors in ASD (Cook, 1998).

Many genetic candidates have now been identified for ASD, but few are specific to it; rather, they appear to contribute to increased risk of a range of psychiatric, neurodevelopmental and neurological

disorders (Geschwind, 2011). This suggests that dysfunction in these genes is implicated in the disruption of key biological pathways. Great progress has been made in establishing the role of genetic factors in ASD over the past two decades. However, further progress hinges on developing a better understanding how a specific genetic risk leads to changes in neural circuitry and function, which ultimately relates to cognitive and behavioural abilities and thus the clinical features of ASD (Geschwind, 2011).

2.1.2 The neurobiology of ASD

In recent years the anatomical and functional brain structure of those with ASD has been increasingly studied, to investigate how altered neural functioning might be implicated in the clinical features of the disorder (Eigisti & Shapiro, 2003). There are now a number of new technologies that provide non-invasive approaches for studying the brain. To examine brain structure, researchers often use computerised tomography (CT: cross-sectional x-ray of the brain) or magnetic resource imaging (MRI: magnetic fields and radio waves pass through the brain to form images); the most common methods to examine brain functioning are positron emission tomography (PET: creation of images through tracking a tracer injected into the bloodstream), functional MRI (fMRI: blood oxygenation levels are tracked to indicate areas of relative activity and inactivity) or electroecephalography (EEG: recordings of neuronal electrical activity). One further method is diffusion tensor imaging (DTI) which allows for imaging of the white matter tracts of the brain, and is also used to examine connectivity of the brain.

Measuring overall brain volume provides a general measure of neural development. Studies of individuals with ASD indicate that a small proportion have associated brain enlargement, but a significant subset have enlarged head circumferences (Piven et al., 1995), which may be the consequence of atypical head growth in childhood (Lainhart et al., 1997). CT and MRI studies have not indicated any gross structural abnormalities in ASD, and although there is some evidence for sub-cortical neuropathology such as cerebellar abnormalities this is not thought to be central to the disorder (Eigsti & Shapiro, 2003). DTI studies of ASD have also consistently indicated a decrease in the size of the corpus callosum, a structure responsible for communicating information between the two halves of the brain (see Eigsti & Shapiro, 2003). A "disconnection" between the two spheres of the brain may account for some of the symptoms of ASD (Alexander et al., 2007).

The limbic system, and more specifically the amygdala, has been widely studied in ASD based on its putative role in social learning and emotion regulation. MRI studies indicate an enlarged amygdala volume in high-functioning individuals with ASD, and relatively decreased volume in other limbic areas (Howard et al., 2000; Abell et al., 1999). The temporal lobes and ventricular system has also been widely studied and a range of abnormalities have been reported in ASD (see Eigisti & Shapiro, 2003). The frontal lobes are thought to be the seat of executive functions, which may be impaired in ASD. The term

executive functions refer to high-level cognitive processes, including working memory, reasoning, mental flexibility, problem-solving and planning (Stuss & Knight, 2013). Consistent with this, MRI studies that indicate enlarged total brain volume in ASD tend not to find a corresponding increase in the frontal lobe, suggesting that relative to healthy controls the frontal lobes may be reduced in volume in ASD (Eigisti & Shapiro, 2003).

Studies of neural functioning in ASD using EEG have indicated electrophysiological abnormalities, consistent with sensory disturbance seen in the disorder Eigisti & Shapiro, 2003). A recent review examined seventy-nine studies of ASD using PET and SPECT (single-photon emission computer tomography) (for a review see Zurcher, Bhanot, McDougle & Hooker, 2015). The authors of the review identified consistent evidence for localised dysfunction in the bilateral temporal lobes, with corresponding reduced blood-flow in this region in ASD when performing socio-cognitive tasks such as emotion and language processing. Other studies found atypical glucose metabolism in the cingulate, occipital and parietal cortices when performing memory tasks in individuals with ASD. The review also discusses how PET has been used to explore the role of possible neurotransmitters abnormalities in ASD. Differences between individuals with ASD and healthy controls have been identified in a number of neurotransmitters and their functions, including the serotonergic system (thought to be involved in the inhibition of sensory input and behavioural output), dopamine (involved in social reward and motivation) and gamma-aminobutryic acid (GABA: central to synaptic pruning and other early developmental processes).

More recently, fMRI has been increasingly used to investigate ASD, examining a range of populations, and using a variety of tasks and methods of analysis (Phillip, Dauvermann, Whalley, Baynham, Lawrie & Stanfield, 2012). A review of ninety articles examining fMRI in ASD identified a number of consistent differences in comparison to control groups (Phillip et al., 2012). During motor tasks, participants with ASD displayed enhanced activation of the bilateral precentral gyri and the inferior/middle frontal gyri in comparison to controls. For visual processing tasks, the ASD groups showed greater activation of the thalamus and the medial frontal gyrus, whilst control groups showed more activation in the cingulate and occipital regions. Studies exploring executive functions found that individuals with ASD activated the left, as compared to the right, middle-frontal gyrus more than controls. Control participants also tended to display greater activation in other prefrontal and subcortical regions. In auditory and language tasks participants with ASD showed greater activation in the right precentral gyrus and left declive, whereas controls showed greater activation in the bilateral superior temporal gyri. In social processing tasks there was a complex pattern of differences between the groups; in particular the ASD participants showed a mixture of over- and under- activation of the left superior temporal gyrus compared to controls.

Interestingly, few studies included in the review reported behavioural differences between the groups (Phillip et al., 2012), suggesting that differences in observed activation patterns are not simply due to performance differences. The authors of the review note that this is an unexpected finding, especially given that performance differences on such tasks are well established in the existing neuropsychological literature. However, they also highlight that the tasks used in fMRI scanners tend to be simpler, because of the constraints of what can be assessed within a scanner, and thus "ceiling effects" (when a test is too easy to discriminate between conditions or groups) may account for the lack of difference. It is of note that the more complex social cognition tasks did reveal behavioural differences between the groups, which might be more sensitive to the specific deficits seen in ASD. This is consistent with other work that identifies both impoverished behavioural performance in ASD on a classic social cognition task and corresponding reduced activation in critical "social brain" regions (including the medial prefrontal cortex, the tempo-parietal junction and the temporal poles) (Castelli, Frith, Happé & Frith, 2002).

In summary, a range of abnormalities in brain structure, connectivity and function have been identified in those with ASD, which correspond to the clinical features associated with the disorder. It seems likely that there are that abnormalities across the neural network including the fronto-temporo-parietal cortex, limbic system, and cerebellum. Some have suggested that these differences may underlie the pathophysiology of ASD, and that such changes could result from abnormal brain development during early life. For instance, if synaptic pruning is abnormal in ASD, then excessive white matter connections could lead to "sticky brain networks" that fail to learn readily (Cohen, 1994). However, it is difficult to draw causal links about the differences seen in ASD on the basis of the relatively mixed and limited evidence available. Most of the work discussed in this section utilised cross-sectional designs to compare the biology and performance of individuals with ASD to healthy matched-controls. Large-scale longitudinal designs that explore brain development in healthy individuals in comparison to those with neurodevelopmental conditions such as ASD would provide better 'causal' evidence. This is particularly important for ASD, a hugely heterogeneous condition, where cross-sectional comparisons are likely to be hampered by large inter-participant variability (Zurcher et al., 2015). However, longitudinal designs are typically challenging to design and manage, and very resource intensive.

2.1.3 Prenatal, perinatal and neonatal risk factors in ASD

Although the specific neuropathology of ASD remains elusive, a range of associated brain abnormalities has been identified. This suggests that neurodevelopmental processes may be aetiologically involved in ASD, perhaps as early as in utero or in early infancy (Gardener, Spiegelman & Buka, 2011). A substantial amount of research has examined the potential role of obstetric complications and neonatal exposures as risk factors for ASD. A meta-analysis examined forty studies that explored the relationship

between perinatal and neonatal factors and risk of ASD (Gardener et al., 2011). The review concluded that although many studies find that obstetric and neonatal complications increase the risk of ASD, the body of work is inconsistent and that overall the associations are not statistically significant. The authors identified some factors that had stronger evidence for increased risk of ASD, including abnormal fetal presentation, umbilical-cord complications, fetal distress, birth injury or trauma, multiple births and low birth weight. The authors also identified some factors that had stronger evidence *against* associated increased risk for ASD, including use of anaesthesia during delivery, assisted vaginal delivery, post-term birth and high birth weight.

The studies were often inconclusive when considered as a whole body of evidence, and for factors where results were more consistent, the effect sizes tended to be very heterogenous (Gardener et al., 2011). The authors noted that variability in study design may account for some of this heterogeneity, and highlighted the potentially important role of prenatal complications. A previous meta-analysis identified a few prenatal factors that were associated with increased risk of ASD, including advanced maternal and paternal age at birth, maternal gestational bleeding, gestational diabetes, maternal prenatal medication use, and being first-born or fourth or later born (Gardener, Spiegelman & Buka, 2009). Prenatal factors may not have been recorded or accounted for in the studies examined in the perinatal and neonatal meta-analysis, which may account for some of the variability across studies. It is also possible that the presence of some prenatal risk factors may have knock-on effects, with increased risk of subsequent perinatal and neonatal risk factors, but this is not well understood. Overall, the authors concluded that it was most likely that exposure to multiple perinatal and neonatal complications was associated with an increased risk of ASD, over and above any one single complication. They also proposed that complications might only have an impact for individuals who were also genetically vulnerable to ASD. The authors recommended that further research should explore the joint and independent effects of adverse conditions throughout the conception to early infancy time course, including prenatal factors.

2.1.4 The role of oxytocin in ASD

Oxytocin has been implicated in a range of social behaviours in animal studies (e.g. Insel & Hulihan, 1995) and more recently in humans (Lam, Aman & Arnold, 2006). It is synthesised in the paraventricular nucleus and supraoptic nucleus of the hypothalamus, which diffuses throughout the brain and brainstem (Sofroniew & Weindl, 1981), to oxytocin receptors in the limbic system and subcortical regions of the brain (Barberis & Tribollet , 1996). Oxytocin nerve fibers are evident in a variety of brain regions, including those though to be associated with social perception and cognition, and emotion regulation, such as the amygdala (Anagostou et al., 2014).

Oxytocin is thought to play a crucial role in maternal behaviour, infant separation distress and the development of social attachments in animals (Lam, Aman & Arnold 2006). Studies examining the effects of intranasal oxytocin in humans have found alterations in social decision-making, processing of social stimuli and social memory (MacDonald & MacDonald, 2010. Since ASD is primarily characterised by pervasive social impairments researchers have begun to investigate whether the oxytocin system is dysfunctional in individuals with ASD. The first study to examine blood levels of oxytocin found that children with ASD had significantly lower levels of oxytocin compared to controls (Modahl et a., 1998). Interestingly, higher levels of oxytocin were positively related to socialisation skills in controls, but negatively associated with socialisation skills in children with ASD. This suggests that the measurement of lower oxytocin in the participants with ASD reflects underlying abnormalities in receptors or substances upstream of the oxytocin system. This indicates that a postulated oxytocin deficit in ASD is overly simplistic and rather points towards secondary oxytocin dysregulation. There is also some evidence for a genetic association between oxytocin and ASD, with studies reporting single nucleotide polymorphisms of the oxytocin gene (e.g. Wu et al., 2005; Jacob et al., 2007) rare genetic variations of the gene in ASD (Liu et al., 2015), and potential epigenetic modification of the oxytocin gene in ASD (Gregory et al., 2009).

Further research has examined the social effects of administering oxytocin intra-nasally to individuals with ASD. This has been found to improve recognition of others' emotions (Guastella et al., 2010), and also face processing and social perception abilities (Andari, et al., 2010). Another study explored the effects of intranasal oxytocin on performance on a face-matching task during fMRI in individuals with ASD (Domes et al., 2013). Under the placebo condition the ASD participants showed decreased activity in the right amygdala, fusiform gyrus and inferior occipital gyrus during face processing, compared to a control group. After oxytocin treatment right amygdala activity to facial stimuli increased in the ASD group. However, this effect did not hold for task performance: there was no improvement in the ASD or control groups' face processing abilities in the oxytocin versus placebo condition. The authors of a review of the social effects of oxytocin in humans highlighted the discrepant findings across studies, and suggested that instead of simply asking whether oxytocin improves social cognition, research should focus on examining under what circumstances oxytocin acts to improve functioning (Bartz, Zaki, Bolger & Ochsner, 2011). A more nuanced understanding of oxytocin's effects should provide insight into the basic mechanisms underlying how it works, and clarify the therapeutic expectations with respect to the alleviation of social deficits.

Despite the relatively recent insights into the mechanisms of oxytocin, and the limited evidence for its possible role in ASD, there is a great deal of enthusiasm for its potential therapeutic benefits. There is a paucity of medications targeting the core symptoms of ASD, and oxytocin may prove a good candidate for targeting the associated social and emotional deficits (Anagnostou et al., 2014). The evidence discussed

above is limited to single-dose intranasal oxytocin studies, which do not provide a good understanding of the therapeutic potential of administering intranasal oxytocin over time (MacDonald & Feifel, 2013). Multi-dose studies that can evaluate the compound's long-term therapeutic role are required, and the small number that have been undertaken thus far are proving promising (see Anagnostou et al., 2014 for a review). However, larger studies that examine the safety and efficacy of the role of oxytocin as a therapy for ASD are required.

2.2 Psychological models of ASD: Non-social deficits

Psychological theories of ASD are useful in conceptualising the disorder and in drawing links between observed abnormal behaviours and postulated cognitive impairments. Several major psychological theories have been proposed to account for the symptoms associated with ASD, two of which, 'Weak Central Coherence Theory' (Frith, 1989; 2003; Happé, 1999; Happé & Frith, 2006) and 'Executive Function Theory' (Ozonoff, Pennington & Roger, 1991; Ozonoff, South & Provencal, 2005), are more commonly used to explain the non-social symptoms of ASD.

2.2.1 'Weak Central Coherence'

'Weak Central Coherence' (WCC) theory (Frith, 2003; Happé, 1999; Happé & Frith, 2006) suggests that individuals with ASD tend to process information in a piecemeal manner, involving an enhanced focus on local features at the expense of contextual details (Joliffe & Baron-Cohen 1999; Shah & Frith, 1993; Witkin, Oltman, Raskin & Karp, 1971). Typically developing individuals tend to display strong central coherence and a preference towards wholes rather than parts, whereas those with ASD seem to concentrate more on details. A range of evidence has explored perceptual processing in ASD and has broadly found support for WCC (see Happé & Frith, 2006 for a review). For instance, within the auditory modality there is evidence of more stable memory of exact pitches (Bonnel, Mottron, Peretz, Trudel & Gallun; 2003) less susceptibility to interference to the melodic structure (Foxton et al., 2003), and a reduced McGurk effect (less influence from visual to auditory speech perception; DeGelder, Vroomen & Van der Heide, 1991). Within the visual domain, individuals with ASD show an enhanced ability to perceive coherent motion (Bertone, Mottron, Jelenic & Faubert, 2003), superior visual search (Plasited, O'Riordan & Baron-Cohen, 1998) and a reduced susceptibility to visual illusions (Happé, 1996).

Whilst the majority of work exploring perceptual processing in ASD has found support for the WCC theory, there have been some inconsistent findings. The authors of a review of the evidence for WCC in ASD concluded that this might reflect differences in task demands (Happé & Frith, 2006). They suggest a reduction in global processing may only be evident in tasks where participants with ASD are not deliberately directed to attend to global information. This highlights the sensitivity of open-ended tasks,

and has led to a revision of the theory, with ASD characterised by superiority in local processing, as opposed to a deficit in global processing (Happé & Frith, 2006). This conceptualises autistic symptoms as the result of a certain processing style rather than of a specific deficit, lending itself to a continuum approach. The authors argued that the WCC seen in ASD represents the extreme end of a normal distribution of processing styles (Happé & Frith, 2006).

WCC is typically used to address the characteristic repetitive behaviours and narrow interests of ASD, and the theory does not provide a coherent account of the full range of associated social and communication impairments. However, it may have implications for social functioning via a failure to integrate or process contextual factors (Lawson, Baron-Cohen & Wheelwright, 2004). Socially skilled behaviour requires an appreciation of the social context, involving interpreting and predicting others' behaviour and adapting own's own accordingly. According to WCC theory, people with ASD may fail to comprehend the social context of a situation and thus behave in a socially inappropriate and unskilled manner (Lawson et al., 2004). For example, De Martino, Harrison, Knafo, Bird, & Dolan (2008) found that individuals with ASD displayed enhanced logical consistency during a gambling game, but were less susceptible to social framing. Whilst the participants with ASD performed more successfully on this particular task, this pattern was interpreted as reflecting a lack of integration of emotional and contextual cues into the decision-making process, which may be linked to impoverished social abilities in ASD.

2.2.2 Executive dysfunction

A theory of executive dysfunction in ASD has also been proposed (Ozonoff et al., 1991; Ozonoff et al., 2005). The term 'executive function' encompasses a wide range of skills and abilities that are involved in higher order control of behaviour, such as planning, working memory, impulse control and inhibition, mental flexibility, and initiating and monitoring action (White, Burgess & Hill, 2009). This theory makes an explicit link to the observed frontal dysfunction seen in ASD, specifically to the prefrontal cortex and medial temporal lobes, corresponding with the impaired executive functions of patients with frontal lobe damage (Hill, 2004; White et al., 2009).

Children, adolescents and adults with ASD have been found to be impaired on a range of executive function tasks, particularly where long sequences of moves are involved (for a review see Hill, 2004). Tasks examining 'planning' involve executing a sequence of actions that must be constantly monitored, re-evaluated and updated. For instance, in the 'Tower of London' task individuals must move disks from a pre-arranged sequence to match a new configuration, in as few moves as possible and following several specific rules. Individuals with ASD have consistently been found to show impaired performance on this task, in relation to several other neurodevelopmental conditions and to controls (Hill, 2004).

Individuals with ASD are also well documented to show impairment on tasks examining 'mental flexibility', particularly at the more complex levels (Ozonoff et al., 1991; Hughes, Russell & Robins, 1994; Heaton, Chelune, Talley, Kay & Curtiss, 1993). For instance, people with ASD consistently display difficulties with the Wisconsin Card Sorting Task. This requires participants to work out the rule for sorting cards according to one of three dimensions (colour, shape or number), update it frequently and shift their performance accordingly. Impairment on tasks that tap mental flexibility has been attributed to perseverative problems, where individuals with ASD may get 'stuck in set', lacking the ability to shift to a different thought or action. It has also been linked to problems in self-monitoring behavior (Hill, 2004). Deficits in planning and cognitive flexibility have also been implicated in the broader phenotype in parents and siblings of those with ASD (Piven & Palmer, 1997; Hughes, Leboyer & Bouvard, 1997; Hughes, Plumet & Leboyer, 1999).

Executive dysfunction theory is again mainly applied to the repetitive and stereotyped patterns of behaviour seen in ASD, but may also account in part for the characteristic deficits in social functioning and communication (Happé, Booth, Charlton, & Hughes, 2006a). Effective social functioning is postulated to involve several cognitive process, including evaluating relevant aspects of social situations, planning responses and generating potential courses of action, and appreciating and comparing the social consequences of each (Channon, Charman, Heap & Crawford & Rios, 2001). People with ASD have been found to display impairments in generating socially appropriate solutions to social problems, which have been suggested to be related in part to executive function impairments (Channon et al., 2001). More work is required to examine how executive functions may account for other aspects of social functioning in ASD.

2.3 Psychological models of ASD: Social deficits

The classic non-social accounts of ASD (WCC and executive dysfunction theories), struggle to fully explain the associated social and communication impairments. This is particularly true for individuals of high intelligence who display unimpaired, or indeed superior, reasoning in non-social areas (Happé, Ronald & Plomin, 2006b). The most dominant social theory of ASD is impaired empathic processes; empathy is thought to involve two distinct components, namely cognitive and emotional. Cognitive empathy refers to the ability to attribute and infer the content of another's thoughts, emotions and intentions by 'putting oneself into another's shoes' (Premack & Woodruff, 1978). Whilst cognitive empathy refers to the ability to represent another's mental state (Blair, 2008; Hoffman, 2000; Riggio, Tucker & Coffaro, 1989), emotional empathy refers to the ability to mirror or resonate with their emotional state (Eisenberg & Miller, 1987; Hoffman, 2000). It has been posited that those with ASD posses impaired cognitive empathy (synonymous with the terms 'theory of mind', 'perspective-taking' or 'mentalising'), but preserved

emotional empathy. This is a well-supported account that can provide a full explanation of the social, emotional and communication impairments characteristic of ASD (Spek, Scholte, Van Berckelaer–Onnes, 2010; Happé et al., 2006b). However, it has also been proposed that high-functioning people with ASD may rely more heavily on learnt social knowledge and rules than typically developing individuals to support social performance. Another proposition is the 'Extreme Male Brain' theory of ASD (Baron-Cohen, 2002). Baron-Cohen proposes that the rigid behaviours, social and communication difficulties associated with the disorder can be viewed as an extension of the relative strengths and weaknesses of the male versus female brain. The evidence for these different accounts will now be reviewed.

2.3.1 Impaired cognitive empathy

Impairment in cognitive empathy is posited in those with ASD, often termed 'mind-blindness', which is thought to account for the social and communication impairments (Baron-Cohen, 1995; Baron-Cohen, Leslie & Frith, 1985). This is consistent with neuroimaging work (discussed in Section 2.1.2), which demonstrates that individuals with ASD show reduced activation in the 'social brain' network (encompassing the medial prefrontal cortex, the tempo-parietal junction, and the temporal poles; Blakemore, 2008), when performing tasks involving cognitive empathy (Castelli et al., 2002; Happé et al., 1996).

The 'false-belief' paradigm assesses cognitive empathy by examining the ability to take another's perspective (Dennet, 1978; Premack & Woodruff, 1978). These tasks involve inferring another's false belief about a situation, in order to accurately predict their corresponding behaviour. For instance, in the example shown below (see Figure 1), to arrive at the correct answer – that Maxi will look in the blue cupboard, because this is where he falsely believes the chocolate is – participants are required to appreciate three things. Firstly, Maxi has an independent mind, and his experience of the world is autonomous from that of his mother and the participant. Secondly, Maxi will hold a 'false-belief', because he did not observe nor was he informed, that his mother moved the chocolate. Finally, Maxi is likely to act according to his false-belief.

Figure 1: False-belief paradigm example

> "Maxi has some chocolate and puts it into a blue cupboard. Maxi goes out. Now his mother comes in and moves the chocolate to a green cupboard. Maxi comes back to get his chocolate.
> Where will Maxi look for the chocolate?"

These tasks have since been widely used to identify the stage at which a child can predict another person's behaviour on the basis of their false-belief (Wimmer & Perner, 1983). By four years of age, children begin to understand and explain false belief scenarios (Happé, 1995). By five years of age, over ninety percent of typically developing children can understand such tasks (Baron-Cohen et al., 1985; Perner, Leekham & Wimmer, 1987), and at age six all typically developing children can answer them correctly. Children with ASD are slower to develop these abilities (Happé, 1995), and are found to struggle with more abstract and less explicit perspective-taking tasks (Heavey, Phillips, Baron-Cohen, Ruetter, 2000). In adults with ASD, difficulties with cognitive empathy are often tested at a more subtle level, such as failure to understand and infer the motives, intentions and emotions of characters in stories (Happé, 1994; Spek, Scholte, Van Berckelaer–Onnes, 2010). Failure to detect faux-pas in social situations has also been identified, demonstrating an inability to appreciate and predict the responses of others (Stone, Baron-Cohen & Knight, 1998; Spek et al., 2010).

Cognitive empathic abilities have been linked to measures of social and interpersonal skills (Dawson & Fernald, 1987). Individuals who fail false belief tasks (Baron-Cohen et al., 1985) have been found to show less insightful social behaviour and poorer verbal communication skills (Frith, Happé & Siddons, 1994). Furthermore, performance on measures of cognitive empathy have been associated with skills required for appropriate social behaviour, including conversational abilities, such as maintaining conversation and responding appropriately (Hale & Tager-Flusberg, 2005). It has also been associated with the ability to contribute novel information (Capps, Kehres, & Sigman, 1998), understand non-literal language (Martin & MacDonald, 2004), and to understand and identify others embarrassment (Hiller & Allinson, 2002). People with ASD have been found to be impaired in all of these areas.

Impairment in cognitive empathy can clearly explain why social and communicative difficulties might develop, and it is a widely accepted psychological account of ASD. However, the extent to which this theory can explain the rigid and repetitive behavioural features of the disorder is limited (Happé et al., 2006b). Some have argued that these behavioural features might be a downstream effect of primary impairment in social communication, occurring in reaction to social distress or anxiety as an attempt to calm the individual. However, this argument has not been terribly convincing, since these features occur even in high-functioning individuals with ASD who have some insight into others' thoughts and feelings and tend to experience more subtle social deficits. Furthermore, there is evidence that they serve a self-stimulatory effect, in addition to their proposed soothing function (Turner, 1999; Happé et al., 2006b).

2.3.2 Preserved emotional empathy

Both cognitive and emotional empathy are thought to contribute to successful social functioning (Blair, 2008; Rogers, Dziobek, Hassenstab, Wolf, & Convit, 2007). Although it is likely that these two

abilities are used in concert when studying other people's intentions, beliefs and feelings, evidence from patient populations with marked deficits in empathic abilities suggests these two components are dissociable. For instance, a double dissociation is proposed for the empathic deficits of ASD and psychopathy. Psychopathy is a personality disorder that is characterised by antisocial and impulsive behaviour, and a lack of remorse (Hare, Hart & Harpur, 1991). A core deficit in emotional but not cognitive empathy is proposed for psychopathy, whereas intact emotional empathy with impaired cognitive empathy is posited in ASD (Blair, 2008). Thus, people with ASD are posited to be able to feel for others but not to understand them, and conversely, psychopathic individuals are purported to understand others but not to resonate with their emotional state. Studies with incarcerated psychopathic populations confirm that these individuals do not feel distress when watching others in pain and exhibit reduced physiological arousal (Lykken, 1957; Patrick, Cuthbert, & Lang, 1994; Blair, 1995; Blair, Jones, Clark & Smith, 1997), and that they pass classic tasks of cognitive empathy (Dolan & Fullam, 2004).

Although the theory of impaired cognitive empathy in ASD is well supported, the notion of intact emotional empathy remains disputed. Neuroimaging work has found that people with ASD do not exhibit common activation of neuronal regions, for both the experience and observation of others' pain (Minio-Paluello, Baron-Cohen, Avenanti, Walsh & Aglioti, 2009; Singer, Seymour, O'Doherty, Kaube, Dolan & Frith, 2004). The co-activation of these regions in healthy individuals has led to consideration of their possible role in emotional empathy (Minio-Paluello et al., 2009). A lack of neural overlap in ASD may suggest reduced resonance, and hence impaired emotional empathy. However, a study examining self-reported distress in response to viewing others pain did not support this, as people with ASD were not found to differ from controls (Minio-Paluello et al., 2009). This suggests that those with ASD are capable of emotionally resonating with another's pain but may do so from a 'self' stance due to impairments in understanding others' perspectives (Bacon, Fein, Morries, Waterhouse, Allen, 1998; Eisenberg, 2007; Frith & de Vignemont, 2005; Minio-Paluello et al., 2009). It seems likely that individuals with ASD are capable of resonating emotionally with others, but only if it is made explicit what another is thinking or feeling. The literature is also very mixed with respect to the types of tasks that are used to examine emotional empathy, and given the heterogeneous pool of people who fit under the ASD umbrella, it is not all that surprising that different studies yield different results (Blair, 2008).

Recent work suggests a link between empathy and the mirror neuron system (MNS). The term MNS is used to describe the regions in the inferior parietal and inferior frontal cortex, that respond both when an individual performs an action, and when observing another's action (Rizzolatti & Craighero 2004). The MNS is postulated to facilitate matching actions of the self to those of others', allowing for inference of others' intentions (Hamilton & Grafton 2006). In turn, the ability to understand others' actions and goals might be fundamental to more complex social abilities and cognitive empathy (Gallese & Goldman

1998; Gallese et al. 2004). It has been speculated that a 'broken' MNS may be the cause of the poor social skills and cognitive difficulties characteristic of ASD (Iacoboni & Dapretto 2006; Oberman & Ramachandran 2007), and psychopathic personality traits have been linked with intact or indeed superior functions of the MNS (Fecteau, Pascual-Leone, & Théoret, 2008). Whilst the MNS has received a lot of attention, the extent to which it is able to mediate more complex social abilities and to explain disorders of empathy remains poorly understood (Southgate & Hamilton 2008).

Another emerging area of research is exploring the extent to which co-occuring alexithymia may account for some of the emotional symptoms of ASD. Alexithymia is a sub-clinical condition characterised by difficulties with identifying and describing emotions of the self (Nemiah, Freyberger & Sifneos, 1976) and corresponding impairments in social cognition (e.g. Parker, Taylor & Bagby, 1993). These features may overlap with the social and communication deficits seen in ASD, which is known to frequently co-occur with alexithymia (Bird & Cook, 2013). Some work has explored the extent to which a difficulty in recognising one's own emotional state might lead to difficulties in understanding others' states (Moriguchi et al., 2006), and there is some supporting evidence to suggest that the self- versus other- functions of the MNS are impaired in alexithymia (Moriguchi et al., 2009).

2.3.3 Compensatory social knowledge

Notwithstanding the abundance of evidence supporting impairment in cognitive empathy in ASD, there is also evidence to suggest that some high-functioning individuals may show intact performance. For example, high-functioning children and adolescents with ASD were found to pass a series of social stories examining understanding of non-literal language (e.g. sarcasm), other's intentions (e.g. false belief) or inappropriate behaviour (e.g. faux pas) (Scheeren, Rosnay, Koot & Begeer, 2013). Begeer, Malle, Nieuwland & Keysar (2010) similarly found that high-functioning adolescents and adults with ASD correctly performed a communication game requiring taking another player's perspective. A further study indicated that adults with high-functioning ASD were also able to correctly infer characters' emotional responses upon the receipt of gifts, feigning a positive response for homemade gifts versus a genuinely positive response for chocolate gifts (Cassidy, Ropar, Mitchell & Chapman, 2014). Finally, a number of studies have failed to identify impairment in high-functioning individuals' performance on a classic task involving inferring mental and emotional states on the basis of pictures of people's eyes (Reading the Eyes in the Mind: Ponnet et al, Roeyers, Buysse, De Clercq Van Der Heyden, 2004; Roeyers, Buysse, Ponnet & Pichal, 2001).

In order to account for this discrepancy, researchers have proposed that high-functioning people with ASD might rely on learnt social knowledge and rules in their performance, and that this might help them to mask their deficits, at least to some extent. There is some limited experimental work to support

this theory. Firstly, high-functioning adults with ASD have been found to correctly infer the thoughts and feelings of others when they are familiar with the script of the social situation (e.g. an initial 'chit-chat' conversation with an acquaintance: Ponnet, Buysse, Roeyers & De Corte, 2005). It was suggested that intact performance in the ASD group was supported by the availability of cues that triggered the retrieval of information from memory about similar situations, social scripts or other relevant social knowledge. Secondly, in a follow up study, increasing the structure of a conversation was found to support individuals with high-functioning ASD ability to infer the thoughts and feelings of the other person (Ponnet, Buysse, Roeyers & De Clercq, 2008). Lastly, clinical reports using qualitative methods to examine the everyday problems of those with ASD indicate that high-functioning people may find it helpful to deliberately observe others in order to learn about social rules and principles (Carrington, Templeton, & Papinczak, 2003; Müller, Schuler, & Yates, 2008).

If one applies this pattern of findings to the examples discussed above of intact performance on classic cognitive empathy tasks, it could be argued that that high-functioning people with ASD use their previous experience of social knowledge and rules to 'hack' such tasks (e.g., Frith & Happé, 1999; Hill & Frith, 2003). Thus, the intact performance of high-functioning people with ASD on cognitive empathy tasks is not thought to reflect intuitive awareness or understanding of others' mental states, but rather to be driven by the conscious and laborious application of static and explicit rules and principles (Dewey, 1991; Frith & Happé, 1999). This may serve to disguise the social difficulties of high-functioning people (Kasari, Chamberlain, Bauminger, 2001), provided that the demands are relatively low such as in structured laboratory tasks. However, over-reliance on social knowledge may lead to clumsy or inflexible patterns of behaviour in more complex situations that are typical of the 'real world'.

2.3.4 The 'Extreme Male Brain' hypothesis

Turning to an alternative conceptualisation of ASD, it has been proposed that the condition may represent an extreme version of the natural cognitive differences that exist between men and women (Baron-Cohen, 2002). This proposition was motivated by the observation that a greater proportion of males than females are diagnosed with ASD (Wing, 1981), and a larger proportion of males are found to have high levels of autistic traits in comparison to females (Baron-Cohen et al., 2001). Baron-Cohen has suggested that the male brain is more suited to systemising (a tendency to analyse and deconstruct systems, and to focus on rules and structures), whereas females are more likely to spontaneously engage in empathising (encompassing both cognitive and emotional aspects of empathy). This is supported by evidence that individuals with ASD tend to score better than neurotypical individuals (neurologically typical i.e. not on the autism spectrum) on tests of systemising, regardless of gender, and that neurotypical males tend to score better than neurotypical females on systemising measures (Baron-Cohen, 2009).

Furthermore, when males and females with ASD are compared, there is little evidence of gender differences on measures of cognitive or emotional empathy (Auyeung et al., 2009; Lai et al., 2011; Wheelwright et al., 2006), whereas neurotypical females often score higher than neurotypical males on tests of both emotional and cognitive empathy (Auyeung et al. 2009; Baron-Cohen & Wheelwright, 2004; Mestre, Samper, Frias, & Tur, 2009).

This evidence has led to the suggestion that ASD constitutes an exaggeration of the typical male cognitive profile, known as "The extreme male brain theory" (Baron-Cohen, 2002). Recently, this has been extended to evidence from the fields of neuroanatomy and neuroimaging, with some research suggesting that females with ASD may be physiologically and neurologically more similar to males than their neurotypical counterparts. A crucial role seems to be played by pre-natal levels of testosterone; for instance, this has been found to be positively correlated with scores on a measure of systemising, the Systemising Quotient (Auyeung, et al., 2006), and negatively correlated with a measure of empathy, the Empathising Quotient (Chapman, Baron-Cohen, Auyeung, Knickmeyer, Taylor, & Hackett, 2006). Consistent with this, women with ASD have been found to have higher serum testosterone levels (Schwarz et al., 2011) and a higher prevalence of androgen-related medical conditions such as polycystic ovary syndrome (Ingudomnukul, Baron-Cohen, Wheelwright, & Knickmeyer, 2007). A recent MRI study suggested that women with ASD may have neuroanatomical features that overlap with sexually dimorphic structures in controls, suggesting a neural 'masculinization' of their brains (Lai et al., 2013).

Whilst there is a substantial body of biological and psychological evidence to support the extreme male brain theory of ASD, critics suggest that this theory implies that gender imbalance is an inevitable feature of the disorder. There is increasing acknowledgement that the gender imbalance observed in ASD may be an artifact of societal gender roles. Some evidence suggests the deficits seen in females with ASD are more subtle, and/or that females with ASD are better at masking their impairments, which may mean the difficulties they experience are less likely to reveal themselves and result in a diagnosis (Wing 1981; Ehlers & Gillberg 1993; Gould & Ashton-Smith, 2011). Whether this reflects protective factors associated with a female versus male brain, or under-diagnosis of females, or overdiagnosis of males is not well understood (Halladay et al., 2015).

2.4 A single psychological model of ASD?

There is no single explanation that is satisfactory for explaining the range of cognitive and behavioural impairments associated with ASD. Some attempt has been made to explore the relationship between these different contributions. WCC appears to be independent of both cognitive empathy (Frith & Happé, 1994) and executive functioning (Booth, Charlton, Hughes & Happé, 2003). However, it has been argued that the development of executive functions may be essential for the maturity of subsequent mechanisms that are involved in cognitive empathy (Ozonoff et al., 1991), and that poor performance on false belief tasks may arise as a result of a failure to inhibit a pre-potent response (Frye, Zelazo & Palfai, 1995). Moreover, it has been suggested that the capacity to represent mental states and recognise one's own intentions is essential for executive functions, such as the ability to plan ahead (Perner, 1998). On the other hand, some authors have argued that future work should not seek to account for the full range of impairments, since the features of ASD may be fractionable and relatively independent (Happé et al., 2006b). Rather, they argue that psychological models should strive to find a good account for each of the distinct domains: social/ communication difficulties and behavioural features.

2.5 Chapter summary

This second literature Chapter explored the different aetiological models of ASD. First, a summary of the evidence for various biological accounts of ASD was presented, including genetic candidates and studies exploring differences in anatomical and functional brain imaging. Research examining complications during birth and in early infancy as risk factors for ASD, and putative biochemical associations were also briefly summarised. In conclusion, there is very little consistent evidence for any one biological factor, and the way in which biological risk factors may interact with environmental processes, and one another, to present increased risk for ASD, is poorly understood. This body of work suggests that ASD is a heterogeneous disorder, with multiple possible biological aetiologies. It may involve the disruption of several key biological pathways, many of which are implicated in other neurodevelopmental, psychiatric and neurological disorders.

Secondly, this Chapter examined psychological accounts of both the non-social and social features of ASD. It is thought that individuals with ASD may experience global cognitive differences in their ability to integrate information and execute actions, which may account for the repetitive and restricted features of the disorder. Furthermore, they are thought to be capable of resonating with others (emotional empathy), but to be impaired in their ability to represent others' internal states (cognitive empathy), which accounts for the observed social and communication impairments. There is also some evidence to suggest that the cognitive and behavioural differences in ASD may reflect an 'extreme male brain'.

This thesis does not investigate aetiological models of ASD, but rather seeks to better understand the profile of social difficulties associated with the autistic spectrum and to extend this to the general population. However, the psychological models discussed in this Chapter provide a useful framework for conceptualizing ASD and for interpreting the meaning of experimental results. The biological models discussed provide a further layer of understanding, regarding the extent to which the social difficulties in ASD are primary to the disorder, and afford important considerations when developing clinical interventions.

Chapter 3: Social skill training and pro-social functioning in ASD

3.1 Understanding everyday social functioning in ASD

Despite an abundance of work examining the cognitive and emotional performance of those with ASD, there is a paucity of literature exploring how the profiles identified translate into everyday social functioning. Much of the experimental work examining social performance in individuals with ASD has focused on identifying deficits in fundamental skills, and there is less work exploring the understanding of complex social situations in ASD (Loveland, 2009). Furthermore, the vast majority of evidence comes from laboratory studies that tend to use simple stimuli, such as inferring the mental states of characters from pictures of their eyes, and to explore simple situations, such as false-belief paradigms. Whilst this approach has proved fruitful for generating and testing the predictions of theories about the observed social and communication deficits in ASD, tasks of this nature are often criticised for lacking ecological validity, as they tend to focus on narrow samples of behaviour that can be more easily experimentally manipulated (Loveland, 2009). Thus, this approach lacks the capacity to fully capture the clinical manifestations of the disorder, and to provide useful information about how to best manage social and communication difficulties.

By contrast, clinical accounts tend to focus on documenting the everyday experiences of people with ASD when they navigate the social world. For instance, Müller et al., (2008) individually interviewed adults with ASD to discuss their social and communication challenges. The descriptions of difficulties included trouble following with 'chit-chat' conversations that tend not to have predictable sets of rules, and problems in following unstructured dialogue that require improvised responses. A further study (Jones, Zahl & Huws, 2001) explored the emotional accounts of individuals with high-functioning ASD, and identified several themes: a sense of alienation and frustration, the experience of depression, and a sense of fear or apprehension. This implies that emotional issues are important for people with ASD, and that some experience predominantly negative emotions. This is in stark contrast to the traditional conceptualisation of a lack of emotional experience in ASD that is generated from laboratory-based studies. Another clinical study also found a pattern that contradicts common perceptions of the disorder. Mazurek (2013) administered questionnaires to those with ASD, which found that the experience of loneliness was common, but that it depended upon the quantity and quality of relationships. Accounts such as these provide a rich and detailed information about the difficulties that people with ASD face in everyday life. For example, difficulty with developing and maintaining friendships may be central to the disorder, but many people on the spectrum experience reward from relationships, and may feel low in mood when they

lack social contact. This type of understanding can provide insights that have clear and practical implications for helping to ameliorate difficulties. However, the nature of information garnered tends to be at the descriptive level, and thus, unlike experimental work, this approach does not typically yield an explanatory account.

Some limited experimental work has used more complex social stimuli, such as real-life-type social scenarios, to attempt to bridge the gap between abstract laboratory tasks and detailed clinical accounts. This has helped to highlight the relative rather than absolute nature of social deficits in ASD, in particular in high-functioning individuals who tend to show proficiency in simple structured tasks, but difficulties with more advanced and naturalistic tasks (Loveland, Pearson, Tunali-Kotoski, Ortegon, & Gibbs, 2001; Channon et al, 2001; Channon, Crawford, Orlowska, Parikh & Thoma, 2014). For instance, when asked to make judgments about subtle social behaviours (e.g. faux-pas or deception), participants with high-functioning ASD may provide correct answers, but not be able to correctly explain the behaviour (Channon, et al., 2001; 2014). Work that explores everyday behaviour in a context where different factors and environments can be systematically manipulated may be valuable in aiding the clinical assessment of ASD, particularly for high-functioning people whose difficulties may be more elusive. Studies such as these could also be used to identify key difficulties for problem–solving in complex scenarios in real life, which could in turn inform target behaviours for social skills training (Channon et al., 2001; 2014; Klin & Volkmar, 2000; Oznoff, 1998).

There is also a lack of work exploring social performance in relation to autistic traits, with most continuum studies focusing on extending classic laboratory findings to the general population. Individuals who are high in autistic traits drawn from the general population are likely to be high-functioning with comparatively subtle social difficulties. Therefore, real-life-type tasks that provide a more fine-grained understanding of social behaviour, as opposed to laboratory-based tasks that test fundamental skills, are most likely to be valuable for studying this population.

3.2 Pro-social behaviour

'Pro-social behaviour' refers to intentional acts designed to aid another in need or distress (Eisenberg & Miller, 1987), such as helping, sharing, donating, co-operating and volunteering (Brief and Motowidlo 1986). It is thought to be important for both society and the individual. Behaving pro-socially has been found to facilitate social bonding, to have a positive impact on social adjustment, self-esteem, academic achievement and outcomes at work, and to contribute towards psychological wellbeing and physical health (Eisenberg, Fabes & Spinard, 1998; Coie, Dodge & Kupersmidt, 1990, Yogev & Ronen, 1982; Puffer, 1987; Osguthorpe & Scruggs, 1986). Within a group, pro-social action is thought to maximise benefits for the 'greater good' (Hoffman, 2001).

Classic psychological studies have examined the role of various situational factors that may influence the propensity to help others. In 1968, Darley & Latane coined the term the 'bystander' effect, which refers to the phenomenon whereby the presence of others inhibits the likelihood a person will help. Darley & Latane (1968) were intrigued by media claims that followed the murder of Kitty Genovese; these stated that many neighbours heard Kitty scream and that some even witnessed the attack, but that they collectively failed to offer any help. Although the facts of this account have since been disputed (e.g. Manning, Levine & Collins, 2007), this event inspired a new line of social psychological research which has now established that large numbers of bystanders decrease the likelihood that any individual will step forward and help another. Further work has found that when participants overheard confederates discussing an emergency situation, the more salient the confederate made the emergency appear, the more likely the participant was to offer their help Bickman (1972). People have also been found to be more likely to help others if they are familiar with the environment they are in (Latané & Darley, 1970a; 1970b), and have greater similarity to (Levine, Prosser, Evans, Reicher, 2005) or familiarity with (Ruthowski, Gruder & Romer, 1983) the individual in need of help. The bystander effect might also be mitigated by the relationship between bystanders; a study found that intervention in a street violence situation was inhibited by the presence of strangers but encouraged by the presence of friends (Levine & Crowther, 2008).

Other work has explored the role of values (also termed 'goals' or 'norms', and conceptualised as the principles or standards that guide our behaviour) in motivating pro-social behaviour. Schwartz (2010) proposed that the propensity of a person to help others is dependent on values that promote positive social relationships, such as 'benevolence' (the belief that one should preserve and enhance the welfare of others). Schwartz (2010) also proposed that individual may be motivated to engage in pro-social behaviour as a means of self-preservation, driven by values that seek to avert negative consequences for the self, such as 'conformity' (the belief that one should restrain actions and impulses that are likely to harm others and violate social norms).

Values or norms have been linked to affect (Shwartz, 2010), and in turn, various authors have emphasised the role of empathy in motivating socially sensitive behaviour (Eisenberg, 2007; Minio-Paluello et al., 2009). Although the distinction has not always been clearly delineated, empathy as a motivating force for pro-social behaviour is postulated to involve both emotional and cognitive mechanisms (Eisenberg & Miller, 1987). Vicariously invoked feelings of distress or discomfort and increased physiological arousal when witnessing someone in need may play a motivating role in pro-social behaviour (Batson, Fultz & Shoenrade, 1987; Coke, Batson, & McDavis, 1978; Eisenberg, 2003; Eisenberg & Miller, 1987). Thus, acting on behalf of another in need is a 'self'-orientated action, which reduces vicarious empathic arousal (Schaller & Cialdini, 1988). On the other hand, acting on mental state apprehension of others' needs on the basis of perspective-taking, is a cognitive 'other'-orientated process.

A range of experimental work has found support for the theoretical assertion that empathy is positively associated with engagement in pro-social behaviour. A number of studies have used the interpersonal reactivity index (IRI), a self-report measure of empathy. High scores on this measure have been positively correlated with charitable giving (Davis, 1983) and volunteering (Carlo, Allen & Burhman, 1999), and greater reported concern for the welfare of others (Batson, 1998). Another approach has been to use cardiovascular or electrodermal indices of empathy. For example, a reduction in heart rate (an index of vicariously induced sadness) and facial indicators of sadness are both associated with an increased willingness to help others (Eisenberg, McCreath & Ahn, 1988; Eisenberg, et al., 1989). Other studies have explored how social rewards might facilitate pro-social behaviour. For instance, expressions of gratitude have been found to increase pro-social behaviour (Grant & Gino, 2010), and the more others cooperate, the more an individual is inclined to do so as well (Frey & Meier, 2004).

Finally, a developmental approach has been applied. Eisenberg, Guthrie, Murphy, Shepard & Cumberland (1999) used a multi-method longitudinal research design to assess the development of pro-social behaviour over ages 4-20, and to establish whether measures of pro-sociality at aged 4 predicted later measures. The findings indicated that there are dispositional factors that emerge early, which consistently predict pro-social behaviour over time. For instance, spontaneous sharing in the classroom predicted subsequent behaviour up to 17 years later. Sze, Gyurak, Goodkind & Levenson (2012) assessed empathy and pro-social behaviour in young, middle-aged and older adults. They identified age-related linear increases in empathic responding, which was associated with greater pro-social behaviour.

3.3 Pro-social functioning in ASD

If empathy is thought to motivate pro-social behaviour then impairment in either cognitive or emotional empathic mechanisms may impinge on pro-social behaviour. Difficulties in identifying others' intentions and emotional states might translate into impaired understanding of others' needs, leading to inappropriate responses. This is consistent with the notion of impaired cognitive empathy resulting in less appropriate pro-social behaviour. Furthermore, correctly understanding another's needs is not sufficient for pro-social behaviour, as one must also be motivated to respond appropriately. It is likely this occurs via reduced emotional distress or improved emotional wellbeing on the part of the other and that this is vicariously experienced by the pro-social actor. Thus, impairments in emotional empathy may also lead to reduced pro-social behaviour. However, very little work has explored pro-social behaviour in clinical populations that are characterised by impairments in empathy to provide support for this assertion. For instance, whilst a lot of work has examined anti-social behaviour in psychopathy, to the authors' knowledge, very little has systematically explored how psychopathic tendencies and a postulated deficit in

emotional empathy may impact upon pro-social behaviour. Some work has explored everyday social problem-solving in neuropsychological populations with damage to the frontal lobes, a region thought to be involved in cognitive empathy abilities. Two studies found that participants with brain lesions experienced difficulty in assessing and generating appropriate solutions (Channon & Crawford, 1999), understanding characters actions, and detecting sarcastic comments or awkward aspects of social situations (Channon & Crawford, 2010). One study also explored social strategy usage in groups high and low in self-reported social skill in response to awkward requests (Channon, Collins, Swain, Young & Fitzpatrick, 2012), reporting that those higher in social skill tended to use more sophisticated strategies that considered all parties' perspectives. By contrast, those lower in social skills used more simplistic strategies, such as simple agreement or refusal to complete the request, and conveyed their responses less politely. The identified difficulties, such as the failure to understand subtle behaviours and to respond appropriately to them, suggests that pro-social behaviour might also be impaired in such groups, but this has not been directly studied.

Turning to ASD, if perspective-taking skills (part of cognitive empathy) are necessary to mediate the adequate identification of others' needs, then it would be expected that this group would also be less successful in behaving pro-socially. Adherence to values, or norms, that guide behaviour might also motivate pro-social action (Schwartz, 2010). Furthermore, a range of social cues have been found to influence pro-social behaviour, including the presence of others, the relationship with the individual in need of help, the degree of gratitude expressed by the recipients of help, and the extent to which others are co-operating (Levine & Crowther, 2008; Levine et al., 2005; Ruthowski et al., 1983; Grant & Gino, 2010; Frey & Meier, 2004). The understanding of social norms and cues in those with ASD is not very well understood, but it is likely that their influence upon social behaviour might be altered (e.g. failure to identify faux-pas; Baron-Cohen, O'Riordan, Jones, Stone, & Plaisted, 1999). Yet little work has explored pro-social behaviour in ASD, which is particularly surprising given the key role that it is thought to play in everyday functioning (Eisenberg et al., 1998; Coie et al., 1990, Yogev & Ronen, 1982; Puffer, 1987; Osguthorpe & Scruggs, 1986).

Some limited work has explored pro-social behaviour in children with ASD, and sought to ameliorate difficulties. Parental reports indicate that children with ASD tend to display significantly less pro-social behaviour compared to typically developing children, e.g. being kind and considerate to others, sharing or offering practical help (Meyer, Mundy, Van Hecke, & Durocher, 2006; Allik et al. 2006). Several small-scale interventions using social stories have attempted to promote pro-social behaviour in children with ASD (Crozier and Tincani, 2007; Leaf et al., 2009). Another line of work has explored charitable giving, finding that participants with ASD donated less and showed reduced preference for organisations

benefiting other people (Lin, Tsai, Rangel & Adolphs, 2012), and were less influenced by the presence of an observer as compared to controls (Izuma, Matsumoto, Camerer, & Adolphs, 2011).

3.4 Brief review of social skill training in ASD

Since difficulties with social interaction are central to ASD a great deal of work has focused on developing interventions for training social skills (Reichow, Steiner & Volkmar, 2013), which have employed a range of methods and approaches (Bishop-Fitzpatrick, Minshew & Eack, 2014). A number of different behavioural approaches have been used to either build up positive social behaviours, such as conversational strategies, use of eye gaze or facial expression, or to reduce inappropriate behaviours, such as perseverative speech or abnormal prosody (Howlin & Yates, 1999). The most dominant model is 'Applied Behavioural Analysis' (ABA) that draws on classical psychological principles of learning, namely that when a behaviour is followed by a reward it is more likely to be repeated, and when it is followed by the withdrawal of a reward or a punishment it is more likely to be discontinued. This is known as positive and negative reinforcement, and is used in ABA training to encourage appropriate behaviours and discourage inappropriate behaviours. This style of intervention is very intensive, generally requiring 20-40 hours of weekly instruction, and is often delivered in a one-to-one format (Virues-Ortega, 2010). Most published reports of such interventions find positive gains for the individuals involved. However, since they typically comprise individually tailored case-study designs, it is hard to evaluate the extent of the benefits and to generalise results from one programme to another, or to a larger population (Bishop-Fitzpatrick et al., 2014). It has been suggested that there is inadequate evidence to support this approach and that large multi-site randomised-controlled-trials (RCTs: this is a type of experimental design which is seen as the 'gold standard' for establishing the efficacy or effectiveness of medical treatments since it includes both a control condition and the random assignment of participants) are needed to improve understanding of ABA's efficacy in ASD (Sprekley & Boyd, 2009).

Another widely-used approach for improving social functioning is social cognition training. This draws on the postulated impairment of cognitive empathy in ASD, and targets component skills thought to underpin observed behavioural deficits, such as difficulties in appreciating others' perspectives (see e.g. Gray 1998) and in recognising their emotional states (see e.g. Golan et al., 2010). In recent years this approach has begun to utilise computer-based training, but there is a lack of evidence indicating that this is more effective than non-computer-based programmes (Bishop-Fitzpatrick et al., 2014). However, there does appear to be more robust evidence for social cognition training approaches as compared to ABA; the published reports tend to compare groups of individuals with ASD with a control group and/or condition, and often utilise a RCT design (Bishop-Fitzpatrick et al., 2014).

Whilst there is a great deal of literature exploring social skill training and psycho-social interventions in ASD, only a few systematic reviews and meta-analyses have been conducted to summarise the collective evidence for the different approaches outlined. Initial attempts indicate that the quality of the evidence is low, and reveals inconsistent methods and findings, with a huge variety of approaches and target outcome measures (Bishop-Fitzpatrick et al., 2014; Reichow et al., 2013; Lounds Taylor et al 2012). Relatively little is known about the efficacy of such programmes or the key ingredients for success (Mueser and Bellack 2007; Rao, Beidel & Murray, 2008; Schreiber 2011). Furthermore, it is difficult to achieve generalisation beyond the specific materials and environments used during the training programme (Howlin & Yates, 1999). Thus, even if targeted skills improve, they do not easily translate to other settings, including the real-world environment.

3.5 Supporting adults with ASD

Most of the programmes developed to ameliorate social difficulties in ASD have concentrated on children and/or younger adolescents, and are thus often unsuited to high-functioning adults (Howlin & Lounds-Taylor, 2015; Channon et al., 2012), and yet they often have considerable difficulties in functioning independently. Although they may have the intellectual capacity to engage with more demanding work and social environments, they may lack the social skills required to navigate these more complex social situations. For instance, at work they might need to learn to adapt to being part of a team. They also may need support to develop understanding of the etiquette and boundaries for specific types of encounters, both formal and informal. The National Autistic Society (NAS) conducted a survey and found that one in three adults with ASD say they have experienced a serious mental health problem that might have been avoided with appropriate support (Rosenblatt, 2008). The Autism Act was passed in 2009, and successfully led to the development of the first ever strategy for adults with ASD in England (Department of Health, 2010). In 2014 a three-year review was completed, when the government asked adults with ASD, their parents or carers, and relevant professionals how well they thought the strategy had been implemented thus far (Department of Health, 2014). This highlighted that local authorities and NHS bodies have not sufficiently addressed the recommended steps for improvement. Thus, there remain gaps in service provision for high-functioning adults with ASD, including access to diagnostic services, employment opportunities and support in the community. Early identification of an individual's difficulties is not only associated with greater improvement (see e.g. Garcia-Villamisar & Hughes, 2007; Garcia, Villamisar & Datilo, 2010; Zwaigenbaum, Bryson, & Garon, 2013), but may also reduce the burden on the NHS, as they are less likely to require on-going support from health services. There is also a need to examine the impact of changes to the diagnostic criteria (see Chapter 1.1), to ensure that vulnerable people do not slip through the gaps (Kaland, 2011).

In addition to health service provision, there is also a need for support in the workplace for people with ASD in view of the social interaction and communication challenges that they face. The National Autistic Society (NAS) emphasises that despite the costs involved in designing, implementing and delivering such systems, in the long run this will in fact save money; increased employment for those with ASD could reduce reliance on welfare benefits and improve overall wellbeing (Howlin & Lounds-Taylor, 2015; NAS, 2004, 2011). Adults with ASD also struggle to get the support they need in the community, for example with social housing or welfare benefits. This is particularly true for high-functioning individuals, whose social impairments may be somewhat masked by their intellectual capabilities; for instance they may be comparatively articulate, but struggle to understand others, or present with challenging behaviour.

One approach that is more appropriate for older adolescents and adults is community-based interventions, such as supported employment schemes or group recreation programmes. There is some limited evidence to suggest that enrolment in such schemes is associated with better executive functioning and quality of life for individuals aged 18 years or more (see e.g. Garcia-Villamisar & Hughes, 2007; Garcia-Villamisar & Datilo, 2010). Bishop-Fitzpatrick et al., (2014) examined 1217 studies of psychosocial interventions in ASD, but found that only 13 met the inclusion criteria for participants of aged 18 or above. Studies have indicated that individuals with ASD have challenges and difficulties as they transition to adulthood (Howlin, Goode, Hutton & Rutter, 2004) that are not being sufficiently met by the available treatments and services. It is therefore important for research to address this gap and develop an evidence base for psychosocial interventions for adults with ASD, in order to inform future research and treatment.

3.6 Chapter summary

This third literature Chapter explored the different approaches to understanding social functioning in ASD. Experimental studies provide a useful method for systematically investigating specific aspects of social cognition and testing the predictions of psychological theories of observed social dysfunction in ASD. However, these studies tend to use laboratory-based tasks with simple or abstract stimuli to test fundamental skills, and thus have relatively low ecological validity. On the other hand, clinical accounts of ASD give rich and detailed information about the everyday manifestations of the disorder, which is important for informing social skill interventions. Although these studies have comparatively high ecological validity, they do not provide predictions or give an explanatory account of the disorder. There is some limited experimental work using more complex and real-life-type stimuli, which attempts to reconcile the disparity between these two approaches. This may prove a useful approach for examining relatively subtle aspects of social behaviour in high-functioning individuals with autistic traits drawn from the general population.

Pro-social behaviour was also introduced, a crucial aspect of social behaviour thought to be motivated by empathic processes. Surprisingly little work has explored pro-social behaviour in disorders where empathic dysfunction is implicated. Some tentative evidence suggests that pro-social behaviour is reduced in individuals with ASD, but this has not been systematically explored, and the extent to which empathic processes versus social learning may underpin this pattern has not been investigated. A more fine-grained understanding of what drives pro-social behaviour and how this might be altered in ASD may be useful for informing social skill interventions in ASD.

Finally, the literature examining different social skill training programmes was briefly reviewed. In conclusion, despite considerable attention to social skills training in children and adolescents with ASD, there is very little work focused on older, and in particular higher-functioning individuals. This group may be most in need of support since they navigate increasingly complex social environments independently, and there appears to be a lack of appropriate services.

As highlighted above, there is a need for work that affords an appreciation of the nuances of social deficits experienced by those with ASD, which can in turn be used to guide the focus of future interventions, particularly for a high-functioning population. This thesis seeks to address gaps in the extensive literature exploring social behaviour in ASD by developing real-life-type scenario-based tasks that provide a nuanced understanding of the difficulties experienced, and extending this to a population of people with higher versus lower levels of autistic traits. This may help to shed light on the extent to which social behaviour is altered in those with high numbers of autistic traits, and may inform future social skill interventions. Pro-social behaviour provides a good candidate for exploration, owing to the fact that: i) it has been relatively under-researched, ii) it is a central part of everyday functioning, and iii) it is theoretically motivated by empathy and thought to be influenced by various contextual cues and social norms.

3.7 Introduction to Experiments A: Chapters 5-7

The first three experimental Chapters of this thesis describe a series of studies that explore pro-social behaviour in people with higher versus lower levels of autistic traits drawn from a student population. Each experiment addresses a different facet of pro-social behaviour to systematically investigate factors that might facilitate or impinge upon successful social functioning, and how this might vary across people with higher and lower levels of autistic traits.

Chapter Five explores the central question: is pro-social behaviour indeed reduced in those with high versus low levels of autistic traits? It also explores whether there are group differences in the extent to which pro-social behaviour is motivated by self-benefit (a sense of personal satisfaction for offering help) versus benefit for others (perceived alleviation of their distress/ satisfaction at having needs met). Chapter Six builds on this first study to explore the extent to which the clarity of social rules that incline us to help

others has an influence on actual helping behaviours, and whether it differentiates those with higher versus lower levels of autistic traits. Chapter Seven examines a related driver of pro-social behaviour, the characters' 'deservingness of help' and how this might differentially influence those with higher versus lower levels of autistic traits.

Chapter 4: Moral judgment and reasoning in ASD

4.1 The role of cognitive and emotional processes in morality

The term 'morality' refers to the principles that constitute 'right or wrong', or discriminate 'good from bad' behaviour (Long & Sedley, 1987). Various psychologists have proposed theories for how morality might develop, and two different approaches have become well established, with some emphasising the role of cognitive processes and others the role of emotional features. Jean Piaget (1932) developed a cognitive theory of moral development, which contended that morality is contingent upon deliberate reasoning. This theory was then expanded by Lawrence Kohlberg (1958), who proposed that people move through a number of discrete and sequential stages of moral development across the life-course, which can be grouped into three levels. Level 1 is known as 'pre-conventional' and is common in children. Individuals at this level judge an action by its direct consequences, and focus on questions such as "How can I avoid punishment?" to guide their behaviour. Level 2 is called 'conventional', which is typical of adolescents, and some adults may never graduate beyond this level. Reasoning at this stage involves judging the morality of actions by comparing them to society's views and expectations. Behaviour is motivated by a 'good boy/ good girl attitude' (e.g. "I want to be liked, and not doing that/ doing this makes people like me") and by social or legal norms, rather than by consideration of the direct consequences for obedience or disobedience. Level 3 is termed 'post-conventional', and it is thought that not all individuals will reach this. It is also known as the 'principled level' and is marked by the realisation that the individual is a separate entity from society whose personal perspective might differ from, or take precedence over, society's views. At this level individuals may disobey rules that are inconsistent with their own principles since they do not view rules as absolute, but are able to apply these flexibly.

Kohlberg (1958) created a case study moral dilemma for assessment of an individual's developmental stage. In this dilemma participants are told that there is a woman dying of a rare type of cancer and that doctors think a recently discovered drug might save her. The drug cost $200 to make, but the pharmacist wants to charge 10 times that ($2000). The woman's husband, Heinz, does not have enough money to buy it and asks everyone he knows if he can borrow money, but only manages to get about half (~$1000). Heinz asks the pharmacist if they could sell it to him cheaper or let him pay the rest later, but the pharmacist says no. Heinz becomes desperate and breaks into the pharmacist's store to steal the drug for his wife. Participants are asked "Should Heinz have done that?". Kohlberg was not concerned about whether the participants argue that Heinz should or should not have stolen the drug, but rather he was interested in the type of justification they used to argue their case, believing that this indicated their stage of moral development. For instance, participants might argue that Heinz should not have stolen the drug

because a) it would lead to him ending up in prison, which would be worse than his wife dying (consistent with reasoning at Level 1), or b) because stealing is against the law and would mean he was a bad person (consistent with reasoning at Level 2), or c) because it is up to the scientist to charge what he wants/ other people might also be in need of the drug (consistent with reasoning at Level 3). On the other hand, participants might argue that Heinz should have stolen the drug using justifications such as, a) he would be happier if his wife survived (consistent with reasoning at Level 1), or b) because he wanted to be a good husband (consistent with reasoning at Level 2), or c) because saving a human life is more important than respecting the property rights of another person (consistent with reasoning at Level 3).

Whilst Piaget and Kohlberg contended that moral behaviour is the result of an elaborate process of reasoning, others have emphasised the role of intuitive and emotional processes in moral development. Prinz (2006) argued that it is 'emotionally taxing to violate social and moral rules', and thus we tend to behave in the 'right' way or engage in 'good' behaviours in order to avoid distress and conserve energy. There is evidence from functional neuroimaging studies to suggest that emotions co-occur with moral judgment. For instance, when participants evaluated moral versus factual sentences, brain regions associated with emotional responses were found to be more active (Moll, de Oliveria-Souz & Eshlinger, 2003). In another study, when participants considered social rule violations, e.g. spitting out food at a dinner party without apology, similar engagement of emotional brain areas was found (Berthoz, Armony, Blair & Dolan). However, this evidence cannot tell us the type of role that emotions play in moral judgment and reasoning. Nonetheless, Prinz (2006) argued that emotions do not simply co-occur with moral judgment, but actively influence it. There is some tentative evidence to suggest that negative emotional states or new negative associations can lead to more harsh moral appraisals (Schnall, Haidt, Clore & Jordan, 2008; Wheatley & Haidt, 2005). Prinz (2006) argued that examining how parents instruct children regarding what is 'right or wrong' might provide more convincing evidence for the necessity of emotions for moral judgment. Analysis of parents' moral instructions to their children suggested that three main techniques are used; i) power assertion, such as threat of punishment which elicits fear, ii) induction, which involves orienting a child to harm they have caused to another person e.g. "Look you made your little brother cry!', and iii) the withdrawal of positive reinforcement, which elicits sadness. Each of the techniques results in the child experiencing negative emotions when they do something 'wrong', which should over time result in the child learning to avert doing the 'wrong' thing in favour of doing the 'right' thing, in order to avoid a negative outcome.

Haidt (2001) tried to reconcile the different arguments put forward regarding the role of cognitive versus emotional processes in morality. He argued that there was a distinction between knowing an action to be morally wrong and being able to reason about why this is the case. For instance, Haidt, Bjorklund & Murphy (2000) found that people could not suspend their sense of 'it just feels wrong', even when their

basis for arguing that an action was 'wrong' was challenged. Participants were shown a scenario in which a brother and sister willingly and knowingly engaged in incest. Participants tended to view this behaviour as morally irreprehensible, and argued that the experience might be emotionally damaging to the siblings. However, when participants were told that the siblings enjoyed the experience and that neither of them experienced distress or any negative repercussions, they failed to provide an alternative reason for why in this case incest was 'wrong' although they maintained this view. Neither an intuitive emotional account, such as that proposed by Prinz (2006), or a deliberate cognitive account, such as that proposed by Kohlberg (1958), can account for the discrepancy between knowing something is 'wrong' but not being able to justify why, as demonstrated in this example. Haidt (2001) thus argued that intuitive emotional responses may lead to moral judgments such as 'incest is wrong', but that when needed they may be required to rationalise this post-hoc via deliberate reasoning, and thus both affective and cognitive processes are central to moral decision-making. Greene, Nystrom, Engell, Darley & Cohen (2004), built upon this idea, arguing that when faced with a moral dilemma, an emotional route provides a fast and instinctive response on the basis of minimising any potential distress. By contrast, a cognitive route involves processes that are slower and more deliberate, such as reasoning about the costs versus benefits to those involved.

In summary, it seems likely that both cognitive and emotional routes are used in concert when making moral judgments and reasoning about these. Unrestricted access to both routes might therefore be necessary for typical development of morality. It is possible that moral judgment on the basis of affective processes alone might lead to black and white judgments about whether it is 'right or wrong'. Such a pattern might theoretically be seen in populations with impairment in cognitive empathy such as ASD or in those with anterior brain damage. On the other hand, utilising a cognitive route alone might enable a more flexible assessment that acknowledges and potentially exploits the shades of grey lying in between 'right' or 'wrong'. This pattern might be seen as consistent with the manipulative, callous and immoral behaviour seen in psychopathy where an impairment in emotional empathy is proposed.

4.2 Exploring different aspects of morality in ASD

4.2.1 Normative versus descriptive morality

Philosophers have long made a distinction between the 'normative' and 'descriptive' aspects of morality (Mill, 1863). 'Normative' morality refers to the absolute rules regarding how people should act, and could be considered to be represent the etiquette underpinning moral behaviour e.g. 'You should not hit someone because you are angry'. On the other hand, descriptive morality describes what people think is 'right or wrong' in a given situation and is concerned with questions such as 'Does this person deserve my help?' or 'Is it okay to violate a moral norm in this situation?'. It is also concerned with how people weigh

up the arguments for and against different possible courses of action (Anscombe, 1958; Kohlberg, 1971; Nichols & Mallon, 2006).

Accordingly, some psychological work on morality has focused on addressing absolute questions regarding people's understanding of what is 'right or wrong', such as in the moral versus conventional paradigm (see Section 4.2.2 for a description of this). This could be said to take a 'normative' approach, uncovering the moral rules underlying human behaviour. Other work has examined how humans use these moral rules flexibly when making decisions in novel or conflicting situations and reasoning about them. This is generally the approach adopted by theoretical frameworks that explore how morality develops across the lifespan, such as in the Heinz dilemma (i.e. should Heinz steal the medicine for his dying wife?) described by Kohlberg (1958) (See Section 4.1 for a description). Such paradigms could be said to provide an insight into the relative or 'descriptive' use of moral norms in relation to various contingencies, rather than uncovering their absolute nature.

Over the past few decades a range of different paradigms have been used to probe different types of questions about morality and how we apply moral rules or norms. Some studies have examined relatively simple questions of the 'normative' type such as "What is permissible and under what circumstances?". Other studies have focused on pinpointing different types of moral judgments such as "Who is deserving", "Who is responsible?", "Who is to blame?" or "What punishment is justified?" which are likely 'descriptive' in nature and require a flexible integration of moral norms with other contextual factors. Some further work has explored the reasoning and justifications behind moral judgments asking questions such as "Why is this the case?" or "How could this have been different". In the next sections, some of the types of paradigms that have been used to explore morality will be reviewed, and the theoretical underpinnings and the strengths versus limitations of each approach will be briefly considered. As discussed, since empathy is thought to relate to moral development, then impairment in cognitive empathy could theoretically result in atypical judgments or difficulty in reasoning about these (Leslie, Mallon & DiCorcia, 2006). Therefore, the next section will also highlight what aspects of moral judgment have been investigated in ASD.

4.2.2 Black and white judgments of 'right and wrong'

A classic approach to examining morality is the use of the moral versus conventional distinction, which is believed to be a key marker of moral development (e.g. Nucci, 1985; Smetana & Braeges, 1990; Turiel, 1983; Wainryb & Turiel, 1994). This paradigm uses scenarios where a character either violates a moral or a conventional rule. Moral rules tend to prohibit actions that might cause others physical or emotional harm (e.g. hitting, stealing or bullying), whereas conventional rules tend to prohibit actions that are socially inappropriate or defy cultural norms (e.g. talking in a library or licking one's plate clean at a

dinner party) but do not cause others distress. This 'normative' account of morality highlights the differences between social versus moral norms.

Typically developing children as young as three years old are thought to be able to distinguish between acts that violate moral versus conventional rules along a number of dimensions. For instance, they can discern that conventional transgressions are less serious than moral transgressions (e.g. Nucci, 1985; Smetana & Braeges, 1990; Turiel, 1983; Wainryb & Turiel, 1994). They also understand that conventional transgressions are contextually contingent (e.g. it might be 'ok' to talk in a library but only after hours or if a special event was taking place there), whereas as moral rules tend to be absolute (e.g. it is never 'ok' to hit someone). They can also comprehend that conventional transgressions are contingent on authority (e.g. when the teacher says you can talk in the library it is 'ok') whereas moral rules are again more absolute (e.g. hitting someone is universally wrong, even if the teacher says it is 'ok') (Nucci & Turiel, 1978; Smetana, 1981; Nucci & Nucci, 1982). When people are asked to justify their distinctions they tend to refer to the potential for harm and abstract concepts for moral norms such as justice, whereas for conventional norms they tend to reference the need to maintain social order and preserve local customs; the ability to distinguish between these types appears to be universal (Turiel, 1983).

The extent to which moral reasoning is intact in ASD has predominantly been explored via the conventional/moral distinction. Children with ASD are assessed for their ability to distinguish between conventional and moral transgressions and to explain why they are wrong, and compared to typically developing children of different ages. The first study to explore this compared high-functioning children with ASD who were divided into two groups: those who passed false-belief tasks (a measure involving predicting another's actions on the basis of their false belief, which is thought to draw upon cognitive empathic skills; see Chapter 2.3.1) and those who failed false-belief tasks. Their performance was compared to two control groups of children: those with mild learning difficulties and typically developing children. Interestingly, all groups appeared to draw successful distinctions between moral and conventional rule breaches (Blair, 1996). This was contrary to predictions, since it was postulated that this ability might draw upon cognitive empathy, and thus children who failed the false-belief task were expected to fail the task. It was suggested that the children with ASD who failed the false belief tasks might thus have relied upon distress cues to make accurate distinctions, which are more salient for moral than for conventional transgressions since the former involve a victim, whereas the latter do not.

In order to test out this hypothesis, Leslie et al., (2006) added a new story to the standard moral versus conventional transgression format, known as the "cry baby". In this scenario two children each have one cookie, but one of the children wants to eat both of the cookies. The other child proceeds to eat her own cookie, whereupon the "cry baby" is distressed and bursts into tears. Participants were asked to judge whether the action that led to the cry baby's distress was "bad" or not. In fact, the child who proceeds to

eat her own cookie has committed no transgression whatsoever, and thus the distress of the cry-baby is a) not justified and b) there is no 'victim' in this scenario, as with conventional but unlike moral transgressions. However, according to Blair (1996), if the participant has impaired or under-developed moral reasoning skills they may simply react to the distress of the cry-baby and judge that the child who ate her own cookie did something bad, using this as a signal to judge the action as a moral transgression. Nevertheless, Leslie et al., (2006) found that both typically developing children, and children with ASD correctly differentiated this scenario and did not view the cry-baby's excessive distress as justified, nor as a signal to judge the act of eating the cookie as morally wrong. Thus, it appears that accurate responding to distress cues cannot alone account for the intact ability of children with ASD to distinguish between moral and conventional transgressions.

At first glance, the finding of an intact ability to make moral/conventional distinctions in ASD may seem surprising. However, people with ASD are thought to have intact physiological correlates of emotional empathy, such as fear responses to aversive stimuli, that may come into play when learning about the rules of what is right versus wrong' in moral development (Prinz, 2006). Thus, it follows that children with ASD may have developed a reasonably good understanding of the difference between breaking a moral versus a conventional rule, and hence they can distinguish between genuine distress and that of a "cry-baby". However, a further study found that whilst adults with ASD made intact judgments of moral/conventional transgressions, they provided more basic justifications than controls, which made reference to social rules rather than to the characters' welfare (Zalla, Barlassina, Buon & Leboyer, 2011). The authors concluded that the difficulty observed by those with ASD might indicate an impaired cognitive appraisal system. They postulated that, whilst people with ASD might be responsive to rule violations, they might fail to integrate this with relevant information about the agent's intentions, and to evaluate the affective impact of the action upon victims.

Other studies have explored 'normative' judgments of 'right and wrong' by asking participants to rate simple stimuli along different dimensions and scales that involve judging the extent to which something is 'good or bad' or 'appropriate or inappropriate' (e.g. Mendez, Anderson & Shapira, 2005; Moll et al., 2002). This approach was recently used to explore the extent to which emotions shape moral acceptability judgments in those with and without ASD. Brewer et al., (2015) compared individuals with and without ASD who were matched for alexithymic symptomatology on judgments of how acceptable it is to make emotion-evoking statements to others e.g. "I bought you a present" (happiness), "I do not want to be friends any more" (sadness), "I could easily hurt you" (fear), "I never wash my hands" (disgust), and "I broke your phone on purpose" (anger) (Please see Chapter 2.3.2 for a discussion of the 'alexithymia hypothesis'). Contrary to the study's predictions, individuals with ASD did not differ from those without ASD in their moral judgments. However, in individuals without ASD alexithymic symptomatology was

associated with atypical moral acceptability judgments, whereby those with more severe symptomatology considered it less acceptable to induce happiness in others, and more acceptable to induce sadness, fear, disgust, and anger. Conversely, in individuals with ASD, alexithymic symptomatology did not predict judgments of moral acceptability. Furthermore, when asked to identify the emotion expressed by the different statements, this was correlated with moral acceptability judgments in those without ASD but not in those with ASD. The authors suggested that individuals without ASD based the moral acceptability of emotion-evoking statements on the emotion likely to be evoked, and that the presence of alexithymic symptomatology impaired this process. Conversely, for individuals with ASD the presence of alexithymic symptomatology did not appear to impair judgments of the acceptability of moral statements. This was argued to suggest that although those with ASD gave intact moral judgments they may have used different strategies to assess the acceptability of the statements, which did not involve drawing on emotional information.

These types of 'normative' approaches have been invaluable in examining the distinction between social and moral norms, the age at which the ability to identify moral norms occurs and the extent to which these are intact in ASD. However, such studies only provide a relatively simple test of morality that is not able to simulate the dynamic nature of the real-world environment. When venturing outside of the boundaries of black and white judgments there are many shades of grey, and thus it is also important to understand how people use their knowledge of 'absolute' moral norms flexibly to navigate complex and novel scenarios. It is possible that exploring 'descriptive' types of morality might provide a more sensitive test of possible impairment in individuals with ASD.

4.2.3 Judgments of causality and intentionality

Other researchers have examined judgments of causality, which could be argued to provide 'descriptive' accounts of morality. Here people are thought to use their knowledge of moral norms to colour their subsequent judgments of intentionality, responsibility, blame and consequence (please see Alicke, Mandel, Hilton, Gerstenberg & Lagnado, 2015 for an overview of research in this field). For instance, Cushman (2008) explored judgments of permissibility and punishment for a scenario in which a character is taking a welding class and either accidentally or intentionally burns her friend's hand. The pattern of results suggested that when people make judgments about whether an action is permissible and consider what punishment they deserve, they take into account whether the agent believed they would cause harm and desired to do so, rather than simply assessing whether harm was caused and the extent of it (Cushman, 2008).

In a further study exploring how intentionality affects moral judgment, participants were shown a scenario in which a chairman of a company ignores advice that an initiative will either harm or help the

environment (Knobe 2003; Zalla & Leboyer, 2011). In both variants the chairman decides to proceed regardless, claiming that he is focused only on profits and does not care about the environment. Almost all participants were likely to judge the initiative as intentional when it harmed the environment, but were likely to judge it to be unintentional when it helped the environment. Thus, here the blameworthy action was judged as more intentional than the praiseworthy action.

Some work has looked at moral reasoning in ASD under circumstances that require differentiating between intentional and unintentional transgressions. For example, Moran and colleagues (2011) presented adults with ASD and a control group with two sets of scenarios. In one set, an agent harmed a victim on purpose (e.g., they invited them to enter an unsafe toxic room), whereas in the other set they did so by accident (e.g., they put poison in their coffee falsely believing that the white powder was sugar). Unlike the control group, who judged accidental harm as less morally wrong than intentional harm, the ASD group did not reliably judge accidental and attempted harm as morally different. In judging accidental harm, ASD participants appeared to show an under-reliance on information about a person's innocent intention, and thus an over-reliance on the action's negative outcome. These findings support the notion of impaired cognitive empathy leading to difficulty in using information about the characters' intentions to inform moral judgment.

Similarly, Buon et al., (2013) compared adults with and without ASD for performance on a series of non-verbal cartoon scenarios depicting a victim being harmed by an agent, either intentionally or by accident. In this case, the stories did not contain any written or verbal information, and thus required participants to infer the agents' intentions on the basis of the images provided. Consistent with the findings by Moran et al. (2011), those with ASD again rated the agent's intention to harm the victim as higher, considered the agent to be more responsible and punished them more severely than control participants in the unintentional condition. However, another study found that people with ASD could correctly differentiate between culpability for intentional and unintentional actions, at least to some extent. Grant et al., (2005), compared children with ASD and controls for performance on a range of stories where a main character did something harmful to someone else (e.g., a boy burned his little brother's hand) either intentionally or by accident. In contrast with the findings of Moran et al., (2011) and Buon et al., (2013), all participants evaluated deliberate harm as more culpable than accidental harm. Participants were also asked to explain their culpability judgments in order to assess whether correct judgments derived from appropriate, adult-like reasoning, or resulted from idiosyncratic reasoning which might not have involved an appreciation of motive. This measure differentiated the groups, since the majority of responses provided by the children with ASD reiterated the story, rather than serving to explain or justify the judgments of culpability made. This again suggested that although the children with ASD were able to

identify that unintentional harms should not be judged as severely as intentional harms, they did not fully understand why this should be the case.

Interestingly, in the task used by Grant et al., (2005) intentionality was manipulated via two variants of the same scenario with the same outcome, whereas in the tasks used by Moran et al., (2011) and Buon et al., (2013) different sets of scenarios were used to compare intentional versus unintentional harms with different outcomes. The contrast between different variants of the same set of scenarios involving the same outcomes but different actions may have made the difference between intentional and unintentional actions more obvious for participants in the Grant et al., (2005) study compared to the tasks used by Moran et al., (2011) and Buon et al., (2013). Thus, the use of task cues regarding intentionality may have resulted in the pattern of correct identification of culpability judgments, but more limited justifications by the ASD group.

A further study by Channon et al., (2011) presented adults with and without ASD with a range of stories depicting complex causal chains of events, all leading to an unpleasant outcome for a victim. As with the task used by Grant et al., (2005), there were variants of the same story which matched in all aspects other than the intentionality of the human action. For instance, a man took an overdose of his medication because his wife gave him too many tablets either deliberately (intentional condition) or accidentally (unintentional condition). This time they also included a further control variant where there was no human error implicated, i.e. the medication had a wrong label (physical condition), and thus it was not the wife's fault. There were no group differences for judgments of causality, but surprisingly the ASD group showed enhanced sensitivity to intentionality when making judgments of blame. Thus, here the ASD group were not only able to differentiate between intentional and unintentional actions, but in comparison to controls blamed characters for their actions more when these were intentional and less when they were unintentional. The contrast might have been made even more explicit in this task than in that used by Grant et al., (2005) via the combination of both matching the scenarios for outcomes and including a physical control condition; the ASD participants might therefore have been over-sensitive to this contrast.

There is likely a complex relationship between the cognitive abilities involved in the understanding of others' intentions and how this relates to moral judgments of causality. Whilst a great deal of work has examined the role of intentionality in moral judgment, and some work has explored this in ASD, researchers have used different types of moral judgments interchangeably. Cushman (2008) conducted a series of careful studies that revealed that different types of moral judgment depend upon different types of mental state and causal information. This has important methodological implications, and suggests that researchers should consider the particular type of moral judgment when designing, interpreting and comparing studies. It is also unclear to what extent intentionality may influence moral judgment in ASD, and this might also reflect differences in the design of the tasks used.

4.2.4 Judgments of deservingness

Judgments of deservingness shape our perceptions of the 'justice' of events and outcomes, and can also influence how we behave towards others (Feather, 1999). For instance, in everyday life we might concede that a student did or did not deserve a top class grade, and this will influence the extent to which we are pleased for them. Or we might believe that the perpetrator of an offence did or did not deserve the harsh penalty they received, and this will influence the extent of sympathy we have for their predicament. When making judgments about the 'deservingness' of an outcome there is evidence to suggest that people consider the intentions behind the behaviour as well as its consequences (Falk, Fehr, Fischbacher, 2008). For instance, did the student who received the top grade work hard or not, and did the perpetrator of the offence do so by accident or on purpose?

Judgments of 'deservingness' potentially provide a means of understanding how we use moral norms to make decisions in everyday life, since they require the use of both our fixed knowledge of what is 'right or wrong' (e.g. normative morality), and consideration of a number of contextual factors (e.g. descriptive morality) (Feather, 2006; Feather & Sherman, 2002). However, there is very little work in this area, and to the author's knowledge, no studies have directly explored judgments of 'deservingness' in people with ASD. One recent study explored 'deservingness' in people with low versus high levels of psychopathic traits drawn from a student population (Vyas, Jameel & Channon, in preparation). When comparing judgments of 'deservingness' for outcomes, this study found that participants with high versus low psychopathic traits were more likely to rate people as deserving, even if the outcomes were undeserved (i.e. they obtained a positive outcome unfairly, or a negative outcome resulted from misfortunate rather than wrong-doing). It was suggested that the high psychopathic trait group might have discriminated less on the grounds of 'deservingness' than the low psychopathic trait group, because they were less sensitive to rule violations (i.e. you shouldn't get something you don't deserve), or a had a lack of regard for a 'fair' or 'just' world. Thus, it seems that the postulated deficit in emotional empathy might lead to atypical perceptions of 'deservingness' or 'fairness'. What might we expect if cognitive rather than emotional empathy is the primary impairment? This task will be explored in relation to people with high versus low levels of autistic traits in Chapter 8 of this thesis, in order to establish how impairment in cognitive versus emotional empathy might impact upon perceptions of 'deservingness'.

4.2.5 Counterfactual thinking

Counterfactual thinking is a psychological concept used to describe the tendency to think about alternative outcomes that are contrary to things that have already occurred. It refers to thoughts such as, "What if…?" or "If only…" that hypothesise about how events in the past might have turned out differently. It is thought to involve both cognitive processes such as executive functioning (e.g. Beck,

Riggs & Gorniak), and emotional processes (e.g. Camille et al., 2004). Counterfactual thinking has also been developmentally linked to false belief understanding (a component of cognitive empathy), since both counterfactual thinking and false belief are based on the capacity to simulate different worlds, either from another's viewpoint, or by going back in time and changing the antecedents (e.g. Drayton, Turley-Ames & Guajardo, 2011). Counterfactual thinking is also thought to be predominantly activated by negative events (Epstude & Roese, 2008), and there is some evidence to suggest that engaging in counterfactual thinking might be related to the experience of self-conscious emotions, such as shame, regret or guilt (Nidenthal, Tangeney & Gavanski, 1994; Mandel & Dhami, 2005; Branscombe et al., 2003). Thus, people engage in counterfactual thinking as a functional process in order to reflect upon and solve problems, and to evoke appropriate behaviours that will avoid future trouble (Epstude & Roese, 2008).

Very little work has directly assessed the role of counterfactual thinking in relation to moral judgment, although it seems likely that the two abilities might rely on similar cognitive and emotional processes. For instance, moral judgment often requires consideration of factors such as intent, responsibility and ascriptions of cause and blame, which is similar to when we hypothesise about how things could have turned out differently (Esptude & Roese, 2008). It has even been postulated that counterfactual thinking about an outcome might illuminate the causal link between an antecedent behaviour and an outcome (Roese, 1997). Thus, counterfactual thinking might provide a useful tool for probing questions about 'descriptive' morality in relation to negative events, such as "What could they have done differently?", or "If only they had…". However, there is mixed evidence regarding the extent to which counterfactual reasoning draws on causal attributions, and little is known about how this relates to scenarios involving cases of immoral action (Zultan, Gerstenberg & Lagnado, 2012; Lagnado, Gersternberg & Zultan, 2013; N'gbala & Branscome, 1995; Wells & Gavanski, 1989).

Children with ASD have been found to demonstrate impaired counterfactual thinking, which is thought to be related to failure to understand others' false beliefs (e.g. Grant et al., 2004). However, very little work has explored counterfactual thinking in adults with ASD, which is pertinent since many high-functioning adults pass false belief tasks despite showing difficulties in other aspects of cognitive empathy. These studies have also typically explored counterfactual thinking in relation to un-doing or avoiding mistakes (e.g. taking of your muddy shoes when entering inside next time OR avoiding breaking a glass) (e.g. Begeer, Meerum Terwogt, Lunenburg, Stegge, 2009), as opposed to rectifying wrong-doing (e.g. having a party in someone's house without their permission and causing damage). Thus, it remains unclear if counterfactual thinking is also impaired in adults with ASD, and whether or how it might be related to atypical moral judgment and reasoning in ASD has not been explored. In Chapters 9 and 10 of this thesis novel tasks will be used to probe whether counterfactual thinking is impaired in people with high versus low AQ traits. These will describe scenarios where there is either social or practical harm caused to an

individual by self or other's wrong-doing. Moral judgments of blame and causality will be explored, as well as examining the emotional consequences of either perpetrating wrong-doing or being on the receiving end, in order to examine the role of cognitive versus emotional factors.

4.2.6 Moral dilemmas

An alternative approach to exploring moral judgment and reasoning is to present participants with moral dilemmas. These describe situations in which there is a moral conflict between two incompatible courses of action and their subsequent outcomes, such as a conflict between the participants' own personal interests and moral norms (e.g. I need to hurt someone to save someone else), a conflict between the participants' different responsibilities (e.g. If I can only help one person, who do I choose?), or a conflict in ensuring the best outcome (e.g. Is helping x better than helping y?). One of the earliest moral dilemmas was Kohlberg's Heinz dilemma, described above (See Section 4.1). Here participants are required to resolve a conflict between the characters' own interests (e.g. save his dying wife) and an absolute moral norm (stealing is wrong). Another example was proposed by Haidt (2001), as discussed above, where participants are presented with scenarios that challenge their perceptions about the absoluteness of moral norms (e.g. is incest always wrong?) (See Section 4.1).

Differences in the way people with ASD reason about morality have been identified using moral dilemmas. One such study investigated spontaneous understanding of moral dilemmas using film clips of four emotionally charged situations in individuals with ASD versus controls (Barnes, Lombardo, Wheelwright & Baron-Cohen, 2009). For instance, a doctor informs his patient that she is a match for organ donation, but cannot donate because she is pregnant. Participants with ASD and a control group were asked to write film-based narratives. Whilst the task did not require individuals to correctly attribute mental states to film characters, there were differences between groups in the spontaneous willingness and frequency with which they included mental states in their retellings. Those with ASD produced narratives that were significantly shorter and included fewer mental state terms, which might reflect a lack of depth of understanding.

A similar approach has been to explore utilitarian judgments, which involve making difficult decisions revolving around making a sacrifice for the benefit of the 'greater good'. The most cited of these paradigms is the now infamous 'Trolley' dilemma (Foot, 1967), where a runaway trolley (or train carriage) is hurtling towards five railway workers. The driver of the trolley has the option to divert it onto another track where only one worker will be killed. In this scenario, participants tend to view choosing to divert the trolley to another track where only one person will be killed as the most justifiable course of action. Research conducted in clinical populations typically characterised by empathic disturbances, such as those with lesions to the anterior regions of the brain or psychopathy, has revealed increased rates of utilitarian

judgment in such dilemmas (Anderson, Barrash, Bechara & Tranel, 2006; Anderson, Bechara, Damasio, Tranel & Damasio, 1999; Blair, 1995; Eslinger, Grattan, Damasio & Damasio, 1992; Gleichgerrcht, Torralva, Roca, Pose & Manes, 2010; Koenigs et al., 2007; Mendez, 2006; Mendez et al., 2005; Miller et al., 2010; Young et al., 2010).

One study has applied the classic utilitarian trolley scenario to people with high functioning ASD versus controls (Gleichgerrcht, Torralva, Rattazzi, Marenco, Roca, & Manes, 2013). Participants' performance was compared for variants of the trolley dilemma that varied in respect to the personal involvement (e.g. low = pull a lever to divert the train; high = push a man of a bridge to block the train), but resulted in the same outcome (e.g. saving five and killing one). Whilst people with ASD did not differ from controls in their utilitarian choices when the degree of personal involvement was low (e.g. pull the lever, saving five and killing one), they more frequently choose the utilitarian course of action when the personal involvement was high (e.g. to push the man, saving five and killing one). This pattern of enhanced utilitarian action in ASD was associated with a decreased ability to infer other people's thoughts and to understand their intentions (e.g. cognitive empathy), but also with an intact understanding of the 'appropriateness' of these actions and similar scores on a test of moral knowledge. This indicates a dissociation between moral knowledge and moral judgment in those with ASD.

Moral dilemmas challenge people's perceptions of what they think is 'right or wrong', and therefore provide a means with which to explore the more flexible use of relative moral rules (or 'descriptive morality'). Thus, the use of dilemmas has become very popular, since it allows for the systematic manipulation of different parameters and examination of how these shape moral judgment and reasoning (Hauser, Cushman, Young, Kang-Xing Jin & Mikhail, 2007). For instance, different permutations of the trolley scenario produce different answers. Exploring the results of a number of the trolley variants has been useful in a) unpicking the types of factors that influence people when making moral judgments, and b) further exploring the contribution of cognitive and emotional processes to moral judgment (e.g. Cushman, 2008; Greene et al., 2004; Greene, Somerville, Nystrom, Darley & Cohen, 2001; Moll, Oliveria-Souza & Zahn, 2008; Moll, Oliveria-Souza, Zahn & Grafman, 2008). However, this scenario provides a very circumscribed understanding of utilitarian decisions, and it is not clear to what extent these findings might be generalised to other situations. There is also a lack of analysis or comparison of the relevant parameters involved, and research exploring moral dilemmas and utilitarian judgment has proceeded in a piecemeal fashion, making comparability across studies very difficult (Christenesen & Gomilia 2012). Moral dilemmas also tend to represent extreme scenarios of physical or emotional harm such as killing, stealing and incest. Thus, whilst they provide dynamic materials that can explore reasoning as well as knowledge, in contrast they tend to depict extreme scenarios that might make their relevance to everyday dilemmas and problem-solving more limited.

4.3 Chapter summary

A number of theories of moral development have been proposed, with some emphasising cognitive factors and others stressing the importance of emotional factors. It seems likely that both processes are used in concert, with different types of moral judgment relying on different routes. A number of different approaches to moral judgment and reasoning have been employed. These paradigms typically have limitations with regard to understanding moral judgment in an everyday context. Many studies focus on simplistic 'normative' questions, such as the distinction between moral and conventional transgressions or distinguishing moral from non-moral statements. These types of studies are useful in informing us about what is viewed as wrong from a societal perspective and the stage at which various aspects of moral knowledge develops. However, since the ability to distinguish between moral and conventional norms develops at around age three, it provides a rather simplistic test of moral judgment. These types of studies cannot tell us about how we use moral norms and rules to navigate everyday issues and more advanced forms of moral judgment.

Some studies have used more graded types of problems to explore 'descriptive' questions where there is no black and white 'right or wrong' answer. Some of these have explored more graded or sophisticated types of judgments, such as the role of intentionality in establishing who is to blame. Moral dilemmas provide more dynamic stimuli where there is some kind of moral conflict. Whilst these studies provide a more nuanced account of moral judgment, they have classically used extreme scenarios which are not representative of everyday life, such as utilitarian decisions involving whether or not to kill one person to save many more. Some more recent work has used more everyday situations, such as judging the deservingness of receiving an exam grade, but this body of work is more limited in scope. Most experiments have focused exclusively on judgment, probing questions about what is right, or how someone should act in a certain situation. Less attention as been paid to examining reasoning about or providing justifications for moral decisions, which might give deeper insights into the understanding of morality than examining judgment alone. Crucially, there has not been a systematic review of the different types of questions or paradigms used, and it remains unclear how different types of moral judgments are navigated and what factors influence them.

Morality has been surprisingly under-researched in ASD, and it is currently unclear to what extent different aspects of moral judgment and reasoning are intact or impaired. There remain inconsistencies and gaps in the existing literature. Discrepant findings might reflect the relatively simplistic or extreme scenarios used, which appear to represent easier 'tests' of moral judgment that are less sensitive to subtle difficulties that may be present in 'non-black-and-white' everyday dilemmas. The discrepancies might also reflect the mixture of measures used, since different types of moral judgment and task demands are likely to draw on different processes, with some drawing more heavily on emotional versus cognitive resources.

Despite this, the current pattern of findings suggests that people with ASD might 'know' what is 'right or wrong' but face difficulties in applying this knowledge.

4.4 Introduction to Experiments B: Chapters 8-11

'Experiments A' of this thesis presents three experimental Chapters that describe a series of studies exploring questions about when or whether to help others. These all broadly focus on "What would you do…?" -type questions, which are combined with a number of manipulations that are thought to be relevant to distinguishing those with high versus low levels of autistic traits. 'Experiments B' of this thesis presents four further experimental Chapters that shift in focus from examining different aspects of participants' own hypothetical behaviour to exploring questions of morality. These focus on more abstract questions probing judgments and reasoning, such as "Is this person deserving of/ responsible for this outcome?" and "What is the right thing to do?" or "How could this situation have been avoided?", in people with high versus low levels of autistic traits.

Chapter 8 focuses on reasoning about positive or negative outcomes for others', and follows on from the final study in Experiments A, continuing the exploration of whether the notion of 'deservingness' is intact in those with high levels of autistic traits. Chapter 9 uses counterfactual thinking to explore judgments of cause and blame in relation to self- versus other-intentions and actions. Chapter 10 extends the exploration of counterfactual thinking in relation to emotional judgments of regret and guilt. Chapter 11 examines utilitarian decision-making in a novel variant of the classic trolley task, exploring both physical and social/emotional harm.

PART 2: EXPERIMENTS A

CHAPTERS 5-7

PRO-SOCIAL FUNCTIONING

Chapter 5: Self- versus other-satisfaction for pro-social behaviour

Chapter 6: Examining the influence of the social expectation to help others on pro-social behaviour

Chapter 7: Examining the influence of characters' justifications for help on pro-social behaviour

A note on experimental methods

All the experiments described in this thesis consisted of a screening phase in which a large sample of participants completed a questionnaire examining autistic traits, and a subsequent experimental phase comparing sub-samples of those with high versus low levels of autistic traits. This approach was used across all experiments; for succinctness, this is described here but is not repeated in subsequent experimental Chapters.

Screening Phase

The Autism-Spectrum Quotient

In the screening phase a large number of university students were asked to complete the AQ (Baron-Cohen et al., 2001). This is a brief, self-administered questionnaire that measures personality traits associated with the autistic spectrum in adults of normal intelligence. It consists of 50 statements rated on a four point Likert scale (1 = definitely agree; 4 = definitely disagree) covering five different aspects of autistic symptomatology (APA, 1994; Rutter, 1978; Wing & Gould, 1979); social skill, attention switching, attention to detail, communication and imagination. Total AQ trait scores range from a minimum of 0 to a maximum of 50. Approximately half the items are worded to produce a 'disagree' response, and half an 'agree' response, in a high-trait individual. It has been found to have good internal consistency and construct validity (Cronbach's alpha coefficients are modest to high for all five domains; range .63 - .77), strong test–retest reliability (r = 7, p = .002), and robust self versus parent report reliability (mean difference score = 2.8 points, standard deviation = 0.6) (Baron-Cohen et al., 2001).

Selection of high and low AQ participants

The method for selecting high and low AQ participants was also consistent across all experiments. All participants in all experiments provided informed consent before completing the AQ. As an incentive they were entered into a prize draw and informed that they might be invited to take part in the second phase of the study, for which they would be paid. Once total AQ scores were calculated for the whole sample, participants from the top and bottom deciles of the total sample were invited to come and take part in the experimental phase of the relevant study for individual testing on the various tasks. Since AQ traits are more common in males than in females (Baron-Cohen et al., 2001), matched numbers of males and females were recruited to control for any effects of gender. Participants within the highest-scoring and lowest-scoring 10% of males and the highest-scoring and lowest-scoring 10% of females were contacted via email or telephone and invited to take part in the second stage. These formed the high AQ and low AQ

participant groups for the experimental phase of the studies. Sample sizes of 20+ for each group were selected as adequate for alpha set at .05 and power at .80 (Cohen, 1992).

Experimental Phase

Participants and procedure

Certain aspects of the experimental phase were also consistent across all experiments. All participants described in the experimental phase of each experiment were selected from the screening sample according to the method described above. They were invited to come in and take part in the second stage of the study at a time convenient to them, for which they would be paid a fixed sum; their right to withdraw at any stage was made clear. Before taking part in the experiment they were required to provide informed consent and to complete a brief health-screening questionnaire that asked about any psychiatric or neurological conditions that might affect responses; in practice no exclusions were required.

Construction and nature of scenario-based tasks

All of the experiments described in this thesis involve the use of novel scenario-based tasks, which comprise a range of scenarios that reflect common everyday situations and attempt to approximate the demands and circumstances of real-life, as much as possible. Scenario-based measures have been shown by previous research to have higher ecological validity compared to more abstract, traditional laboratory measures, but to allow for the exploration of participants' thought and reasoning processes in more a rigorous way than observational experiments (see e.g. Anderson & Anderson, 1951; Hughes & Huby, 2001). Thus, this approach was adopted in order to allow for the detailed assessment of social performance, whilst adhering to the fundamental principles of experimental control.

The information contained within the scenarios was always clearly defined and standardised across scenarios, ensuring that a naturalistic range of situations was explored, but that the various factors inherent in the scenarios were well-balanced or deliberately controlled. Firstly, each task included scenarios with either similar numbers of male or female characters, or their gender was deliberately not specified. Secondly, the nature of the character's relationship to the participant was either deliberately controlled within each task, or was well-balanced across scenarios (e.g. a range of characters with similar numbers of friends, siblings, colleagues or housemates). Similarly, the type of setting was either deliberately controlled within each task, or was well-balanced across scenarios (e.g. a range of scenario settings with similar numbers set at home, work, or in public). Finally, any other task-specific factor that might have varied between scenarios was either deliberately controlled, or carefully counterbalanced within each task (e.g. the clarity of the social rule in Chapter 6, or the nature of harm in Chapter 11).

Experimental pairings

Please note that with the exception of Chapter 6 all experiments were carried out in counterbalanced pairs. The pairs were as follows: Chapters 5 and 9; Chapters 7 and 8; Chapters 10 and 11; Chapter 6 was conducted independently. Thus, the participants described in Chapter 5 are the same as for Chapter 9, and so on for each pairing.

Ethics

All experiments presented were granted ethical approval from the UCL Research Ethics Committee.

A note on statistical analysis

Assumptions of parametric tests

Variables across all experimental studies were initially examined for skewness and outliers using the methods described by Tabachnick and Fidell (1983). When variables approximated a normal distribution, parametric analyses were performed, since parametric tests are more robust than non-parametric ones (e.g., Howell, 1997), and allow testing for interactions between multiple variables. When assumptions for normality were not met, transformations were performed to meet assumptions of normality where possible, and when variables could not be transformed to normality, non-parametric tests were carried out; where this applies it is indicated in each experimental Chapter.

Significance levels

A significance level of p=.05 was adopted across all experimental Chapters. A stricter significance level of p=0.05/number of comparisons was used to control for multiple comparisons; where this has been applied, it is indicated in the relevant Chapter.

Chapter 5: Self- versus other- satisfaction for pro-social behaviour

5.1 INTRODUCTION

As previously discussed, pro-social behaviour is a key component of everyday social functioning (e.g. Eisenberg et al., 1998; Coie et al., 1990; Yogev & Ronen, 1982; Puffer, 1987; Osguthorpe & Scruggs, 1986), which is considered to be primarily driven by empathic processes (Eisenberg, 2007). Although very little work has explored pro-social behaviour in connection to disorders of empathy, it is postulated that impairment in either route might lead to reduced helping behaviours, or inappropriate responses to others' requests for help. Difficulty identifying others' mental and emotional states (i.e. impaired cognitive empathy) might lead to a lack of *understanding* of others' needs. This can be conceptualised as a deficit in the other-oriented empathic route that drives pro-social behaviour. Acting to help another may also be guided by a desire to alleviate vicariously induced feelings of distress. Theories of helping behaviour have debated whether any act can ever be truly considered altruistic, since the actor is likely to receive an intrinsic reward (i.e. a sense of personal satisfaction) that may motivate helping behaviours (e.g. Schaller & Cialdini, 1988). Thus, impaired ability to *resonate* with another's emotional state (i.e. emotional empathy) might lead to a lack of self-oriented motivation to behave pro-socially.

Since ASD is thought to be characterised by impaired cognitive empathy, with mixed evidence regarding whether emotional empathic abilities are intact, it seems likely that those high in autistic traits might show reduced sensitivity to others' needs, and thus display fewer pro-social behaviours. In the present experiment, this was investigated using a novel scenario-based task, 'Above and Beyond' (please note this task has now been described in this publication: Jameel, Vyas, Bellesi, Roberts & Channon, 2014). Participants were presented with a series of everyday situations involving a character in need of their help. In each of the scenarios the high (high AQ group) and low (low AQ group) were first asked to respond freely regarding how they would help the character in need. They were then presented with three possible courses of action, representing low, medium and high pro-social behaviours, and asked to choose which they would be most likely to follow. In order to explore how self- versus other- orientated empathic routes might be implicated, participants were also asked to rate how much satisfaction either they (self-) or the character (other-) would derive from each of the different courses of action.

5.1.1 Hypotheses

5.1.1.1 Pro-social behaviour: generation versus selection

There were two methods of assessing participant's pro-social behaviour, which intended to explore spontaneous versus cued responding in relation to high AQ versus low AQ scores. This allowed for examination of pro-social behaviour under two different conditions; the ability to generate one's own response is relatively demanding, whereas selecting from alternatives is comparatively low-demand. In two previous studies examining real-life-type problem solving using social scenario-based tasks, individuals with ASD displayed difficulty in generating problem solutions, but not in judging alternatives (Channon et al. 2001, 2014). On this basis, it was hypothesised that the high AQ group may have been able to identify which was the best option when presented with alternatives, but not to produce it spontaneously.

5.1.1.2 Self- versus other-satisfaction ratings

Participants were also asked to give satisfaction ratings for each possible course of action from both their own perspective (self) and that of the character (other). The self- versus other-satisfaction ratings were expected to reveal potential difficulties in taking the characters' perspectives, whereby the high AQ group would give lower estimates than the low AQ group of the characters' satisfaction when they performed actions of high pro-social value, and conversely, would give higher estimates of the characters' satisfaction when they performed actions of low pro-social value. In addition, it was predicted that the high AQ group may experience less personal satisfaction for going 'above and beyond' (performing actions of high pro-social value) than the low AQ group.

5.2 METHODS

5.2.1 Screening phase

5.2.1.1 Screening participants and procedure

An opportunistic sample of 573 full-time university students (43% male) who were fluent in English and aged 18 or over (mean age 20 years old) was recruited for the screening phase of the study. All participants completed the AQ (Baron-Cohen et al., 2001) and total scores were calculated for the whole sample. Participants within the highest-scoring and lowest-scoring 10% of males and females were contacted and invited to take part in the second stage. For more details on the AQ or other aspects of the Screening Phase procedure please see the note on page 65-66.

5.2.2 Experimental phase

5.2.2.1 Design

There was one between-participants factor of AQ group (high versus and low AQ participants), and one within-group factor (self- versus other- satisfaction ratings).

5.2.2.2 Experimental participants and procedure

Of those contacted from the screening phase, 27 (14 female, 13 male) individuals from the upper range and 24 (12 female, 12 male) individuals from the lower range agreed to take part in the experimental phase of the study, forming two groups of high AQ and low AQ participants. AQ scores ranged from 25 to 43 in the high AQ group (25–43 for male participants, and 26–37 for female participants), and 3–10 in the low AQ group (4–10 for male participants, and 3–9 for female participants). A *t* test confirmed that AQ scores differed significantly between groups, $t(49) = 24.42$, $p < .0001$; mean AQ scores were 30.70 (SD = 4.33), and 6.83 (SD = 2.16) for the high and low AQ groups respectively. The groups did not differ significantly in age, $t(49) = .495$, $p = .623$; mean age was 20.37 (2.71) and 20.79 (3.36) for the high and low groups respectively. All participants were tested individually, and provided written informed consent before completing all measures from the 'Above and Beyond' task.

5.2.3 The 'Above and Beyond' Task

This task was designed to assess individuals' propensity to behave pro-socially in everyday situations, and the lengths to which individuals are willing to go to help others. A range of scenarios was devised and piloted with healthy volunteers in order to refine the items and develop the scoring system. The final set consisted of ten brief scenarios each describing social situations involving a main character in need of help, where only the participant was potentially available to help them. Each scenario required a difficult social judgment with respect to balancing the needs of the character against their own interests. The character was male in half the scenarios, and female in the other half, and the type of relationship and social context varied across scenarios to reflect a natural range of situations. To control for order effects, two different scenario orders were created and counterbalanced within each group.

Participants were shown a set of task instructions and asked to answer as quickly, truthfully and accurately as possible. They were taken through an example scenario with the researcher before completing the ten experimental items. Scenarios and questions were presented in separate paper booklets such that relevant scenarios remained on display throughout task performance in order to minimise any memory demands. Each scenario was followed by four questions (see Figure 2). Participants were first asked to generate responses for what they would do in the situation, and were then asked which course of action they would be most likely to follow when presented with a choice of three. These were designed to

represent low, medium and high pro-social actions, requiring increasing effort on the participant's part. Since these two measures were intended to examine how a reduction in task demands (generation vs. selection of pro-social responses) may relate to pro-social behaviour, the presentation of the options was systematic in order to make the pro-social values as salient as possible (low pro-social actions first and high pro-social actions last). Participants were also required to rate satisfaction with each of the three alternative actions from their own and the characters' perspective.

Figure 2 - Example scenario from the 'Above and Beyond' task

Example Scenario:
"You are walking down an empty side street when a man trips over in front of you and falls down heavily on the pavement. You are in a rush to get to work on time for a meeting."

Example Questions:
Generation of Pro-Social Response: What would you do in this situation?

Selection of Pro-Social Action: Which of the following would you most likely to do?
(Low): Carry on walking.
(Medium): Help him up and carry on walking.
(High): Help him up and offer to take him to sit down on a nearby bench.

Self-perspective Satisfaction:
On a scale of 1-10 how 'pleased' would you feel if you chose to do the following?
1= not at all pleased, 10 = very pleased.
[rate low, medium and high actions]

Other-perspective Satisfaction:
On a scale of 1-10 how 'pleased' do you think he would feel if you chose to do the following?
1 = not at all pleased, 10 = very pleased.
[rate low, medium and high actions]

5.2.3.1 Scoring

5.2.3.1.1 Generation of pro-social responses

Scoring of verbal responses for each scenario was in accordance with their pro-social value: one point for low, two points for medium, and three points for high pro-social value. Low pro-social actions were those involving little effort on the participant's part, since they tended to prioritise their own needs over others. Medium pro-social actions involved making significant effort to help another, but within limits as to the personal cost involved. High pro-social actions went 'above and beyond' in helping others to their own disadvantage. In the example shown above, where a man has fallen over, a response classified as low pro-social effort involved making little or no attempt to stop and help the man (e.g. "Continue rushing to

work on time, assume someone else will help him."). A response classified as medium pro-social effort described stopping to help the man up and some attempt to offer further assistance, but made it clear that the participant was not prepared to be late for their meeting (e.g. "Check if he is okay and if I can call for him first of all. Try to keep in mind that I am in a rush."). A response classified as high pro-social effort indicated that the participant was prepared to be late for their meeting if required (e.g. "Stop and help the man up, see if he needs medical attention. My meeting can't be that important—probably phone to say I might be a bit late.").

The responses were classified by a rater who was not blind to group membership, and by a second, blind independent rater. There was an inter-rater agreement rate of 94.23%; all disagreements were resolved by a third party adjudicator (also blind to group membership). Participant scores were then summed across all 10 scenarios (range 10–30).

5.2.3.1.2 Selection of pro-social actions

Participants were awarded a score of 1 for choosing the lowest pro-social actions, 2 for choosing medium pro-social actions and 3 for choosing the highest pro-social actions. Participant scores were then summed across all ten scenarios (range 10-30).

5.2.3.1.3 Self- versus other- satisfaction ratings

For each scenario, participants gave satisfaction ratings from both their own (self) and the characters' (other) perspectives, on a scale of 1 and 10, where higher scores indicated greater satisfaction. Scores were summed across all ten scenarios (range 10-100) creating 6 total scores; satisfaction for low, medium and high pro-social options for self-perspective; and satisfaction for low, medium and high pro-social options for other-perspective. An overall, self-perspective-taking satisfaction difference score was then calculated (high pro-social actions satisfaction score minus low pro-social actions satisfaction score); an overall other-perspective-taking satisfaction difference score was calculated on the same basis.

5.3 RESULTS

5.3.1 Data analysis

Means and standard deviations (SD) for each of the measures below are presented in Table 1. A significance level of .05 was adopted, with a stricter level (.05/3 = .017) for post hoc tests to control for multiple comparisons. The assumptions of normality were met and thus parametric analyses were performed.

5.3.2 The 'Above and Beyond' task

5.3.2.1 Generation of pro-social responses

A t-test was used to compare the high and low AQ groups on the total score for generation of pro-social responses. The high AQ group scored significantly lower than the low AQ group, $t(49) = 5.332$, $p < .0001$, suggesting that their responses were less often classified as pro-social. Post-hoc *t* tests were conducted to examine the pattern underlying this overall difference in score. The groups did not differ in their generation of medium pro-social responses, $t(49) = 2.081$, $p = .043$, but did significantly differ in their generation of low and high pro-social responses, whereby the high AQ group generated fewer high pro-social responses, $t(49) = 2.64$, $p = .013$, and more low pro-social responses, $t(49) = 3.97$, $p < .0001$.

5.3.2.2 Selection of pro-social actions

The high and low AQ groups were compared on total scores for selection of pro-social actions. The high AQ group was found to behave significantly less pro-socially overall than the low AQ group, $t(49)=4.392$, $p < .0001$, suggesting that they chose fewer high pro- social actions and more low pro-social actions. Further *t* tests were carried out to examine choices of low, medium and high pro-social actions separately, summed across scenarios. Using a strict significance level of .017, the groups did not differ on the medium pro-social actions, $t(49) = 3.49$, $p = .037$; the high AQ group was found to choose significantly more low pro-social actions, $t(49) = 3.49$, $p < .0001$, and significantly fewer high pro-social actions, $t(49) = 4.07$, $p < .0001$.

5.3.2.3 Self- versus other-satisfaction ratings

A repeated measures ANOVA was conducted, using the overall high-low satisfaction difference scores to compare groups for self-perspectives (participant) versus other-perspectives (character). There was one between-group factor (high vs. low AQ), and one within-group factor (self-satisfaction difference score vs. other-satisfaction difference score). There were significant main effects of perspective, $F(1,49) = 84.82$, $p < .0001$, and group, $F(1,49) = 17.08$, $p < .0001$, and a significant perspective by group interaction $F(1,49) = 7.43$, $p = .009$.

Post-hoc *t* tests were conducted to compare the two groups for overall self- and other-satisfaction difference scores separately, using a strict significance level ($p = .017$). The groups did not differ significantly for other-satisfaction difference scores, $t(49) = 1.86$, $p = .070$, but did show a significant difference for self- satisfaction difference scores, $t(49) = 3.94$, $p < .0001$. Comparison of mean scores revealed that for self-satisfaction scores the high AQ group differentiated very little between low and high courses of action; they also tended to rate satisfaction for high pro-social actions lower than the low AQ

group, $t(49) = 2.98$, $p = .005$, and rated satisfaction for low pro-social actions higher than the low AQ group, $t(49) = 2.87$, $p = .006$.

Table 1: Mean percentage scores and standard deviations for all measures for the 'Above and Beyond' task.

	Low AQ group (N = 24) M (SD)		**High AQ group** (N = 27) M (SD)		**Significance** (p = .05)	**Effect Size** (d)
Generation of pro-social response (%)						
Total quality	81.94	(6.59)	67.40	(11.82)	.0001	1.52
Low pro-social	10.42	(9.08)	28.15	(20.20)	.0001	1.13
Medium pro-social	30.83	(14.42)	40.37	(17.86)	.043	0.53
High pro-social	56.36	(13.64)	41.00	(18.53)	.013	0.95
Selection of pro-social action (%)						
Total score	83.89	(7.39)	72.22	(10.97)	.0001	1.25
Low action	6.25	(6.47)	18.51	(16.10)	.0001	1.82
Medium action	36.25	(15.82)	46.29	(17.35)	.037	0.60
High action	57.50	(17.99)	35.19	(20.82)	.0001	1.15
Self- perspective ratings (%)						
High-Low satisfaction difference	27.46	(14.54)	7.74	(20.33)	.0001	1.12
Low action	45.46	(10.44)	54.22	(11.26)	.006	0.81
Medium action	70.67	(9.52)	66.52	(11.19	___	___
High action	72.92	(7.30)	61.96	(16.64)	.005	0.85
Other- perspective ratings (%)						
High-Low satisfaction difference	44.91	(8.61)	39.89	(10.50)	.070	0.52
Low action	40.33	(7.38)	45.04	(9.40)	___	___
Medium action	70.21	(7.30)	70.14	(6.72)	___	___
High action	85.25	(6.24)	84.93	(7.18)	___	___

5.4 DISCUSSION

5.4.1 Summary of findings

The present study examined how high and low levels of autistic traits translate into everyday pro-social behaviour. It employed a novel task describing real-life-type scenarios in which a main character required help, to assess the generation and selection of pro-social responses. The pattern of results supported the prediction that the high AQ group would behave less pro-socially overall, since high AQ participants generated verbal responses that were significantly less pro-social in quality than those of their low AQ counterparts. It was also hypothesised that any group differences might be ameliorated when the more demanding task of generating responses was removed, and participants were simply required to select responses from a choice of three possible courses of action. However, this was not supported since high AQ participants were significantly less pro-social both in their spontaneous generation of responses and in their selection of actions from alternatives. In addition, participants rated satisfaction from their own perspective and from those of the main characters. Contrary to predictions, the high AQ group did not differ from the low AQ group in ratings of the characters' satisfaction. However, they did differ in self-satisfaction ratings, where they tended to express greater satisfaction for performing low pro-social actions and lesser satisfaction for performing high pro-social actions.

With respect to the generation of pro-social content, the high AQ group's verbal responses contained fewer classified as high pro-social, and more classified as low pro-social relative to the low AQ group, with similar numbers of responses that were of medium pro-social value. For instance, in one of the scenarios participants were asked to decide what they would do if a friend rang at an inconvenient time, upset that her partner had just broken up with her. The high AQ group was less successful at generating responses of high pro-social value that went 'above and beyond', failing to prioritise others' needs over consideration of their own (e.g. "Talk to her for as long as she wanted and offer to go round. I would try and make sure she is okay—me having a quiet night in isn't as important."). The high AQ group also made more verbal responses that were low in pro-social value and gave little help to the main character (e.g. "Try to end the phone call as soon as possible, or wait for the answer machine to get it."). The groups did not differ for their usage of medium pro-social responses, which although effective in responding to the characters' needs did not involve not incurring significant personal costs (e.g. "Calm her down, help her and make her feel better over the phone."). The finding that the high AQ group tended to respond to the characters' needs in a less pro-social and possibly inappropriate manner is broadly consistent with previous work indicating that those lower in self-reported social skill used more simplistic strategies to respond to others' requests and conveyed their responses less politely (Channon et al., 2012). It also extends previous work that has identified difficulties with pro-social behaviour in children with ASD (Allik et al., 2006;

Meyer et al., 2006), and less sensitivity to others' needs when deciding whether to donate to charity in adults with ASD (Lin et al., 2012).

5.4.2 Non-social accounts of the findings

A number of different psychological accounts of ASD might be pertinent to the present findings. Namely, performance deficits in ASD have been attributed to impaired cognitive empathy with preserved emotional empathy. Whilst this is the most dominant theory of social and emotional functioning in ASD, the possible roles of WCC and executive dysfunction will first be briefly considered, although these theories are more commonly used to explain the non-social symptoms of ASD. WCC theory stipulates that people with ASD may pay enhanced attention to details at the expense of integrating information (e.g. Happé, 1996; 1999). On social scenario-based measures like the 'Above and Beyond' task this may manifest as a lack of appreciation of the wider social context. Thus, with respect to the present study, the high AQ participants may have paid increased attention to the details of the scenarios, particularly to those aspects that were relevant to their own perspectives, resulting in less pro-social behaviour overall. For instance, the high AQ participants may have focused on details like the fact that they would be late for work if they stopped to help the man, or that their friend happened to ring at an inconvenient time. The high AQ participants may also have failed to consider the characters' perspectives and weigh up the relative merits of the aspects of the scenarios that related to the characters' versus their own perspective. For instance, it is possible that they did not understand that it would be acceptable to be late for work under exceptional circumstances. Nor might they have considered that the character who was upset because her boyfriend had broken up with her, might be in great need of comfort, whereas in comparison they would only be mildly inconvenienced by engaging in a phone conversation.

Turning now to the theory of executive dysfunction, an impaired ability to evaluate relevant aspects of the social scenarios, to generate and plan appropriate responses, and to appreciate the social consequences of these, could account for the reduced pro-social behaviour observed in the high AQ group. However, this explanation is not satisfactory, since the high AQ group displayed reduced pro-social behaviour on both the more demanding measure of generating their own response, and the less demanding measure requiring judgments of possible alternative responses. If executive dysfunction was an exhaustive account, then once the task demands were reduced and participants provided with cues (via systematic presentation of the low, then medium, and then high alternative courses of pro-social action), differences between the groups would not be expected. Executive difficulties could, however, also manifest as reduced inhibition or capacity to control impulse reactions resulting in less pro-social responses on the 'Above and Beyond' task. Individuals with high AQ might have failed to use appropriate strategies to search knowledge stores for relevant experience and to evaluate accurately possible future outcomes of different courses of

action (Channon et al., 2001). Thus, the high AQ group may have acted impulsively in their own interests, at the expense of considering the possible benefits of pro-social behaviour for the characters in the short-term and for themselves in the long-term. This explanation is consistent with some evidence indicating that inability to consider others' perspectives may result from an executive failure to inhibit one's own perspective (Samson, Apperly, Braithwaite, Andrews & Bodley Scott, 2010; Samson, Apperly, Kathirgamanathan & Humphreys, 2005).

5.4.3 Impaired cognitive empathy?

Impaired cognitive empathy (Baron-Cohen, 1995) is perhaps the most relevant psychological account for interpreting the present findings, since it focuses directly on the social deficits associated with ASD. Failure to appreciate the characters' needs and feelings in the current scenarios might have operated to reduce motivation to act pro-socially, leading in turn to fewer pro- social verbal responses and choices of actions. The most direct evidence for cognitive empathy in the present study comes from the self- versus other-satisfaction ratings. As hypothesised, the groups differed significantly in their self-ratings, where the range was narrower for the high versus the low AQ group. The high AQ group rated their satisfaction higher for performing low pro-social actions of little benefit to the character (e.g. for the falling over scenario: "carry on walking"; for the break up with partner scenario: "hang up as quickly as possible"), and rated their satisfaction lower for high pro-social actions, which went 'above and beyond' the social expectation to help the character (e.g. for the falling over scenario: "stop to help the man up and offer him additional aid"; e.g. for the break up with partner scenario: "offer to go and visit your friend"). However, in contrast with predictions, the high AQ group was *not* found to differ from the low AQ group when rating satisfaction from the perspective of the characters. Both groups judged low pro-social actions to be the least satisfactory, and high pro- social actions to be the most satisfactory for the characters. The lack of a group difference on the other-satisfaction ratings may indicate intact ability in the high AQ group to fully appreciate their perspectives, but this seems unlikely in the context of the well-documented difficulties with cognitive empathy in the literature exploring ASD (e.g. Happé, 1994; Spek et al., 2010).

Alternatively, and more convincingly, the current task may not have been sufficiently sensitive to reveal any potentially subtle perspective-taking difficulties in the high AQ group, and was thus unable to detect impairment in cognitive empathy. As discussed, people are likely to draw upon their social knowledge when considering others' perspectives and empathising with their needs, thereby facilitating a flexible response to novel situations. The scenarios in the 'Above and Beyond' task were designed to examine how a reduction in task demands (generation vs. selection of pro-social responses) may relate to pro-social behaviour. Thus, for the selection component of the task, the layout systematically presented the low pro-social actions first and the high pro-social actions last, to make the pro-social values salient. It is

therefore conceivable that the perspective-taking difficulties of the high AQ group were 'masked' as they may have used task cues and other deliberately learned social rules to accurately assess the main characters' satisfaction. However, it is possible that they were not able to fully benefit from or apply this knowledge, resulting in reduced pro-social behaviour nonetheless. Non-intuitive social knowledge may fail to support effective social interaction, as it is over-reliant on rigid rules and tends to be slowly and clumsily applied (Bowler, 1992). This has been implicated in ASD, and a similar explanation could account for the pattern of apparently intact understanding of the characters' expectations but reduced pro-social behaviour displayed by the high AQ group.

The finding of diminished personal gratification for going 'above and beyond' is also consistent with this interpretation: without an intuitive appraisal of the characters' needs, the high AQ participants may not have identified with them emotionally and may thus have experienced less satisfaction for helping them. Difficulties in understanding how the characters would view their own actions may have influenced their action choices and ratings of their own satisfaction, reflecting prioritisation of their own interests over those of the characters, even when they could readily gauge that a different action might be more beneficial to the main character. A reliance on salient task cues may have obviated the need for emotional identification with the characters, highlighting the difficulty of dissociating cognitive from emotional aspects of empathy.

5.4.4 Impaired emotional empathy?

On the other hand, impaired emotional empathy for could hypothetically account for the pattern of impaired self- and intact other-satisfaction ratings displayed by the high AQ group. However, in the light of the substantial body of literature pointing towards the opposite pattern, tending to find impaired cognitive empathy with intact emotional empathy in those with ASD (Blair, 2008; Singer et al., 2004), this seems unlikely. More plausible as an explanation of the findings is the notion of reduced capacity to experience or recognise their own emotions in the high AQ group. At the simplest level, reduced capacity to experience emotions could lead to a narrower range of self-satisfaction ratings across the three levels of pro-social action. There is also evidence that individuals with ASD have difficulty identifying and describing their own emotions (Hill, Berthoz & Frith, 2004) Whilst higher-functioning individuals on the spectrum show capacity to recognise and express basic emotions (e.g. happiness, sadness, and anger) difficulty with more complex or self-conscious emotions (e.g. pride and embarrassment) has been reported, and has been linked to the well-documented impairments with taking others' perspectives (Capps, Yirmiya & Sigman, 1992). Furthermore, a recent study found that as compared to the low AQ group, high AQ scorers experienced selective difficulty in recognising emotions, and required expressions of higher intensity to do so correctly (Poljac et al. 2012). Thus, personal experience of satisfaction for behaving pro-

socially on the 'Above and Beyond' task may represent a self-conscious emotional experience involving an appreciation of the social context, including appraisal of the characters' needs and responses to help.

5.4.5 The role of social knowledge

Everyday-type tasks such as 'Above and Beyond' may involve drawing on previously acquired social knowledge. For instance, taking the scenario where someone falls over as you are walking by, the unwritten 'rule' could be stated as "you should stop and help someone who might be injured". For the scenario where a friend has broken up with her partner, the unwritten rule might be said to be "you should comfort a friend who is upset". It has been suggested that knowledge stores relating to prior social experience may be more limited in those with ASD (Channon et al., 2001). This may result from a lack of exposure to relevant social situations because individuals with ASD actively avoid social encounters (Richer, 1976), which is often attributed to a sense of anxiety associated with such experiences (White, Ollendick & Bray, 2011). Whilst this study did not measure social engagement specifically, the AQ includes various statements that are likely to elicit agreement and disagreement respectively in high trait individuals (e.g."I would rather go to a library than to a party." and "I prefer to do things with others rather than on my own."). Thus, it is possible that the high AQ group engaged less in social interaction, and had fewer opportunities to gain relevant social knowledge.

Various authors have explored how social norms might motivate pro-social action and a range of social cues have been found to influence pro-social behaviour (e.g. Bickmen, 1972; Darley & Latane, 1968; Levine et al., 2005; Ruthowski et al., 1983; Schwartz, 2010). Reduced capacity to acquire such knowledge may account for the performance differences observed, since those with ASD are well known to be impaired in skills including pretend play (Travis & Sigman 1998), which offer children opportunities to engage in complex social negotiations and to practice social roles. With respect to the 'Above and Beyond' task, the low AQ group appeared to show greater compliance with social expectations by acting more pro-socially overall, often inconveniencing themselves in the process. It is possible that reduced social knowledge in the high AQ participants meant that they were less aware of these expectations or may have felt less pressure to comply, resulting in behaviour that was less pro-social. Furthermore, even if social knowledge was intact in the high AQ group, they may have been less motivated to apply it. It is well established that individuals with ASD show diminished responses to social rewards, and this has been related to reduced social learning (Zeeland et al. 2010). In the present study, insensitivity to reward may account in part for the high AQ group's reduced pro-social behaviour. Potential sources of reward include possible reciprocal future actions by the characters in need, and intrinsic reward through satisfaction gained by helping the characters.

5.4.6 Conclusion

In summary, the present study explored how pro-social behaviour might differ in those with high versus low AQ scores. Participants with high AQ scores were found to behave less pro-socially both on measures involving generating their own responses to characters' needs, and when selecting from alternatives. This indicated that task demands alone cannot account for the performance differences observed between the groups. When the groups' ability to take others' perspectives via ratings of characters' satisfaction with different courses of action was examined, no differences were found; this was in contrast to predictions. However, the high AQ group reported diminished personal satisfaction for helping others. This pattern of results is complex with many possible interpretations. Nevertheless, it seems most likely that the high AQ group utilised task cues, such as the systematic presentation of the alternative pro-social actions, to compensate for their difficulties in estimating the characters' perspectives. This non-intuitive method of social problem-solving may have resulted in a lack of emotional identification with the characters, leading to diminished personal reward for helping others and a corresponding reduction in the lengths to which they were willing to go.

Reliance on social knowledge to solve social problems may facilitate successful performance in the absence of an available mentalistic route. However, there is also evidence to suggest that this may result in rigid and clumsy patterns of behavior; the successful application of social knowledge requires flexibility and integration with emotional factors that may be diminished in ASD. The next Chapter will directly examine the role of contextual norms in guiding behaviour by varying the clarity of the social rule underpinning the expectation to help others, across those with high versus low AQ traits in a new scenario-based task: 'Social Expectations'.

Chapter 6: Examining how the clarity of the social expectation to help others influences pro-social behaviour

6.1 INTRODUCTION

Social norms define what types of behaviour are considered acceptable and how people are expected to act in certain situations. These can be viewed as rules of thumb and may serve as everyday heuristics that guide successful social interaction in constantly evolving and novel encounters. Crucially, social norms are not absolute, but should be applied flexibly and integrated with other information that may inform how one should behave in a particular situation. It is essential to utilise contextual factors when applying social norms, since behaviours that are considered acceptable may differ according to the circumstances. For instance, before greeting someone with a big hug you might wish to consider if the person is a close friend or a stranger. Or if you heard a funny joke that you'd like to share, you should first consider if the setting is appropriate, such as a family party, or not e.g. a work meeting.

There is little experimental work examining the understanding and application of social norms in ASD. However, some have proposed that the relative preservation of certain aspects of social performance in ASD, particularly in high-functioning people with more subtle difficulties, may reflect reliance on learned social rules and knowledge during social problem-solving tasks (e.g. Frith & Happé, 1999; Hill & Frith, 2003). For instance, in a study examining performance on a classic cognitive empathy task, high-functioning participants with ASD showed reduced activation in brain regions typically associated with this type of activity, even when correct mental state attributions were made. Conversely, these individuals tended to show greater activation in areas typically associated with more general problem-solving skills (Happé et al. 1996). White, Frith, Rellecke, Al-Noor & Gilbert (2014) subdivided participants with ASD into two groups, depending on whether they were able to pass false belief tasks or not. Regardless, all participants with ASD showed reduced activation of the areas typically associated with cognitive empathy, and thus even when individuals with ASD pass tasks involving taking others' perspectives they may rely on alternative routes to do so.

Although high-functioning people with ASD are able to learn about the rules underpinning everyday social behaviour, which may to some extent mask their difficulties, they may not be able to apply this knowledge flexibly. Clinical observations indicate that children with ASD struggle to appreciate the nuances of social norms, instead viewing them as absolute codes of conduct (Howlin & Yates, 1999). A recent experimental study by Callenmark, Kjellin, Ronnqvist, & Bolte (2014) also found that adolescents with ASD were able to judge the appropriateness of socially inappropriate behaviours accurately (e.g., asking a stranger in the lift for a comb). However, when they were required to explain why the characters'

behaviours were socially unacceptable, they tended to refer to practical aspects of the stories (e.g. "the comb may have lice"), or to make simplistic statements which indicated that they knew it was not appropriate behaviour but could not explain why (e.g. "it's wrong"). These tentative findings suggest that people with ASD may apply social rules rigidly. Therefore, even though reliance on knowledge of social rules may somewhat disguise social difficulties, clumsy and inflexible patterns of social behaviour may occur in more complex, unpredictable social circumstances.

Social norms do not only influence how we behave, but also what we expect of others. As discussed in Chapter 5, it is possible that the lower levels of pro-social behaviour displayed by the high AQ group reflected their limited perception of the social norms and the corresponding expectations of the characters. The lack of personal reward reported by those with high AQ levels is also suggestive of a reduced sense of pressure to comply with such expectations, which again might indicate a lack of appreciation of the norms inherent in the situations. The present study sought to explore this directly, by examining the role of societal expectations in guiding pro-social decision making. A novel scenario-based task, 'Social Expectations', was developed whereby the clarity of social norms inherent in everyday situations that may guide helping behaviours was manipulated (please note this task has now been described in this publication: Jameel, Vyas, Bellesi, Cassell & Channon, 2015). Participants with high versus low AQ scores were presented with scenarios featuring a character in need. Each scenario had two variant endings: a clear-cut versus ambiguous social rule. In the 'clear-cut' condition there was a strong social rule guiding the participants to behave pro-socially (i.e. offer to give up your seat to an elderly woman walking with a stick). In the 'ambiguous' condition the social rule was weaker (i.e. offer to give up your seat to a young woman carrying a heavy parcel). Participants were asked to reason about why someone might act pro-socially in the situation, and to rate characters' expectations of help. They were also asked to rate their own likelihood to offer help. Ratings indicating sympathy for the characters in need were also included. Sympathy refers to feelings of concern about the welfare of others, and is thought to play a motivating role in pro-social behaviour via other-oriented processes (Decety & Michalska 2010). Whilst sympathy and empathy are often conflated, they are in fact distinct concepts; the experience of sympathy is said to be dependent upon cognitive empathic skills such as apprehending another's mental state, but does not necessarily require a vicarious emotional experience (Decety & Chaminade 2003). Therefore, this measure was intended to provide an insight into the extent to which participants identified with the characters' needs when performing the task.

6.1.1 Hypotheses

6.1.1.1 Pro-social behaviour and sympathy ratings

Although little work has examined this, on the basis of the results of the 'Above and Beyond' task used in Chapter 5 and the existing literature it seemed probable that 'likelihood of helping' ratings and 'sympathy' ratings would be reduced in the high AQ group, especially in the ambiguous condition where social rules were less clear. However, it was also considered possible that the high AQ group might differentiate more than the low AQ group between the clear-cut and ambiguous conditions in their ratings of likelihood of helping and sympathy, showing more 'black and white' thinking consistent with a rigid reliance on social rules. In support of this prediction, some previous work with individuals with ASD reported that they showed heightened sensitivity to 'good' versus 'poor' justifications for wrongdoing (Channon, Fitzpatrick, Drury, Taylor & Lagnado, 2010), and greater differentiation between intentional and unintentional actions when assigning blame (Channon, Lagnado, Fitzpatrick, Drury & Taylor, 2011).

6.1.1.2 Understanding of social expectations: rationales versus ratings

It was predicted that the high AQ group would generate rationales for pro-social behaviour that relied upon social rules, rather than engaging with the individual perspectives of the characters, at least in the clear-cut condition where there was a readily available social rule. However, on the lower-demand measure of rating characters' expectations of help, it was postulated that the high AQ group might not differ from the low AQ group, at least for the clear-cut condition.

6.2 METHODS

6.2.1 Screening phase

6.2.1.1 Screening participants and procedure

An opportunistic sample of 645 full-time university students (41.39% male) who were fluent in English and aged 18 or over (mean age 20 years old) was recruited for the screening phase of the study. All participants completed the AQ (Baron-Cohen et al., 2001) and total scores were calculated for the whole sample. Participants within the highest-scoring and lowest-scoring 10% of males and females were contacted and invited to take part in the second stage.

6.2.2 Experimental phase

6.2.2.1 Design

There was one between-participants factor of AQ group (high versus and low AQ participants), and one within-group factor (clear-cut versus ambiguous social rule).

6.2.2.2 Experimental participants and procedure

Of those contacted from the screening phase, 21 (11 male, 10 female) individuals from the upper range and 20 (10 male, 10 female) individuals from the lower range agreed to take part in the experimental phase of the study, forming two groups of high AQ and low AQ participants. AQ scores ranged from 26 to 46 in the high AQ group (28–37 for male participants, and 26–46 for female participants), and 2–9 in the low AQ group (2–9 for male participants, and 4–8 for female participants). A *t* test confirmed that AQ scores differed significantly between groups, $t(39) = 23.23$, $p < .0001$; mean AQ scores were 30.52 (SD = 4.40), and 6.15 (SD = 1.66) for the high and low AQ groups respectively. The groups did not differ significantly in age, $t(39) = .064$, $p = .950$; mean age was 21.11 (2.62) and 21.06 (2.57) for the high and low groups respectively. All participants were tested individually, and provided written informed consent before completing all measures from the 'Social Expectations' task.

6.2.3 The 'Social Expectations' task

This task was designed to examine pro-social behaviour in relation to some of the unwritten social rules that govern everyday interactions, comparing scenarios based on both clear-cut and ambiguous rules. A range of scenarios were devised and piloted, in order to refine the items and develop the scoring systems. The final task consisted of ten hypothetical scenarios involving the participant and an unfamiliar character. These consisted of everyday social situations where the participant had the opportunity to engage in pro-social behaviour aiding the character, in line with a social rule. They were designed to make it clear that the participant was the only individual who could aid the character, and that engaging in the pro-social behaviour would be inconvenient to the participant (e.g. offering to give up your seat and stand for someone) (see Figure 3 for an example). The character was male in half the scenarios, and female in the other half, and the social context varied across scenarios to reflect a range of natural situations. To control for order effects, two different scenario orders were created and counterbalanced within each group.

All participants first read a sheet of instructions about the task. This explained that they would see short scenarios about everyday situations and would respond verbally to questions, supplying either ratings or free responses. Participants were requested to answer as quickly and as truthfully as possible. The scenarios were presented on paper, and participants were taken through an example before completing the ten experimental items. Scenarios and questions were presented in separate booklets such that relevant scenarios and endings remained on display throughout task performance, in order to minimise any memory demands.

Each scenario stem had two endings, manipulating the strength of the social rule guiding pro-social behaviour in the situation. 'Clear-cut' endings implied a strong social rule (e.g. "she is elderly and walking with a stick"), cueing an appropriate response (i.e. you should offer to give up your seat). 'Ambiguous'

endings still referred to a character that would benefit from help, but did not rely on such strongly endorsed social rules (e.g. "she is a young adult and is carrying a large parcel"). All participants responded to both scenario endings, and these were presented in their pairs in counterbalanced order; some participants saw the 'clear-cut' endings first and other participants the 'ambiguous' endings first, and this was counterbalanced within group. The order of the scenarios was also counterbalanced within group. Please note that the scenarios were not described as 'clear-cut' or 'ambiguous' to participants, rather they were simply told they would see two endings for each scenario.

Each scenario ending was followed by four questions; participants were first asked to rate how likely they would be to help the characters, and then to rate how sympathetic they felt towards the characters. Participants were then asked to indicate the strength of the characters' expectations for help, and finally to provide a rationale explaining why they might offer to help the character.

6.2.3.1 Scoring

6.2.3.1.1 Likelihood of helping ratings:

For each scenario, participants rated the likelihood of offering to help the characters on a scale of 1–10, where higher scores indicated greater pro-social behaviour. Ratings were then summed across all ten scenarios to create a total score for each condition (range 10–100), creating 2 scores: (1) clear-cut rule pro-social behaviour rating, (2) ambiguous rule pro-social behaviour rating.

Figure 3 - Example scenario from the 'Social Expectations' task

Example Scenario:
"You are sitting in a crowded waiting room with a small bag, waiting for a delayed train. All the other seats are taken by passengers with lots of luggage. A woman enters the waiting room looking for a seat."

Clear-cut Ending: She is elderly and walking with a stick.

Ambiguous Ending: She is a young adult and is carrying a large parcel.

Example Questions:

Likelihood of Helping Ratings:
On a scale of 1-10 how likely is it that you would offer her your seat?
1= not at all likely, 10 = very likely

Sympathy for Character Ratings:
On a scale of 1-10 how sympathetic do you feel towards her?
1= not at all sympathetic, 10 = very sympathetic

Strength of Character Expectation Ratings:
On a scale of 1-10 how much do you think she expects you to offer her your seat?
1 = not at all, 10 = very much

Verbal Rationales of Societal Expectation Understanding:
Why might you offer her your seat?

6.2.3.1.2 Sympathy for character ratings

For each scenario, participants rated the degree of sympathy they experienced for the characters on a scale of 1–10, where higher scores indicated greater sympathy. Ratings were then summed across all 10 scenarios to create a total score for each condition (range 10–100), creating 2 scores: (1) clear-cut rule sympathy rating, (2) ambiguous rule sympathy rating.

6.2.3.1.3 Strength of character expectation ratings

For each scenario, participants rated how much they thought the characters' expected their help on a scale of 1–10, where higher scores indicated a greater expectation to help. Ratings were then summed across all 10 scenarios to create a total score for each condition (range 10–100), creating 2 scores: (1) clear-cut rule character expectation for help rating, (2) ambiguous rule character expectation for help rating.

6.2.3.1.4 Verbal rationales: understanding of social expectations

Verbal responses were categorised according to two dimensions: rule-based or person-based rationales. The person-based rationales reflected responses that referred to the characters' needs, and/or conveyed a sense of self- sacrifice on the part of the participants, in order to meet the characters' needs. Rule-based rationales reflected responses that made explicit reference to a social rule guiding an expectation to help, or implied a social rule by simply referring to the facts of the scenario. Scoring of the example scenario is shown below in Figure 4. Responses could only score for one of the two dimensions; if both dimensions were met then the best answer was taken, and thus participants scored for person-based rationales.

The responses were classified by one blind independent rater, and one rater who was not blind to group membership. There was an inter-rater agreement rate of 90.73 %; all disagreements were resolved between the raters via discussion. Once all responses had been classified and disagreements resolved, participants' scores were summed across all 10 scenarios, and the percentage of person-based versus rule-based responses was calculated for both the clear-cut and ambiguous conditions.

Figure 4: Scoring of example scenario from the 'Social Expectations' task

Description of criteria

Person-based rationales
A response that either referred to the characters' needs, and/or conveyed a sense of self-sacrifice on the participants' part in order to meet the characters' needs.

Rule-based rationales
A response that made explicit reference to a social rule guiding an expectation to help, or implied a social rule by simply referring to the facts of the scenario.

Example responses:
Clear-cut ending: "She is elderly and walking with a stick"

Person-based rationale examples
e.g. "I would feel sorry for her and it would be difficult for her to stand in a crowded waiting room"
e.g. "She needs it more than I do"

Rule-based rationale examples
e.g. "You should always offer your seat to women, elderly and the disabled"
e.g. "She is walking with a stick"

Ambiguous ending: "She is a young adult and is carrying a large parcel"

Person-based rationale examples
e.g. "She must be feeling very tired"
e.g. "I think she needs it more than I do because she is carrying a large parcel"

Rule-based rationale examples
e.g. "To be polite"
e.g. "It is common courtesy to offer your seat"

6.3 RESULTS

6.3.1 Data analysis

Means and standard deviations (SD) for each of the measures are presented below in Table 2. A significance level of .05 was adopted, with adjustment for post hoc *t* tests (.05/2) to control for multiple comparisons. The assumptions of normality were met and thus parametric analyses were performed.

6.3.2 The 'Social Expectations' task

6.3.2.1 Likelihood of helping ratings:

A repeated measures 2 x 2 ANOVA was conducted to examine ratings for likelihood of helping for all scenarios. There was one between-participant factor (AQ group: high versus low AQ) and one within-participant factor (ambiguity of social rule: clear-cut versus ambiguous). There were significant main effects of condition, $F(1,39) = 491.87$, $p < .0001$, and of group, $F(1,39) = 6.79$, $p = .013$. The condition by group interaction was not significant, $F(1,39) = .274$, $p = .604$.

Inspection of the mean scores (presented in Table 2) revealed that all participants were more likely to behave pro-socially when the social rule was clear-cut versus ambiguous. This is in line with the prediction that a clear-cut rule would enhance the characters' expectation for help, and thus also the likelihood of complying with it. The high AQ group was less pro-social overall; the lack of condition of social rule by group interaction suggests that the groups were not, however, differentially affected by the strength of the social rule.

6.3.2.2 Sympathy for character ratings:

A repeated measures 2 x 2 ANOVA was also conducted to examine ratings for sympathy for all scenarios. There were significant main effects of condition, $F(1,39) = 356.63$, $p < .0001$, and group, $F(1,39) = 11.4$, $p = .002$. The condition by group interaction was not significant $F(1,39) = .486$, $p = .490$.

The pattern of results suggest that all participants were more sympathetic when the rule was clear-cut versus ambiguous, and that high AQ participants were less sympathetic towards characters overall. The lack of condition of social rule by group interaction suggests that the groups were not, however, differentially affected by the strength of the social rule.

6.3.2.3 Strength of character expectation ratings

A repeated measures 2 x 2 ANOVA was also conducted to examine ratings for the strength of the characters' expectation for help across all scenarios. There was a significant main effect of condition, $F(1,39) = 175.41$, $p < .0001$. However, the main effect of group, $F(1,39) = 1.24$, $p = .272$, and the condition by group interaction, were not significant $F(1,39) = .30$, $p = .864$. This confirms that the social

rule manipulation operated as intended: all participants identified a stronger character expectation to help when the rule was clear-cut versus ambiguous. In line with predictions, the high AQ group were able to identify a stronger character expectation for help in the clear-cut versus ambiguous condition, on this lower demand measure of social expectation understanding.

6.3.2.4 Verbal rationales: understanding social expectations

Finally, the high and low AQ groups were compared for their verbal responses outlining why one might choose to help the character in each scenario. A repeated measures 2 x 2 ANOVA was conducted to examine the percentage of rationales classified as rule-based versus person-based for the two conditions. The main effect of condition was not significant, $F(1,39) = .114$, $p = .738$; nor was the main effect of group $F(1,39) = 2.161$, $p = .150$. However, there was a significant condition by group interaction, $F(1,39) = 5.57$, $p = .023$. Post-hoc t tests, using a strict significance level, showed that as predicted the high AQ group used significantly more rule-based versus person-based rationales than the low AQ group in the clear-cut condition, $t(39) = 2.327$, $p = .025$; there was no significant group difference in the ambiguous condition, $t(39) = .269$, $p = .790$.

Table 2: Mean percentage scores and standard deviations for all measures for the 'Social Expectations' task.

	Low AQ group (N = 20) M (SD)		High AQ group (N = 21) M (SD)		Significance (p = .025)		Effect Size (d)
Likelihood of helping (%)					Condition	*	
					Gp	*	
					Gp x condition	NS	
Clear-cut	86.95	(9.26)	78.42	(9.68)	__		0.90
Ambiguous	55.10	(10.16)	48.05	(12.73)	__		0.61
Sympathy (%)					Condition	*	
					Gp	*	
					Gp x condition	NS	
Clear-cut	80.25	(9.25)	68.81	(13.09)	__		1.01
Ambiguous	43.85	(9.55)	35.00	(12.62)	__		0.79
Strength of expectation (%)					Condition	*	
					Gp	NS	
					Gp x condition	NS	
Clear-cut	76.05	(9.01)	73.24	(2.17)	__		0.43
Ambiguous	54.15	(8.77)	50.76	(11.09)	__		0.34
Verbal rationale classification (%)					Condition	NS	
					Gp	NS	
					Gp x condition	*	
Clear-cut							
Rule	33.00 (18.38)		49.05 (25.08)		0.025*		
Person	67.00 (18.38)		50.95 (25.08)		0.079		
Ambiguous							
Rule	40.5 (20.64)		39.05 (13.38)		__		
Person	59.5 (20.64)		60.95 (13.38)		__		

** Significant at p = .025* *NS = not significant*

6.4 DISCUSSION

6.4.1 Summary of findings

The present study examined the role of social rules in guiding pro-social behaviour, and how this might be influenced by autistic traits by comparing everyday situations ending with either a clear-cut or an ambiguous social rule, in which a character required help. Ratings of likelihood of complying with these societal expectations revealed that, as expected, all participants were less likely to offer help in the ambiguous versus clear-cut condition. The high AQ group was less pro-social overall compared to the low AQ group, but was not differentially affected by the clarity of the social rule. Ratings of sympathy for the character showed that, as predicted, all participants were more sympathetic in the clear-cut versus ambiguous condition. Relative to the low AQ group, the high AQ participants expressed less sympathy for characters overall, but again they were not differentially affected by the clear-cut versus ambiguous manipulation. Participants' understanding of the social expectation to help was also assessed. On ratings of the strength of the characters' expectations of help the groups did not differ in performance; contrary to predictions the high AQ group showed an intact understanding of the expectation to help in both conditions. However, when producing verbal rationales to justify why one would help, the high AQ group revealed a more simplistic and rule-bound understanding of the social expectations inherent in the scenarios.

As expected, the high AQ participants were less likely to help characters, but the groups were not differentially affected by the strength of social rule manipulation. Thus, all participants were more compliant with the expectation to help characters in the clear-cut versus ambiguous condition. For instance, in the example scenario, participants were much more likely to help when the character was elderly and walking with a stick, but far less likely to help when she was young and carrying a heavy parcel; the high AQ group gave lower likelihood of helping ratings than the low AQ group across conditions. With respect to the sympathy ratings, all participants were more sympathetic in the clear-cut versus ambiguous condition. The high AQ group was thus less likely to feel sympathy for the characters, regardless of the clarity of the social rule.

Did those with high AQ traits understand the societal expectations inherent in the scenarios? One way of estimating this was by asking participants to rate the extent to which the characters expected help. Interestingly, here the groups did not differ; the high AQ group was equally able to identify the stronger expectation to behave pro-socially in the clear-cut versus ambiguous condition. However, when asked to complete the more demanding task of providing verbal rationales outlining why one might behave pro-socially, the picture was more complex. In the clear-cut condition the low AQ group used more person-

based rationales (e.g. she needs the seat more than me) than rule-based rationales reflecting societal expectations (e.g. you should always give up your seat for the elderly), whereas the high AQ group used these equally. In the ambiguous condition, both groups used slightly more person-based than rule-based rationales. The findings of the present study corroborate that of Chapter 5, providing further evidence of reduced pro-social behaviour in individuals with high numbers of autistic traits. The two studies taken together also reveal some hint of preserved social knowledge in the high AQ group, but differences between groups in the socio-emotional processes thought to motivate pro-social behaviour.

6.4.2 Can non-social and social models of ASD account for the findings?

Various theoretical accounts associated with ASD may be relevant for explaining the pattern of findings. As discussed above, executive dysfunction is typically used to account for the non-social symptoms of ASD, such as repetitive and restricted interests, but it may also have implications for social functioning (see Chapter 2.2 and 2.4 for a discussion of this). In relation to the present pattern of findings, mild executive difficulties may again have resulted in failure of the high AQ group to put their relatively preserved knowledge of social rules into practice and could account for the reduced pro-social behaviour observed in the high AQ group. The high AQ group may have acted impulsively in their own interests, and/or experienced difficulty in evaluating relevant aspects of the social scenarios. Similarly, WCC theory, which proposes that people with ASD may pay enhanced attention to details at the expense of integrating information (e.g. Happé, 1996), can best account for non-social deficits, but may also have implications for social behaviour. With respect to the dominant theory that deficits in cognitive empathy with intact emotional empathy are the primary source of social difficulties in ASD, this will be briefly reviewed with respect to the current experiment.

As discussed in Chapter 2.3.2 there is some evidence that emotional empathy is also impaired in ASD (Blair, 2008). Thus, it is possible that an imbalance in the priorities placed on the participants' versus characters' needs resulted in less motivation to behave pro-socially, regardless of whether the high AQ group could correctly identify and understand the characters' needs. As with the 'Above and Beyond' task in Chapter 5, it could be argued that the high AQ participants may have focused on themselves and prioritised their own interests at the expense of helping others, and experiencing reduced resonance with the characters' points of view may have compounded this. Regardless, impairment in emotional empathy could not readily account for the greater tendency of the high AQ group to use rule-based rationales in the clear-cut condition when reasoning about why one should behave pro-socially. Impairment in emotional empathy in the context of intact cognitive empathic abilities would not be expected to affect the high AQ group's capacity to understand how they should behave and why; rather, it should selectively influence their actual behaviour. Nonetheless, a possible contribution of emotional empathy cannot be dismissed,

since it is difficult to disentangle the relative contributions of this versus cognitive processes in relation to measures designed to explore everyday social behaviour.

Turning now to the potential contribution of cognitive empathy, it is contended that pro-social behaviour is dependent upon appreciating others' perspectives, and acting accordingly (Eisenberg et al., 1989; Eisenberg, 2007). Failure to do so may have reduced the high AQ participants' incentives to help the characters, resulting in reduced compliance with the societal expectation to behave pro-socially. This explanation is consistent with evidence from studies of people with high autistic traits showing differences on tasks tapping cognitive empathy, including social attention (Freeth, Bullock & Milne, 2013) and false belief tasks (Best, Moffat, Power, Owens, & Johnstone, 2008). The sympathy ratings in the present study may provide the most direct measure of cognitive empathy, since sympathy refers to feelings of concern about the welfare of others, and is thought to play a motivating role in pro-social behaviour via other-oriented processes (Decety & Michalska, 2010). The experience of sympathy is said to be dependent upon the ability to apprehend another's mental state, but it does not necessarily require a vicarious emotional experience (Decety & Chaminade, 2003). On this basis, sympathising with characters in the 'Social Expectations' task requires participants to put themselves in others' shoes and imagine how they would feel in such a situation, and is potentially mediated by cognitive empathic abilities. Lower sympathy ratings by the high AQ group are therefore consistent with the evidence of impaired cognitive empathy in those with ASD.

Some previous work has reported 'black and white' sympathy ratings in those with ASD, with heightened sensitivity to 'good' versus 'poor' justifications for wrongdoing (Channon et al. 2010), and greater differentiation between intentional and unintentional actions when assigning blame (Channon et al. 2011). No evidence of such 'black and white' thinking was found in the present AQ study, since sympathy ratings were lower overall in the high versus low AQ group, and the groups were not differentially affected by the strength of the social rule. However, the nature of the present study is different to those previously used, since in the Channon et al. tasks, lack of sympathy in the group with ASD related to characters acting for reasons that generally contravened societal, and indeed legal, expectations (e.g. drunk driving after a party, or intentionally giving a spouse an overdose of their medication). By contrast, in the present AQ study, the characters were all deserving of help and thus of sympathy, regardless of the condition.

6.4.3 The role of social knowledge

One potential caveat for an interpretation consistent with a deficit in cognitive empathy is that those with high AQ traits did not differ in their ratings of the characters' expectations for help. At first glance, it might appear that an appraisal of the characters' mental state is required to complete this

measure, suggesting the high AQ group displayed intact perspective-taking abilities (part of cognitive empathy). However, reliance on knowledge of societal norms about how to behave in such situations may have obviated the need to apprehend the characters expectations directly and lead to similar judgments. It has been suggested that individuals with ASD may not rely on intuitive socio-emotional processes for solving social and moral dilemmas (Greene & Haidt, 2002). Rather, it may be a more laborious process in which they learn about and apply social rules, especially when these are readily available. This interpretation is also consistent with the finding that in the clear-cut task condition where the social rules were more salient, the high AQ group appeared to draw upon these rules, providing more rule-based and fewer person-based justifications for the reasons surrounding why one should act pro-socially. On the other hand, when salient rules were not readily available in the ambiguous condition, the high AQ group was able to produce person-based rationales as often as the low AQ group. This may indicate a stylistic preference for rule-based reasoning rather than a deficit in cognitive empathy, since engagement in processes such as perspective-taking may be possible, but more effortful for them. This is also consistent with previous literature indicating that individuals with ASD might provide correct answers, but that the reasoning behind their judgments is often more limited (Moran et al., 2011). For instance, Zalla, Sav, Stopin, Ahade, & Leboyer (2009) found that adults with ASD were able to use compensatory strategies to carry out social judgments regarding faux-pax, but failed to justify their responses adequately. In the present experiment, rating the characters' expectations could be argued to be a relatively low-demand task in which knowledge of simplistic rules may suffice, whereas producing rationales is comparatively high-demand and requires a more detailed understanding of social situations. Thus, the lower-demand measure of ratings may not have been sufficiently sensitive to reveal performance differences, and the impact of the clarity of the social expectation upon rationale production may have only become apparent when social cues to support performance were provided (i.e. clear-cut condition).

It is likely that healthy individuals integrate social knowledge stores and socio-emotional processes when assessing social situations and considering how to respond. In order to make good use of social rules to deal with complex social stimuli in a range of contexts, flexibility is required to learn about the contingencies for applying rules appropriately to different conditions (Nelson & Guyer 2011; Bunge, 2004). This is likely to draw upon executive skills, which as previously discussed are thought to be impaired in individuals diagnosed with ASD (see Chapter 2.2.2 for a discussion). From a developmental perspective, as children become adolescents, and in turn adults, the complexity of the social environments they navigate increases dramatically. In typically developing children, this should be supported by gradually enhanced social knowledge, as a result of exposure to more challenging and ambiguous social environments. However, in children and adolescents with ASD, who tend to avoid social engagement (Richer, 1976), and/or experience the social world in an 'atypical' fashion (Hughes & Leekam, 2004), such learning may be

deficient. Thus, the acquisition of knowledge about social rules may be further protracted or stunted in those with ASD. With respect to the current study, more limited social learning experiences and/or an impaired application of acquired social knowledge could potentially account for the pattern of findings. The high AQ group may have a more superficial understanding of social rules and/or struggle to apply their knowledge flexibly to the particular context.

6.4.4 Conclusion

In summary, the present study explored how pro-social behaviour might differ in those with high versus low AQ scores in the context of a clear-cut or an ambiguous social rule. Participants with high AQ scores were found to behave less pro-socially and to show reduced sympathy for characters across conditions. Understanding of the social expectation to help was also assessed via ratings of the characters' expectations and the generation of rationales justifying why one should help in each scenario. These measures revealed intact ratings of characters' expectations, but differences in the rationales produced by the high AQ group. It seems most plausible that well documented difficulties in cognitive empathy (and possibly emotional empathy), combined with altered social learning opportunities and/or impaired use of social knowledge, predominantly drove performance differences on this task. Reliance on social knowledge as a compensatory mechanism may have circumvented the need to employ empathic processes when estimating the characters' expectations. However, this may be at the expense of resonating with someone emotionally, hence explaining why the high AQ group reported less sympathy with the characters, and were less likely to help them. Furthermore, social knowledge in the absence of emotional processing, may be applied more rigidly and only disguise social difficulties to a certain extent, accounting for the performance differences seen on the more demanding measure of social understanding (rationale production).

The first two experiments of this thesis explored situations in which all the characters depicted were deserving of help; it is unclear if such stark group differences would be detected if the characters were not all deserving. Thus, the next experiment explores this by comparing scenarios in which the main characters are more or less deserving of help, to see how this affects pro-social behavior in those with high versus low AQ traits. In addition, the next study included measures of executive function, cognitive and emotional empathy, to explore the extent of any group differences on these measures.

Chapter 7: Examining the influence of characters' justifications for help on pro-social behaviour

7.1 INTRODUCTION

Previous work indicates that people are more pro-social when another's need for help or distress is more salient (e.g. Bickman, 1972). Consistent with this, the 'Social Expectations' task in Chapter 6 revealed that all participants were more pro-social and sympathetic when the expectation to help others was clear-cut versus ambiguous. Although the high AQ group was less pro-social overall and expressed less sympathy for the characters in this task, the manipulation of clear-cut versus ambiguous rules did not differentiate the groups, and they demonstrated some basic understanding of the relative expectation to help. One possible consideration is that the characters presented in the scenarios were always deserving of the participants' help, regardless of whether the rule underpinning this expectation was clear-cut versus ambiguous. Previous work has shown that people are influenced by characters' intentions when making judgments about their culpability for bad events, and that individuals with ASD show heightened sensitivity to good versus bad justifications for characters' violations of rules (Channon et al., 2010; 2011). On the basis of the previous findings it seems likely that people will be more likely to help others if they have a strong versus weak justification for why they are deserving of it. It also seems plausible that sensitivity to deservingness might be influenced by autistic trait levels, in line with the previous findings of enhanced sensitivity in ASD.

The present study was designed to investigate how pro-social behaviour in those high and low in autistic traits varied as a function of perceived deservingness using the 'Social Favours' task (Vyas, Jameel & Channon, in preparation a). In this task, deservingness is directly manipulated, in order to compare compliance with requests for favours when a strong versus weak justification is provided. In a previous study this task was found to differentiate those high and low in psychopathic traits, whereby high trait participants showed reduced sensitivity to the deservingness manipulation (Vyas et al., in preparation a). In the present study a number of additional measures were also included, tapping both executive functions and different aspects of empathy, to examine group differences on these measures. As discussed in Chapter 2.2.2 and 2.4, executive dysfunction might play a role in comprehension of social scenarios and in formulating appropriate responses. Reduced pro-social behaviour has been linked to a lack of understanding of others' needs and lowered empathic responsivity to others' distress. It seems possible that if either cognitive or emotional empathic routes are impaired, this might not only affect the extent to which people choose to help others, but also judgments of their deservingness.

7.1.1 Hypotheses

7.1.1.1 The 'Social Favours' task

7.1.1.1.1 Pro-social behaviour

Deservingness was directly manipulated, in order to compare compliance with requests for favours when there was a strong versus weak justification provided. Participants rated how likely they would be to comply with each request. It was predicted that all participants would be more compliant in the scenarios with a strong versus weak justification for helping. Relative to the low AQ group, it was expected that the high AQ group would be less compliant overall, in line with the findings of reduced pro-social behavior on the 'Above and Beyond' task and 'Social Expectations' task used in Chapters 5 and 6. However, it also seemed possible that any group differences would be exacerbated by the strength of the justification provided in the scenario. Thus, an alternative prediction was also considered, whereby the high AQ group might show 'black and white' thinking as described by Channon et al., (2010; 2011), providing more help for requests with stronger justifications, and less help for requests with weaker justifications than the low AQ group.

7.1.1.1.2 Acceptability of request ratings

Participants also rated how acceptable it was for the characters to make each request. It was predicted that all participants would rate scenarios with a strong versus weak justification for helping as more acceptable. It was considered possible that this measure might not differentiate the groups, on the basis that on the 'Social Expectations' task the high AQ group demonstrated intact understanding of other contextual factors that guide social behavior, at least to some extent. However, again it also seemed possible that group differences for acceptability ratings might be dependent on the type of request. Thus, an alternative prediction was also considered, whereby the high AQ group might over-estimate the acceptability of requests when stronger justifications were provided, and underestimate when weaker justifications were provided, relative to their low AQ counterparts. This pattern would again be consistent with the 'black and white' thinking found by Channon et al., (2010; 2011).

7.1.1.1.3 Personal sacrifice ratings

Finally, participants gave ratings of how much of a personal sacrifice they would find it to agree to the characters' requests. It was expected that all participants would perceive compliance with requests as more of a personal sacrifice in scenarios with a weak versus strong justification for helping, even though the acts themselves were the same in both conditions. Relative to the low AQ group, it was predicted that

the high AQ group would perceive the personal sacrifice for complying with requests overall to be greater, at least in the weak justification condition. This is in line with the finding described in Chapter 5 of reduced satisfaction in the high AQ group for helping characters on the 'Above and Beyond' task.

7.1.1.2 Additional measures

7.1.1.2.1 Executive functions

In order to explore executive functions, two tests from the Delis-Kaplan Executive Function System (D-KEFS) battery (Delis, Kaplan & Kramer, 2001) were administered: inhibition and verbal fluency. These were expected to differentiate the groups, whereby the high AQ group would demonstrate poorer performance than the low AQ group on both tests. This prediction was made on the basis of a range of work indicating impairment in executive functions in ASD (for a review please see: Hill, 2004).

7.1.1.2.2 'Interpersonal Reactivity Index' ('IRI')

The 'IRI' is a self-report questionnaire measure that explores different dimensions of empathy (Davis, 1980). It was predicted that the high AQ group would score significantly lower for this measure overall, at least for items exploring cognitive aspects of empathy. This is consistent with the classic finding of impairment in cognitive empathy in ASD (e.g. Baron-Cohen, 1985), and recent work exploring empathy in groups with sub-clinical autistic traits (Gökçen, Petrides, Hudry, Frederickson & Smillie; Lockwood, Bird, Bridge & Viding, 2013). The questionnaire contains four subscales, two of which tap predominantly cognitive aspects of empathy, and the remaining two chiefly emotional aspects. Given the evidence of impaired cognitive with some evidence of preserved emotional empathy in ASD, it was expected the high AQ group would receive lower scores for the cognitive, but not necessarily the emotional subscales.

7.1.1.2.3 'Mentalistic Interpretation Test' ('MIT')

The 'MIT' is a scenario-based measure that explores understanding of characters' intentions via interpretation of their sarcastic remarks or actions (Channon, Pellijeff & Rule, 2005; Channon et al., 2007). It was predicted that the high AQ group's performance would be significantly poorer on this measure, consistent with the well-established deficit in cognitive empathy in ASD, including taking others' perspectives and understanding their mental and emotional state. However, it was not clear whether sarcasm- or action-related items would be more sensitive to potential group differences.

7.2 METHODS

7.2.1 Screening phase

7.2.1.1 Screening participants and procedure

An opportunistic sample of 662 full-time university students (65.60% female) who were fluent in English and aged 18 or over (mean age 20 years old) was recruited for the screening phase of the study. All participants completed the AQ (Baron-Cohen et al., 2001) and total scores were calculated for the whole sample. Participants within the highest-scoring and lowest-scoring 10% of males and females were contacted and invited to take part in the second stage.

7.2.2 Experimental phase

7.2.2.1 Design

There was one between-participants factor of AQ group (high vs. low scorers) and one within-participants factor of deservingness (strong vs. weak justifications for requests).

7.2.2.2 Experimental participants and procedure

Of those contacted from the screening phase, 20 (9 male, 11 female) individuals from the upper range and 23 (12 male, 11 female) individuals from the lower range agreed to take part in the experimental phase of the study, forming two groups of high AQ and low AQ participants. AQ scores ranged from 26 to 38 in the high AQ group (26–38 for male participants, and 27–38 for female participants), and 4–13 in the low AQ group (4–13 for male participants, and 5–13 for female participants). A *t* test confirmed that AQ scores differed significantly between groups, $t(41) = 25.03$, $p < .0001$; mean AQ scores were 31.85 (SD = 3.41), and 9.69 (SD = 2.41) for the high and low AQ groups respectively. The groups did not differ significantly in age, $t(41) = .147$, $p = .800$; mean age was 20.52 (2.71) and 20.33 (2.71) years for the high and low groups respectively. All participants were tested individually, and provided written informed consent before completing the 'Social Favours' task and all additional measures. The battery of tasks was administered in counterbalanced order to ameliorate any potential order effects upon performance, and this was counterbalanced within each group.

7.2.3 The 'Social Favours' Task

This task (Vyas et al., in preparation a) was designed to examine compliance with requests for favours with strong versus weak justifications. The task consisted of ten short scenarios (see Figure 5 for example), describing an interaction with a character known to the participant (such as a friend, relative, colleague or flatmate). Each scenario describes a situation in which the character makes a request of the participant, compliance with which would incur a loss of time, effort or money (for example carrying a large parcel upstairs). The participant is then asked a) to rate how likely they would be to comply with this request, b) to rate how acceptable the request is, and c) to rate the extent to which they would find it a personal sacrifice to comply with the request. Scenarios and questions were presented in separate booklets such that relevant scenarios and endings remained on display throughout task performance, in order to minimise any memory demands.

Each scenario had two variant endings; in one ending, the character provided a strong justification for making the request (for example, having an injury that would make carrying the parcel difficult). In the alternative ending, the character provided a weak justification for making the request (for example, not wishing to ruin their shirt). Characters were counterbalanced across the scenarios for gender and proximity of relationship. All participants responded to both scenario endings, and these were presented in their pairs in counterbalanced order; some participants saw the scenarios with strong justifications first and other participants saw the scenarios with weak justifications first, and this was counterbalanced within group. There were two different scenario orders, which were counterbalanced within each group. Please note that the justifications were not described to participants as 'strong' or 'weak', rather they were simply told they would see two endings for each scenario.

Figure 5: Example scenario from the 'Social Favours' task

Example Scenario:
"You run into your neighbour one day, and he has just had a large parcel delivered. He asks you a favour."

***Strong Justification*: "I've strained my back. Would you carry it upstairs for me?"**

***Weak Justification:* "I'm going out and don't want to get my shirt dirty. Would you carry it upstairs for me?"**

Example Questions:
Likelihood of Helping Ratings:
On a scale of 1-10 how likely are you to agree to your neighbour's request?
1= not at all likely, 10 = very likely

Acceptability of Request Ratings:
On a scale of 1-10 how acceptable is it of your neighbour to make this request?
1= not at all sympathetic, 10 = very sympathetic

Personal Sacrifice Ratings:
On a scale of 1-10 how much of a personal sacrifice would you find it to agree to your friend's request?
1 = not at all, 10 = very much

7.2.3.1 Scoring

7.2.3.1.1 Likelihood of helping ratings:

For each scenario, participants rated the likelihood of offering to help the characters on a scale of 1–10, where higher scores indicated more pro-social behaviour. Ratings were then summed across all ten scenarios to create a total score for each condition (range 10–100), creating two scores: (1) strong justification pro-social behaviour rating, (2) weak justification pro-social behaviour rating.

7.2.3.1.2 Acceptability of request ratings:

For each scenario, participants rated the acceptability of the request on a scale of 1–10, where higher scores indicated greater acceptability. Ratings were then summed across all ten scenarios to create a total score for each condition (range 10–100), creating two scores: (1) strong justification acceptability rating, (2) weak justification acceptability rating.

7.2.3.1.3 Personal sacrifice ratings:

For each scenario, participants rated how much of a personal sacrifice they would find it to comply with the request on a scale of 1–10, where higher scores indicated a greater sacrifice. Ratings were then summed across all ten scenarios to create a total score for each condition (range 10–100), creating two scores: (1) strong justification sacrifice rating, (2) weak justification sacrifice rating.

7.2.4 Additional measures

7.2.4.1 Executive functions

Two neuropsychological tests from the D-KEFS battery that explore different aspects of executive functions were administered (Delis et al., 2001) The 'Colour-Word Interference' Test involves inhibiting habitual responses by naming the ink colour of colour words written in conflicting colours (e.g. RED printed in green ink). There were also two control conditions, colour naming (naming blocks of colours) and word reading (reading colour words written in black ink). The 'Verbal Fluency' test involves generating words within a time limit according to certain rules, beginning with either a particular letter or belonging to a semantic category. The scores were converted into age-scaled score equivalents using the manual, which gives a range of 1-19 for each measure.

7.2.4.2 'IRI'

This is a self-report questionnaire designed to measure the global concept of empathy (Davis, 1980). It consists of 28 statements on a 5-point scale, where a rating of 1 represents "does not describe me well", and a rating of 5 represents "describes me well". Items are scored accordingly: 0 for a rating of 1, 1 for a rating of 2, 2 for a rating of 3, 3 for a rating of 4 and 4 for a rating of 5. In order to ensure consistent responding the questions are counterbalanced such that for some, higher scores corresponded to higher empathy and for others higher scores corresponded to lower empathy; items where higher scores = lower empathy were then reversed. before being added together.

There are four subscales that examine different dimensions of empathy. The first two subscales were designed to assess components of cognitive empathy; 'Perspective-taking' was designed to assess spontaneous attempts to adopt the perspectives of others (8 items: range 0-32), and 'Fantasy' was designed to assess the tendency to identify with characters in movies, plays and other fictional situations (7 items: range 0-28). The final two subscales were designed to assess people's emotional empathy via their emotional reactions to others; 'Empathic Concern' refers to respondents feelings of warmth, compassion and concern for others (7 items: range 0-28) and 'Personal Distress' refers to personal feelings of anxiety and discomfort resulting from observing others' negative experiences (6 items: range 0-24). Please see

Figure 6 for an example item from each subscale. All four subscales of the 'IRI' have been found to have good test-retest reliability (.61-.81) and good internal consistency (Cronbach's alpha .70-.78).

Figure 6: Example items from 'IRI'

Perspective-taking

Before criticizing somebody, I try to imagine how I would feel if I were in their place.

1----------2----------3----------4----------5

Does not describe me well *Describes me well*

Fantasy

When I watch a good movie, I can very easily put myself in the place of a leading character.

1----------2----------3----------4----------5

Does not describe me well *Describes me well*

Empathic Concern

I often have tender, concerned feelings for people less fortunate than me.

1----------2----------3----------4----------5

Does not describe me well *Describes me well*

Personal Distress

In emergency situations, I feel apprehensive and ill-at-ease.

1----------2----------3----------4----------5

Does not describe me well *Describes me well*

7.2.4.3 'MIT'

This test was adapted from Channon et al., (2005; 2007). It was designed to assess mentalising skill, a crucial aspect of cognitive empathy, and consisted of two different item sets: sarcastic remarks or human actions. The sarcastic scenarios described a social context that ended with a sarcastic remark by one of the characters. The human actions also consisted of brief social scenarios, ending with an action by one of the characters, where it was necessary to take account of the mental state of the character to interpret the action. There were a total of ten items, with five in each category. The item sets were matched in length and presented in pseudo-randomised order. In order to reduce memory load, each story remained on view throughout.

After participants read the stories, they were asked to explain verbally what the characters meant by their remarks, or why they carried out the actions. If a broadly correct but inadequate response was given,

participants were prompted to explain their answer more fully. A score of two points was given for responses giving a clear correct explanation of the remark or action, a score of one point when the answer was not incorrect, but was not adequately explained, and a score of zero points when the answer was incorrect or irrelevant. Please see Figure 7 for an example of a sarcastic and an action item with scoring. Two scores were calculated for the (1) 'MIT Action': mentalistic action items and for the (2) 'MIT Sarcasm': sarcastic items, both with a range of 0-20.

Figure 7: Example scenarios and answers from the 'MIT'

Example of a Sarcastic Item:

"Liz and her friend often played tennis. Her friend always wanted to be best at everything. One day they were playing tennis in the local park. Liz knew that her friend expected to win the game. However, that day her friend did not win. Liz said: "I suppose you'll say there's a hole in your racket!"

***Question*: "What did Liz mean when she said that?"**

Scoring:

(2) Any answer indicating that Liz thought her friend was a bad loser or would make an excuse
e.g. "She is making an excuse for losing"
e.g. "You won't admit you played badly"
e.g. "She expected her friend to blame her equipment"

(1) Any partial answer that does not quite clarify the issue
e.g. "You're not as good as you think you are"
e.g. "Her friend was disappointed/upset/surprised to lose"

(0) Any clearly incorrect or irrelevant answer
e.g. "She thought there was a hole in her racket"
e.g. "Her friend is a bad tennis player"

Example of an Action Item:

"Dave wanted to impress his new girlfriend Marie. He was cooking her a meal, but had never cooked before. Marie hoped it would be successful. Dave told her he had spent all day preparing it. When it came out of the oven it was badly burnt. Marie ate all her meal. Afterwards she took a second helping of the food."

***Question:* "Why did Marie take a second helping?"**

Scoring:

(2) Any answer indicating that she wanted to support Dave or not hurt his feelings
e.g. "She was impressed with his effort."
e.g. "She didn't want to upset him."
e.g. "To be kind/polite/supportive to Dave."

(1) Any partial answer that does not quite clarify the issue, including those that do not show that she did not like the food
e.g. "To help him out."
e.g. "To be polite"

(0) Any clearly incorrect or irrelevant answer
e.g. "She hoped the second helping would be better than the first"
e.g. "She liked burnt food"

7.3 RESULTS

7.3.1 Data analysis

Means and standard deviations (SD) for each of the measures of the 'Social Favours' Task are presented below in Table 3, and for each of the additional measures these are shown in Table 4. A significance level of .05 was adopted, with a stricter level for post hoc tests to control for multiple comparisons. The assumptions of normality were met and thus parametric analyses were performed.

7.3.2 The 'Social Favours' task

7.3.2.1 Likelihood of helping ratings:

A repeated measures 2 x 2 ANOVA was conducted to examine ratings for likelihood of helping for all scenarios. There was one between-participant factor (AQ group: high versus low AQ) and one within-participant factor (strength of the justification for help: strong versus weak). There was a significant main effect of condition, $F(1,41) = 143.51$, $p < .0001$, whereby all participants were more likely to behave pro-socially when the justification was strong versus weak. This is in line with the prediction that the strength of the justification to help would enhance the likelihood of participants complying with it. However, there was no significant main effect of group, $F(1,41) = .201$, $p = .657$, nor a significant condition by group interaction, $F(1,41) = .001$, $p = .993$. This pattern of findings goes against predictions that the high AQ group would either behave less pro-socially overall, or that their behaviour would be differentially affected by the strength of the justification for help.

7.3.2.2 Acceptability ratings

A repeated measures 2 x 2 ANOVA was also conducted to examine ratings for acceptability for all scenarios. Again, there was a significant main effect of condition, $F(1,41) = 260.22$, $p < .0001$, whereby all participants rated requests as more acceptable when the justification was strong versus weak. This is in line with the predication that the strength of the justification to help would enhance the likelihood of participants complying with it. However, again there were no group differences, since there was no significant main effect of group, $F(1,41) = .360$, $p = .552$, nor a significant condition by group interaction, $F(1,41) = .184$, $p = .670$. Once more, this pattern of findings goes against predictions that the high AQ group would either find requests less acceptable overall, or that their judgments would be differentially affected by the strength of the justification for help.

7.3.2.3 Personal sacrifice ratings

Finally, a repeated measures 2 x 2 ANOVA was conducted to examine the ratings of personal sacrifice incurred for complying with characters' requests across all scenarios. As before, there was a significant main effect of condition, $F(1,41) = 43.33$, $p < .0001$, whereby all participants rated compliance with requests with weak justifications as incurring a higher degree of personal sacrifice. This is in line with the prediction that the strength of the justification to help would affect how much participants perceived a personal sacrifice for complying with it. However, this measure again failed to differentiate the groups; there was no significant main effect of group, $F(1,41) = 1.12$, $p = .296$, and the condition by group interaction was also not significant $F(1,41) = 3.23$, $p = .080$. This goes against predictions that the high AQ group would perceive a great degree of personal sacrifice, at least in the condition where justifications were weaker.

Table 3: Mean percentage scores and standard deviations for all measures for the 'Social Favours' task.

	Low AQ group (N = 23) M (SD)		**High AQ group** (N = 21) M (SD)		**Significance** (p = .05)
Likelihood of helping (%)					Condition *
					Gp NS
					Gp x condition NS
Strong	78.18	(9.14)	76.70	(11.60)	—
Weak	55.05	(14.48)	53.60	(13.28)	—
Acceptability (%)					Condition *
					Gp NS
					Gp x condition NS
Strong	80.47	(10.15)	79.65	(11.39)	—
Weak	47.43	(14.25)	44.80	(10.22)	—
Personal Sacrifice (%)					Condition *
					Gp NS
					Gp x condition NS
Strong	43.91	(14.77)	42.40	(10.93)	—
Weak	56.09	(15.83)	49.35	(11.65)	—

** Significant at p= .05* *NS = not significant*

7.3.3 Additional measures

7.3.3.1 Executive functions

'Colour-Word Interference' Test

For comparisons on the three conditions of the 'Colour-Word Interference' test a strict significance level of $p = .017$ ($p = .05/3$) was adopted to control for multiple comparisons. The first two control conditions did not involve any inhibition, but were simply included to gauge overall performance speed and any perceptual difficulties that might impair performance. There was no difference between the groups when they required to simply name blocks of colours, $t(41) = 1.005$, $p = .320$. However, the high AQ group were significantly slower when reading colour words in black ink than the low AQ group, $t(41) = 3.33$, $p = .002$. For the last condition, there was an interference component requiring participants to inhibit responses when naming the colour of the ink rather than a conflicting colour word (e.g. RED printed in green ink). As predicted, the high AQ group were significantly slower to complete this task, $t(41) = 2.63$, $p = .012$.

'Verbal Fluency' Test

For comparisons on the two conditions of the 'Verbal Fluency' test a strict significance level of $p = .025$ ($p = .05/2$) was adopted to control for multiple comparisons. This test involves generating words within a time limit according to certain rules. In contrast with predictions, this test did not reveal any significant group differences in either of the two conditions: letter fluency $t(41) = 1.61$, $p = .116$; category fluency $t(41) = 1.58$, $p = .122$.

7.3.3.2 'IRI'

A series of t-tests were conducted to compare the high and low AQ groups scores for the 'IRI'. Using a strict significance level of .0125 in view of the four empathy measures, the high AQ group gave significantly lower ratings on two of the subscales: 'Fantasy' $t(41) = 2.503$, $p = .016$, and 'Empathic Concern', $t(41) = 2.893$, $p = .001$. However, the groups did not differ on the two remaining two subscales: 'Perspective-taking' $t(41) = .528$, $p = .600$ and 'Personal Distress' $t(41) = 1.680$, $p = .100$.

7.3.3.3 'MIT'

T-tests were conducted to compare the high and low AQ groups' performance on the 'MIT', using a strict significance level of .025 (p = .05/2). The groups differed significantly for 'MIT Action' score, $t(41) = 2.76$, $p = .009$, but not for the 'MIT Sarcasm' score, $t(41) = 2.043$, $p = .048$.

Table 4: Mean scores and standard deviations for additional measures

Measure (range)	**Low AQ group** (N = 23) M (SD)	**High AQ group** (N = 21) M (SD)	**Significance**	**Effect Size** (d)
DKEFS				
Colour Word Interference Scaled Scores (1-19)				
Colour Naming	11.48 (2.63)	10.67 (2.73)	.320	—
Word Reading	12.78 (1.57)	11.00 (1.98)	.002**	0.3
Inhibition	12.95 (1.85)	11.38 (2.13)	.012**	0.84
Verbal Fluency Scaled Scores (1-19)				
Letter Fluency	12.35 (2.79)	10.95 (2.96)	.115	—
Category Fluency	12.43 (2.37)	11.14 (3.03)	.122	—
IRI				
Perspective-taking (0-32)	19.96 (4.23)	19.23 (4.17)	.600	—
Fantasy (0-28)	19.60 (4.88)	15.23 (6.64)	.016**	0.75
Empathic concern (0-28)	22.17 (2.91)	16.14 (6.77)	.001**	1.10
Personal distress (0-24)	10.13 (4.05)	12.57 (5.52)	.100	—
MIT Scores				
Sarcasm (0-20)	8.70 (1.05)	8.05 (9.44)	.048	—
Action (0-20)	9.18 (1.19)	7.95 (1.70)	.009**	0.83

**Significant at p = .05 ** Significant at p = .0125*

7.3.4 Correlations between performance on 'Social Favours' task and additional measures

Within-group Pearson correlations were conducted for additional measures that significantly differentiated the groups with performance on the 'Social Favours' task. A strict significance level of p = .01 was adopted to control for multiple comparisons.

For the executive measures, neither of the two 'Colour-Word Interference' test conditions that previously revealed group differences (word reading or inhibition) were found to significantly correlate with performance on the 'Social Favours' Task for either the high or low AQ groups.

Turning now to associations for 'IRI' subscale scores, within-group correlations were performed for the two subscales that significantly differentiated the groups ('Fantasy' and 'Empathic Concern') and the three ratings of the 'Social Favours' task. In the strong justification condition, likelihood ratings, r = .58, p = .004 and acceptability ratings, r = .61, p = .002, were positively correlated with 'Empathic Concern' for the low AQ group. This pattern was replicated by the high AQ group, whereby 'Empathic Concern' was positively correlated with both likelihood ratings, r = .64, p = .002, and acceptability ratings, r = .55, p .010 in the strong justification condition. There were no significant correlations between subscale score and any of the ratings in the weak justification condition for either the high or the low AQ group. For the 'MIT, the 'MIT Action' score, which previously revealed group differences, was not found to significantly correlate with performance on the 'Social Favours' Task for either the high or low AQ groups.

7.4 DISCUSSION

7.4.1 Summary of findings

The present experiment investigated the extent to which characters' deservingness affected subsequent pro-social behaviour, and social judgments of acceptability and personal sacrifice in participants with high versus low levels of autistic traits. Deservingness was directly manipulated, comparing compliance with characters' requests for favours when there was a strong versus weak justification provided. The manipulation worked as expected; all participants gave ratings indicating that they were more likely to help and perceived requests to be more acceptable when justifications were strong rather than weak. Moreover, all participants indicated that complying with these requests with a strong justification led to a reduced subjective sense of personal sacrifice when compared to weak justifications, even though the actions needed to help characters in both conditions were the same. However, contrary to predictions there were no group differences, nor any group by condition interactions for any of these measures. This is in contrast to the findings from the 'Above and Beyond' task, where the high AQ group was found to be less pro-social and to experience less satisfaction for helping others. Findings from the 'Social Expectations' task also indicated that the high AQ group would be less pro-social and that they might be differentially affected by the characters' deservingness.

A number of additional measures were also included, to explore whether executive functions and different aspects of empathy differed across the groups. As predicted, in the present study the 'Colour-Word Interference' test of inhibition revealed some evidence of executive weakness in the high AQ group relative to the low AQ group. However, the findings from this measure are difficult to interpret, since although the high AQ group performed worse on the experimental inhibitory condition they also performed worse on one of the two control conditions (word reading), suggesting that they may have had a slower reading speed or visual information processing speed overall rather than signaling an executive weakness. Moreover, there was no evidence of executive weakness on a different executive test, 'Verbal Fluency', where the groups performed at similar levels.

Two tests of empathy were also included: a self-report questionnaire that explored four different dimensions of empathy (the 'IRI'), and a test of cognitive empathy that involved interpreting others' intentions from either sarcastic comments or their actions (the 'MIT'). Inspection of the 'IRI' subscales revealed that the pattern was somewhat complicated: the high AQ group had lower scores for only one of the two measures which tap processes related to cognitive empathy (they differed for the 'Fantasy' but not the 'Perspective-taking' subscale), but also had lower scores for one of the two measures that are thought to be driven by emotional empathy (they differed for the 'Empathic Concern' but not the 'Personal Distress' subscale). As predicted, the high AQ group gave less detailed/accurate answers for the 'MIT'

when compared to the low AQ group. However, once again this finding was complicated, since the high AQ group performed worse for items involving action interpretation; the groups did not differ when interpreting sarcastic items.

Within-group Pearson correlations were also conducted between performance on the 'Social Favours' task and additional measures that significantly discriminated the groups, namely the 'Colour-Word' interference test of inhibition, the 'MIT' cognitive empathy measure, and the 'IRI' subscales: 'Fantasy' and 'Empathic Concern'. No significant associations with social performance were revealed for either the executive test of inhibition ('Colour-Word Interference'), nor for the test of cognitive empathy ('MIT'), for either the high or low AQ groups. 'No significant correlations were identified for the "IRI" 'Fantasy' subscale, but the 'Empathic Concern' subscale was positively associated with likelihood of helping ratings and acceptability ratings in the strong justification condition for both groups. This indicates that a higher score for 'Empathic Concern' was correlated with a greater likelihood of helping characters and a perception of the request as more acceptable, when the justification was strong.

7.4.2 Executive functioning in ASD

As discussed above, executive dysfunction has been postulated in ASD, although the evidence relating to this is mixed (see Chapter 2.2.2). This inconsistent pattern of findings is potentially complicated by a lack of agreement regarding the extent to which executive functions are separable. The traditional accounts of executive functions hypothesise that this is managed by a single structure (the central executive; e.g., Baddeley, 1996), but more recent work suggests that these may consist of up to eight discrete processes (Shallice & Burgess, 1996). In support of this, there is evidence of double dissociations for impairment on different executive tasks (e.g., Tsuchida & Fellows, 2013). Functional imaging and lesion studies provide evidence for some degree of functional specialisation in the prefrontal cortex for separable executive processes (e.g. Fletcher & Henson, 2001; Stuss & Levine, 2002).

With respect to the present task, the high AQ participants were found to differ on some conditions of the 'Colour-Word Interference' test: the high AQ group performed worse for the inhibition condition, but also for a control condition. Thus, it is difficult to establish if there is any evidence of impoverished executive functioning in the high AQ group on the basis of this test alone. Turning to previous literature, the inhibition of a pre-potent response has been extensively examined in ASD using the 'Colour-Word Interference' test, and also via other inhibitory measures such as the 'Windows' task (participants can only win a desired object by pointing away from it; Russell, Mauthner, Sharpe, & Tidswell, 1991) and 'Hayling' task (participants have to complete sentences but must make them nonsensical by generating word that are unconnected to the sentence stem; Burgess & Shallice, 1997) yielding mixed results (for a review see: Hill,

2004). This discrepancy may reflect the extent to which tests of inhibition may also tap other types of executive skills such as working memory or verbal fluency (Hill & Bird, 2006).

With regard to the 'Verbal Fluency' test, in the present study the high AQ group were found to show intact performance on this measure. Whilst there is more limited work examining this particular function in people with ASD, this again reveals a mixed picture, with some studies finding evidence for impaired fluency (e.g. Spek, Schatorjé, Scholte & Berckelaer-Onner, 2009; Geurts, Verté, Oosterlaan, Roeyers & Sergeant, 2004) and others finding no group differences (e.g. Boucher, 1998). It has been argued that tests designed to measure verbal fluency may not be 'pure', as they may involve drawing upon other executive skills, e.g. working memory and cognitive flexibility (e.g. Rende, Ramsberger, & Miyake, 2002).

As a result of the conflicting evidence, and the lack of agreement concerning the purity of the measures used, it is difficult to determine the nature and extent of any executive impairment in ASD (Hill, 2004). Moreover, very little work has examined executive function in the broader autistic. Thus, in the present study the findings that the high AQ group differed on one executive test but not another may reflect the different sensitivities of the tests chosen, or relative strengths and weaknesses in this subclinical group. Even if the findings for the 'Colour-Word Interference' test are interpreted as an indication of relative executive weakness, rather than slower psychomotor speed, in the high AQ group, it is noteworthy that the scores of the high AQ group were still within the average range for all conditions on both measures, with scaled scores of 10-13, i.e. between the 50-84th percentile. It is possible that, due to their abstract nature, traditional laboratory tests of executive functions may be unable to recreate effectively the complex and multi-faceted demands of real-life situations (Channon & Crawford, 2010). Thus, the tests of inhibition and verbal fluency used here might not be sufficiently sensitive at capturing any potential deficits in these domains, especially when any impairment is likely to be subtle, such as for the high AQ group.

7.4.3 Cognitive empathy in ASD

As discussed above, ASD is hypothesised to involve impaired cognitive empathy, with some evidence for intact emotional empathy (see Chapters 2.3.1 and 2.3.2). Cognitive empathy can be assessed using a number of possible approaches, and thus in the current study a self-report questionnaire measure ('IRI') and a scenario-based test requiring mentalistic interpretations ('MIT') were both included. The 'IRI' produced a complex set of findings, whereby the groups differed only for the 'Fantasy' subscale but not the 'Perspective-taking', although both are thought to tap cognitive aspects of empathy. The 'Fantasy' subscale is described as examining respondents' "tendencies to transpose themselves imaginatively into the feelings and actions of fictitious characters in books, movies, and plays". Imagination is thought to be impaired in ASD, at least for social stimuli (Eycke & Muller, 2015), and so a lower score on this measure for the high AQ group is consistent with what one might expect to see in the broader autistic phenotype. The

'Perspective-taking' subscale is thought to examine 'the tendency to spontaneously adopt the psychological point of view of others', which is a crucial aspect of cognitive empathy and well known to be impaired in ASD. Previous work examining the 'IRI' has revealed consistently lower scores for the cognitive subscales of this measure in those with ASD versus controls (Rogers, Dziobek, Hassenstab, Wolf & Convit, 2007). A lack of difference on the 'Perspective-taking' subscale may indicate relative preservation of this aspect in the high AQ group, suggesting that the broader autistic phenotype is not associated with perspective-taking difficulties. However, when this is considered in the context of the findings from the 'MIT', where high AQ group performed worse overall when interpreting characters' intentions, it does not seem plausible that perspective-taking is intact in the high AQ group. There is also previous evidence using different measures that suggests that this skill may be poorer in those with high versus low autistic traits (Freeth et al., 2013; Gökçen et al., 2014).

Although the high AQ group performed worse for items involving action-interpretation ('MIT Action' score), the groups did not differ for items involving sarcasm-interpretation items. Several previous studies have used variants of the 'MIT' task to explore sarcastic versus action interpretation in individuals with frontal brain damage, revealing consistent deficits across both sets of items (Channon et al., 2005; 2007; Channon & Crawford 2010). One study explored performance on the 'MIT' task in people with ASD, finding that this group scored significantly lower for both sets of items relative to controls when generating explanations for characters' intentions (Channon, Crawford, Orlowska, Parikh & Thoma, 2014). Thus, when compared with these clinical groups the high AQ group appear to show milder difficulties, performing worse only for action items relative to the low AQ group.

Why might this be the case? Other studies of mentalising have often used mixed item sets including material such as sarcasm, lies, jokes, double bluff and so on (e.g. Happé, 1994; Joliffe & Baron-Cohen, 1999) and have tended to examine understanding of words and facial expressions rather than actions (Mcdonald, Flanagan & Rollins, 2011; Mcdonald, Flanagan, Rollins & Knich, 2003). It is not clear how these processes may differ with respect to executive demand, or which particular skills they tap. It is possible that interpreting other's actions is more demanding, or that these items were more sensitive to any potential differences between the high and low AQ groups. Although sarcasm is a relatively complex, non-literal form of language, understanding others' actions is part of everyday social communication, but being asked to explain actions explicitly may well be less familiar than being asked to explain others' words. Regardless, there appears to be some evidence of difficulty with cognitive empathy in the high AQ group, albeit milder than that reported in a clinical sample by Channon et al., (2014).

7.4.4 Emotional empathy in ASD

Deficits in cognitive empathy have been consistently identified in ASD, whereas there is mixed evidence regarding whether emotional empathy is intact (Blair, 2008). For instance, previous studies examining the 'IRI' in participants with ASD have revealed inconsistencies for the emotional subscales, with some evidence for intact performance and some evidence of lower scores (see Aaron, Benson & Park, 2015 for a discussion of this issue). Thus, it is not altogether surprising that the high AQ group showed differences on one of the two emotional subscales of this measure, scoring lower for the 'Empathic Concern' subscale but not for the 'Personal Distress' subscale of the 'IRI', both of which are thought to tap emotional empathy.

Why might the high AQ group score lower for the 'Empathic Concern' subscale? Interestingly, although both the 'Empathic Concern' and 'Personal Distress' subscales were designed to examine emotional aspects of empathy, the 'Empathic Concern' subscale items tend to relate to other-oriented concern e.g. "Sometimes I don't feel very sorry for other people when they are having problems." (please note this item is reversed), and the 'Personal Distress' scale items to self-oriented concern e.g. "When I see someone who badly needs help in an emergency, I go to pieces." Therefore, it could be argued that the 'Empathic Concern' subscale does not tap 'pure' emotional empathy, since this might involve invoking others' emotional states, which may draw upon cognitive as well as emotional processes. Thus, a lower score for the high AQ group on this measure might reflect cognitive rather than emotional empathic differences.

Whilst cognitive and emotional empathy may in principle constitute separable processes, the extent to which these can be delineated in everyday life is not well understood. It is likely that they are used in concert, and thus fractionating these for the purposes of assessment can be challenging. Mental state calculations are likely to occur spontaneously during performance on tasks such as the 'Social Favours', even when they have not been primed. Moreover, the 'IRI' and other self-report questionnaires of empathic functioning may not fully represent empathic abilities because of their limited ecological validity (Dziobek et al., 2008). A different approach for examining emotional empathy would be to employ physiological measures to examine responsivity when witnessing others in pain or classifying their emotional expressions. These types of measures of emotional empathy are more implicit in nature, negating the risk of spontaneous cognitive contamination in an explicit measure such as the 'IRI' questionnaire. However, relying on physiological measures of responsivity is still an indirect method of assessing emotional empathy that is subject to interpretation, and cannot readily be compared with other psychological measures (Dziobek et al., 2008; Thoma & Bellebaum, 2012).

7.4.5 Contributions of executive dysfunction to performance on the 'Social Favours' task

As outlined in Chapters 5 and 6, executive dysfunction is one possible contributory factor that might lead to a failure by the high AQ group to appreciate and effectively evaluate the relevant details of the scenarios used in these studies, resulting in impulsive, insensitive or inappropriate social behaviour. Any executive difficulties could thus have resulted in the high AQ group behaving in a less pro-social manner and/or making harsher judgments on any of the social tasks examined. With respect to the present study, the groups, however, behaved in a similar manner on all aspects of the 'Social Favours' task, and hence any executive contribution is difficult to evaluate. The executive measures revealed some evidence of executive dysfunction in the high AQ group, although the findings indicated only subtle differences between the groups at best. Thus, even if there is a link between executive functioning and performance on the 'Social Favours' task, any executive difficulties in the high AQ group may have not been sufficient to lead to impaired social performance on this task.

Within-group Pearson correlations also failed to reveal any significant associations between the executive tests and the 'Social Favours' task ratings. However, these findings should be interpreted with caution. Correlations on the basis of small samples are very unstable and may be particularly affected by sources of variability such as measurement error (Shilling, Chetwynd & Rabbitt, 2002). Shilling et al., (2002) found that correlations between performance on neuropsychological tasks became more reliable as the idiosyncratic differences between tasks were reduced. In the present study, the different additional measures used (executive, emotional empathy and cognitive empathy) were not matched to the 'Social Favours' task in terms of either the social or non-social cognitive demands, and it is therefore difficult to extrapolate what a positive or negative association may represent.

Could executive dysfunction have underpinned performance difficulties on the social tasks used in Chapters 5 and 6? Unfortunately, a direct comparison is not possible as the additional executive measures were not included in those studies, and thus establishing the relative strengths and weaknesses of the samples is not possible. However, the samples were all drawn from similar, relatively homogeneous, populations of university students from the same university, and it is therefore unlikely that their profiles would differ substantially. Although it is difficult to rule out mild executive dysfunction as a contributory factor, it seems improbable that differences in the executive functioning of the samples for the three experiments could fully explain the pattern of a lack of group differences on the 'Social Favours' task with significant differences on the 'Above and Beyond' and 'Social Expectations' tasks.

It is also difficult to see how the executive demands of the 'Above and Beyond' or 'Social Expectations' tasks might have differed from those of the 'Social Favours' task, at least for the rating measures. The relationship between executive functioning and social behavior is not well delineated or understood, and rather executive dysfunction has been postulated to account for social deficits in ASD in a

generic rather than specific way, such as at the level of attention to detail. Thus, it is difficult to make clear predictions or provide post-hoc rationales for why one might expect to see differences in executive demands between the tasks. Thus, there is no strong reason to suspect differences in either the samples or the tasks themselves that support an account in terms of executive dysfunction.

7.4.6 Contributions of empathy to performance on the 'Social Favours' task

ASD is thought to be characterised by difficulties in understanding others, but a relatively intact ability to resonate with them (i.e. impaired cognitive with intact emotional empathy) (Blair, 2008). Impaired cognitive empathy is thus another potential contributory factor that might have resulted in a lack of understanding of the characters' needs or expectations on the 'Social Favours' task. This in turn might be expected to lead to reduced pro-social behavior on the part of the high AQ group and/or more 'black and white' distinctions between weak and strong rationales, as a result of failure to fully appreciate the characters' perspectives.

With respect to emotional empathy, as expected, the groups did not differ on one of the two 'IRI' emotional subscales, 'Personal Distress'. There was, however, a significant group diference on the other emotional subscale of the 'IRI' significantly differentiated the groups ('Empathic Concern'), and this measure was also positively correlated with likelihood of helping and acceptability ratings for the 'Social Favours' task. However, as discussed above (see Chapter 7.4.4), it is by no means certain that this indicates altered emotional functioning, since 'Empathic Concern' may in fact reflect differences in cognitive rather than emotional empathy, since the items tend to focus primarily on other- versus self-oriented concern. In contrast the 'Personal Distress' subscale could be argued to have greater face validity as a measure of emotional empathy, since the items seem to tap self- versus other-oriented concern. An interpretation of the 'Empathic Concern' scale in terms of a deficit in cognitive empathy in the high AQ group is consistent with the finding that the 'MIT' significantly differentiated the groups, suggesting impairment in cognitive empathy in the high AQ group, or at least relative weakness in comparison to the low AQ group, although it did not correlate with performance on the 'Social Favours task'. The 'MIT' is potentially the most sensitive and relevant test of cognitive empathy for examining the difficulties associated with everyday functioning that this study aims to explore.

Despite apparent differences in cognitive empathy, the performance of the two groups did not differ significantly across both conditions for all measures on the 'Social Favours' task. Could deficits in cognitive empathy have underpinned performance difficulties on the social tasks used in Chapters 5 and 6? As discussed (see Chapter 7.4.5), a direct comparison is again not possible since these measures were not included in the previous experiments, but there is no reason to believe that these samples would have differed substantially from the present ones. It is likely that the previous samples would also have shown

group differences on cognitive empathy, and this may well have contributed to performance on those social tasks. Overall, it seems likely that the high AQ group did struggle to fully identify the characters' perspectives, but for some reason the present task was not sufficiently sensitive to reveal these difficulties. This possibility will be further explored in section 7.4.7.

7.4.7 Sensitivity of the 'Social Favours' task

When comparing the results of Chapters 5 and 6 to those of the present study, one should also consider how the social task designs differed between the experiments. Were the scenarios used in the 'Social Favours' task less sensitive to potential group differences than the 'Above and Beyond' and 'Social Expectations' tasks? It is possible that the scenarios were not sufficiently demanding to reveal any differences between the groups. For instance, in the example scenario where the participant has to decide whether to carry a parcel for a neighbour, the action itself does not require much effort on the participants' behalf. Thus, regardless of the strength of the justification provided for making the request, the request itself might be too trivial to elicit group differences in pro-social behaviour ratings, or in judgments of the acceptability of the request or sense of personal sacrifice involved in helping characters. People high in AQ traits might only behave less pro-socially or make altered judgements once a certain effort threshold has been reached, and this task may not have been powerful enough to breach it. If this explanation were to account sufficiently for the lack of differences in pro-social behaviour or judgments found in the present study versus those identified in the 'Above and Beyond' and 'Social Expectations' tasks, then the actions required in the scenarios described in previous tasks should be substantially more effortful. However, examination of the scenarios across tasks suggests that in each case, choosing to help the characters involved expending a similar amount of effort to carry out the actions (e.g. 'Above and Beyond' task: pausing to help a man who had fallen over in the street; e.g. 'Social Expectations' task: giving up your seat for someone who might be more needy).

On the other hand, whilst the scenarios might have been matched in terms of the effort required to carry out the actions themselves, the extent to which there was a personal cost incurred by the participant if they decided to comply may have varied across tasks. In the 'Social Favours' task, compliance with any of the requests appeared to carry only a small personal cost (e.g. take a few minutes from whatever they were doing to carry a parcel for a neighbour). By contrast, in the 'Above and Beyond' and 'Social Expectations' tasks used in Chapters 5 and 6 the scenarios more explicitly spelt out that there would be some kind of cost incurred for complying, and how this would impact on the participant. For instance, in the 'Above and Beyond' task it was made clear that if the participant chose to stop and help a man who had fallen over in the street, they would be late for a work meeting. Moreover, in the 'Social Expectations' task it was implied that giving up your seat to another passenger would incur a cost, since it was made clear that there were no

alternative seats available. It could be argued that in the 'Social Favours' task, carrying the parcel for the character might also incur a cost – e.g. carrying it could cause an injury to the participant if it was very heavy – but this was not made so clear within the scenario. The costs incurred in the other tasks typically involved either a negative consequence for the participant that was clearly spelt out (e.g. being late for a meeting), or it was made clear that they had to sacrifice something concrete (e.g. their seat). Occasionally the cost incurred was simply lost time, as in the example given from the present experiment, but in the 'Above and Beyond' or 'Social Expectations' task (but not the present one) the scenarios made it obvious that losing time would incur a further cost, rather than simply being inconvenient (e.g. one scenario referred to losing one's place in the queue for the loo to help a character, thereby missing the beginning of a movie). Therefore, in the present task the cost incurred might not have been considerable or explicit enough to reveal the high AQ group's less pro-social tendencies. This explanation could also account for the intact ratings of the acceptability of the request; without a considerable personal cost, the high AQ participants may have felt that both the strong and weak requests were acceptable to some extent, and distinguished between them appropriately. It could also explain the lack of differences observed on the personal sacrifice ratings; if the task did not sufficiently spell out the extent of the cost incurred or indicate a cost that was of a high enough threshold then the high AQ participants may simply have responded appropriately to the mild inconvenience experienced. Had more substantive or explicit costs been involved the high AQ group may have found complying with the requests less acceptable and perceived them to incur a bigger personal sacrifice.

In addition to the extent of effort and/or cost involved in helping characters across the different tasks, one further consideration is the manner in which their need for help was made clear to the participant. In the 'Social Favours' task, the characters made direct requests for help to the participant. For instance, in the example scenario the participant is asked "Would you carry [this parcel] upstairs for me?" with either a strong justification, "I've strained my back", or weak justification "I'm going out and don't want to get my shirt dirty" for doing so. By contrast with this, in the 'Above and Beyond task' the participant is never directly asked a request by the character. For instance, in the example scenario it is implied that the character is in need of the participants' help after 'falling down heavily on the pavement', but a direct request is not made. The manner of the request is mixed in the 'Social Expectations' task, since the character's request for a favour is direct in some of the scenarios but not in others, where the favour is implied but not directly requested in words. For instance, in one scenario the participant is directly asked by a character if they can borrow a mobile phone to make a call, either because the character needs to call his friend to change their arrangements for tonight, or because the character needs to call his brother to let him know their mother is in hospital (ambiguous versus clear-cut expectation to help). By comparison, in the example scenario it is implied that the character needs a chair, either because they are 'a young adult

carrying a large parcel' or because they are 'elderly and walking with a stick' (ambiguous versus clear-cut expectation to help). Direct requests may have either drawn the participants' attention to the characters' needs, making them appear more salient, or made it harder for participants to refuse. This may have cued the high AQ participants, resulting in similar behaviour to their low AQ counterparts.

One final factor that was not examined directly in this task is the extent to which the high AQ group may have used social knowledge to conceal any performance differences. As discussed in Chapter 6, people with ASD may have more limited social knowledge stores than others, but may also rely upon explicit knowledge about social rules and norms and apply it deliberately to compensate for their less intuitive behaviour. This may to some extent mask social difficulties, although perhaps only at a surface level. In tasks where a deeper level of understanding is required, such as generating one's own responses or providing justifications for choices, people with ASD have been found to show difficulties, despite showing intact performance on more simple measures of understanding such as rating measures, selecting between alternatives or making yes/no judgments (Channon et al., 2001; Channon et al., 2014; Loveland et al., 2001; please see Chapter 2.3.3 and 3.1 for a discussion of this). Thus, it is possible that the high AQ group used task cues to access knowledge they had built up over time to behave in a similar manner to the low AQ group. On the 'Social Favours' task, the contrast between the two conditions was made salient since the extent to which the character was deserving of help was intended to be clear to the participants. However, the manipulation of strong versus weak justifications for helping others may have been so salient that it primed the high AQ group to behave in an appropriate way. It is possible that the inclusion of a larger range of more graded types of justifications, with nuanced rather than absolute differences between them, then group differences might have been revealed. Furthermore, on the 'Social Expectations' task, high AQ participants gave similar ratings to the low AQ group for character expectations, but produced rationales that were of a poorer quality when they were asked to explain these expectations. Accordingly, had the 'Social Favours' task included a qualitative measure to probe deeper into participants' understanding of the 'acceptability' of the request, differences between the groups might have been revealed, despite the fact they showed intact performance on the ratings measure. Another possible way of improving task sensitivity might be to frame the personal sacrifice rating from the perspective of the character, rather than from that of the participant, i.e. "How much would the character would appreciate your help?".

7.4.8 Is the understanding of 'deservingness' intact in ASD?

It is also possible that the high AQ group in fact had an intact understanding of character 'deservingness', reflected in their similar judgments to the low AQ group on the 'Social Favours' task. The 'Social Favours' task explored how the strength of characters' justifications of 'deservingness' influenced participants' pro-social behaviour and judgment. Although this task primarily explored social behaviour, the manipulation concerned a central question that was essentially moral in nature, "Does this character deserve my help?". In contrast, the 'Above and Beyond' and 'Social Expectations' tasks used in Chapters 5 and 6 probed more socially-oriented questions, such as "How far are you willing to go to help others?" and "Is there a clear social rule underpinning whether I should help this person?", exploring the lengths to which participants might go to help others, and how the clarity of the social rule might impact upon pro-social behaviour and judgment. Notably, all the characters in the 'Above and Beyond' and 'Social Expectations' tasks were deserving of the participants' help, and the tasks focused on how and when people should or would prioritise others' needs over their own rather than questions of merit in relation to receiving help.

Moral judgment and reasoning have not been widely investigated in ASD (Leslie et al., 2006), with most tasks focusing on simple 'wrong versus right' judgments in children and adolescents. This limited evidence is mixed regarding whether this moral judgment and reasoning is intact or not (e.g. Grant et al., 2005). The relationship between cognitive versus emotional processes and moral judgment is also not well delineated, and it is likely that both executive skills and empathic processes are involved. Indeed there is some evidence that children with ASD who fail false-belief tasks (thought to draw upon cognitive empathic skills) make accurate moral judgments (Blair, 1996; Leslie et al., 2006). Thus, it is entirely possible that those with high AQ traits are able to make perform in a similar manner to others when called upon to make judgments but not to explain the underlying rationales, such as in the 'Social Favours' task, regardless of the whether executive functioning and cognitive empathy are intact or impaired. Therefore, the finding of intact performance on this 'deservingness' task versus the 'Above and Beyond' and 'Social Expectations' tasks may be meaningful, and this will be further explored in the following experimental Chapters addressing moral judgment and reasoning.

7.4.9 Conclusion

In summary, the present study explored how deservingness might impact upon the likelihood of helping others and social judgment in people with high versus low AQ scores. This was examined using a scenario-based task in which characters made requests of participants with either a strong or a weak justification for doing so. All participants were influenced by the manipulation in the expected direction, whereby in scenarios with strong versus weak justifications they were more likely to help the characters and to find their requests more acceptable. Interestingly, participants rated requests with weak justifications as incurring a greater personal sacrifice than those with strong justifications (e.g. not wanting to get their shirt dirty versus having a strained back), despite the fact that the favours requested did not differ (e.g. carry parcel).

Contrary to predictions, this task did not differentiate the groups, with similar performance on all ratings measures for both of the justification conditions. This is in contrast to the findings of the 'Above and Beyond' and 'Social Expectations' tasks, where high AQ participants were less pro-social and differed on a number of measures exploring their judgments of characters and contextual factors. Participants with high AQ scores were found to display mild impairment on tests of executive functions and cognitive empathy relative to their low AQ counterparts. Whilst differences in the participants included across the different experiments might plausibly explain the lack of consistency of the results, after careful consideration it seems more likely that this can be explained by differences in the nature of the tasks used in Chapter 5 and 6 versus 7.

It also remains possible that people with high levels of autistic traits do in fact have intact understanding of 'deservingness'. The literature exploring moral judgment and reasoning in ASD is much more limited than that examining social functioning. The next study will seek to explore further the role of 'deservingness' with a new task that probes judgments for good and bad outcomes where the participant is not personally implicated.

PART 3: EXPERIMENTS B

CHAPTERS 8-11

MORAL REASONING

Chapter 8: Judgments of deservingness for positive and negative outcomes

Chapter 9: Counterfactual thinking for 'self' versus 'other' and the experience of regret and blame

Chapter 10: Counterfactual judgments for others' mistakes and the experience of regret, blame and guilt

Chapter 11: Utilitarian decision-making for physical and social harm

Chapter 8: Judgments of deservingness for positive and negative outcomes

8.1 INTRODUCTION

As discussed, very little work has previously explored perceptions of 'deservingness' or how it might be mediated by empathic processes, and it is therefore unclear how these issues might be influenced by autistic traits. In Chapter 7, contrary to predictions, no group differences were found when people with high versus low autistic traits had to decide how likely they would be to help someone who was either more or less deserving. There were also no group differences in participants' judgments of the acceptability of the request, or their ratings of the expected personal sacrifice they would incur for helping others. In order to further investigate this, the present Chapter focused on how deservingness influences social decision-making in situations that have either positive or negative outcomes for the characters involved. This experiment shifted the focus from exploring what participants would do if they were in the situation (behaviour), to examining judgments from the characters' versus participants' perspectives. This was intended to tap cognitive (i.e. appraising the characters' needs) versus emotional empathic processes (i.e. the extent to which the participant engages with their emotional state) that are postulated to be involved in judgments of deservingness.

Perceptions of social justice, where good deeds are rewarded and bad deeds are punished, are thought to be important for maintaining a stable environment (Furnham, 1998). Thus, people may be less likely to endorse or facilitate good outcomes for those who did nothing to deserve it. In contrast, when someone suffers a negative outcome that is undeserved, people are thought to be more motivated to help them and to alleviate their distress. There is evidence suggesting that people experience a range of emotional reactions in response to witnessing outcomes for others (Feather, 2006). When witnessing others receiving positive outcomes that are deserved, observers experience pleasure, but if the outcome is underserved, the observer experiences resentment (Feather, 2006). On the other hand, when witnessing others obtaining negative outcomes that are deserved, observers experience 'schadenfreude' (pleasure at another's misfortune), but if the outcome is underserved, then observers experience sympathy (Feather, 2006).

A previous experiment examined judgments of deservingness in both positive and negative outcomes, finding support for the theory that observers' emotions depended upon 'deservingness', as described above (Lupfer & Gringrich, 1999). However, this study used extreme situations, such as winning the lottery or receiving a diagnosis of cancer. The scenarios were mixed regarding whether the outcomes

were primarily for the main character or other characters, and it was also not always clear whether the characters' actions played a role as an antecedent to the outcome. A recent study explored how deservingness might influence social decision-making in people with high versus low levels of psychopathic traits. This used novel scenario-based-tasks involving real-life-type situations, such as receiving a high grade on a piece of work (positive task), or getting ill whilst travelling (negative task) (Vyas, Jameel & Channon, in preparation b). This study also resolved potential confounds identified in the experimental design of the study by Lupfer & Gringrich (1999), since it was made clear that the outcome was always for the main character, and the extent to which the characters' actions were causally linked to the outcome was carefully controlled. On these tasks, participants who were high versus low in psychopathic traits were found to differentiate less between outcomes for characters who were more or less deserving, regardless of whether the outcome was positive or negative. The authors suggested that this pattern of findings might reflect reduced sensitivity to social rule violations in participants high in psychopathic traits.

The present study was designed to investigate perceptions of deservingness and the extent to which witnessing others receive positive versus negative outcomes might or might not evoke personal reward in people with high versus low autistic traits. The tasks described by (Vyas et al., in preparation b,) were used, which manipulate deservingness via four conditions where the characters' actions were either congruent or incongruent with positive or negative outcomes.

8.1.1 Hypotheses

8.1.1.1 Character deservingness

Deservingness was directly manipulated, in order to compare judgments for scenarios when there was an outcome that was either deserved (congruent condition: e.g. pass driving test because you drove well and made few errors) or undeserved (incongruent condition: e.g. pass driving test because you flirted). Participants rated how much the character deserved the outcome, and it was expected on the basis of the findings of Vyas et al., (in preparation b) that participants would give higher ratings in the scenarios where characters' actions were congruent versus incongruent. It was also postulated that the groups might differ in their ratings of deservingness, either because the high AQ group might be less sensitive to the concept of deservingness as a result of limited understanding of intentionality, or alternatively that they might differentiate between the two conditions more than the low AQ group, showing 'black and white' thinking as described by Channon et al., (2010; 2011).

8.1.1.2 Participant satisfaction

Participants also rated how much personal satisfaction they themselves would derive from outcomes for the character. Here it was also expected that all participants would give higher ratings in scenarios where characters' actions were congruent versus incongruent, indicating that they were more pleased with the outcome when the character was 'deserving' of it. With respect to group differences, it was predicted that the high AQ participants would demonstrate lower personal satisfaction in relation to the outcomes for others, at least when these were positive. This prediction was made on the basis of the findings on the 'Above and Beyond' task in Chapter 5 where high AQ participants demonstrated reduced reward for helping others.

8.2 METHODS

8.2.1 Screening phase

8.2.1.1 Screening participants and procedure

The participants for this study also took part in the 'Social Favours' task and are described in Chapter 7.2.1.1. Thus, an opportunistic sample of 662 full-time university students (65.60% female) was recruited for the screening phase of the study. All participants were fluent in English and aged 18 or over (mean age 20 years old). All participants completed the AQ (Baron-Cohen et al., 2001) and total scores were calculated for the whole sample. Participants within the highest-scoring and lowest-scoring 10% of males and females were contacted and invited to take part in the second stage.

8.2.2 Experimental phase

8.2.2.1 Design

There was one between-participants factor of AQ group (high vs. low scorers) and one within-participants factor of deservingness (congruent vs. incongruent) for the two sets of materials (positive versus negative).

8.2.2.2 Experimental participants and procedure

The participants for this study also took part in the 'Social Favours' task and are described in Chapter 7.2.2.2. There were thus 20 (9 male, 11 female) participants in the high AQ group and 23 (12 male, 11 female) participants in the low AQ group. All participants were tested individually, and provided written informed consent before completing the two tasks: 'Deservingness: Positive Outcomes' and 'Deservingness: Negative Outcomes'. For each task there were two different counterbalanced orders for

the scenarios which were counterbalanced within each group; the positive and negative tasks were presented in counterbalanced order within each group.

8.2.3 The 'Deservingness' Tasks

8.2.3.1 Development of the 'Deservingness' tasks

The two tasks were designed, administered and analysed separately, since studying positive and negative outcomes raises different issues with respect to a) the causal chain of events leading to the outcome and b) the desirability of the outcomes (Vyas et al., in preparation b). The extent to which the chain of events leading to the positive or negative outcome could be readily related to the characters' actions varied across the task sets. In scenarios with positive outcomes, passing one's driving test stemmed from either good driving (good action, congruent with the outcome) or flirtation with the examiner (bad action, incongruent with the outcome). The congruent outcomes in the negative task were also based on the character's bad actions, for instance being reprimanded because of failure to complete the homework. However, there was no equivalent for incongruent outcomes on the negative task, since good actions are not expected to lead to negative outcomes i.e. completing one's homework would not reasonably be expected to lead to being reprimanded. Therefore, incongruent items for the negative tasks instead described situations in which negative outcomes arose from bad fortune rather than bad actions of the character's behalf (e.g. being reprimanded because the instructor was in a bad mood).

The second factor that differentiated the positive and negative tasks was the desirability of the outcomes from the viewpoint of the participant. Whilst the positive outcomes for both congruent and incongruent variants were desirable to the main character, only the congruent ones were likely to be desirable to the participant. Incongruent items, where bad actions led to positive outcomes, might well be considered undesirable to the participant, since the character did not 'deserve' the outcome and it might be viewed as unfair e.g. to pass their driving test first time by driving badly but flirting with the instructor. By contrast, in scenarios with negative outcomes both the congruent and incongruent actions may have been considered undesirable by the participant, since these always involved suffering or disappointment for the main character. The differences between the tasks are outlined in Figure 8 below.

Figure 8: Comparison of the Deservingness Tasks: Positive and Negative Outcomes

	Deservingness: Positive Outcome	**Deservingness: Negative Outcome**
Congruent Outcome	**Example**: Passing driving test; making few errors and driving well **Desirability**: High **Basis:** Good action	**Example**: Reprimanded in front of class, failed to do homework **Desirability**: Low **Basis:** Bad action
Incongruent Outcome	**Example**: Passing driving test; making many errors and flirting with instructor **Desirability**: Low **Basis:** Bad action	**Example**: Reprimanded in front of class, instructor in a bad mood **Desirability**: Low **Basis:** Bad fortune

8.2.3.2 'Deservingness: Positive Outcomes'

'Deservingness: Positive Outcomes' (Vyas et al., in preparation b) was designed to examine how people evaluate positive outcomes on the basis of deservingness. The task consisted of five short scenarios describing a positive outcome for a character known to the participant, such as passing a driving test on their first attempt. Each scenario was presented twice; in the first variant, the characters' actions were either congruent with the outcome and it was evident that they had earned the positive outcome through their own efforts, or incongruent with the outcome and it was evident that they had achieved a positive outcome without earning it (see Figure 9 for an example). In both positive and negative tasks, participants were asked to rate the extent to which the character deserved the outcome and how pleased they would feel with the outcome for the character for each scenario variant. For both tasks the character was always referred to as "One of your friends", and their gender was not specified.

Figure 9: Example scenario from 'Deservingness: Positive Outcomes'

STORY STEM: "One of your friends passes their driving test first time around."

CONGRUENT VARIANT: They made very few errors and drove very well.

INCONGRUENT VARIANT: They made lots of errors but flirted with their driving instructor.

Questions

1. How much does your friend deserve to pass their driving test first time around?

1 _____ 10

Not at all Very much

2. How would you feel about the fact that your friend passed their driving test first time around?

1 _____ 10

Extremely displeased Extremely pleased

8.2.3.3 'Deservingness: Negative Outcomes'

'Deservingness: Negative Outcomes' (Vyas et al., in preparation b) had an identical design to that of 'Deservingness: Positive Outcomes' (see section 8.2.3.2), but the outcomes depicted were negative rather than positive (see Figure 10 for an example). The only other differences between the positive and negative tasks are those outlined in Section 8.2.3.1.

Figure 10: Example scenario from 'Deservingness: Negative Outcomes'

STORY STEM: "One of your friends from an evening class is told off in front of the class."
CONGRUENT VARIANT: They did not do the required homework.
INCONGRUENT VARIANT: The instructor is in a bad mood

Questions

1. How much does your friend deserve to be told off in front of the class?

1 _____ 10
Not at all Very much

2. How would you feel about the fact that your friend was told off in front of the class?

1 _____ 10
Extremely displeased Extremely pleased

8.2.3.4 Administration of both tasks

After reading the instructions, participants were given an instruction sheet, shown an example item, and allowed to ask questions. All scenarios and corresponding questions were then presented one at a time, in a paper booklet. Participants responded verbally to all questions. The scenarios remained on display until participants had completed the relevant questions in order to reduce the confounding effects of memory load. There were also two different scenario orders, which were counterbalanced within each group.

8.2.3.5 Scoring of 'Deservingness: Positive Outcomes'

8.2.3.6.1 Character deservingness ratings:

For each scenario, participants rated how much the character deserved the outcome on a scale of 1-10, where higher scores indicated that they were more deserving. Ratings were then summed across the five scenarios to create a total score for each condition (range 10–50), creating two scores: (1) character deservingness: congruent, (2) character deservingness: incongruent.

8.2.3.6.2 Participant satisfaction ratings:

For each scenario, participants rated how pleased they were with the outcome for the character on a scale of 1-10, where higher scores indicated that they were more pleased. Ratings were then summed across the five scenarios to create a total score for each condition (range 10–50), creating two scores: (1) participant satisfaction: congruent, (2) participant satisfaction: incongruent.

8.2.3.6 Scoring of 'Deservingness: Negative Outcomes'

The processes describes in section 8.2.3.5 were repeated for the 'Deservingness: Negative Outcomes task.

8.3 RESULTS

8.3.1 Data analysis

Means and standard deviations (SD) for each of the measures of the 'Deservingness' tasks are presented below in Table 5. A significance level of .05 was adopted for all comparisons. The assumptions of normality were met and thus parametric analyses were performed.

8.3.2 Deservingness: Positive Outcomes

8.3.2.1 Character Deservingness Ratings

A repeated measures 2 x 2 ANOVA was conducted to examine ratings for characters' deservingness of the positive outcome for all scenarios. There was one between-participant factor (AQ group: high versus low AQ) and one within-participant factor (congruence of characters' actions with the outcome: congruent versus incongruent). As expected, there was a significant main effect of congruence, $F(1,41)=.446.30$, $p < .0001$, whereby all participants rated characters as more deserving of the positive outcome when their actions were congruent with it. However, there was no significant main effect of group, $F(1,41)=.022$, $p = .883$, nor a significant group by congruence interaction, $F(1,41)=2.55$, $p = .118$. Inspection of the mean scores revealed that the high and low AQ groups gave similar character deservingness ratings, whereby they both correctly differentiated between characters' good and bad actions leading to positive outcomes.

8.3.2.2 Participant Satisfaction Ratings

A repeated measures 2 x 2 ANOVA was also conducted to examine ratings for participant satisfaction with the positive outcome for all scenarios. As expected, there was a significant main effect of congruence, $F(1,41)=321.98$, $p < .0001$, whereby all participants were more satisfied with the positive outcomes when the characters' actions were congruent with them. However, once again there was no significant main effect of group, $F(1,41)=.070$, $p = .793$, nor a significant group by congruence interaction, $F(1,41)=1.59$, $p = .215$. Inspection of the mean scores revealed that again the high and low AQ groups gave similar participant satisfaction ratings, whereby they both correctly differentiated between characters' good and bad actions leading to positive outcomes.

8.3.3 Deservingness Negative Outcomes

8.3.3.1 Character Deservingness Ratings

As with the positive items, a repeated measures 2 x 2 ANOVA was conducted to examine ratings for characters' deservingness of the negative outcome for all scenarios. There was one between-participant factor (AQ group: high versus low AQ) and one within-participant factor (congruence of characters' actions with the outcome: characters' bad actions or bad fortune). As with the positive scenarios, there was a significant main effect of the congruence of the characters' actions with the outcome, $F(1,41)=322.56$, $p < .0001$, whereby all participants rated characters as more deserving of the negative outcome when their actions were congruent versus resulting from bad fortune. However, once more there was no significant main effect of group, $F(1,41)=1.24$, $p = .272$, nor a significant group by congruence interaction, $F(1,41)=.435$, $p = .513$. Inspection of the mean scores revealed that the high and low AQ groups again gave similar character deservingness ratings, whereby they both correctly differentiated between characters' actions and bad fortune.

8.3.3.2 Participant Satisfaction Ratings

As with the positive items, a repeated measures 2 x 2 ANOVA was conducted to examine ratings for participant satisfaction with the negative outcome for all scenarios. As expected and as with the positive scenarios, there was a significant main effect of the congruence of the characters' action with the outcome, $F(1,41)=122.426$, $p < .0001$, whereby all participants were more satisfied with the negative outcomes when the characters' actions were congruent versus it resulting from bad fortune. However, once more there was no significant main effect of group, $F(1,41)=.638$, $p = .429$, nor a significant group by congruence interaction, $F(1,41)=.938$, $p = .338$. Inspection of the mean scores revealed that again the

high and low AQ groups gave similar participant satisfaction ratings, whereby they both correctly differentiated between characters' actions and bad fortune.

Table 5: Mean percentage scores and standard deviations for the 'Deservingness' tasks

	Low AQ group (N = 23) M (SD)		**High AQ group** (N = 20) M (SD)		**Significance** (p = .05)
Deservingness: Positive Outcomes (%)					
Character Deservingness					Condition *
					Gp NS
					Gp x condition NS
Congruent	92.69	(6.70)	89.00	(10.61)	—
Incongruent	35.56	(12.53)	39.90	(12.66)	—
Participant Satisfaction					Condition *
					Gp NS
					Gp x condition NS
Congruent	86.09	(11.19)	82.50	(12.27)	—
Incongruent	42.70	(11.72)	44.80	(12.01)	—
Deservingness: Negative Outcomes (%)					
Character Deservingness					Condition *
					Gp NS
					Gp x condition NS
Congruent	56.82	(14.44)	58.00	(11.70)	—
Incongruent	17.13	(5.68)	21.40	(9.93)	—
Participant Satisfaction					Condition *
					Gp NS
					Gp x condition NS
Congruent	30.70	(9.66)	31.40	(8.81)	—
Incongruent	16.87	(5.25)	19.80	(9.08)	—

** = significant at p = .05* *NS = not significant*

8.4 DISCUSSION

8.4.1 Summary of findings

The present study was designed to investigate how autistic personality traits influenced reasoning about characters' deservingness in relation to positive versus negative outcomes using two related 'Deservingness' tasks. These compared scenarios in which a main character's good actions (congruent) or bad actions/ misfortune (incongruent) led to either a positive or a negative outcome. Both groups were significantly influenced by whether the characters' actions were congruent with the positive or negative outcomes, assigning higher ratings of deservingness when the outcome resulted from congruent versus incongruent character actions (positive) or bad fortune (negative). This pattern was in line with the predictions, which were made on the basis of previous findings by Vyas et al., (in preparation b). Both groups also gave significantly higher ratings of their own satisfaction when the outcome was congruent with the characters' actions, for both positive and negative outcomes, which was also in line with the predictions made on the basis of a previous study (Vyas et al., in preparation b).

On the basis of findings from previous work Channon et al., (2010; 2011), it was thought that the high AQ group might differ in their ratings of deservingness. However, no group differences were found for ratings of character deservingness on either the positive or negative tasks. With respect to participant satisfaction ratings, it was expected that the high AQ group would give lower ratings at least for positive outcomes, in line with the findings of the 'Above and Beyond' task in Chapter 5 where high AQ participants gave lower ratings of personal reward for helping others. However, no group differences were identified for either the participant satisfaction ratings on either the positive or negative tasks.

8.4.2 The effect of congruence on judgments of deservingness

Broadly speaking, the hypotheses concerning the congruence manipulation were confirmed: both groups were more likely to judge characters as deserving if they acted in a manner that was congruent with the outcome, for both positive and negative tasks. For instance, in the scenario where the character receives a high grade on a piece of work, participants were more likely to judge this as deserving when the character had put in a lot of independent work (congruent), rather than when they had got a friend to write it (incongruent). Participants were also more likely to experience personal satisfaction for positive outcomes when they were deserved by the character (congruent). This pattern was mirrored in the negative scenarios. For example, when a character fell ill whilst travelling, participants were more likely to judge this as deserving when the character decided not to get the recommended vaccination (congruent), rather than when the recommended vaccination did not work (incongruent). Again, they were more likely to experience personal satisfaction for this outcome when it was deserved by the character (congruent). This

replicated the findings of Vyas et al., (in preparation b), and is consistent with previous work demonstrating that people experience pleasure in response to others receiving rewards if they are deserved, but resentment to those which are undeserved; conversely, they experience pleasure at others' punishment, but only when they have got 'what they deserved' (Feather, 2006).

Although the present study was not primarily designed to explore differences between positive and negative outcomes, tasks exploring both of these were included. However, inspection of the mean scores reveals a serendipitous finding: mean scores showed that both groups gave higher ratings of characters' deservingness and of their own satisfaction for scenarios with positive versus negative outcomes. Since the design of these two tasks differs (see Section 8.2.3.1), they are not directly comparable and it is not appropriate to perform statistical tests that consider the valence as a factor in the task design. Nonetheless, the pattern of results potentially indicates that participants were less likely to judge characters as deserving of negative versus positive outcomes, regardless of whether the characters' actions justified either outcome. This is in contrast to findings from previous work which suggest that people are more likely to judge an outcome as resulting from the characters' actions if it is negative versus positive. For instance, one study found that when characters committed actions that had negative (e.g. shooting someone) versus positive consequences (e.g. hitting the bull's eye), they were more likely to be judged as culpable for negative consequences, regardless of their intent (Knobe, 2003). A further study explored a scenario in which the chairman of a company ignored advice that an initiative would either harm or help the environment (Zalla & Leboyer, 2011). In both variants the chairman decided to proceed regardless, claiming that he was only focused on profits. However, almost all participants were likely to judge that the chairman had intentionally caused the negative versus positive side effects to the environment. In fact, the effect on the environment was always unintentional (positive or negative) since the chairman claimed he was only interested in profits, and thus any effect on the environment resulting from his decision was incidental. Again these findings suggested that people made harsher judgments when ascribing responsibility for negative versus positive outcomes, which appears to be the opposing pattern to that indicated in the present study.

This discrepancy might have reflected variation across the tasks described, with respect to the extent that the scenarios reflected consequences for the agent versus others, and how tightly the agents' actions could be attributed to the outcomes. The tasks also explored different aspects of moral judgment, such as deservingness versus responsibility, and manipulated different factors, such as the congruence of characters' actions with the outcomes versus the characters' intentions behind their actions. The way in which such factors might influence moral judgment is not well understood, but it has been posited that different types of moral judgment might draw on different cognitive processes (Cushman, 2008), which thus might thus lead to different types of attribution biases. Whilst no definitive conclusion can be drawn on the basis of the present data, previous work examining how people ascribe intent, responsibility and

deservingness to positive and negative events, suggests this to be a multifactorial process. Understanding the factors that influence this has potential implications for decision-making in a number of real-life-contexts (e.g. rewarding or disciplining employees), and thus these tentative findings of different biases for positive versus negative outcomes warrant further attention.

8.4.3 Why did the 'Deservingness' tasks fail to differentiate the high versus low AQ groups?

Contrary to predictions, neither the positive nor the negative 'Deservingness' tasks differentiated the high versus low AQ groups. This may reflect intact judgment in relation to deservingness, although it also raises the question of whether the task was sufficiently sensitive to detect any possible group differences. It is also necessary to consider whether the high AQ group displayed the postulated difficulties in empathy and/or executive functioning.

8.4.3.1 Characteristics of the high and low AQ participants

Could difficulties in empathy and/or executive functioning account for the pattern of findings across these studies, which showed group differences on the social tasks used in Chapters 5 and 6, but no differences on those used in Chapters 7 and 8? Separate samples of participants were recruited for the studies reported in Chapters 5 and 6, and a further sample took part in the two studies reported in Chapters 7 and the present Chapter. As highlighted in Chapter 7.4.5, a direct comparison of the three samples is impossible as neither the additional executive or empathy measures were included in the studies for Chapters 5 or 6. However, since the samples were all drawn from similar, relatively homogeneous, populations of university students from the same university, there is no reason to suspect that their profiles would differ substantially. As discussed in Chapter 7 (see sections 7.4.2 – 7.4.6), the additional measures showed that the high AQ sample in the study reported in Chapters 7 and the present study had difficulties with cognitive empathy on the task examining understanding of others' intentions (MIT – see Chapter 7, sections 7.2.4.3, 7.3.3.3 and 7.4.6), and they may also have had mild difficulties in emotional empathy and executive skills. It is likely that the samples used in Chapters 5 and 6 would have also have shown a similar profile of difficulties, and that this may have underpinned the differences identified on the social tasks used in these experiments. Please see Chapter 7, Sections 7.4.5 and 7.4.6 for a more thorough discussion of how executive dysfunction and deficits in empathy may or may not relate to performance on social tasks, and issues of comparability of the samples across tasks. Overall, it seems most plausible that the high AQ group did struggle to identify fully with the characters in the present study, but that the 'Deservingness' tasks were not sufficiently sensitive to reveal these difficulties. This possibility will be further explored in the next section, 8.4.3.2.

8.4.3.2 Sensitivity of the 'Deservingness' tasks

It therefore seems probable that the failure to find group differences on the 'Deservingness' tasks used in Chapter 7 and the present study arose from a lack of sensitivity in the tasks. It was hypothesised that difficulties with cognitive empathy in the high AQ group would result in a lack of understanding of the characters' needs, leading to atypical judgments on the 'Deservingness' tasks in the form of either harsher judgments and greater differentiation between the conditions through the inflexible 'black and white' use of moral rules in the absence of empathy, or an empathic failure to pick up the cues for deservingness, leading to less differentiation between the conditions. Whilst the evidence regarding possible deficits in emotional empathy in ASD is mixed, it was also thought possible that difficulties with emotional empathy might impair performance through failure to resonate with others' feelings and needs, leading to less sympathy and atypical judgments which either over- or under-emphasised the importance of congruence. Any executive difficulties could also potentially have resulted in either the high AQ group acting impulsively and making harsher judgments, or failing to appreciate the differences between the conditions and distinguishing less between the congruence conditions in their judgments. However, no differences between groups were identified for any of the measures on either the positive or negative 'Deservingness' tasks.

As discussed in Chapter 7.4.7, informal comparison of the different sets of materials used for the social tasks in Chapters 5, 6, 7 and the present study suggests that they did not differ substantially with respect to the types of characters or situations depicted. Nor could differences in the pattern of findings be readily attributed to the types of questions asked, since Chapters 5, 6 and 7 all addressed willingness to help others, but did not all find group differences. One important difference between the two studies that did differentiate the groups (Chapters 5 and 6) and those that did not (Chapter 7 and the present study) is the extent to which the characters could be viewed as deserving. In Chapters 5 and 6, all characters could be said to have been deserving of help, although the task conditions varied in the extent to which this was emphasised. By contrast, in both Chapter 7 and the present study, the characters were more deserving (either of help or of an outcome) in one condition than the other. It was suggested in Chapter 7.4.7 that the contrast between the conditions of the 'Social Favours' task might have been very salient, and the same issue may well have applied here. In the present task, the differences between conditions were deliberately made clear, in order to signal to participants differences in how deserving the characters were. This might have been sufficient to allow the high AQ group to distinguish levels of deservingness at a similar level to the low AQ group, at least in their ratings of character deservingness, regardless of whether they understood why the character was deserving or not. This interpretation is consistent with previous work exploring the sensitivity of people with ASD to intentionality when making moral judgments. Consideration of previous tasks used in the literature (Channon et al., 2011; Moran et al., 2011; Buon et al.,

2013; Grant et al., 2005) revealed (see Chapter 4.2.3) that the salience of the contrast between different conditions might have played a role in the inconsistent pattern of findings resulting from these studies. Whilst the present study did not directly examine intentionality, making judgments of characters' deservingness on the basis of whether their actions were congruent or incongruent with the outcome might reasonably be expected to draw upon the ability to infer others' intentions. Grant et al., (2005) found intact understanding of intentionality in those with ASD, whereas others found heightened or reduced sensitivity to intentionality (Channon et al., 2011; Moran et al., 2011; Buon et al., 2013).

Another factor that might be said to distinguish the different tasks used in this thesis is the extent to which a personal cost was incurred on the part of the participant. As highlighted in Chapter 7, complying with characters' requests in the 'Social Favours' task might not have involved a substantial enough personal sacrifice on the part of the participant to differentiate the groups. It is possible that people with the high AQ traits might only display less socially sensitive behaviour once there is a conflict of interest regarding their own versus others' needs. In the 'Above and Beyond' and 'Social Expectations' tasks used in Chapters 5 and 6 this conflict was made explicit, but it was not as clearly spelt out in the 'Social Favours' task in Chapter 7. Similarly, in the 'Deservingness' tasks used in the present study, the outcomes for the character did not affect the participant directly, but rather asked them to make an abstract judgment regarding whether the outcome was 'just' or not. Indeed, the participants might be described as mere observers in the 'Deservingness' tasks, as opposed to active agents in the previous tasks used in this thesis. The extent of personal involvement has been found to influence people's decisions (e.g. Levine & Crowther, 2008; Levine et al., 2005; Ruthowski et al., 1983; Grant & Gino, 2010; Frey & Meier, 2004) and might prove to be an important determinant in mediating the social behavior of those with ASD, or in the broader autistic phenotype. This will be manipulated in Chapter 11 using a classic moral decision-making task.

8.4.3.3 A preserved understanding of 'what is right and wrong' in ASD?

One hypothesis arising from the studies to date is that people with high AQ traits might 'know' what is 'right', but fail to put this into practice once a threshold for personal involvement has been breached. This hypothesis is consistent with the findings of reduced pro-social behaviour in Chapters 5 and 6 versus 7, and the intact judgment of deservingness in Chapter 8 by high AQ participants. It could also potentially account for the conflicting findings across these tasks from ratings of personal reward/ sacrifice; thus, if the threshold for personal cost has not been breached then those with high versus low AQ traits are likely to make similar judgments of whether something is personally rewarding or incurs a sacrifice, as in Chapter 7 and the present study. Direct manipulation of the level of personal involvement or cost would be needed to examine this issue further.

Both Chapter 7 and the present study asked for ratings of behaviour or judgments of others' actions, but did not probe understanding of these issues. This is consistent with findings from previous work demonstrating intact moral/conventional distinctions in participants with ASD, whereby simple judgments of what is 'wrong or right' seem to be intact (e.g. Blair, 1996), although this is contingent upon the depth of understanding probed in the particular study (Zalla et al., 2011). Thus, one way to improve the sensitivity of tasks such as those used in the present study would be to include a measure that required participants to explain or justify their choices, since these seem to require a deeper level of understanding that often differentiates participants with ASD.

8.4.5 Conclusion

The present study explored how the congruence of characters' actions with an outcome was judged with respect to the characters' 'deservingness', and the participants' personal satisfaction for that outcome. In line with predictions, when characters' actions were congruent with an outcome, participants were more likely to judge the character as deserving of it, and to rate that they were satisfied with the outcome for both positive and negative events. Contrary to predictions, this task did not differentiate the high versus low AQ groups. After careful consideration of the task designs and demands, it seems possible that a basic understanding of what is 'fair' or 'right' may be intact in the broader autistic phenotype. This, however, may not be put into action by the high AQ group if there is a conflict of interest, which might be related to difficulties in appreciating both others' perspectives and the long-term consequences of their own actions for themselves. It remains possible that although basic moral judgments of 'deservingness' seem to be intact, people with high levels of autistic traits may have a more limited ability to reason about this.

In the next study, the focus will turn to examining the causes and consequences of wrong-doing, and how negative outcomes could be avoided via counterfactual thinking. This shifts the emphasis from weighing up one's own versus another's needs to attributions of responsibility. It examines 'justice' from a different angle, asking how and whether we should hold people to account for their actions. The next task not only asks participants to make judgments, but also to produce their own counterfactual solutions for problems. Thus, this might provide a more sensitive tool for discriminating between people with high versus low AQ traits, and examining any potential ramifications of the broader autistic phenotype for moral judgment and reasoning.

Chapter 9: Counterfactual thinking for self- versus other- and the experience of regret and blame

9.1 INTRODUCTION

Counterfactual thoughts are mental representations of alternatives to past events (McNamara, Durso, Brown and Lynch, 2003), which mostly occur in relation to negative outcomes. They are a frequent, spontaneous occurrence in everyday life, and are thought to play an important role in social learning via reflection upon mistakes and consideration of how to solve future problems (Epstude & Roese, 2008). Counterfactual thinking is also closely related to decision-making, planning, problem-solving, experience-driven learning, and the ability to pass false belief tasks (Zago et al., 2014). These are all cognitive processes that draw upon the prefrontal lobes (Van Hoeck et al., 2013), and correspondingly impairment in counterfactual thinking has been linked to a number of conditions where frontal lobe dysfunction is implicated, including Parkinson's disease (McNamara et al., 2003), Huntington's disease (Solca et al., 2015) and patients with lesions to this region (Beldarrain, Garcia-Monco, Astigarraga, Gonzalex & Grafman, 2005). By contrast, a recent study revealed an intact ability to generate and reason about counterfactuals in individuals with Tourette's Syndrome, which is a neurological condition where the orbito-frontal pathway is impaired and associated subtle cognitive deficits are observed (Zago et al., 2014). The authors suggested that participants with Tourette's Syndrome might have used compensatory or alternative strategies to successfully solve the counterfactual tasks. This explanation could potentially account for this apparently conflicting finding, but further work is required to fully understand the contribution of the frontal lobes to counterfactual thinking and how this is altered in disease.

As previously discussed, people with ASD demonstrate a range of cognitive difficulties that are thought to be underpinned by executive dysfunction and deficits in cognitive empathy (e.g. Happé, 1999). These areas of impairment are thought to involve cognitive processes that might overlap with those drawn upon when engaging in counterfactual thinking, and to activate the same neuronal networks (Van Hoeck et al., 2013; Zago et al., 2014). Thus, some work has explored whether counterfactual thinking is also impaired in the condition, but this has produced a mixed set of findings, and to date only children with ASD have been examined. For instance, Grant, Riggs & Boucher (2004) found that children with ASD demonstrated impairment on a counterfactual reasoning task, and that this was correlated with lower performance on a false-belief task. However, other work suggested that the performance of those with

ASD on counterfactual reasoning tasks was matched to their typically developing peers. For instance, Begeer et al., (2009) found no difference in the number of counterfactual thoughts generated by those with high-functioning ASD, although it was suggested that the children with ASD likely used different strategies for counterfactual reasoning, since they tended to generate more subtractive (e.g. "If only I hadn't") and fewer additive (e.g. "If only I had") counterfactuals. This preference for counterfactual reasoning that 'un-does' 'bad' actions, rather than providing novel alternative actions that could have prevented negative outcomes, was considered to be linked to difficulties in imagination and executive skills that are required to generate new ideas (Begeer et al., 2009). This is consistent with other work identifying difficulties in generating information in children with ASD (Peterson & Bowler, 2000), and which suggests that counterfactual reasoning is only impaired when imagination is required (Scott, Baron-Cohen & Leslie, 1999). As with Tourette's Syndrome, ASD is a neurodevelopmental condition, and thus a greater capacity for compensation might be afforded in these groups in comparison to those with acquired (i.e. lesions) or neurodegenerative conditions affecting the frontal lobes (e.g. Parkinson's or Huntington's disease), where there is a loss as opposed to an alteration of functioning.

In addition to helping people to explain past events and prepare better for future events, counterfactual thinking might also facilitate emotional reactions (e.g. Roese, 1997). Work in healthy adults suggests that counterfactuals are an important antecedent in the adaptive role of 'self-conscious' emotions such as shame, guilt, regret and envy (Coricelli & Rustichini, 2010). Self-conscious emotions are a subset of moral emotions that are evoked by direct experience from, or anticipation of others' evaluations. They are thought to be supported by three distinct but related cognitive processes: self awareness, other awareness and awareness of social norms (Janowski & Takahashi, 2014). Basic emotions, such as happiness or sadness, emerge within the first nine months of life, but self-conscious emotions emerge much later in life; it is only at the age of seven that children are reliably able to appreciate the emotional consequences stemming from counterfactual thinking, such as contentment or relief at avoiding a negative outcome or disappointment at incurring one (Beck & Guthrie, 2011).

One study compared children with and without ASD for their ability to identify facial expressions of self-conscious and basic emotions, and their cognitive empathy was measured via a classic 'theory of mind task' (Heerey, Keltner & Capps, 2003). The children with ASD performed more poorly at self-conscious emotions, but displayed no difference in their recognition of basic emotions. However, once performance on the theory of mind task was controlled for, these differences disappeared, which suggests that self-conscious but not basic emotions are reliant upon the ability to understand and predict others' mental states. A further study explored how high-functioning children with ASD understood emotions based on counterfactual reasoning, finding that they were poorer at explaining them in comparison to their typically-developing peers (Begeer, De Rosnay, Lunenburg, Stegge, & Terwogt, 2014). Follow-up analyses

revealed that IQ score was associated with the ASD group's task performance, but that this association was not found for control participants. This suggests that those with ASD relied more on their general intellectual abilities to complete the tasks, perhaps in lieu of the abilities required to understand the characters' thoughts and intentions.

Counterfactual thinking might also support moral judgments such as blame. Whilst a substantial body of work has addressed the relationship between counterfactual thinking and causal ascriptions, less work has addressed the relationship to blame. These two types of judgments, cause and blame, are thought to be related but also to be dissociable; for instance, if an infant plays with a loaded gun and shoots someone they are the cause of the incident but are not to blame for it (see e.g. Lagnado & Channon, 2008). There is some evidence to suggest that the availability of counterfactuals influences blame judgments (Wells & Gavanski, 1989), but other work found no evidence to support this assertion (N'gbala & Branscombe, 1995). Furthermore, whilst some work has explored judgments of cause and blame in people with ASD (e.g. Channon et al., 2011), how counterfactual thinking might mediate such judgments in this group has not yet been explored. This merits investigation, given their well-documented difficulties with tasks that require considering others' intentions and some evidence of problems with generating counterfactual alternatives.

In view of the postulated importance of counterfactual thinking for learning about how to avoid future mistakes and to appreciate the consequences of ones' actions, impaired counterfactual thinking might account for some of the difficulties seen in those who struggle to navigate social situations. Exploring counterfactual thinking for social 'wrongs' in relation to the broader autistic phenotype might provide a more fine-grained understanding of the ramifications of autistic traits for everyday social functioning. Thus, the present study was designed to extend the limited literature on counterfactual thinking in the autistic spectrum by exploring whether those with sub-clinical levels of autistic traits showed a reduced ability to generate counterfactual alternatives for everyday mistakes. A novel task, "Reflective Counterfactual Thinking" (RCFT), was designed to assess individuals' ability to generate additive counterfactual thoughts in relation to negative everyday events.

It was expected that participants would find generating counterfactuals from the perspective of another, and considering the consequences of these for others, to be more demanding than when considering their own perspective only, since this would require both imagining alternative 'worlds' and 'putting themselves in others' shoes'. Given the well-established difficulties with cognitive empathy, including perspective-taking, in those with ASD and high levels of autistic traits, it was expected that this would be particularly difficult for the high AQ group. The identity of the agent (participant or main character) was therefore manipulated to investigate this.

It was also thought that participants might generate more counterfactuals in situations where there was a mismatch between the agents' identity and the recipient of the consequences (incongruent scenarios), since causing negative consequences for others might be deemed a more 'serious' transgression than causing negative consequences for oneself (congruent scenarios). Incongruent scenarios might therefore be expected to trigger greater counterfactual reflection than congruent scenarios. It was thought that the high AQ group might differentiate less between these two scenarios, because they might be less concerned about the self-other distinction. Two sets of scenarios were therefore created, congruent and incongruent, and each set had two versions of each scenario, with the agent as participant or main character.

The extent to which counterfactual thinking might mediate, or be mediated by, individuals' perceptions of personal responsibility and regret for bad actions leading to negative events was also of interest. Thus, ratings of blame and regret were also acquired for each scenario.

9.1.1 Hypotheses

9.1.1.1 Counterfactual thinking fluency

On the basis of previous work indicating impairment in people with ASD on 'additive' counterfactual thinking tasks (e.g. Beeger, et al., 2009), it was hypothesised that those with high versus low AQ scores would be less fluent in producing counterfactual alternatives. It was also hypothesised that this might interact with agent identity (participant or other agent), and also with the congruency of the agent identity and the recipient of the consequences (agent's action affected self or other), since those with high AQ scores were expected to generate fewer counterfactual alternatives in relation to others than in relation to themselves.

9.1.1.2 Ratings of regret

Very little work has explored the experience of moral emotions in those with ASD, and thus there was limited evidence to support a particular hypothesis for this measure. However, the experience of regret for ones' mistakes is a self-conscious emotion, and thus is likely to draw upon understanding of others' experiences and expectations, which is known to be impaired in those with ASD. Furthermore, as discussed some work suggests that engaging in counterfactual thinking might activate emotions such as regret, which also might be impaired in those with ASD. On this basis, it was hypothesised that those with high versus low AQ scores would give lower ratings of regret, and that these ratings might also interact with agent identity and congruence.

9.1.1.3 Ratings of blame

As with regret, very little work has explored ratings of blame in those with ASD, but some work has how examined how intentionality affects moral judgment, finding a mixed pattern of results (see Chapter 4.3.2). Furthermore, one study found that prisoners who thought counterfactually about their 'wrong-doing' reported feeling more blameworthy, than those who thought about it factually (Mandel & Dhami, 2005). Since both understanding of intentionality and counterfactual thinking might be impaired in those with ASD it seemed likely those with high versus low AQ traits would make atypical judgments, and that these ratings might again interact with agent identity and congruence.

9.2 METHODS

9.2.1 Screening phase

9.2.1.1 Screening participants and procedure

The participants screened for this study also took part in the study in Chapter 5 and are described in Chapter 5.2.1.1. Thus, an opportunistic sample of 573 full-time university students (43% male) who were fluent in English and aged 18 or over (mean age 20 years old) was recruited for the screening phase of the study. All participants completed the AQ (Baron-Cohen et al., 2001) and total scores were calculated for the whole sample. Participants within the highest-scoring and lowest-scoring 10% of males and females were contacted and invited to take part in the second stage.

9.2.2 Experimental phase

9.2.2.1 Design

There was one between-participants factor of AQ group (high vs. low scorers), and two within-participants factors, the agent's identity (participant or main character), and the congruence between the agent's actions and the consequences of these (congruent or incongruent i.e. either for agent or not for agent).

9.2.2.2 Experimental participants and procedure

The participants tested for this study also took part in the 'Above and Beyond' task and are described in Chapter 5.2.3.2. There were thus 27 (14 male, 13 female) participants in the high AQ group

and 24 (12 male, 12 female) participants in the low AQ group. All participants were tested individually, and provided written informed consent before completing the 'RCFT' task.

9.2.3 The 'Reflective Counterfactual Thinking' (RCFT) Task

This task was primarily designed to assess individuals' ability to generate counterfactual thoughts in relation to negative everyday events, both from the participants' own perspective and from that of another agent. A range of scenarios was devised and piloted with healthy volunteers in order to refine the items. The final set consisted of eight brief scenarios each describing a negative event involving the participant and a main character. Half of the items described interactions where the agent's actions had negative consequences for themselves (congruent) and the other half described interactions where the agent's actions had negative consequences for another person (incongruent). Each scenario also had two versions: in one version the participant was the agent, and in the other version the main character was the agent. This gave a total of sixteen vignettes: four congruent scenarios with a) the participant as the agent, and b) the main character as the agent; and four incongruent scenarios with c) the participant as the agent, and d) the main character as the agent. Please see Figure 11 for an outline of the study design.

Figure 11: Design of 'RCFT' Task

	Participant Agent	**Character Agent**
Congruent	*Participant Consequence*	*Character Consequence*
	You are with a friend on the train and you forget your laptop.	You are with a friend on the train and they forget their laptop.
Incongruent	*Character Consequence*	*Participant Consequence*
	You are looking after your sibling's house. You forget to water the plants and they die.	Your sibling is looking after your house. They forget to water the plants and they die

The scenario main characters included a range of relationships such as siblings, friends, colleagues and flatmates, and their gender was not specified. The scenarios reflected a range of everyday situations where the agent acted in a way that caused a negative event with either practical or social consequences. These issues were all counterbalanced across the items in accordance with the task factors. Each scenario

was followed by three questions. Participants were first asked to generate as many counterfactual thoughts as they could to complete the relevant stem: "If only I/they had…". Participants were then asked to give ratings of regret (i.e. how much the agent would regret their actions) and of blame (i.e. how responsible the agent was for the negative consequence). Please see Figure 12 and 13 for full scenario examples.

All participants first read a sheet of instructions about the task. This explained that they would see short scenarios about everyday situations and would respond verbally to questions, making free responses and giving ratings. Participants were requested to answer as quickly and as truthfully as possible. The scenarios were presented on paper, and participants were taken through an example before completing the eight experimental items. Scenarios and questions were presented in separate booklets such that the relevant scenario remained on display throughout task performance, in order to minimise any memory demands. There were two different scenario orders, which were counterbalanced within each group.

Figure 12: Example congruent scenario from the 'RCFT' task

Example Scenarios and Question

a) **Congruent: participant agent**

"You are going swimming with your housemate. You lead the way to get changed but are not paying attention and you accidentally walk into the changing room of the opposite sex."

Counterfactual thinking fluency:
Please list as many "If only I HAD" thoughts that you might experience in this situation**.**

Ratings of regret:
On a scale of 1-10 how much would YOU regret not paying attention when leading the way to get changed?
1 = regret very little; 10 = regret very much

Ratings of blame:
On a scale of 1-10 how much are YOU to blame for walking into the wrong changing room?
1 = not at all to blame; 10 = very much to blame

b) **Congruent: character agent**

"You are going swimming with your housemate. They lead the way to get changed but are not paying attention and they accidentally walk into the changing room of the opposite sex."

Counterfactual thinking fluency:
Please list as many "If only THEY HAD" thoughts that you might experience in this situation**.**

Ratings of regret:
On a scale of 1-10 how much would THEY regret not paying attention when leading the way to get changed?
1 = regret very little; 10 = regret very much

Ratings of blame:
On a scale of 1-10 how much are THEY to blame for walking into the wrong changing room?
1 = not at all to blame; 10 = very much to blame

Figure 13: Example incongruent scenario from the 'RCFT' Task

Example Scenarios and Question

c) Incongruent: participant agent

"In a meeting your colleague puts forward a proposal for a project. You laugh thinking it was a joke and their proposal is not taken seriously by the rest of the team."

Counterfactual thinking fluency:
Please list as many "If only I HAD" thoughts that you might experience in this situation**.**

Ratings of regret:
On a scale of 1-10 how much would YOU regret thinking their suggestion was a joke?
1 = regret very little; 10 = regret very much

Ratings of blame:
On a scale of 1-10 how much are YOU to blame for your colleague's suggestion not being taken seriously?
1 = not at all to blame; 10 = very much to blame

d) Incongruent: character agent

"In a meeting you put forward a proposal for a project. Your colleague laughs thinking it was a joke and your proposal is not taken seriously by the rest of the team."

Counterfactual thinking fluency:
Please list as many "If only THEY HAD" thoughts that you might experience in this situation**.**

Ratings of regret:
On a scale of 1-10 how much would THEY regret thinking your suggestion was a joke?
1 = regret very little; 10 = regret very much

Ratings of blame:
On a scale of 1-10 how much are THEY to blame for your suggestion not being taken seriously?
1 = not at all to blame; 10 = very much to blame

9.2.3.1 Scoring

9.2.3.1.1 Counterfactual thinking fluency

Counterfactual responses were scored in accordance with the number of alternatives that were generated for each scenario. This gave four different fluency scores, one for each of the four types of scenario: a) congruent scenario: participant agent, b) congruent scenario: character agent, c) incongruent scenario: participant agent, d), incongruent scenario: character agent. Fluency scores were calculated by a rater who was not blind to group membership, and by a second, blind independent rater. There was an inter-rater agreement rate of 96.21%; all disagreements were resolved by discussion between the two raters.

The first rater developed both general and item specific scoring guidelines for the 'RCFT'. One point was allocated per counterfactual alternative; thus if a participant generated three alternatives for a scenario then they received a score of three, if they generated four they received a score of four, and so on. However, in order to score one point, a counterfactual alternative needed to meet three criteria. Firstly, each counterfactual alternative should correctly complete the relevant "If only I/THEY had…" stem provided. Secondly, each counterfactual alternative should be independent and separable from any others that were also generated in response to the scenario. For instance, if a response to the scenario where the participant walked into the wrong changing room contained three ways of saying "If only I had been careful", "e.g. been more careful/ been less careless/ taken more care", it would only score one point. However, if it also contained two other independent and separable ideas, "e.g. changed at home before coming swimming or asked my friend to lead the way", it would score a further two points, creating a total score of three points. Finally, counterfactual alternatives were not allowed to simply repeat or reiterate information contained in the scenario. For instance, in the changing room example the alternative, e.g. "walked into the correct changing room", would not score a point, as it only attempts to 'undo' the agent's 'bad' actions and does not provide an alternative that could have avoided the mistake.

9.2.3.1.2 Ratings of regret

For each scenario, participants rated how much the agent would regret their actions on a scale of 1–10, where higher scores indicated greater regret. Ratings were then summed across scenarios to create a total score for each of the four types of scenario (range 4–40): a) congruent: participant agent, b) congruent: character agent, c) incongruent: participant agent, d), incongruent: character agent. These were converted to percentages for ease of interpretation and comparison.

9.2.3.1.3 Ratings of blame

For each scenario, participants rated how much the agent was to blame for the negative consequences on a scale of 1–10, where higher scores indicated more blame. Ratings were then summed across scenarios to create for each of the four types of scenario (range 4–40): a) congruent: participant agent, b) congruent: character agent, c) incongruent: participant agent, d), incongruent: character agent. These were converted to percentages for ease of interpretation and comparison.

9.3 RESULTS

9.3.1 Data analysis

Means and standard deviations (SD) for each of the measures of the 'RCFT' task are presented below in Table 6. A significance level of .05 was adopted, with a stricter level (.05/2 = .025) for post hoc tests to control for multiple comparisons. The assumptions of normality were met and thus parametric analyses were performed.

9.3.2 The 'RCFT' task

9.3.2.1 Counterfactual thinking fluency

A repeated measures 2 x 2 x 2 ANOVA was conducted to examine the counterfactual thinking fluency scores. There was one between-participants factor of AQ group (high versus low scorers), and two within-participants factors: congruence between the agent's action and the recipient of the consequences of these (congruent versus non-interactive), and the agent's identity (participant versus character). There were no significant main effects of congruence, $F(1,49) = .443$ $p = .509$, nor agent identity, $F(1,49) = 1.26$, $p = .267$, nor a significant congruence by agent identity interaction, $F(1,49) = 2.01$, $p = .162$. There were also no significant congruence by group interaction, $F(1,49) = .078$, $p = .782$, nor a significant agent identity by group interaction, $F(1,49) = 1.43$, $p = .238$, nor a significant three-way interaction $F(1,49) = .018$, $p = .895$. However, there was a significant main effect of group, $F(1,49) = 5.168$, $p = .027$. Inspection of the mean scores suggests that, in line with predictions, the high AQ group achieved significantly lower fluency scores than the low AQ group for all types of four scenarios.

9.3.2.2 Ratings of regret

A repeated measures 2 x 2 x 2 ANOVA was also conducted to examine ratings of regret. There was a significant main effect of agent identity, $F(1,49) = 10.55$, $p = .002$, and mean scores showed that

participants across groups gave higher regret ratings for scenarios where the participant rather than the character was the agent. There was also a significant main effect of congruence, $F(1,49) = 4.05$, $p = .050$, and mean scores showed that participants gave higher regret ratings when the consequences of the actions were congruent versus incongruent. There was also a significant congruence by agent identity interaction, $F(1,49) = 48.61$, $p < .0001$, and mean scores suggested that the difference between the two agent identities was greater for incongruent versus congruent scenarios.

With respect to group, there was no main effect $F(1,49) = 3.03$, $p = .088$. There was also no significant congruence by group interaction $F(1,49) = .018$, $p = .894$, nor a significant agent identity by group interaction, $F(1,49) = .092$, $p = .762$, nor a significant three-way interaction, $F(1,49) = .272$, $p = .605$. This is contrary to predictions, and suggests that the groups gave similar ratings of regret in relation to the different factors manipulated.

9.3.2.3 Ratings of blame

A repeated measures 2 x 2 x 2 ANOVA was also conducted to examine ratings of blame. There was a significant main effect of agent identity, $F(1,49) = 60.13$, $p < .0001$, and mean scores showed that across groups, participants again gave higher ratings for scenarios where the participant rather than the character was the agent. There was a significant main effect of congruence, $F(1,49) = 106.48$, $p < .0001$, and mean scores showed that they again gave higher ratings when the consequences of the actions were congruent versus incongruent. However, unlike regret, there was no significant congruence by agent identity interaction, $F(1,49) = .365$, $p = .548$.

With respect to group, there was no main effect of group $F(1,49) = .157$, $p = .694$, nor a significant three-way interaction, $F(1,49) = .201$, $p = .656$. However, there was a significant congruence by group interaction, $F(1,49) = 16.12$, $p < .0001$; a post-hoc t-test using a strict significance level (p=.05/2=.025) showed that the difference between the congruent and incongruent scenarios was less for the high AQ group than for the low AQ group, $t(49) = 4.015$, $p < .0001$. There was also a significant agent identity by group interaction, $F(1,49) = 9.80$, $p = .003$, and a post-hoc t-test showed that the difference between scenarios where the participant versus the main character was the agent was greater for the high AQ group than for the low AQ group, $t(49) = 3.145$, $p = .002$.

Table 6: Mean percentage scores and standard deviations for the 'RCFT' task

	Low AQ group (N = 24) M (SD)		High AQ group (N = 27) M (SD)		Significance (p = .05)
Counterfactual Fluency					
Congruent Consequence					
Participant Agent	8.04	(0.52)	6.89	(0.49)	__
Character Agent	7.96	(0.69)	6.18	(0.65)	__
Incongruent Consequence					
Participant Agent	8.12	(0.75)	6.33	(0.71)	__
Character Agent	8.25	(0.69)	6.19	(0.65)	__
					Congruence = NS Agent = NS Group = .027* Congruence*Agent = NS Congruence *group = NS Agent*group = NS Congruence*agent*group = NS
Regret %					
Congruent Consequence					
Participant Agent	73.65	(10.78)	77.77	(13.14)	__
Character Agent	74.37	(13.85)	79.90	(11.29)	__
Incongruent Consequence					
Participant Agent	74.90	(12.4)	80.19	(11.49)	__
Character Agent	66.45	(16.01)	71.67	(11.48)	__
					Congruence = .050* Agent = .002* Group = NS Congruence*agent < .0001* Congruence*group = NS Agent*group = NS Congruence*agent*group = NS
Blame %					
Congruent Consequence					
Participant Agent	84.59	(9.11)	83.89	(9.10)	__
Character Agent	80.00	(13.67)	71.67	(10.38)	__
Incongruent Consequence					
Participant Agent	65.21	(14.18)	75.19	(10.90)	__
Character Agent	59.38	(11.55)	62.78	(11.55)	__
					Congruence < .0001* Agent < .0001* Group = NS Congruence*agent = NS Congruence*group < .0001* Agent*group = .003* Congruence*agent*group = NS

** = significant at p = .05* *NS = not significant*

9.4 DISCUSSION

9.4.1 Summary of findings

The present study examined whether those with high versus low autistic traits showed impairment in counterfactual thinking, using a task where they were required to generate counterfactual alternatives to correct either their own or others' mistakes. In line with expectations, the high AQ group was found to produce fewer counterfactual alternatives overall. Whilst there is some conflict in the existing literature regarding whether counterfactual thinking is impaired in ASD, this pattern appears to be broadly consistent with the notion that those with ASD struggle to produce counterfactuals that involve generating novel alternatives, as opposed to simply retracting mistakes (e.g. Begeer et al., 2009). The high AQ group was also expected to have greater difficulty generating counterfactuals relating to others than to themselves. However, neither the manipulation of agent identity, nor of the congruence between the agents' actions and the recipient of the consequences, differentiated the groups. Indeed, both groups seemed to perform fairly consistently in their ability to generate counterfactual alternatives across the four scenario types.

Regret ratings were also obtained; however, contrary to predictions this measure did not differentiate the groups. In contrast to the fluency measure, both groups differentiated in a similar manner between the four scenario types in their regret ratings. The groups gave higher regret ratings for scenarios that were congruent versus incongruent, and higher ratings for scenarios where the agent was the participant versus the main character. Furthermore, there was an interaction between the congruence and agent identity factors, where the difference between ratings of regret for the participant versus the main character was bigger in scenarios that were incongruent. This suggests that the influence of the agents' identity for the experience of regret was more pronounced when the agents' actions had consequences for others than for themselves.

Finally, blame ratings were obtained, whereby participants across groups again gave higher ratings for scenarios which were incongruent versus congruent, and for scenarios where the participant versus the main character was the agent. However, in contrast to regret ratings, there was no significant congruence by agent identity interaction. Partial support was provided for the prediction that the high AQ group would give atypical ratings, since there was no significant main effect of group, but there were significant interactions between group and congruence, and between group and agent identity. Post-hoc t-tests revealed that the high AQ group differentiated **less** between the congruence of the scenario type, and were **more** sensitive to the manipulation of agent identity. This suggested that participants with high AQ traits were less concerned about the type of scenario interaction, differentiating less between the culpability of agents who harmed others versus themselves. In contrast, they appeared to be more concerned with

whether they were the agents of the actions or not, judging themselves to be more culpable for causing negative consequences than main characters.

9.4.2 The role of congruence and agent identity

Since counterfactual thinking is postulated to facilitate learning from one's mistakes and appreciating the consequences of these (Esptude & Roese, 2008), it was expected that counterfactual thinking would vary as a function of the two factors manipulated. Turning first to congruence, it was considered possible that causing negative consequences for others might be deemed a more 'serious' transgression than for oneself, triggering greater reflection upon one's actions. Thus, in scenarios where there was a mismatch between the agent and the recipient of the consequences (i.e. incongruent scenarios), it was expected that participants might generate more counterfactual thoughts. Secondly, the 'RCFT' task investigated the extent to which agent identity might be associated with the ability to think counterfactually, since adopting the characters' perspective might have been more effortful for them. It was therefore predicted that participants would generate more counterfactual thoughts for situations where they were the agent of the 'bad' actions. However, neither of these possibilities was supported with respect to counterfactual fluency, since across both groups' fluency scores did not vary as a function of either the congruence or the agent identity condition.

Whilst these should be treated as exploratory findings and interpreted with caution, it nonetheless seems surprising that the generation of counterfactual thinking might be unaffected by whether the consequence is for the self or other, or whether the agent of the action is the self or other. One explanation of this might be that the task focused on the fluency of generation of counterfactual thoughts, but did not investigate the content of the thoughts produced. It may be that any effects of congruence or agent identify could be detected in counterfactual content, if not fluency.

Another possible explanation is that counterfactual thinking might play a similar role in learning about how to correct mistakes and their consequences, both for oneself and for others. This notion is consistent with the idea of 'simulation theory', which posits that one strategy for inferring the mental states of other people is to imagine one's own thoughts, feelings or behaviours in a similar situation (see Mitchell, Banaji & MacRae, 2005 for a review). Thus, self-reflection is postulated to provide a tool for predicting how others might be thinking or feeling in a given situation. Furthermore, there is currently a great deal of attention paid to the role of mirror neurons, which fire both when performing an action and observing others performing actions. Mirror neurons are potentially thought to underlie our empathic abilities by matching the actions of the self to those of others, which might then facilitate the inference of others' intentions (Hamilton & Grafton, 2006: see Chapter 2.3.2 for more detail), although this is highly speculative at this stage. The relationship between this basic self-other matching function and higher-level

forms of self-other representation, such as via counterfactual thinking, might provide an interesting avenue for further investigation.

In contrast to the counterfactual fluency measure, the ratings of regret and blame did prove sensitive to the manipulations of congruence and agent identity. For both sets of ratings, participants across groups gave higher estimations of regret and blame when the scenarios were incongruent, and in scenarios where they adopted their own perspective. This suggested that perpetrating actions that cause negative consequences for others rather than for oneself is perceived to be more serious, and that participants were harsher on themselves than on other agents. Whilst some previous work has looked at the dissociation between empathy for one's own pain versus that of others (e.g. Singer et al., 2004), surprisingly little work has explored emotions such as regret and guilt, or judgments of blame for self- versus other-actions and consequences. Some work has suggested that the experience of regret and the perception of blame are coupled, and these have jointly been referred to as 'counterfactual emotions' (Kahneman & Miller, 1986). The ratings of regret and blame in the present study showed similar patterns, although for ratings of regret, but not blame, there was an interaction between the scenario's congruence and the agent's identity; the difference between adopting one's own versus the main character's perspective was exacerbated in incongruent scenarios. This complexity may reflect the importance of considering both the role of the identity of perpetrator and of the recipient when interpreting counterfactual judgments. One study presented participants with scenarios in which 'Lucy' got food poisoning after visiting a restaurant (Macrae & Milne, 1992). Participants were asked to consider the scenario from either the restaurant's (agent's) or Lucy's (victim's) perspective, and then to rate how much of a fine should be levied against the restaurant, how much compensation Lucy deserved, and how much sympathetic they felt towards Lucy. As was expected, those considering the 'victim's' position made more generous suggestions of compensation and demonstrated greater sympathy for Lucy; conversely, they gave harsher suggestions of fines to the restaurant. The reverse was found for participants who were instructed to consider the scenario from the restaurant's point of view. Whilst this study did not examine counterfactual thinking per se, it also highlighted the interesting inter-play between perpetrators and victims, and the importance of considering both the role of the agent and that of the recipient of any consequences. The findings of the present study pose some interesting questions for further work to address: Is the experience of regret heightened for interactions where agents commit transgressions that have consequences for others, and does this correspond to judgments of blame? Do people judge their own mistakes more harshly than others, and if so why?

9.4.3 Counterfactual thinking, regret, blame and autistic traits

Although the predicted effects of the agent identity and congruence manipulations did not emerge for the counterfactual fluency measure, either as main effects or interactions with group, the main hypothesis of the study was confirmed: those with high versus low autistic traits produced fewer counterfactual alternatives across scenario types. This finding in adults with sub-clinical levels of autistic traits is consistent with previous work indicating impairment in the generation of counterfactual alternatives in children with ASD (e.g. Begeer et al., 2009). For the ratings of blame there was no main effect of group, but there were group by scenario condition interactions, which supports predictions regarding the role of both congruence and agent identity in differentiating the groups. On the other hand, the high versus low AQ groups did not differ their ratings of regret, either overall or in response to different scenario types. This is in contrast to the previous work findings impaired recognition and understanding of self-conscious emotions, including regret, in those with ASD (e.g. see Heerey et al., 2003).

What might account for the overall impaired ability to generate counterfactual alternatives in those with high versus low autistic traits? Moreover, how can we explain the atypical ratings of blame, but apparently preserved experience of regret in those with high versus low autistic traits? A number of possibilities will be considered in the following sections, including the possibility of executive dysfunction, impaired imagination, the role of empathic processes and social knowledge.

9.4.3.1 The role of executive functioning and imagination

As previously highlighted, executive dysfunction in a number of domains has been postulated in ASD which might account for the non-social differences observed in the disorder (e.g. repetitive or restricted interests), and also contribute to associated social difficulties (Hill, 2004). Consistent with this, recent work indicated that individuals with high levels of autistic traits displayed impairments on both a global measure of executive functioning, and on sub-components of executive functioning assessing behavioural regulation and metacognition (Christ, Kanne & Reiersen, 2010).

Counterfactual thinking is a complex cognitive skill that is likely to draw upon a number of executive functions. For instance, Beck, Riggs & Gorniak (2009) investigated the relationship between counterfactual thinking and executive functioning in typically-developing children aged three and four. They found that the children's performance on different types of counterfactual thinking tasks was strongly correlated. Regression analysis revealed that scores on a measure of inhibitory control predicted performance on all three of the counterfactual thinking tasks, but that working-memory measures did not. Inhibitory control and working memory are both aspects of executive functioning, and thus the authors suggested that executive difficulties on this task might have specifically manifested as problems with

inhibitory control, whereby children who experienced difficulties with the counterfactual thinking tasks might thus have struggled to inhibit the truth: what they knew to be true might have competed with their ability to think about what might have been. However, the extent to which counterfactual thinking is mediated by specific aspects of executive functioning, such as planning, mental flexibility, inhibition, generativity or self-monitoring or domain-general executive functioning remains unclear.

Counterfactual thinking not only requires an inhibition of the 'truth', but also the capacity to imagine novel alternatives. Some previous work has indicated that children with ASD only show difficulties with counterfactual thinking when the task requires the production of new ideas to counteract reality (additive counterfactual thinking), rather than simply retracting mistakes (subtractive counterfactual thinking) ("e.g. If only I had done… instead" as opposed to "If only I hadn't done …"). Whilst children with ASD might also display difficulties with inhibition, their ability to 'un-do' mistakes by engaging in subtractive counterfactual thinking appears to be relatively intact. Rather, it is the more demanding task of creating novel alternatives that seems to pose a problem. This is consistent with evidence of a general deficit in imagination in those with ASD (Eycke & Muller, 2015). For instance, Low, Goddard & Mesler (2009) found that people with ASD displayed deficits in imaginative drawing, and that this was associated with lower generativity scores on a task that involved suggesting novel ideas for how to use everyday objects, such as a newspaper or brick.

From previous work examining counterfactual thinking in people with ASD it is not clear whether the difficulties identified reflect an overall deficit in the ability to engage in counterfactual reasoning, or a selective deficit (or more pronounced difficulty) for counterfactual thinking in relation to social/ emotional versus physical/ practical events. For instance, Beeger et al., (2009) identified differences in children's ability to engage in counterfactual thinking about practical events (e.g. getting muddy feet on the kitchen floor), which was argued to reflect an executive failure. Although the scenarios used focused on practical harm, arguably these had social (e.g. upset your parents) as well as practical (e.g. have to clean it up) consequences. Thus, it is not clear whether differences would have emerged between the groups had the events involved more pure forms of practical harm. Similarly, whilst the RCFT included a range of negative outcomes, these were all broadly social/emotional in nature (e.g. embarrassment or offending someone). In future work it would be interesting to explore whether more nuanced group differences emerged, in either the fluency or quality of the counterfactual produced, when comparing such social/emotional harm with physical harm (e.g. break something). This might also help to clarify the extent to which this reflects an inability to appreciate and reflect upon social consequences, or more general executive difficulties.

It is possible that an impairment in executive functioning might have underpinned the overall reduced fluency on the 'RCFT' task seen in those with high versus low autistic traits. Whilst no measures of executive functioning were included in the present study, measures of inhibition and verbal fluency were

administered in a previous study in this thesis. The high AQ participants described in Chapters 7 and 8 displayed some mild inhibitory difficulties (please see Chapter 7.2.4.1 and 7.3.3.1), but performed similarly to those with low AQ traits on the measure of verbal fluency. Since the participants used in the present study were drawn from a similar, relatively homogeneous population of university students, it is likely that their cognitive profiles would be fairly comparable. Thus, it seems possible that the high AQ participants generated fewer counterfactual alternatives as a result of impaired inhibitory processing, leading to a reduced capacity to suppress the 'truth' and consider alternatives. It is also possible that the reduced fluency displayed by participants with high AQ traits represents a general imaginative deficit, which may have contributed. In the 'RCFT' task participants were required to generate novel ideas by engaging in additive counterfactual thinking, and simply 'undoing' mistakes was not sufficient to score a point on this measure. Whilst verbal fluency was found to be unimpaired in the previous sample, the extent to which verbal fluency provides an adequate test of the ability to generate or imagine novel ideas is not clear (e.g. Spek et al., 2009), and it is likely that it would not be sensitive with more subtle difficulties. Verbal fluency tests simply require participants to list words according to various rules, whereas generating novel alternatives on the 'RCFT' task is overall more demanding.

The possible role of executive functions in accounting for the ratings of regret and blame is less clear. One possible factor in differentiating regret and blame ratings is that the experience of regret may depend entirely on the ability to take the perspective of our fictive selves or of others, whereas ratings of blame might involve the integration of this information with factual information about the scenario, such as causation (Nicolle, Ropar & Beck, 2014). Thus, making ratings of regret might have been less executively demanding than ratings of blame, which could potentially account for the apparent similarity between groups in ratings of regret but not blame on the 'RCFT' task.

9.4.3.2 The role of cognitive empathy

Could cognitive empathy also account for the group difference in the generation of counterfactual thoughts? As discussed, ASD is thought to be characterised by problems with taking others' perspectives, or cognitive empathy. The ability to create counterfactuals develops throughout childhood and is thought to contribute to reasoning about other people's perspectives, including their false beliefs - a key milestone in cognitive empathy. Both counterfactual thinking and false belief tasks require mutating aspects of one's own mental representation, either by mentally going back in time and changing antecedents or by considering another's viewpoint (Riggs, Peterson, Robinson & Mitchell, 1998). Thus, these processes both involve the ability to suspend one's current mental representation, to simulate an alternative representation, and to distinguish these and compare them. Several studies have linked performance on false belief and

counterfactual thinking tasks in both typically-developing children and children with ASD (e.g. Grant et al., 2004).

Impaired cognitive empathy in those with high AQ traits might therefore be expected to be associated with poorer performance on the 'RCFT' task fluency measure. This prediction was confirmed, since the high AQ group generated significantly fewer counterfactual alternatives. As discussed above (see section 9.4.2.1), it seems plausible that counterfactual thinking is a device that is simultaneously used to generate insight both into one's own and into others' behaviour. If this is the case, then for those with high AQ traits, the difficulties with engaging in counterfactual thinking might be just as pervasive when examining one's own behaviour as when considering others, since both might rely on the same mechanisms. This interpretation could potentially account for the **overall reduction** in fluency, but further work is needed to understand the relationship between thinking about one's own versus other's thoughts particularly within a counterfactual framework.

Whilst there was a main effect of group, the predicted interactions with congruence and agent identity were not found. The high AQ group was expected to demonstrate particular difficulty with the scenarios that required adopting another agent's perspective, and/or considering the negative consequences for others versus themselves. This might reflect a lack of sensitivity in the task, since no main effects of congruence or agent identity were found for the fluency measure. It is also possible that the participants did not in fact differ in their cognitive empathy skills, although this seems unlikely. Whilst cognitive empathic ability was not directly assessed in the present study, this was previously assessed in a similar sample of participants in Chapter 7. The high AQ group in that study displayed difficulties with a mentalistic inference test, although the differences between groups were somewhat subtle (see Chapter 7.3.4.3 and 7.3.3.3). Therefore, it seems reasonable to assume that the high AQ participants in the present study might have had weaker cognitive empathy skills than those with low AQ traits.

Should impairment in cognitive empathy also lead to group differences on the other task measures, regret and blame? As discussed, the experience of regret for one's actions is considered to be dependent upon the capacity to represent another's emotional state (i.e. cognitive empathy), and also to occur in response to thinking about alternative outcomes and comparing them with reality (i.e. counterfactual thinking). The lack of a group difference on the regret ratings in the present study contrasts with previous work demonstrating difficulties with self-conscious emotions, such as regret. However, the nature of the tasks is very different, since these studies tended to focus on the identification of emotions, typically from pictures of faces (e.g. Heerey et al., 2003), whereas the 'RCFT' task required participants to rate their personal experiences of regret in response to specific events. Thus, it is possible that those with ASD are able to make judgments about emotions such as regret when given contextual information such as in the present study, but did not necessarily recognise them in abstract stimuli, such as pictures of faces.

With regards to blame, making causal attributions about our own and others' actions is crucial to comprehending everyday life and regulating society (Channon et al., 2011). After making mistakes or suffering negative consequences there is a desire to reflect on how these may have happened and who was responsible, in order to hold the relevant individuals accountable (Alicke, 2000; Shaver, 1985). Thus, making judgments about the cause and responsibility of actions is a complex ability that is likely to require both cognitive empathy, to consider the intentions and motives behind actions, and the ability to engage in imagining alternatives. The high AQ group showed reduced sensitivity to the congruence of the scenarios and a greater tendency to blame themselves when making judgments of blame. The high AQ thus appear to have primarily considered the agents' identity, rating themselves to be even more blameworthy than did the low AQ group, and to show less consideration for who the recipients of the agents' actions were. This pattern could potentially be accounted for via a failure of cognitive empathy, whereby the high AQ group did not consider the consequences for the recipient when making their ratings.

9.4.3.3 The role of emotional empathy

What role might emotional empathy potentially have played in the findings? Whilst both cognitive and emotional empathy might potentially be associated with counterfactual thinking, to date there is little work examining the role of emotional empathy in the generation of counterfactual thoughts. Instead this work has tended to focus on the ramifications of engaging with counterfactual thinking for the experience of emotions (e.g. Boninger, Gleicher & Strathman, 1994). There is very little experimental work addressing whether a deficit in emotional empathy might result in a reduced capacity to engage in counterfactual thinking, and the existing work has tended to emphasise the role of cognitive rather than emotional factors in generating counterfactual thoughts. Thus, it remains unclear how emotional empathy might have contributed to performance on the fluency measure task, regardless of whether this was intact in the high AQ group or not.

Experiencing regret involves the capacity to represent another's emotional state, but also an appreciation of the consequences of one's actions. The experience of regret for one's actions thus might be considered to primarily dependent upon emotional rather than cognitive empathic processes, which are thought to be largely unaffected in pure ASD (Blair, 2008). This account would favour the present findings, where regret was found to be intact in those with high AQ traits, regardless of the perspective-taking demands of the scenarios. It is more difficult to argue that blame might be primarily dependent upon emotional rather than cognitive empathic processes, since blame does not reflect a basic or self-conscious emotion, but rather a cognitive appraisal that might be coloured by emotional processes.

9.4.3.4 The role of social knowledge

Performance on the 'RCFT' task might also have depended in part upon prior knowledge and experience of how one should behave or respond in these situations. As discussed, social learning might be impaired in ASD, resulting in less sophisticated or developed knowledge stores to fall back on. Generating alternatives on the 'RCFT' task requires an understanding of what is expected by the main character and also of the wider social context. For instance, in the scenario where the agent makes a mistake and thinks the colleague's proposal is a joke, generating counterfactual alternatives requires a grasp of both the relatively short-term practical implications of their actions (i.e. the colleague's proposal was not taken seriously), and also the wider-ranging social implications both for the colleague (i.e. the colleague might feel humiliated, and fearful of putting forward new ideas in future meetings) and for themselves (i.e. the colleague might not trust them, or others might think that they are insensitive). The high AQ group might have either had a more limited social knowledge-base relating to these scenario aspects, and/or suffered an executive failure to apply this knowledge appropriately. This could manifest as a focus solely on the short-term practical consequences, and a failure to consider the longer-term social ramifications of their own or others' actions. This could hypothetically account for the overall reduced fluency, as a lesser appreciation of the consequences might result in a more limited range of possible alternatives. It is less clear how an account in terms of reduced social knowledge could explain group interactions on the blame ratings, although it is possible that a more limited appreciation of the wider social context and its antecedents in the high AQ group might lead to differences on blame ratings.

9.4.4 The sensitivity of the 'RCFT' task

In what ways might the 'RCFT' task be improved to demonstrate greater sensitivity to the group differences and task manipulations in the present study? One key limitation of the 'RCFT' task is that it only assessed participants' fluency for counterfactual thinking and did not examine the quality of the counterfactuals produced. The 'RCFT' task data were initially inspected for qualitative scoring, but several obstacles emerged. Firstly, participants were instructed to produce as many counterfactuals as possible when responding to each scenario. Since the responses were so unconstrained it was difficult to decide how to proceed in scoring these. Scoring all of the counterfactuals produced in response to each scenario might conflate the fluency of responding with the quality of responding, since a response would potentially score higher for quality if it contained lots of counterfactuals, even if these were of poor quality, than a response containing just one or two high quality counterfactual thoughts. Alternative approaches were considered, including scoring only the first or best counterfactual produced. However, this might undermine the richness inherent in the data, and would require a very clear hierarchical scoring structure, which did not emerge. Secondly, unlike on other tasks presented in this

thesis, clear themes did not emerge consistently across the different scenarios. Thus, it would have been difficult to develop scoring criteria that applied equally across the scenarios. Finally, insights garnered from testing indicated that the most prominent difference emerging between the groups' performance was in relation to the fluency of the response generated, rather than the content of what was produced. Therefore, it was not considered appropriate or worthwhile to try to assess the quality of counterfactuals produced on this task. Despite these limitations, some very broad themes emerged upon inspection of the data, concerning the extent to which thoughts corresponded to more practical, social or emotional focused alternatives. Given the paucity of research exploring the nature of the content of counterfactuals within the autistic spectrum, and how this might mediate subsequent judgments or experiences, this was systematically followed up in the next experimental chapter, using a simpler paradigm (Chapter 10).

A great deal of work has focused on the mechanistic aspects of counterfactual reasoning (e.g. upwards or downwards, i.e. offering better or worse alternatives; additive or subtractive, i.e. suggesting new actions or 'undoing' actions) (Epstude & Roese, 2008). The 'RCFT' task examined counterfactuals that offered novel alternatives which could have prevented the negative consequences, using the stem "If only I/they had…", i.e. these scenarios were designed to prompt upwards additive counterfactuals. This was based on previous work indicating difficulty with these types of counterfactual thoughts in those with ASD, but not necessarily with all types of counterfactual thinking. Nonetheless, it would have been interesting also to explore how the groups might have performed in relation to different types of counterfactual stems.

Furthermore, simple ratings of regret and blame, as used in the present study, may not have been sufficiently sensitive to reveal any potential difficulties with these types of judgments in the high AQ group. However, a deeper probe of such emotional experiences or judgments might have revealed group differences, such as asking participants to describe their own or others' emotional states rather than simply to rate them. Another possible factor is the nature of the negative consequences for the recipient; the scenarios described outcomes that might be said to be irritating but not devastating. Using more emotive scenarios might have revealed stronger group effects in relation to these manipulations, and could potentially have increased both the generation of counterfactual thoughts and the strength of the regret and blame ratings.

9.4.5 Conclusion

This study examined how the ability to engage in counterfactual thinking and to make related ratings of regret and blame might be impaired in those with high versus low autistic traits. Overall, the study's main hypothesis was confirmed: those with high AQ traits produced fewer counterfactual alternatives. Whilst impaired ability to take others' perspectives has long been established in the autistic

spectrum, far less is known about the about the ability to take the perspective of counterfactual versions of oneself or others. This study provides further evidence to support the notion that those on the autistic spectrum deploy counterfactual abilities atypically (Nicolle, Ropar & Beck, 2014). Interestingly, neither of the groups appeared to differentiate between the scenario types, and produced fairly consistent numbers of counterfactuals across the board. However, for ratings of regret and blame a more complex pattern emerged, with both types of ratings influenced by agent identity and the congruence between the perpetrator and the recipient. There were no group differences in regret ratings, whereas for ratings of blame the picture was less clear, since those with high AQ traits differentiated less on the basis of scenario congruence and more on the basis of agent identity.

Counterfactual thinking will be further assessed in Chapter 10, using a new scenario-based task designed to focus on thought content rather than generation. The requirement to produce novel ideas will be removed, and participants will instead be asked to judge counterfactual alternatives, and to rank the alternatives in order of preference. The alternatives provided will correspond to different types of counterfactual thoughts that could occur in the scenarios, comparing those that are more practical and others that are more social or emotional in nature. Ratings of regret and blame will also be re-examined, in order to assess if any group differences relate to blame rather than to regret ratings, as the findings from the present study suggest.

Chapter 10: Counterfactual judgments for others' mistakes and the experience of regret, blame and guilt

10.1 INTRODUCTION

As highlighted (see Chapter 9.4.4), a great deal of work has examined different types of counterfactual thoughts with regard to the mechanistic aspects, such as whether they aim to improve or avoid outcomes, and the circumstances under which these occur. Research also suggests that counterfactual thinking serves both to offer solutions to mistakes and to allow people to reflect upon the chain of events leading up to an outcome. Reflecting on events via counterfactual thinking might have both positive and negative effects, leading to greater self-awareness and perhaps even prompting rumination (Mandel & Dhami, 2005; Branscombe et al., 2003). Some work has also addressed the nature of the content of counterfactual thoughts, assessing which aspects of an event people choose to focus on (see e.g. Kahneman & Miller, 1986; Roese, 1997).

Previous findings suggest that the autistic spectrum is associated with limited counterfactual abilities (e.g. Begeer et al., 2009). However, research in those with ASD has tended to focus on the fluency of counterfactuals produced, and/or the ability to produce different types of counterfactuals according to their mechanistic aspects, executive load, or functional role. The nature of the content of counterfactual thoughts has not been examined within the autistic spectrum; thus, this study will assess how people use more practically-focused or more socially-oriented thoughts, which might evoke and inform related emotions, judgments and decisions, and guide future behaviour, and if this varies in as a function of autistic traits. A novel task, "Counterfactual Judgments", was designed to examine whether those with high AQ traits would also differ in their preference for counterfactual types in terms of their content. This aimed to complement the findings of Chapter 9, demonstrating reduced fluency for counterfactual thinking in participants with high versus low AQ scores, by exploring qualitative as well as quantitative aspects of counterfactual performance.

This task described scenarios in which a main character committed a 'bad' action that led to a negative consequence for the participant. For instance, in one scenario the participant missed their graduation ceremony because their sibling filled up the car with the wrong type of fuel, and they broke down on the motorway. First, participants were asked to rate how much characters would regret their 'bad' actions, and how much they were to blame for the negative consequences. Next, participants were presented with a list of different types of counterfactual thoughts that described alternative actions that could have prevented the negative consequences. These included three types of thoughts: those that were

'practical' in nature, such as 'if only I had been driven by someone else'; those that were 'emotional' in nature, such as 'if only my sibling had been more trustworthy'; and those that were 'punitive' in nature, e.g. 'if only my sibling had been unwell and unable to drive me to the ceremony'. Each of these types of counterfactuals focused on two targets, the character or the participant. Participants were required to judge these in two ways: first they rank-ordered them according to their preference, and then they rated how much guilt experiencing each thought would evoke.

In order to assess any group differences in executive functions and empathy, the same battery of measures that was described in Chapter 7 (see: 7.3.3) was used for the current samples. As discussed, executive dysfunction might play a role both in the comprehension of social scenarios and in formulating counterfactual thoughts and judgments (see Chapter 9.4.3.1). Counterfactual thinking has also been linked to the ability to impute others' mental states as measured by false-belief tasks, which are thought to be a key component of cognitive empathy. The possible role of empathic responsivity to others' distress (i.e. emotional empathy) has not been directly examined in relation to counterfactual thinking (see Chapter 9.4.3.2 and 9.4.3.3). Nonetheless, it seems possible that if either cognitive or emotional empathic routes are impaired, this might influence the types of counterfactual thoughts the high AQ group preferred with respect to both their content (practical, emotional or punitive) or their target (participant or character).

10.1.1 Hypotheses

10.1.1.1 'Counterfactual Judgments' task

10.1.1.1.1 Ratings of regret and blame

In line with work indicating difficulties with taking others' perspectives, it was predicted that the high AQ group would under-estimate the main characters' experience of regret, in comparison to their low AQ counterparts. On the basis of previous work (Channon et al., 2011), it was also expected that the groups would differ in their ratings of blame, with the high AQ group blaming the main characters more for their 'bad actions' than their low AQ counterparts.

10.1.1.1.3 Preferences and ratings of guilt for counterfactual alternatives

In view of the putatively poorer empathic and social skills of the high AQ group, it was expected that the high AQ group would be likely to rank practical-focused counterfactuals higher than the low AQ group, and that they might also rank punitive counterfactuals higher than the low AQ group. It was also expected that the groups would differ with respect to the extent to which they preferred counterfactuals where they themselves versus the main characters were the targets, with the high AQ group preferring

counterfactuals where the main character was the target more than the low AQ group, particularly for counterfactuals that were punitive in nature. For ratings of 'guilt' it was expected that across groups, participants would rate punitive versus emotional/ practical counterfactuals as eliciting more guilt. In view of previous work indicating difficulties with self-conscious and complex emotions in ASD, it was expected that the high AQ group would experience less guilt overall.

10.1.1.2 Additional measures

10.1.1.2.1 Executive functioning

In order to explore whether executive functioning played a role in performance on the 'Counterfactual Judgments' task, two tests from the D-KEFS battery (Delis et al., 2001) were administered: inhibition and verbal fluency. These were expected to differentiate the groups, whereby the high AQ group would demonstrate poorer performance than the low AQ group on the inhibition measure, and also possibly on the verbal fluency measure.

10.1.1.2.2 'IRI'

The 'IRI' is a self-report questionnaire measure that explores different dimensions of empathy (Davis, 1980). It was predicted that the high AQ group would score significantly lower on this measure, at least for items exploring cognitive aspects of empathy.

10.1.1.2.3 'MIT'

The 'MIT' is a scenario-based measure that explores understanding of characters' intentions via interpretation of their sarcastic remarks or actions (Channon et al., 2005; 2007). It was predicted that the high AQ group's performance would be significantly poorer, at least for action-related items.

10.2 METHODS

10.2.1 Screening phase

10.2.1.1 Screening participants and procedure

An opportunistic sample of 828 full-time university students (58.60% female) who were fluent in English and aged 18 or over (mean age 20 years old) was recruited for the screening phase of the study. All participants completed the AQ (Baron-Cohen et al., 2001), and total scores were calculated for the whole

sample. Participants within the highest-scoring and lowest-scoring 10% of males and females were contacted and invited to take part in the second stage.

10.2.2 Experimental phase

10.2.2.1 Design

There was one between-participants factor of AQ group (high vs. low scorers). There were two within-participant factors, type of counterfactual alternative (practical, emotional or punitive) and the targetof the counterfactual (participant or character).

10.2.2.2 Experimental participants and procedure

Of those contacted from the screening phase, 20 (10 male, 10 female) individuals from the upper range and 20 (10 male, 10 female) individuals from the lower range agreed to take part in the experimental phase of the study, forming two groups of high AQ and low AQ participants.[1] AQ scores ranged from 26 to 38 in the high AQ group (27–40 for male participants, and 27–38 for female participants), and 4–13 in the low AQ group (5–11 for male participants, and 5–11 for female participants). A *t* test confirmed that AQ scores differed significantly between groups, $t(39) = 25.95$, $p < .0001$; mean AQ scores were 31.43 (SD = 3.65), and 7.15 (SD = 2.08) for the high and low AQ groups respectively. The groups did not differ significantly in age, $t(39) = .149$, $p = .883$; mean age was 20.34 (2.72) and 20.53 (2.70) years for the high and low groups respectively.

All participants were tested individually, and provided written informed consent before completing the 'Counterfactual Judgments' task (in preparation)[2], and completed some measures of executive functioning and empathy. Participants also completed some further measures that are described in Chapter 11. The battery of tasks was administered in counterbalanced order to ameliorate any potential order effects upon performance, and this was counterbalanced within each group.

10.2.3 'Counterfactual Judgments' Task

This task was designed to examine how people judged different types of counterfactual thoughts in response to negative consequences. The 'Counterfactual Judgments' task was developed by inspecting the responses to the 'RCFT' task (please see Chapter 9), and using these as an informal guide to the different types of counterfactual thoughts that participants generate on tasks of this nature. The 'Counterfactual Judgments' task consisted of nine short scenarios that described a situation in which a character known to

[1] One participant had to be excluded from the low AQ group owing to missing data on the 'Counterfactual Judgments' task. This resulted in 9 male and 10 female low AQ participants.

[2] This task was developed jointly with Ms Karishma Vyas. Ms Vyas used this task to explore counterfactual judgments in people scoring high versus low on a self-report measure of psychopathic personality traits. This study is included in her doctoral thesis (Vyas, 2015).

the participant (i.e. a friend, sibling, housemate or colleague) accidentally did something that inconvenienced the participant. For example, in one scenario, the participant's housemate forgot that it was the funeral of the housemate's grandmother that day and invited friends over to their house. Other situations included having their property damaged or losing time, effort or money. For each situation participants were asked to rate the extent to which they thought the other character would regret their actions, and the extent to which the character was to blame for the consequences. Participants were then presented with a list of different types of counterfactual thoughts, each of which described alternative actions that would have prevented the negative consequences. These varied according to both their type (practical, emotional or punitive) and the target (character or participant). There were thus six counterfactual alternatives for each situation. Participants were required to judge these in two ways: first they rank-ordered them according to their preference, and then they rated how much guilt experiencing each thought would evoke. Please see Figure 14 for a full example.

The type of relationship between the main character and the participant was counterbalanced across items, and gender was not specified. The order of the items and counterfactual alternatives was counterbalanced to create two different presentation versions, and this was counterbalanced within each group.

10.2.3.1 Administration

Participants were given a paper booklet containing instructions, all nine scenarios, and corresponding questions. After reading the instructions, they completed the first item by filling in the booklet. After completing the first item, the experimenter checked that they had filled it in correctly and gave the participant the opportunity to ask questions. The participant then filled in the rest of the booklet alone. The relevant scenario was presented on each page of corresponding questions in order to reduce the confounding effects of memory load.

10.2.3.2 Scoring

10.2.3.2.1 Ratings of regret and blame

For each item participants rated how much the character would regret their actions, or were to blame for the negative consequences on a scale of 1-10, where higher scores indicated greater regret or blame. Ratings were each then summed across scenarios to create a total score for regret and a total score for blame (range 9-90). These were converted to percentages for ease of interpretation and comparison.

10.2.3.2.2 Preferences for counterfactual alternatives

For each item, participants ranked the six counterfactual alternatives in accordance with their preference. The alternative ranked as the first choice received a score of 1 and the alternative ranked as their last choice received a score of 6. These individual counterfactual alternative scores were summed across all items for each participant, resulting in six total preference scores corresponding to each counterfactual alternative: Practical alternative: character target; Practical alternative: participant target; Emotional alternative: character target; Emotional alternative: participant target; Punitive alternative: character target; Punitive alternative: participant target. Each of these scores had a range of 9-54, with lower scores denoting greater preference. These were converted to percentages for ease of interpretation and comparison.

10.2.3.2.3 Ratings of guilt for counterfactual alternatives

For each counterfactual alternative participants rated how guilty they would feel if they had that thought, on a scale of 1-10, where higher scores indicated greater guilt. Guilt ratings were then summed across all items for each participant, resulting in six total guilt scores corresponding to each counterfactual alternative (range 9-90). These were converted to percentages for ease of interpretation and comparison.

Figure 14: Example scenario from the 'Counterfactual Judgments' task

Scenario: "Today, it was your grandmother's funeral. You feel exhausted after it has finished and head home looking forward to a quiet night in. Your housemate has forgotten about the funeral and invited some of their friends over."

Questions:

1. How much do you think your housemate would regret forgetting about the funeral and inviting their friends over? (Please circle one number ONLY)
1----------2---------3---------4----------5---------6---------7---------8----------9---------10
Not regret at all Regret very much

2. How much do you think your housemate is to blame for spoiling your quiet night in?
(Please circle one number ONLY)
1----------2---------3---------4----------5---------6---------7---------8----------9---------10
Not to blame at all Very much to blame

3**. Here is a list of options that would have prevented your housemate inviting some friends over and spoiling your quiet night in.**
Please rank order the options according to your preference, whereby 1= most preferred and 6 = least preferred (Please use **ONE** option per ranking).

a) If you had reminded your housemate that it was your grandmother's funeral today
(Participant target, practical)

b) If your housemate had invited their friends over another night
(Character target, practical)

c) If you had realised your housemate was so selfish
(Participant target, emotional)

d) If your housemate had been more considerate
(Character target, emotional)

e) If you had made your housemate's friends feel unwelcome in your flat
(Participant target, punitive)

f) If your housemate had recently lost a relative and knew how it felt
(Character target, punitive)

4.

a) **How guilty would you feel "if you wished you had reminded your housemate it was your grandmother's funeral today"?**

Please circle one number **ONLY**)
1----------2----------3----------4----------5----------6----------7----------8----------9----------10
Not at all guilty Very guilty

[Repeat for options b-f)

10.2.4 Additional measures

10.2.4.1 Executive functions

Two neuropsychological tests from the D-KEFS battery that explore different aspects of executive functions were administered (Delis et al., 2001). The 'Colour-Word Interference' Test involves inhibiting habitual responses by naming the ink colour of colour words written in conflicting colours (e.g. RED printed in green ink). There were also two control conditions, colour naming (naming blocks of colours) and word reading (reading colour words written in black ink). The 'Verbal Fluency' test involves generating words within a time limit according to certain rules, beginning with either a particular letter or belonging to a semantic category. The scores were converted into age-scaled score equivalents using the manual, which gives a range of 1-19 for each measure.

10.2.4.2 'IRI'

This is a self-report questionnaire designed to measure the global concept of empathy (Davis, 1980). There are four subscales that examine different dimensions of empathy: 'Perspective-taking', designed to assess spontaneous attempts to adopt the perspectives of others; 'Fantasy', designed to assess the tendency to identify with characters in movies, plays and other fictional situations; 'Empathic Concern' refers to respondents' feelings of warmth, compassion and concern for others; 'Personal Distress' refers to personal feelings of anxiety and discomfort resulting from observing others' negative experiences. For example items and more detail about the measure please see Chapter 7.2.4.2.

10.2.4.3 'MIT'

This test was adapted from Channon et al., (2005; 2007). It was designed to assess mentalising skills, a crucial aspect of cognitive empathy, and consisted of two different item sets: sarcastic remarks or human actions. Participants were asked to explain verbally what the characters meant by their remarks, or why they carried out the actions. A score of two points was given for responses giving a clear correct explanation of the remark or action, a score of one point when the answer was not incorrect, but was not adequately explained, and a score of zero points when the answer was incorrect or irrelevant. Thus, the two scores were calculated for (1) 'MIT Action': mentalistic action items and for (2) 'MIT Sarcasm': sarcastic items, both with a range of 0-20 since there were 5 items in each set. Please see Figure 7 for an example of a sarcastic and an action item with scoring. For an example item and more detail about the measure please see Chapter 7.2.4.3.

10.3 RESULTS

10.3.1 Data analysis

Means and standard deviations (SD) for each of the measures of the 'Counterfactual Judgments' Task are presented below in Table 7, and for each of additional measures these are shown in Table 8. Some of the variables were positively skewed (preferences for counterfactual alternatives and ratings of guilt for counterfactual alternatives), and these could not be transformed to normality, non-parametric Mann-Whitney group comparisons showed the same pattern as the parametric tests, and hence parametric statistics were reported here. A significance level of .05 was adopted, with a stricter level for post hoc tests to control for multiple comparisons.

10.3.2 'Counterfactual Judgments' task

10.3.2.1 Ratings of regret and blame

A t-test was conducted to compare the groups' performance on ratings of how much the main characters (the agents) would regret their 'bad' actions. However, the difference between groups did not reach significance, $t(37) = 1.67$, $p = .103$.

A t-test was also conducted to compare the groups' performance on ratings of how much participants (the recipients) blamed the main characters (the agents) for negative outcomes. There was a significant difference between groups, $t(37) = 2.11$, $p = .041$, whereby the high AQ group gave higher ratings of blame.

10.3.2.2 Preferences for counterfactual alternatives

A repeated measures 2 x 3 x 2 ANOVA was conducted to compare the groups' preference scores for each type of counterfactual alternative. There was one between-participants factor of AQ group (high versus low scorers), and two within-participants factors: type of alternative (practical/ emotional or punitive), and target (participant versus character).

There was a significant main effect of type of alternative, $F(2,36) = 2517.75$, $p < .0001$, and a significant main effect of target, $F(1,37) = 52.69$, $p < .0001$. There was also a significant type of alternative by target interaction, $F(2,36) = 49.85$, $p = < .0001$. Inspection of the mean scores suggests that across groups, participants preferred practical alternatives, followed by emotional and then by punitive alternatives (please note that a lower score = greater preference). With respect to target, participants across groups preferred alternatives that targeted the main character, suggesting actions that they could have taken to prevent the negative outcome. When the alternative by target interaction was examined, participants across groups preferred practical and emotional alternatives that targeted the main character as opposed to

themselves, but for punitive alternatives they gave a high preference to those which targeted themselves rather than others.

There was no significant main effect of group $F(1,37) = 0.86$, $p = .360$, nor were there significant types of alternative by group, $F(2,36) = 2.27$, $p = .118$, or target by group interactions $F(1,37) = 0.21$, $p = .650$. However, there was a significant three-way interaction, $F(2,36) = 5.40$, $p = .009$. In order to examine the nature of this interaction, six difference scores were calculated for each group, to compare each pair of counterfactual alternatives ((1) practical - emotional; (2) practical - punitive; (3) emotional – punitive), that targeted either the participant or the main character. Post-hoc *t*-tests were conducted on these difference scores, using a strict significance level ($p = .01$). There was a significant difference between groups on the character practical – punitive difference scores, $t(37) = 3.47$, $p = .001$. No other comparisons reached significance (all $p > .05$).

10.2.3.2.4 Ratings of guilt for counterfactual alternatives

A repeated measures 2 x 3 x 2 ANOVA was also conducted to compare the groups' ratings of guilt for each type of counterfactual alternative. There was a significant main effect of type of alternative, $F(2,36) = 445.91$, $p < .0001$, and a significant main effect of target, $F(1,37) = 34.54$, $p < .0001$. There was also a significant type of alternative by target interaction, $F(2,36) = 84.62$, $p < .0001$. Inspection of the mean scores suggests that across groups participants displayed a similar pattern for ratings of guilt of the counterfactual alternatives as for their rankings; they rated practical alternatives as evoking the least guilt, followed by emotional and then by punitive alternatives (please note that a lower score = less guilt). Furthermore, with respect to target, participants across groups once again gave lower ratings of guilt for experiencing thoughts about the main character as opposed to themselves. When the alternative by target interaction was examined, participants across groups rated practical and emotional alternatives that targeted the main character as evoking less guilt than those that targeted their own perspective, but rated punitive alternatives as evoking more guilt for the main character than for themselves.

There was no evidence of group differences, since there was no significant main effect of group $F(1,37) = 2.92$, $p = .096$. There was also no significant type of alternative by group interaction, $F(2,36) = 1.26$, $p = .297$, nor a target by group interaction $F(1,37) = 0.14$, $p = .713$; nor was there a significant three-way interaction, $F(2,36) = 0.22$, $p = .802$.

Table 7: Mean percentage scores and standard deviations for the 'Counterfactual Judgments' task

	Low AQ group (N = 19) M (SD)	High AQ group (N = 20) M (SD)	Significance	Effect size (d)
Regret (%)	84.15 (6.68)	79.67 (9.70)	.103	0.53
Blame (%)	63.16 (11.66)	71.22 (12.14)	.041*	0.67
Counterfactual preference rankings (%) *(N.B. Lower score = higher ranking, i.e. greater preference)*				
Participant Target				
Practical	36.25 (5.02)	34.63 (5.02)	—	—
Emotional	65.40 (4.28)	65.20 (5.94)	—	—
Punitive	82.75 (4.28)	82.03 (5.80)	—	—
Character Target				
Practical	26.41 (3.94)	30.55 (5.11)	—	—
Emotional	49.22 (5.93)	49.07 (7.19)	—	—
Punitive	89.87 (4.63)	86.57 (6.47)	—	—
			Alternative < .0001** Target < .0001** Group = NS Alternative*target < .0001** Alternative*group = NS Target*group = NS Alternative*target*group = .009**	
Counterfactual guilt (%)				
Participant Target				
Practical	25.90 (6.76)	29.56 (11.21)	—	—
Emotional	51.70 (13.45)	59.28 (13.46)	—	—
Punitive	70.53 (13.64)	73.44 (11.61)	—	—
Character Target				
Practical	15.97 (4.58)	20.61 (10.05)	—	—
Emotional	32.87 (14.72)	42.56 (23.60)	—	—
Punitive	80.64 (12.06)	82.67 (10.08)	—	—
			Alternative < .0001** Target < .0001** Group = NS Alternative*target = NS Alternative*group = NS Target*group = NS Alternative*target*group = NS	

*** p = .025* *** p = .05* *NS = not significant*

10.3.3 Additional measures

10.3.3.1 Executive functions

10.3.3.1.1 Colour-Word Interference Test

For comparisons on the three conditions of the 'Colour-Word Interference' test a strict significance level of .017 (p = .05/3) was adopted to control for multiple comparisons. The first two conditions did not involve any inhibition, but were simply included to gauge overall performance speed and any perceptual difficulties that might impair performance. There was no difference between the groups when they were required to simply name blocks of colours, $t(37) = .333$, $p = .533$, or when they were simply required to read colour words, $t(37) = 0.03$, $p = .973$. For the last condition, there was an interference component requiring participants to inhibit responses when naming the colour of the ink rather than a conflicting colour word (e.g. RED printed in green ink). As predicted, the high AQ group were significantly slower to complete this task, $t(37) = 2.88$, $p = .007$.

10.3.3.1.2 Verbal Fluency Test

For comparisons on the three conditions of the 'Verbal Fluency' test a strict significance level of .025 (p = .05/2) was adopted to control for multiple comparisons. This test involves generating words within a time limit according to certain rules, beginning with either a particular letter or belonging to a semantic category. The groups significantly differed for the letter fluency $t(37) = 2.65$, $p = .012$, measure but not for the category fluency $t(37) = 1.19$, $p = .244$ measure, with the high AQ group performing worse for letter but not category fluency.

10.3.3.2 'IRI"

A t-test was conducted to compare the high and low AQ groups scores for the 'IRI'. Using a strict significance level of $p = .0125$ (p = .05/4) in view of the four empathy measures, the high AQ group gave significantly lower ratings on one of the subscales, 'Perspective-taking' $t(37) = 2.93$, $p = .006$, and significantly high ratings on another, 'Personal Distress' $t(37) = 4.15$, $p < .0001$. However, the groups did not differ on the other two subscales: Empathic Concern', $t(37) = 2.45$, $p = .019$, or 'Fantasy' $t(37) = 1.58$, $p = .122$.

10.3.3.3. 'MIT'

T-tests were conducted to compare the high and low AQ groups' performance on the 'MIT', using a strict significance level of .025 (p = .05/2). The groups significantly differed for 'MIT Sarcasm' score, $t(37) = 2.57$, $p = .014$, but not for the 'MIT Action' score $t(37) = 0.61$, $p = .547$, with the high AQ group performing worse for items involving sarcasm but not action interpretation.

10.3.4 Correlations between performance on the 'Counterfactual Judgments' task and additional measures

Within-group Pearson correlations were conducted for additional measures that significantly differentiated the groups with performance on the 'Counterfactual Judgments' task. A strict significance level was adopted to control for multiple comparisons (p = .01).

For the executive measures, neither the 'Colour-Word Interference' test (inhibition) or the 'Letter Fluency' test that revealed group differences were found to significantly correlate with performance on any measures of the 'Counterfactual Judgments' task for the high or low AQ groups.

For the 'IRI', neither the 'Perspective-taking', or 'Personal Distress' subscales that previously revealed group differences were found to significantly correlate with performance on any measures of the 'Counterfactual Judgments' task for the high or low AQ groups.

For the 'MIT', the 'MIT Sarcasm' score, which previously revealed group differences was not found to significantly correlate with performance on the 'Counterfactual Judgments' task for either the high or low AQ groups.

Table 8: Mean scores and standard deviations for additional measures

Measure (range)	**Low AQ group** (N = 19) M (SD)	**High AQ group** (N = 20) M (SD)	**Significance**	**Effect Size** (d)
DKEFS				
Colour Word Interference Scaled Scores (1-19)				
Colour Naming	10.89 (1.85)	10.50 (2.20)	.553	0.19
Word Reading	11.63 (1.71)	11.65 (1.63)	.973	0.01
Inhibition	12.68 (1.60)	11.15 (1.72)	.007**	0.92
Verbal Fluency Scaled Scores (1-19)				
Letter Fluency	13.21 (2.90)	10.70 (3.01)	.012*	0.85
Category Fluency	11.89 (3.39)	10.60 (3.42)	.244	0.38
IRI Scores				
Perspective-taking (0-32)	19.31 (3.88)	15.20 (4.82)	.006**	0.94
Fantasy (0-28)	17.84 (4.40)	15.50 (4.82)	.122	0.51
Empathic concern (0-28)	18.68 (2.71)	15.60 (4.79)	.019	0.79
Personal distress (0-24)	8.63 (3.64)	13.70 (3.96)	<.0001**	1.33
MIT Scores				
Sarcasm (0-20)	8.95 (0.91)	7.85 (1.63)	.014*	0.83
Action (0-20)	8.73 (1.31)	8.50 (1.31)	.547	0.18

Significant at p = .025* * Significant at p = .0125*

10.4 DISCUSSION

10.4.1 Summary of findings from the 'Counterfactual Judgments' task

The 'Counterfactual Judgments' task described situations in which participants experienced negative outcomes as a result of a main character's actions. Contrary to predictions, there were no group differences when participants were asked how much they thought the character would regret their actions. However, the high AQ group gave higher ratings of blame for the extent to which the main character was responsible for the negative outcome.

Participants were then asked to rank-order practical, emotional and punitive counterfactual alternatives that targeted either the main character or the participant, and to report how guilty they would feel in relation to thinking about each of the counterfactual alternatives. Participants across groups ranked the counterfactual alternatives from best to worst in the following order: practical, then emotional, and then punitive. Participants across groups also preferred counterfactuals where the character rather than the participant was the target. However, there was an interaction between the counterfactual alternative types and the target of the thought: when considering punitive alternatives, participants preferred these types of thoughts to be directed towards themselves rather than towards others. With respect to group differences, whilst there was no main effect of group, or group by alternative or target interactions, a three-way interaction did emerge. Difference scores showed that this primarily reflected a tendency for the high AQ group to differentiate less between practical and punitive counterfactual alternatives when the other character was the target.

This pattern was also reflected in the ratings of guilt, with participants across groups rating practical alternatives as evoking the least guilt and punitive alternatives as evoking the most. They also rated counterfactual thoughts about the character as evoking less guilt than thoughts about themselves. There was again an interaction between the type of counterfactual alternative and the target of the thought, whereby participants experienced more guilt when considering punitive alternatives that were directed at themselves rather than others. With respect to group, there were no differences for the ratings of guilt, with the high AQ group differentiating between the counterfactuals in terms of the alternatives' content type and target similarly to that of the low AQ group.

10.4.2 Performance on the 'Counterfactual Judgments' task

10.4.2.1 Judgments of regret and blame

Contrary to predictions, no group differences were identified on the ratings of regret. Although this conflicts with previous literature, it in fact replicates the finding of intact ratings of regret for self and other displayed by the high AQ group in Chapter 9 (see 9.3.2.2). As discussed, regret is a complex, self-conscious emotion that people with ASD have been found to display difficulty with, despite showing little difficulty with the processing of basic emotions like happiness or sadness (Capps et al., 1992). This is thought to reflect the need to engage with another's perspective and expectations when experiencing emotions like regret or shame, which is likely to draw upon cognitive empathy, but not other more basic emotions. However, as discussed previous work has tended to use tasks that focus on the recognition of emotions (see Chapter 9.4.3.2), and very little work to date has examined the experience of self-conscious emotions like regret. The present task asked people to estimate how much another character would regret their 'bad' actions. If we assume that people with ASD, and thus perhaps those with high levels of autistic traits, have difficulties with experiencing emotions, then it would be expected that they would also experience difficulty in estimating the experience of these emotional states for others. However, if any difficulty is confined to the ability to recognise the emotions of others, (at least on abstract tasks requiring the participant to gauge another's emotional state on the basis of a picture of a facial expression with no contextual information provided) then we might expect the ability to experience such emotions in response to particular events, and to estimate these in others, to be unimpaired. On both the 'RCFT' and the 'Counterfactual Judgments' tasks, people were provided with clear signals that the agent (i.e. character for 'Counterfactual Judgments' and character/participant for 'RCFT') should regret their 'bad' actions to some extent. Thus, it was made relatively explicit what the agent (character or participant) was feeling, removing the need for the participant to put themselves in another's shoes, or to imagine themselves in the scenarios. Future research should clarify the extent to which the personal experience of such emotions and/or the ability to predict others' corresponding emotional states is preserved in those with ASD, and how this might vary in line with different sources of information. It would also be important to investigate the extent to which the experience and interpretation of emotional states is dissociable from the ability to recognise such emotions in others, and the extent to which these abilities are impaired.

In contrast, the groups were found to differ on their ratings of blame, with the high AQ group blaming the characters more for negative consequences in comparison to the low AQ group. This partly supports the findings of the 'RCFT' task presented in Chapter 9, where the high AQ group was found to be less sensitive to whether the agent's actions resulted in negative consequences for themselves or for others, but to be more sensitive to the agent's identity, blaming themselves more than others. It is unclear

why that the high AQ group seemed to blame themselves versus others more on the 'RCFT' task, and to blame characters more on the present task. It is difficult to compare and interpret these findings since the 'RCFT' task involved two sets of items, where for half the character was the agent and for the other half the participant was the agent, whereas the present task only involved scenarios where the character was the agent and the participant the victim. Previous studies have also found mixed results regarding the nature of responsibility and culpability attributions in those with ASD, which can be best explained by considering the clarity of task factors, and the varying demands of the scenarios used (please see Chapter 4.2.3 for a review of these studies and a discussion). Future work should clarify how such judgments might be influenced by both cognitive and emotional factors, and thus correspondingly (un)impaired in people with ASD. Nonetheless, whilst a complex set of findings emerges when considering the two tasks used in the present thesis, it appears that people with high versus low AQ traits seem to blame agents more for their actions, although this might depend on various factors which as yet remain unclear.

10.4.2.2 Counterfactual preferences and guilt

10.4.2.2.1 The type of counterfactual thought

Participants across groups ranked practical alternatives as their most preferred counterfactual option, followed by emotional and then punitive alternatives. This suggests that they were more focused on solving the problem presented rather than blaming others, or making other kinds of judgments surrounding responsibility or culpability which emotional or punitive counterfactual thoughts might mediate. However, when interpreting this pattern of findings, there are a number of additional factors that might have varied across the alternatives in addition to content, and/or influenced participants' behaviour on the 'Counterfactual Judgments' task; these will be considered in turn below.

Firstly, it could be argued that the practical alternatives (e.g. been driven by someone else to my graduation ceremony, or they had invited their friends over another night) were more concrete in nature than emotional or punitive counterfactual thoughts. In comparison, emotional (e.g. my sibling had been more trustworthy, or my friends had been more considerate) and punitive counterfactual alternatives (e.g. my sibling had been unwell and unable to drive me to my graduation, or their friends did not feel welcome in your flat), might require more imagination and thus be more taxing for participants to consider. The emotional and punitive counterfactuals could be said to be less clearly linked to the specific event and to require a more sophisticated form for mental mutation, considering prior relationships and events or a wider set of factors than practical counterfactuals that simply sought to 'undo' or prevent mistakes.

Secondly, although the scenarios presented in the 'Counterfactual Judgments' task described situations that were annoying or distressing for participants, in reality the scenarios had relatively trivial or

short-term consequences (e.g. miss you graduation ceremony, or feel upset that you friend has behaved insensitively). The choice of everyday 'mistakes' was deliberate, since the thesis is concerned with examining real-life-type social problem solving. However, it is possible that more severe events might have triggered a different quality or wider range of counterfactuals, including emotional and punitive thoughts. For instance, getting cancer or being involved in a serious car crash might be more emotive and lead to a greater tendency to reflect on emotional or punitive counterfactuals, in addition to or instead of practical counterfactuals.

Finally, the severity of the event might interact with how precise the locus of its cause was. For instance, being diagnosed with cancer is a very severe event, but there is no one single clear cause or person to blame for this event. On the other hand, being involved in a serious car crash with another vehicle is also a life-threatening situation, where there is a much clearer locus of cause, with either or both the drivers likely to blame. When the locus of an event is clearer, more practical counterfactual thoughts might be generated, regardless of how severe or distressing it is. For instance, if you were involved in a serious car crash you might consider how you or others could have avoided the crash, perhaps by taking a different route or driving more carefully. This situation is likely to prompt a big focus on specific facts of the event that could have been altered, but is less likely to trigger a consideration of wider factors or of longer-term actions or changes you could have made. By contrast, if you are diagnosed with cancer, you might consider how you could have avoided getting cancer, by changing your diet or lifestyle, and/or blaming it on your genetics; such events with less clear loci are more likely to lead to a wider consideration of factors, and/or a tendency to regress further back in the causal chain. The present task contained both less severe scenarios and scenarios where the locus of the event was clear-cut, thus either or both of these factors might have prompted a greater consideration of practical versus emotional or punitive counterfactuals.

The findings of the present task highlight the importance of examining the nature of the content of counterfactuals. There is thought to be a close inter-relation between causality and counterfactuals, with factors like whether there was a clear agent of the consequences, and how important the agent's role was in influencing judgments of responsibility (Lagnado et al., 2013). Future work should seek to examine how such factors might affect the nature of counterfactual thoughts and the purpose that they serve (i.e. practical and solution-focused compared to more punitive and blameworthy). With respect to the mechanistic aspects of counterfactual thoughts, the 'RCFT' and 'Counterfactual Judgments' tasks examined one subset only (please see Chapter 9.4.4 for a discussion of this issue). It would also be interesting to explore how the content of counterfactuals might vary in relation to different types of counterfactual mechanics i.e. upward/ downwards (how things could have been better – to prevent; how things could have been worse – to console) and additive/subtractive (had/ had not).

10.4.2.2.2 The target of the counterfactual thought

Participants across groups preferred counterfactuals that were about the main character rather than themselves. Given that the main character was always the agent and the participant the victim on the 'Counterfactual Judgments' task, interpreting this seems straightforward: since the character was directly responsible for the consequences participants experienced, they were more likely to prefer counterfactuals which emphasised the characters' role rather than their own. However, there was an interaction between the type of counterfactual and its target; although participants preferred practical and emotional counterfactuals that were about others rather than themselves, when it came to rank-ordering punitive counterfactuals they preferred thoughts where they rather than the character were the target. Participants might have felt that punitive thoughts were unfair, inappropriate or disproportionate, and thus they not only ranked these as their least preferred type, but also preferred to inflict such thoughts on themselves rather than others. This suggests there might be a threshold for the extent to which people wish to engage in counterfactual thinking about others' mistakes, at least once it becomes blameworthy and punitive in nature. Whilst counterfactual thinking is thought to be a crucial aspect of learning from our own and others' mistakes (e.g. Mandel & Dhami, 2005; Epstude & Roese, 2008), excessive counterfactual thinking has also been linked to adverse effects, such as rumination, which is a cognitive hallmark of depression (e.g. Branscombe et al., 2003). The findings from the present study suggest that healthy people might have a mechanism that prevents them from thinking too much about aspects of situations that cannot be easily fixed, and a protective bias towards thinking about pro-active strategies that could be used to improve future outcomes. As highlighted above, given the specific nature of the scenarios examined, it is difficult to generalise these findings to other situations. The target of counterfactual thoughts has also surprisingly been neglected in the existing literature, and thus future work should seek to build upon these tentative findings addressing both the content and the target of counterfactual thinking in relation to a number of factors.

10.4.2.2.3 Ratings of counterfactual guilt

The pattern of guilt ratings broadly reflects the rank-ordering of counterfactuals by preference. This suggests that across groups, participants' ratings of guilt broadly corresponded to their preferences; they were more likely to prefer practical counterfactuals (e.g. been driven by someone else to my graduation ceremony, or my housemate had invited their friends over another night) that made them feel less guilty, compared to emotional counterfactuals (e.g. my sibling had been more trustworthy, or my housemate had been more considerate) or punitive counterfactuals (e.g. my sibling had been unwell and unable to drive me to my graduation, or your housemate's friends did not feel welcome in the flat). There was again an interaction between the type of counterfactual thought and its target, whereby participants felt

less guilty about practical and emotional counterfactuals that were about others rather than themselves, but more guilty about punitive counterfactuals which were about others rather than themselves. This again likely reflects the fact that punitive counterfactuals might have been regarded as disproportionate, and thus participants felt particularly guilty about experiencing these for other people rather than themselves.

10.4.2.3 Counterfactual preferences and guilt in people with high versus low autistic traits

The two groups performed fairly similarly in their counterfactual rankings, except that the high AQ group demonstrated a greater preference for punitive and a lower preference for practical alternatives where the character was the target. This combined with the finding of higher ratings of blame suggests that the high AQ group was more resentful towards the main character, and thus favoured a more punitive approach towards the characters. The high AQ group may have been less forgiving of the characters' mistakes and thus demonstrated a greater preference for counterfactual thoughts that emphasised the stupidity of the character, at the expense of those which sought to avoid or improve outcomes. The high AQ group might also have experienced the consequences more intensely, resulting in a greater desire to blame and punish characters for their mistakes. ASD is associated with greater rigidity and less flexible behaviour (e.g. D'Cruz et al., 2013), which might result in experiencing greater anxiety in social situations, especially when things do not go to plan (Richer, 1976). Whilst it is important to note that this study focused on a sub-clinical rather than a clinical sample, people along the autistic spectrum are also thought to experience these difficulties, although perhaps to a lesser degree than a clinical population. Thus, the high AQ group might have found interacting with other people and accepting the 'give and take' of social relationships more stressful.

Alternatively, the high AQ group might have simply been less sensitive to the fact that punitive counterfactuals about the characters might have been a disproportionate or an inappropriate response to the situation. It is possible that both groups in fact experienced punitive thoughts, but that the low AQ group was more likely to inhibit these or differentiate between the counterfactual types more strongly, because they recognised that expressing such punitive thoughts about the characters might have been deemed as overly harsh. In contrast, the high AQ group may have failed to acknowledge this and simply expressed themselves honestly, which is consistent with work suggesting that people with ASD are less concerned with protecting their social reputations than neurotypical individuals (e.g. Izuma et al., 2011).

Overall, participants with high versus low AQ traits appear to differ in the nature of their counterfactual thinking as well as in their fluency, although this might be subtle. This pattern of differences in both quantity and quality of counterfactual thoughts might be reflected in difficulties with social problem-solving, as has been demonstrated in previous work presented in this thesis, such as the 'Above and Beyond' or 'Social Expectations' tasks described in Chapter 5 and 6. People across the autistic

spectrum are thought to find navigating social situations and/or building relationships more challenging, and it is possible that altered ability to produce counterfactuals, both from the point of view of quantity and quality, might in part explain this. Engaging in counterfactual thinking is thought to be a functional process that allows people to reflect upon and solve problems, and to evoke appropriate future responses (Epstude & Roese, 2008). A lack of engagement with counterfactual thinking, or a more limited appreciation of the different aspects of situations, might hamper the capacity to improve and learn from our own and other's mistakes, thus leading to less skilled social performance.

In contrast, the high AQ group did not differ from the low AQ group in their ratings of guilt for experiencing counterfactuals thoughts. This is contrary to predictions, since it was expected that the high AQ group might have either experienced less guilt overall or have differentiated between the types of counterfactuals in a different manner in comparison to the low AQ group. Guilt, like regret, is a complex and self-conscious emotion, and whilst this has not been widely studied in people with ASD, it is again likely that this group struggle with emotions such as these that require taking into consideration others' experiences. As discussed, it is difficult to establish to what extent the findings of preserved regret or guilt on the present task conflict with previous work; the discrepancy between these studies might simply reflect a preserved experience of such emotions, but difficulty in identifying them on the basis of facial expressions.

10.4.3 Limitations of the 'Counterfactual Judgments' task

The 'Counterfactual Judgments' task provides a novel approach to assessing the content of counterfactual thinking. However, there are a number of possible limitations that could be addressed to improve the task. Firstly, participants were required to rank the alternatives in terms of 'preference'. On reflection, it is not clear what 'preference' refers to; thoughts that you would experience, or thoughts that you think you ought to experience? Thus, different participants might have interpreted this term differently. This issue could be addressed by instead asking participants instead how likely they would be to experience each of the counterfactual thoughts.

Secondly, it was considered important to examine the quality as well as the quantity of counterfactual thoughts, in order to complement and extend the findings of the 'RCFT' task used in Chapter 9. Thus in the present study participants were provided with a range of counterfactual thoughts that corresponded to particular types of content categories and asked to make judgments about these. This methodology was chosen as a simple means by which to assess systematically the content of counterfactual thinking, which would have been rather more difficult to achieve by using a more naturalistic approach, such as asking participants to generate their own counterfactual thoughts in response to a scenario. These categories were developed by inspecting the free responses to the 'RCFT' task and using these as an

informal guide to the different types of counterfactual thoughts that might occur in such situations. Nonetheless, this approach might have resulted in providing participants with thoughts that might not have occurred to them spontaneously, and/or it might have failed to capture the true range of thoughts that participants would experience in each scenario. One way to improve the validity of this task might be to ask participants to generate their own counterfactuals in response to different types of content. Participants could be shown various types of counterfactual thoughts that they could experience in a given situation on an example item according to the different categories, practical, emotional and punitive. Then for experimental items they would be asked to generate one thought per category. This would, however, result in a potentially complex and unwieldy dataset to code and analyse. Another 'quick fix' approach to improve the ecological validity of the task would be more rigorous piloting, asking people whether the alternatives provided capture the range of thoughts that they might experience in each scenario, and how they could be more realistic.

Finally, the 'Counterfactual Judgments' task could have been improved to make it more sensitive to possible group differences. The high AQ group displayed relatively subtle differences in their preferences of alternatives. Previous tasks used both in this thesis (see the 'Social Expectations' task used in Chapter 7) and in the wider literature have highlighted the importance of considering both the ability to make judgments about alternatives and to generate one's own response when assessing social performance, with generation measures proving to be more sensitive to group differences. The need to generate counterfactuals was negated on this task, which might have artificially enhanced the high AQ group's performance. It is possible that if participants were required to generate counterfactuals as opposed to judging them, more marked group differences might have emerged. Similarly, the emotional judgments of regret and guilt failed to differentiate the groups, and it is possible that a rating measure might have been too simplistic to reveal any possible deficits that the high AQ group might have experienced. As highlighted in Chapter 9 (see 9.4.4), using a deeper probe such as asking participants to describe their own emotional reactions and/or explain those of the characters might have proved more sensitive. Whilst either of these alterations would improve the task's sensitivity, on the downside they might result in more complex or unwieldy datasets to code and analyse, as discussed above.

10.4.4 Summary of findings from executive and empathy measures

As predicted, the 'Colour-Word Interference' test of inhibition revealed some evidence of executive difficulties in the high AQ group relative to the low AQ group. Moreover, there was some evidence of executive dysfunction on a different executive test, 'Verbal Fluency', where the high AQ group performed worse for 'Letter Fluency' but not 'Category Fluency'. On the 'IRI' questionnaire the high AQ group scored significantly lower for 'Perspective-taking', and significantly higher for 'Personal Distress' then the

low AQ group. For the MIT, the high AQ participants were found to score significantly lower on items involving interpreting sarcastic items, but not those involving action interpretation.

Within-group Pearson correlations were also conducted between performance on the 'Counterfactual Judgments' task and the executive and empathy measures that significantly discriminated the groups, namely the 'Colour-Word' interference test of inhibition, the 'Letter Fluency' measure, the 'MIT Sarcasm' measure, and two of the 'IRI' subscales: 'Perspective-taking' and 'Personal Distress'. No significant associations with performance on the 'Counterfactual Judgments' task were revealed for any of these measures for either the high or low AQ groups.

10.4.5 Contribution of executive functioning and empathic processes

Additional measures of executive functioning and empathic processes were also included. The possible contribution of each these abilities to performance on the 'Counterfactual Judgments' task and evidence for group differences will be discussed in turn. First considering executive functioning, two tests were included in the present study; one of inhibition and one of verbal fluency. These revealed a pattern of poorer inhibition in the high AQ group that is fairly similar to that of the sample compared on these measures in Chapter 7 (see section 7.3.3.1), although differences were found on the verbal fluency test which were not identified in the previous sample. Moreover, the high AQ group performed worse on the letter fluency but not the category fluency measure in the present study; why might this be the case? Letter fluency requires participants to search through their lexical or phonemic memory, whereas category fluency is thought to be more reliant upon semantic knowledge (Monsch et al., 1992; 1994). In addition, there is evidence to suggest that these two abilities rely on partially different brain networks, with studies showing that frontal lobe damage results in a disproportionate impairment to letter fluency, whereas temporal lobe damage impairs category fluency to a greater extent (e.g. Hodges et al., 1999). The pattern of findings is therefore consistent with the notion that ASD is predominantly associated with frontal rather than temporal lobe impairments (please see Chapter 2.1.2).

Thus, it is possible that poorer executive functioning might have influenced both the range of counterfactuals that the high AQ group considered and their preferences for different types of counterfactual content. As discussed, practical versus emotional or punitive counterfactuals might be said to be less executively demanding since they a) tend to focus on the more immediate aspects of the scenario and b) are more concrete in nature, requiring less imagination. However, it is not entirely clear how this account could explain the choice of more punitive and less practical counterfactuals by the high AQ group, or why this difference should occur selectively when considering alternatives where the character rather than the participant was the target.

Now turning our attention cognitive empathy, this was assessed via the 'MIT' and two subscales from the 'IRI', namely 'Perspective-taking' and 'Fantasy'. The 'MIT' again revealed group differences, although in this study the high AQ group performed worse on the items involving interpreting sarcasm but not actions, which is the converse pattern to that found in previous sample described in Chapter 7 (see 7.3.3.3). Given that the samples were recruited in a very similar way and drawn from the same university population, it is not clear why they should differ in this way (see Chapter 8.4.3.1 for a detailed discussion of these issues). However, as previously discussed (see Chapter 7.4.3) whilst sarcasm might be said to be a more complex form of non-literal interpretation, it might also be more familiar than interpreting people's actions, at least when asked to explain these explicitly. It is possible this might have affected the two samples differentially, based on their prior experiences with such situations.

With respect to the 'IRI' subscales, as expected the high AQ group had lower scores for 'Perspective-taking' scale, which taps processes that are predominantly driven by cognitive empathy rather than emotional empathy. However, the groups did not differ for the other cognitive empathic scale, 'Fantasy', which is in contrast to a previous study using these measures in people with high versus low AQ traits which found differences on the 'Fantasy' but not 'Perspective-taking' subscales (reported in Chapter 7: see 7.3.3.2). Once again, it is not clear why the samples should differ in this way, and it is unclear if this is a meaningful finding. Rather, it is likely that this discrepancy reflects the unreliability of using such measures in small samples; both the 'IRI' subscales and the 'MIT' comprised only of a handful of items each, and thus these measures might be very sensitive to any small sources of variability.

Cognitive empathy is thought to be impaired in those with ASD and also in those with high AQ traits, at least to some extent (Blair, 2008; Freeth et al., 2008; Lockwood et al., 2013). Although somewhat inconsistent, the current findings do point to weaker cognitive empathic skills in the high versus low AQ groups. This might have manifested in their performance on the 'Counterfactual Judgment' task, resulting in less consideration for the main character's perspective, and this might potentially underlie their more punitive and blameworthy approach. It would also be expected that poorer cognitive empathic skills might result in a reduced experience of guilt for punitive counterfactual thinking and/or a poorer ability to estimate the characters' experience of regret for their action; however, this was not found to be the case. As discussed (see Chapter 9.4.3.2), it is unclear to what extent the experience of these emotions is intact in those with ASD, and future work should seek to elucidate this.

Turning now to emotional empathy, this was assessed via the remaining two subscales of the 'IRI': 'Empathic Concern' and 'Personal Distress'. The groups were not found to differ on the 'Empathic Concern' subscale, but they were found to differ on the 'Personal Distress' subscale with the high AQ reporting more personal distress. Whilst the study described in Chapter 7 also found differences between the high and low AQ groups on these predominantly emotional subscales, this was a contrasting pattern to

that identified here; the high AQ group scored lower for 'Empathic Concern', but no differences were found on the 'Personal Distress' subscale. Once again this discrepancy is somewhat puzzling, and can probably be best accounted for by the argument presented above regarding comparing small sample sizes on only a number of items.

Furthermore, within-group correlations between the executive and empathic tests and performance on the 'Counterfactual Judgments' task did not reveal any significant associations. It should, however, be noted that correlations in small samples are very unstable (Shilling et al., 2002) and thus it is difficult to determine the importance of this finding.

10.4.6 Conclusion

This study examined the content and target of counterfactual thinking by asking participants with high versus low autistic traits to judge different types of counterfactual alternatives. Overall, the main study hypothesis was confirmed: those with high AQ traits showed a greater preference for more punitive counterfactuals when the character was the target of these thoughts. In line with this more punitive approach, the high AQ group also blamed the character more for their 'bad' actions. However, contrary to predictions, the groups did not differ on ratings of the character's regret for their actions, nor for ratings of own their own experience of counterfactual guilt. Whilst impaired ability to take others' perspectives has long been established in the autistic spectrum, far less is known about the about the ability to use counterfactual thinking to inform judgments about oneself or others and to evoke appropriate emotional responses. This study extends the findings of Chapter 9 and provides further evidence to suggest that those on the autistic spectrum deploy counterfactual abilities atypically (Nicolle, Ropar & Beck, 2014). It also raises the possibility that the experience of complex emotions and the ability to predict others' emotional states might be intact if sufficient scaffolding is provided by the context.

The previous Chapters in part B of this thesis have examined how groups high and low in autistic traits reason about positive or negative outcomes and judge themselves and others in relation to different situational factors. Chapter 8 concentrated on the role of 'deservingness', and Chapters 9 and 10 focused on emotional and blame attributions and counterfactual sanctions. The next and last experimental Chapter will focus more directly on the influence of autistic traits on moral reasoning. Chapter 11 will adapt and extend a classic moral reasoning task, examining utilitarian decision-making in situations with different levels of personal involvement, and with both social and physical harms.

Chapter 11: Utilitarian decision-making for physical and social harm

11.1 INTRODUCTION

Utilitarian decisions are moral judgments that seek to maximise the benefit 'for the greater good' and/or to reduce the impact of any possible costs (Rosen, 2003). These typically involve a conflict between the 'normative' principles of moral behaviour, such as 'you should not hurt someone', and achieving the best outcome, since it might be necessary to hurt an individual in order to save a greater number of people. Utilitarianism has long been of interest to philosophers (e.g. Mill, 1863) and has more recently come to the attention of psychologists and neuroscientists who are interested in exploring how and under what circumstances people make utilitarian decisions (e.g. Greene, Morelli, Lowenbreg, Nystrom & Cohen, 2008).

The most famous paradigm used to examine utilitarian decision-making is the classic 'Trolley Problem' (Foot, 1967). Scenarios are presented in which a train carriage is hurtling towards five railway workers and will kill them. The participant, acting as a bystander in the scenario, is then given the option of flicking a switch, which will divert the train onto another track where there is only one railway worker who will die. The utilitarian decision is to flick the switch and divert the train (resulting in just one rather than five deaths); since the bystander has to decide whether to interfere with the 'natural course of events' or not, this might be thought to be an uncomfortable or unpopular choice. However, most people do choose to divert the train on the basis that they will save more lives overall, and correspondingly judge this action to be morally acceptable (e.g. Petrinovich, O'Neill & Jorgensen, 1993). This is associated with activation of brain regions linked to general reasoning and problem solving, such as the dorsolateral prefrontal and inferior parietal cortices (Greene et al., 2001, 2004), suggesting that a deliberate reasoning process occurs when making such decisions.

An amendment to the scenario, however, results in a very different pattern of findings. In a second scenario, known as the 'Footbridge Problem' (Thomson, 1976) participants are again presented with the same problem, (a train is hurtling towards five people and will kill them) and with an opportunity to intervene, but this time rather than simply flicking a switch they must decide whether to push someone onto the tracks in order to stop the train. Here, a utilitarian course of action results in the same outcome (one will be sacrificed in order to save five), but it requires greater personal involvement on the participant's part, causing direct physical harm to another person. In this version of the dilemma, participants are much more likely to choose **not** to intervene and to judge this to be the most ethical course

of action. In contrast to a utilitarian decision, choosing not to act is known as a 'deontological' decision that adheres to 'normative' moral rules, regardless of the utility or maximal benefit. This more 'personal' version of the dilemma is also associated with greater activation of bran regions implicated in the experience of emotion and social cognition, including the medial prefrontal cortex and the posterior cingulate gyrus (Greene et al., 2001; 2004), suggesting that this dilemma invokes a more emotionally-charged and thus possibly a less rational response. Research conducted with a range of clinical populations characterised by emotional disturbances and problems with social cognition tasks has revealed increased rates of utilitarian judgments in response to moral dilemmas, including in patients with frontal lobe damage (Anderson et al., 2006; Eslinger et al., 1991; Koenigs et al., 2007), frontotemporal dementia (Gleichgerrcht et al., 2010; Mendez et al., 2005; 2006) and people with psychopathy (Bartels & Pizarro, 2011; Blair, 1995) or alexithymia (Patil & Silani, 2014).

As discussed (see Chapter 4.2.2), surprisingly little work has examined moral reasoning in people with ASD, despite the associated difficulties with cognitive empathy and perhaps emotion processing. Only one study has to date examined the classic utilitarian trolley and footbridge dilemmas in people with ASD (Gleichgerrcht et al., 2013). High-functioning participants with ASD did not differ from controls in their utilitarian choices when the degree of personal involvement was low (e.g. impersonal: pull the lever, saving five and killing one), but they more frequently choose the utilitarian course of action when the personal involvement was high (e.g. personal: to push the man, saving five and killing one). This pattern of enhanced utilitarian action in participants with ASD was associated with a decreased ability to infer other people's thoughts and to understand their intentions on a test of cognitive empathy. It was also associated with an intact understanding of the 'appropriateness' of utilitarian actions, and similar scores on a questionnaire of moral knowledge. This suggests that people with ASD might intrinsically known what is the 'right or wrong' thing to do, which corresponds to a range of work showing intact understanding of moral versus conventional transgressions. However, on the trolley dilemmas they tend to apply a more utilitarian and arguably rational approach, which might possibly be underpinned by their reduced ability to empathise with the different characters' perspectives involved.

Such moral dilemmas have become commonly-used tools for assessing and demonstrating psychological and philosophical theories of moral judgment and reasoning. However, their use and interpretation has been criticised extensively on various bases (see e.g. Waldmann & Wiegmann, 2010). One common objection to this body of work is the tendency to use just one or two scenarios, namely the 'Trolley Problem' and the 'Footbridge Problem' (Patil & Silani, 2014), which might potentially reduce the reliability and generalisability of the findings. Greene et al., (2001; 2004) attempted to extend these findings by developing a novel battery of scenarios that are similar in structure to the 'Trolley Problem' and

'Footbridge Problems', but which include a number of small variations. For instance, there are dilemmas in which one person is killed to save many, including the person who commits the moral transgression (as opposed to the scenarios used in the classic task in which the person making the utilitarian decision is merely a bystander), and scenarios where the decision to commit a moral transgression is for one's own selfish benefit rather than for the greater good (Gleichgerrcht et al., 2013). Unfortunately the scenarios were not well balanced and it is therefore difficult to disentangle the influence of different factors. Perhaps most importantly, both the classic trolley/footbridge dilemmas and the battery of scenarios developed by Greene et al., (2001; 2004) described situations resulting mostly in physical injury or death. Although Greene et al. included some incidents of non-physical harm, these were equally extreme scenarios, such as whether a father should sell pictures of his daughter for child pornography in order to feed his family. These types of dilemmas are hypothetical in nature, and are unrepresentative of the types of real-life dilemmas that people face, which might again undermine their potential generalisablity. Finally, these studies tended to involve participants making forced-choice decisions or using Likert-style rating scales to make various judgments. In addition, asking participants to verbalise the justifications that underlie their moral judgments can provide very useful information to better comprehend both typical moral psychology and the source of any differences in clinical populations (Gleichgerrcht et al., 2013).

In order to systematically investigate the influence of autistic traits on utilitarian decision-making, and to address the limitations of the previous stimuli outlined above, a novel task was developed: the 'Utilitarian Judgments' task. This included an adapted version of the trolley/ footbridge problem and seven other novel dilemmas, and included equal numbers of scenarios describing cases of physical or social harm. Each scenario also had two variants, one where the agent's personal involvement was low (impersonal), and one where it was higher (personal). The scenarios were tightly matched for all other factors, including that all decisions involved weighing up the needs of one individual versus that of a group of individuals. Participants always played the role of an impartial agent and had to decide between a utilitarian and non-utilitarian course of action, and to rate how uncomfortable each course of action would make them feel. Participants were also required to provide justifications for why following either a utilitarian or a non-utilitarian course of action might be considered the right thing to do; their responses were classified according to the extent to which they considered the perspectives of the characters affected by the outcome and the role of the agent making the decision. In order to further examine the relationship between autistic traits and moral behaviour, two additional questionnaire measures were also included: the 'Ethics Position Questionnaire' (Forsyth, 1980) and the 'Moral Behaviour Inventory' (Mendez et al., 2005).

11.1.1 Hypotheses

11.1.2 The 'Utilitarian Judgments' task

11.1.1.1 Utilitarian/ non-utilitarian choices

It was expected that participants across groups would display more utilitarian decision-making in situations involving low versus high personal involvement i.e. impersonal versus personal. It was also expected that physical versus social harm might elicit different considerations; moral reasoning in relation to physical harm might be more 'black and white', whereas reasoning about social harm might be more subtle and nuanced. Thus, different patterns of utilitarian decision-making might also emerge with respect to the type of harm, with more utilitarian choices being selected in physical versus social scenarios.

In line with the findings of the study by Gleichgerrcht et al., (2013) it was predicted that the high AQ group would make more utilitarian decisions than the low AQ group, at least in personal scenarios. It was also expected that group differences might interact with the type of harm, with the high AQ group showing less sensitivity to the distinction between physical versus social harm.

11.1.1.2 Discomfort ratings

It was expected that participants across groups would feel more uncomfortable following utilitarian versus non-utilitarian courses of action, particularly in personal scenarios. It was also expected that physical scenarios might result in participants feeling more uncomfortable, since this represented a more extreme form of harm. With respect to group, given the putatively poorer empathic skills of the high AQ group, it was expected that they would feel less uncomfortable than the low AQ group, at least for personal scenarios and in cases of social harm.

11.1.1.3 Verbal responses

It was expected that utilitarian courses of action would invoke a greater consideration of both the agent's and the various characters' perspectives than non-utilitarian courses of action, and that this might be exacerbated for personal and/or physical scenarios. It was also expected that the high AQ group would not take the characters' perspectives into consideration as much as the low AQ group, and perhaps would also consider less the agent's perspective, and that this might interact with the factors of personal involvement and type of harm.

11.1.3 Moral questionnaires

The 'Ethics Position Questionnaire' is a self-report questionnaire that has two sub-scales: one corresponding to 'Relativism', the idea that moral rules are dependent upon contextual factors, and another to 'Idealism', the idea of absolute moral rules. In view of their putatively rigid understanding and/or use of social rules, it was expected that the high AQ group might score more highly for 'Idealism' and thus lower for 'Relativism'.

The 'Moral Behaviour Inventory' is a self-report questionnaire that asks people to rate the acceptability of various transgressions. In view of work finding intact understanding of moral and conventional transgressions in children with ASD, and the findings of Gleichgerrcht et al., (2013), it was expected that the groups would not differ on this measure.

11.2 METHODS

11.2.1 Screening phase

11.2.1.1. Screening participants and procedure

The participants for this study also took part in the 'Counterfactual Judgments' task and are described in Chapter 10.2.1.1. Thus, there was an opportunistic sample of 828 full-time university students (58.60% female) who were fluent in English and aged 18 or over (mean age 20 years old). All participants completed the AQ (Baron-Cohen et al., 2001) and total scores were calculated for the whole sample. Participants within the highest-scoring and lowest-scoring 10% of males and females were contacted and invited to take part in the second stage.

11.2.2. Experimental Phase

11.2.2.1 Design

There was one between-participants factor of AQ group (high vs. low scorers). There were two within-participant factors, the personal involvement of the agent in the scenario (low vs high: personal vs. impersonal) and the type of harm (social vs. physical).

11.2.2.2 Experimental participants and procedure

Please note that the participants included here are the same as those described in Chapter 10. Thus, there were 20 (10 male, 10 female) participants in the high AQ group and 20 (10 male, 10 female) participants in the low AQ group. All participants were tested individually, and provided written informed

consent before completing the 'Utilitarian Judgments' task (in preparation)[3], and completing two questionnaire measures; the 'Ethics Position Questionnaire' and the 'Moral Behaviour Inventory'. Participants also completed some further measures of executive functioning and empathy, which are described in Chapter 10.

11.2.3 The 'Utilitarian Judgments' task

This task was designed to investigate moral decision-making in situations where the needs of an individual were weighed up against the needs of a group. The task consisted of eight scenarios in which an 'agent' is required to make a decision that will either favour the best interests of one character at the expense of a group of characters, or vice versa. For four of the scenarios, the agent's dilemma related to situations involving physical harm, for instance injury or death either for one character or for a group of characters. Please note that one of the four physical harm scenarios was an adapted version of the classic 'trolley problem'; please see Figure 15. For the remaining four scenarios, the agent's dilemma related to incidents of social harm, for instance social exclusion, emotional distress or inconvenience either for one character or for a group of characters; please see Figure 16 for an example.

Each of the scenarios had two versions, which varied with respect to the personal involvement of the agent in the situation (high vs low); one version involved an impersonal dilemma (i.e. pulling the lever in the trolley problem), and the second version involved a personal dilemma (i.e. pushing someone in the footbridge problem). For both social and physical harm scenarios, participants had to 1) decide whether they would choose the non-utilitarian or the utilitarian course of action, 2) rate how uncomfortable they would feel with each course of action (non-utilitarian and utilitarian), and 3) provide a verbal response as to why each course of action might be the right thing for the agent to do.

There was no incentive for the participant to choose one course of action over the other, since the agent did not stand to receive any personal gain; nor would they incur any loss or sanction by making a particular decision. The task instructions made it clear that the agent in the scenarios would not be punished for their decision, even if such decisions would normally have led to legal consequences. For instance, pushing a man to his death would typically lead to punitive consequences, but here it would not. It was also made clear that the outcomes were absolute; either the individual or the group would experience harm, according to the agent's decision, since there was no available course of action that could prevent harm for both parties. The main characters were always known to the agent, and the nature of their relationship to the agent (i.e. siblings, friends or colleagues) was counterbalanced across items, as was their gender. The order of the scenarios was counterbalanced to create two different presentation versions;

[3] This task was developed jointly with Ms Karishma Vyas. Ms Vyas used this task to explore utilitarian judgments in people scoring high versus low on a self-report measure of psychopathic personality traits. This study is included in her doctoral thesis (Vyas, 2015).

half of the participants saw the physical scenarios first, and the other half saw the social scenarios first. This was counterbalanced within groups.

11.2.3.1 Administration

Participants first read the task instructions, and then they were shown an example item and allowed to ask clarification questions. All scenarios and corresponding test questions were then presented one at a time, in a paper booklet; participants responded verbally. The scenario remained on display until participants had completed the relevant questions in order to reduce the confounding possible effect of memory load.

11.2.3.2 Scoring

11.2.3.2.1 Scoring of utilitarian/ non-utilitarian choices

Participants were awarded a score of 0 if they chose the non-utilitarian course of action and a score of 1 if they chose the utilitarian course of action. These scores were then summed across all eight scenarios, with higher scores denoting more utilitarian choices. This created a total score for each scenario type (range 0-4), creating four scores; (1) physical scenario choice: impersonal, (2) physical scenario choice: personal, (3) social scenario choice: impersonal, (4) social scenario choice: personal.

Figure 15: Example physical scenario from the 'Utilitarian Judgments' task

SCENARIO: "Rachel is standing on a railway bridge and an empty train is quickly approaching. As the result of a vicious attack, five of her friends are tied to the tracks and are unable to move."

IMPERSONAL VARIANT:
On the bridge beside Rachel, there is a lever. Pulling this lever will cause the train to switch onto a different track. Another friend, Darren, is trapped on this track. If Rachel doesn't pull the lever, Darren will remain alive. However, the five friends on the other track will die. If Rachel pulls the lever, Darren will die. However, the five friends on the other track will be saved.

PERSONAL VARIANT:
On the bridge beside Rachel, there is another of Rachel's friends, Darren, who is very large. Pushing Darren would cause the train to stop. If Rachel doesn't push Darren, he won't be harmed. However, the five friends on the track will die. If Rachel pushes Darren, he will die. However, the five friends on the track will be saved.

Questions

1. **If you were Rachel, which of the following options would you choose?**
   - **a.** Do not pull the lever/Do not push Darren (*Non-utilitarian decision*)
   - **b.** Pull the lever/Push Darren (*Utilitarian decision*)
2. **If you were Rachel, how uncomfortable would you feel if you did NOT pull the lever/ did NOT push Darren?**
   1 _____10
   Not at all uncomfortable Very uncomfortable
3. **If you were Rachel, how uncomfortable would you feel if pulled the lever/ pushed Darren?**
   1 _____10
   Not at all uncomfortable Very uncomfortable
4. **Why would pulling the lever/pushing Darren NOT be the right thing for Rachel to do?** (*Non-utilitarian decision)*
5. **Why would pulling the lever/pushing Darren be the right thing for Rachel to do**? (*Utilitarian decision)*

Figure 16: Example social scenario from the 'Utilitarian Judgments' task

SCENARIO: "Ellie has invited a group of friends to stay for the weekend. Amir is not coping very well after breaking up with his girlfriend. The last few times he has socialised with the group, he has become tearful and talked about nothing but his break-up. Although the group were initially sympathetic, they are now fed up with Amir, since he and his girlfriend broke up a long time ago."

IMPERSONAL VARIANT:
Over coffee, Amir mentions that he's feeling very lonely and would like to do something with the group to take his mind off things. Ellie feels sorry for him and wonders whether she should invite him to stay for the weekend. If Ellie invites Amir, he will be happy. However, the rest of the group will be annoyed and won't enjoy the weekend. If Ellie doesn't invite Amir, he will be left out lonely. However, the rest of the group will enjoy the weekend.

PERSONAL VARIANT:
Amir finds out through Facebook that the group are staying with Ellie for the weekend and asks her if he can come. He says he has been feeling very lonely and would like to do something with the group to take his mind off things. If Ellie lets Amir come, he will be happy. However, the rest of the group will be annoyed and won't enjoy the weekend. If Ellie doesn't let Amir come, he will be left out and lonely. However, the rest of the group will enjoy the weekend.

Questions

1. **If you were Ellie, which of the following options would you choose?**
   a. Invite Amir/ Let Amir come (*Non-utilitarian decision*)
   b. Do NOT invite Amir/ Do NOT let Amir come (*Utilitarian decision*)

2. **If you were Ellie, how uncomfortable would you feel if you invited Amir/ let Amir come?**
   1 _____ 10
   Not at all uncomfortable — Very uncomfortable

3. **If you were Ellie, how uncomfortable would you feel if you DID NOT invite Amir/ DID NOT let Amir come?**
   1 _____ 10
   Not at all uncomfortable — Very uncomfortable

4. **Why would inviting Amir/ letting Amir come be the right thing for Ellie to do?** (*Non-utilitarian decision*)

5. **Why would NOT inviting Amir/NOT letting Amir come be the right thing for Ellie to do?** (*Utilitarian decision)*

11.2.3.2.2 Scoring of discomfort ratings

Participants' ratings of how uncomfortable they would feel with the non-utilitarian and utilitarian courses of action respectively were summed across all eight scenarios, with higher scores denoting greater discomfort. This created a total discomfort rating score for each scenario type, creating eight scores: four for physical scenarios, and four for social scenarios (range 4-40):

Physical scenario: impersonal (1) non-utilitarian discomfort rating, (2) utilitarian discomfort rating;
Physical scenario: personal (3) non-utilitarian discomfort rating, (4) utilitarian discomfort rating;
Social scenario: impersonal (5) non-utilitarian discomfort rating, (6) utilitarian discomfort rating;
Social scenario: personal (7) non-utilitarian discomfort rating, (8) utilitarian discomfort rating.

11.2.3.2.3 Scoring of verbal responses

Participants' responses were classified according to the extent to which they reasoned about the characters' and/or the agent's perspectives. The criteria for these two categories are outlined below; please see Figure 17 for an illustrated example.

Reasoning about characters

Participants' verbal responses were firstly classified according to whether or not they made reference to the characters affected by the decisions (either the individual or the group). Responses were given a score of 1 if they showed sympathy for the characters, considered their perspectives, or to referred to guiding ethical principles. Responses were given a score of 0 if they simply reiterated the consequences for the characters stated in the scenario without elaboration or made no reference to the characters.

This created a total character reasoning score for each scenario type, creating eight scores: four for physical scenarios, and four for social scenarios (range 1-4):

Physical scenario: impersonal (1) non-utilitarian character reasoning, (2) utilitarian character reasoning;
Physical scenario: personal (3) non-utilitarian character reasoning, (4) utilitarian character reasoning;
Social scenario: impersonal (5) non-utilitarian character reasoning, (6) utilitarian character reasoning;
Social scenario: personal (7) non-utilitarian character reasoning, (8) utilitarian character reasoning.

Reasoning about the agent

Participants' verbal responses were also classified according to whether or not they made reference to the agent making the decision (Rachel or Ellie in the examples given above). Responses were given a score of 1 if they showed sympathy for the agent, considered their perspective, or referenced their responsibility. Responses were given a score of 0 if they simply reiterated the role of the agent as stated in the scenario without elaboration or made no reference to the agent.

This created a total score for each scenario type, creating eight scores: four scores for physical scenarios, and four scores for social scenarios (range 1-4):

Physical scenario: impersonal (1) non-utilitarian agent reasoning, (2) utilitarian agent reasoning;

Physical scenario: personal (3) non-utilitarian agent reasoning, (4) utilitarian agent reasoning;

Social scenario: impersonal (5) non-utilitarian agent reasoning, (6) utilitarian agent reasoning;

Social scenario: personal (7) non-utilitarian agent reasoning, (8) utilitarian agent reasoning.

Inter-rater reliability

In order to ensure consistency of scoring, participants' verbal responses were then coded by one rater who was not blind to group membership and by a second blind, independent rater. There was an inter-rater agreement rate of 96% and all disagreements were resolved by discussion.

Figure 17: Scoring of verbal responses from the 'Utilitarian Judgments' task

REASONING ABOUT CHARACTERS:

Non-utilitarian option
"Darren must be feeling very afraid/Darren has a right to life" *(Physical harm)*
"Amir would feel betrayed"/"friendship is more important than a fun weekend"*(Social harm)*

Utilitarian option
"The friends are not to blame/it is right to prioritise the majority" *(Physical harm)*
"The friends would feel frustrated and disappointed if Amir came along" *(Social harm)*

REASONING ABOUT THE AGENT:

"Rachel must be in a difficult position" *(Physical harm)*
"Ellie should be a supportive friend" *(Social harm)*

11.2.4 'Ethics Position Questionnaire' ('EPQ')

This is a self-report questionnaire designed to measure the extent to which people adhere to two ethical perspectives: 'Relativism' and 'Idealism' (Forsyth, 1980). 'Relativism' represents the view that ethical principles are not right or wrong in an absolute sense, but rather that they are dependent upon contextual factors. For instance, "whether a lie is judged to be moral or immoral depends on the circumstances surrounding the action". By contrast, 'Idealism' represents the view that ethical principles are unconditional and give rise to an absolute sense of what is right or wrong. For instance, "people should make certain that their actions never intentionally harm another even to a small degree".

The EQP comprises twenty ethical statements rated on a nine point scale (1 = completely disagree; 9 = completely agree) Total EPQ scores for the respective 'Relativism' and 'Idealism' subscales ranged from 9-90 and the measure has been shown to have good internal consistency (Cronbach's alpha .073-.080) and test-retest reliability (.66-.67) (Forsyth, 1980; Forsyth, Nye & Kelley, 1988).

11.2.5 'Moral Behaviour Inventory' ('MBI')

This is a self-report questionnaire describing twenty-four behaviours, for example "refuse to help people who don't deserve it", "take the last seat on a crowded bus" or "drive out the homeless from your community" (Mendez et al., 2005). Participants rated these behaviours on a scale of 1 to 4, where 1 represents 'Not Wrong', 2 represents 'Mildly Wrong', 3 represents 'Moderately Wrong' and 4 represents 'Severely Wrong'; total MBI scores ranged from 24-96. The measure has been shown to have good split-half reliability (Cronbach's alpha .072-.076).

11.3 RESULTS

11.3.1 Data analysis

Means and standard deviations (SD) for each of the measures of the 'Utilitarian Judgments' task are presented below in Table 9, and for each of moral questionnaires these are shown in Table 10. The assumptions of normality were met and thus parametric analyses were performed.

11.3.2 The 'Utilitarian Judgments' task

11.3.2.1 Utilitarian/ non-utilitarian choices

A repeated measures 2 x 2 x 2 ANOVA was conducted to compare the groups' choices of utilitarian or non-utilitarian courses of action. There was one between-participants factor of AQ group

(high versus low scorers) and two within-participants factors: type of harm (physical versus social), and the personal involvement of the agent (personal versus impersonal).

As expected, there was a significant main effect of type of harm, $F(1,38) = 16.83$, $p = <.0001$, and a significant main effect of the personal involvement of the agent $F(1,38) = 87.18$, $p < .0001$. Inspection of the mean scores confirmed that participants across groups were more utilitarian in situations involving physical versus social harm, and in impersonal rather than personal dilemmas. There was also a significant type of harm by personal involvement interaction, $F(1,38) = 9.71$, $p = .003$, whereby the effect of personal involvement was exacerbated in scenarios involving physical harm.

Contrary to predictions, there was no significant main effect of group, $F(1,38) = 1.11$, $p = .300$, nor were there significant type of harm by group, $F(1,38) = .001$, $p = 1.0$, or personal involvement by group interactions, $F(1,38) = 2.06$, $p = .159$, nor a significant three-way interaction, $F(1,38) = .001$, $p = 1.0$.

11.3.2.2 Discomfort ratings

A repeated measures 2 x 2 x 2 x 2 ANOVA was conducted to compare the groups' ratings of how uncomfortable they would feel in relation to the non-utilitarian and utilitarian courses of action presented in each scenario. There was one between-participants factor of AQ group (high versus low scorers) and three within-participants factors: type of harm (physical versus social), the personal involvement of the agent (personal versus impersonal), and the course of action (non-utilitarian versus utilitarian).

In line with expectations, there was a significant main effect of type of harm, $F(1,38) = 31.58$, $p <.0001$, whereby participants across groups rated situations involving physical versus social harm as more uncomfortable. However, contrary to predictions there was no significant main effect of personal involvement, $F(1,38) = .114$, $p = .738$, nor was there a significant personal involvement by type of harm interaction, $F(1,38) = 1.04$, $p = .314$.

As expected, there was a significant main effect of course of action, $F(1,38) = 54.56$, $p <.0001$, whereby participants across groups rated making utilitarian courses of actions as more uncomfortable than non-utilitarian courses of actions. However, there was no significant course of action by type of harm interaction $F(1,38) = .626$, $p = .434$. There was a significant course of action by personal involvement interaction, $F(1,38) = 48.89$, $p <.0001$, and a significant three-way interaction between course of action, personal involvement and type of harm, $F(1,38) = 11.72$, $p = .001$. Inspection of the mean scores confirms that participants across groups rated making utilitarian versus non-utilitarian courses of actions as more uncomfortable when the personal involvement was high (i.e. personal versus impersonal dilemmas), and that this effect was exacerbated in situations involving physical harm.

As expected, there was a significant main effect of group, $F(1,38) = 6.26$, $p = .017$, whereby the

high AQ groups rated the scenarios as less uncomfortable. However, there were no significant type of harm by group, $F(1,38) = .249$, $p = .621$, personal involvement by group $F(1,38) = 3.69$, $p = .062$, nor course of action by group interactions, $F(1,38) = .262$, $p = .612$. There were also no significant three-way interactions involving group; type of harm by personal involvement by group, $F(1,38) = .831$, $p = .368$; type of harm by course of action by group $F(1,38) = .001$, $p = .986$; personal involvement by course of action by group, $F(1,38) = .123$, $p = .727$. Finally, there was no significant four-way interaction, $F(1,38) = 1.05$, $p = .323$. Thus, the high versus low AQ group rated both utilitarian and non-utilitarian courses of actions as evoking less discomfort, regardless of the scenario conditions.

11.3.2.3 Verbal responses

Reasoning about characters

A repeated measures 2 x 2 x 2 x 2 ANOVA was also conducted to compare the groups' verbal responses with respect to their ability to reason about the scenario characters affected in each scenario. There was again one between-participants factor of AQ group (high versus low scorers) and three within-participants factors: type of harm (physical versus social), the personal involvement of the agent (personal versus impersonal), and the course of action (non-utilitarian versus utilitarian).

Contrary to predictions there was no significant main effect of type of harm, $F(1,38) = .210$, $p = .650$, nor was there a significant personal involvement by type of harm interaction, $F(1,38) = .001$, $p = 1.0$. However, there was a significant main effect of personal involvement, $F(1,38) = 26.14$, $p < .001$, whereby participants across groups were more likely to consider the characters' perspectives in dilemmas with lower personal involvement (i.e. impersonal versus personal scenarios).

As expected, there was a significant main effect of course of action $F(1,38) = 28.99$, $p < .0001$, whereby participants across groups were more likely to consider the characters' perspectives when reasoning about utilitarian versus non-utilitarian courses of actions. However, there was no significant course of action by personal involvement interaction, $F(1,38) = 1.73$, $p = .196$, nor a significant three-way interaction between course of action, personal involvement and type of harm, $F(1,38) = .14$, $p = .710$. There was a significant course of action by type of harm interaction $F(1,38) = 20.07$, $p < .0001$, whereby the effect of greater consideration for characters when reasoning about utilitarian versus non-utilitarian courses of actions was exacerbated in situations involving physical harm.

There was no significant main effect of group, $F(1,38) = .326$, $p = .572$, nor were there significant personal involvement by group, $F(1,38) = .001$, $p = 1.0$, or course of action by group interactions, $F(1,38)$

=.10, p = .753. There were also no significant three-way interactions involving group; type of harm by personal involvement by group, $F(1,38)$ =.001, p = 1.0; type of harm by course of action by group, $F(1,38)$ =.01, p = .921; personal involvement by course of action by group, $F(1,38)$ = 2.59, p = .116. There was also no significant four-way interaction, $F(1,38)$ = 2.59, p = .116. However, in line with expectations there was a significant type of harm by group interaction, $F(1,38)$ = 10.27, p = .003. Inspection of the mean scores revealed that the high AQ group was more likely to reason about the characters' perspectives in situations involving physical versus social harm, whereas the low AQ group were more likely to consider the character's perspectives in situations involving social versus physical harm. However, these group differences did not reach significance when post-hoc t-tests collapsed across the course of action and personal involvement distinctions were performed for physical harm, $t(38)$ = 1.08, p = .287, and social harm, $t(38)$ = 1.72, p = .094 independently.

Reasoning about the agent

A repeated measures 2 x 2 x 2 x 2 ANOVA was also conducted to compare the groups' verbal responses with respect to their ability to reason about the agent's perspective in each scenario. There was once again one between-participants factor of AQ group (high versus low scorers) and three within-participants factors: type of harm (physical versus social), the personal involvement of the agent (personal versus impersonal), and the course of action (non-utilitarian versus utilitarian).

Contrary to predictions, there was no significant main effect of type of harm, $F(1,38)$ = 1.93, p = .173, nor was there a significant personal involvement by type of harm interaction, $F(1,38)$ = .298, p = .589. However, there was a significant main effect of personal involvement, $F(1,38)$ = 11.38, p = .002, whereby participants across groups considered the agent's perspective more when their personal involvement was high (i.e. personal versus impersonal scenarios).

In line with expectations there was a significant main effect of course of action $F(1,38)$ = 139.96, p <.0001; however, this was in the opposing direction, whereby participants considered the agent's perspective more for non-utilitarian versus utilitarian courses of actions. There were also significant course of action by personal involvement $F(1,38)$ = 8.06, p = .007, and course of action by type of harm interactions $F(1,38)$ = 75.18, p < .0001. Inspection of the mean scores suggested that the greater tendency to consider the agent's perspective when reasoning about non-utilitarian courses of actions was exacerbated in situations involving physical harm, and also in situations involving personal harm. However, there was no significant three-way interaction between course of action, personal involvement and type of harm, $F(1,38)$ = .001, p = 1.0.

Contrary to predictions, there was no significant main effect of group, $F(1,38) = .113$, $p = .830$, nor were there significant personal involvement by group $F(1,38) = 2.39$, $p = .130$, nor course of action by group interactions, $F(1,38) = .396$, $p = .533$. There were also no significant three-way interactions involving group: type of harm by personal involvement by group, $F(1,38) = .019$, $p = .892$, type of harm by course of action by group, $F(1,38) = .294$, $p = .591$, or personal involvement by course of action by group $F(1,38) = .224$, $p = .639$. There was also no significant four-way interaction, $F(1,38) = .016$, $p = .901$. However, in line with expectations, there was once again a significant type of harm by group interaction, $F(1,38) = 5.22$, $p = .028$. Inspection of the mean scores revealed that the high AQ group was more likely to consider the agent's perspective in situations involving social versus physical harm, whereas the low AQ group were more likely to consider the agent's perspective in situations involving physical versus social harm. However, these group differences did not reach significance when post-hoc t-tests collapsed across the course of action and personal involvement distinctions were performed for physical harm, $t(38) = 1.62$, $p = .114$, and social harm, $t(38) = 1.25$, $p = .219$ independently.

Table 9: Mean percentage scores and standard deviations for scenarios from the 'Utilitarian Judgments' task

	PHYSICAL HARM		SOCIAL HARM	
	Low AQ group (N = 20) M (SD)	**High AQ group** (N = 21) M (SD)	**Low AQ group** (N = 20) M (SD)	**High AQ group** (N = 21) M (SD)
Choice (0-4)				
Impersonal	2.55 (1.54)	3.00 (1.12)	1.40 (1.09)	1.85 (1.26)
Personal	1.3 (1.34)	1.45 (1.28)	1.00 (0.92)	1.15 (0.81)
Discomfort rating				
Non-utilitarian (4-40)				
Impersonal	32.50 (1.56)	28.25 (1.56)	26.35 (1.49)	22.55 (1.49)
Personal	27.95 (1.70)	24.70 (1.70)	25.00 (1.45)	22.55 (1.45)
Utilitarian (4-40)				
Impersonal	34.85 (1.38)	30.34 (1.38)	31.70 (1.25)	29.40 (1.25)
Personal	36.90 (1.70)	35.60 (0.98)	33.20 (0.88)	31.15 (0.88)
Reasoning about characters				
Non-utilitarian (0-4)				
Impersonal	1.35 (0.21)	1.45 (0.21)	2.25 (0.24)	1.70 (0.24)
Personal	0.80 (0.19)	1.25 (0.19)	1.85 (0.28)	1.50 (0.28)
Utilitarian (0-4)				
Impersonal	2.50 (0.25)	2.85 (0.25)	2.40 (0.26)	2.00 (0.26)
Personal	2.15 (0.31)	2.15 (0.31)	1.90 (0.25)	1.30 (0.25)
Reasoning about agent				
Non-utilitarian (0-4)				
Impersonal	2.65 (0.20)	2.15 (0.20)	1.15 (0.24)	1.40 (0.24)
Personal	3.10 (0.21)	2.85 (0.21)	1.55 (0.26)	1.95 (0.26)
Utilitarian (0-4)				
Impersonal	0.60 (0.17)	0.30 (0.17)	1.25 (0.25)	1.40 (0.25)
Personal	0.55 (0.20)	0.60 (0.20)	1.10 (0.26)	1.60 (0.26)

11.3.3 'EPQ'

The high and low AQ groups were compared with respect to their 'Relativism' and 'Idealism' and scores on the EPQ, adopting a strict p value of .025 (p = .05/2). Contrary to expectations the groups did not differ on either the 'Relativism' $t(38) = 1.02$, $p = .315$, or the 'Idealism' scales $t(39) = .190$, $p = .851$.

11.3.4 'MBI'

The high and low AQ groups were also compared with respect to their judgments of the morality of various misdemeanours on the MBI. As expected, this did not reveal any significant differences between groups, $t(38) = .247$; $p = .806$.

Table 10: Mean percentage scores and standard deviations for the 'EPQ' and 'MBI'

	Low AQ group (N = 20) M (SD)	**High AQ group** (N = 21) M (SD)	**Significance**	**Effect size** (d)
EQP				
Relativism (9-90)	56.15 (10.98)	59.95 (12.55)	.315	.032
Idealism (9-90)	61.40 (11.59)	62.20 (14.88)	.851	.059
MBI (24-96)	58.65 (8.60)	59.25 (6.62)	.806	.078

11.4 DISCUSSION

11.4.1 Summary of findings

11.4.1.1 Utilitarian/ non-utilitarian choices

The 'Utilitarian Judgments' task replicated the classic trolley/ footbridge dilemmas, finding a main effect of personal involvement on people's tendency to make utilitarian decisions: participants across groups were less likely to choose utilitarian choices of action when the agent's personal involvement was high (i.e. personal versus impersonal scenarios). For instance, participants were much less likely to choose to try to stop the train when it involved pushing Darren off the bridge rather than simply pulling a lever. Furthermore, this effect of personal involvement interacted with the type of harm, whereby this was exacerbated in situations involving physical rather than social harm. There was also a main effect of harm, since participants across groups were much less likely to select utilitarian courses of actions in situations involving social versus physical harm. However, contrary to expectations, there was no main effect of group, nor were there any interactions between group and the factors of personal involvement or type of harm.

11.4.1.2 Discomfort ratings

Participants made separate ratings for their experience of discomfort in following the utilitarian and non-utilitarian course of actions. There was a three-way interaction between course of action, personal involvement and type of harm, whereby utilitarian decisions were found to be more uncomfortable when the agent's personal involvement was high (personal versus impersonal, i.e. push Darren), particularly for situations involving physical harm. There was a two-way interaction between course of action and personal involvement. There were also main effects of course of action and of type of harm, with participants across groups rating utilitarian versus non-utilitarian courses of action and physical versus social scenarios as more uncomfortable. However, there was no main effect of personal involvement. With respect to group, the high AQ group experienced less discomfort than the low AQ group. However, this was confined to a main effect and there were no interactions with course of action, personal involvement or type of harm.

11.4.1.3 Verbal responses: Considering the characters' and the agent's perspectives

Participants were asked to provide verbal justifications for why following each course of action might be considered the 'right' thing to do. These were classified according to references made to the various characters' perspectives and/or to the agent's perspective. For characters' perspectives, there was a main effect of course of action, whereby participants across groups reasoned more about characters' perspectives when justifying utilitarian (e.g. "You will be able to save five of your friends, even though Darren would die. You would hope that Darren would be understanding.") versus non-utilitarian courses of action (e.g. "Darren is a bystander and this means that he shouldn't be involved in the decision"). There was no main effect of type of harm, but there was an interaction between type of harm and course of action, whereby groups reasoned more about characters' perspectives when justifying utilitarian courses of actions in physical versus social scenarios. There was also a main effect of personal involvement, whereby participants across groups were more likely to consider the characters' perspectives when the personal involvement was low (impersonal rather than personal). With respect to group, contrary to expectations, there was no main effect nor any interactions with the course of action or personal involvement factors. However, there was an interaction with type of harm, whereby the high AQ group showed greater consideration of the characters' perspectives for physical scenarios, and the low AQ group greater consideration of the characters' perspective for social scenarios.

For the agent's perspective there was a main effect of course of action, whereby participants across groups reasoned more about the agent's perspective when justifying non-utilitarian (e.g. "It would be more difficult to have five friends than one friend die.") versus utilitarian decisions (e.g. "You would be leaving things as they are, which might save you some guilt."). There was also a main effect of personal involvement whereby participants across groups were more likely to consider the agent's perspective when the personal involvement was high (personal rather than impersonal). However, there was no main effect of type of harm. Whilst, there was no three-way interaction, there were significant two-way interactions, whereby personal involvement and type of harm exacerbated the effect of course of action with participants more likely to reason about the agent's perspective for non-utilitarian choices when the personal involvement was high, or when the harm was physical. With respect to group, there was again no main effect nor any interactions with course of action or personal involvement. However, there was an interaction with type of harm, whereby the high AQ group considered the agent's perspective more for social versus physical scenarios, whereas the low AQ group considered the agent's perspective more for physical versus social scenarios. Thus, taken together, for social scenarios the high AQ group paid more attention to the agent and less attention to the characters than the low AQ group; for physical scenarios, the high AQ group paid more attention to the characters and less attention to the agent than the low AQ group.

11.4.1.4 Moral questionnaires

In order to assess both moral knowledge and ethical beliefs, and whether these might differ across the groups, two questionnaires were also administered, the 'MBI' and the 'EPQ'. As expected, the groups did not differ on the 'MBI', which assesses moral knowledge. However, contrary to expectations, the groups also did not differ on the 'EPQ', which assesses ethical beliefs and principles.

11.4.2 Performance on moral questionnaires

Whilst a number of studies have found evidence of intact 'black and white' moral judgments in people with ASD (e.g. Blair 1996; Leslie et al., 2006; Brewer et al., 2015), other work indicates a less sophisticated understanding of morality. For instance, Zalla et al., (2011) found that although adults with ASD made accurate distinctions between moral and conventional transgressions (please see Chapter 4.2.2 for a discussion of this), they provided much more basic justifications for this than controls, largely referencing social rules rather than considering the implications of such transgressions for the characters' welfare. Furthermore, Grant et al., (2005) found that whilst children with ASD were able to specify that unintentional harm should not be judged as harshly as intentional harm, their explanations for this tended to simply reiterate the facts of the story rather than explaining the reasoning behind their culpability judgments. Overall, it seems that those with ASD, and possibly those with high AQ traits, have an intact knowledge of moral rules/principles, but that differences might emerge when we assess the understanding of these at a deeper level.

Thus, on the present study it was expected that the groups might show similar performance on the 'MBI', reflecting the high AQ participants understanding of basic moral 'rights and wrongs'. In line with expectations, no group differences were identified on the 'MBI', extending the findings of Gleichgerrcht et al., (2013) where people with ASD were not found to differ from controls on this measure. In contrast, it was thought that the high AQ group might express their understanding of moral rules on the 'EPQ' differently; they were expected to score more highly for the 'Idealism' and lower on the 'Relativism' sub-scales, indicating a more idealistic ethical standpoint and reflecting their putatively more simplistic and rule-bound approach to moral behaviour. However, this was not found to be the case, since group differences did not emerge on the 'EPQ'.

It seems likely that whilst the 'EPQ' might provide a deeper test than the 'MBI', which simply asks people to judge the extent to which something is right or wrong, it still relies predominantly on explicit knowledge of moral rules. The 'EPQ' is a simple questionnaire measure, and does not require people to explain or justify their ratings. The studies by Zalla et al., (2011) and Grant et al., (2005) highlight the importance of considering not only the quantitative or forced-choice aspects of moral behaviour, but also the reasoning behind judgments and choices that might reveal why such beliefs might be justified. A more

sensitive measure was, however, provided in the 'Utilitarian Judgments' task where participants were required to justify both utilitarian and non-utilitarian courses of action, and this revealed group differences. This suggests that although the high AQ group might have had a fairly good understanding of moral rules, neither the 'MBI' nor the 'EPQ' were sensitive enough to reveal the more subtle differences between the groups that emerged on the 'Utilitarian Judgments' task.

11.4.3 Performance on the 'Utilitarian Judgments' task

11.4.3.1 The influence of personal involvement and the type of harm involved for utilitarian decisions

This present study replicated the classic finding that the degree of personal involvement mitigates utilitarian choices; although participants across groups were more likely to deliver utilitarian choices in the original trolley dilemma and equivalent scenarios (impersonal), they were less likely to deliver utilitarian choices when their personal involvement was increased (personal), such as in the footbridge dilemma (e.g. Greene et al., 2001). Furthermore, participants across groups also gave higher ratings of discomfort when considering utilitarian choices in scenarios involving high versus low personal involvement (personal versus impersonal), e.g. participants experienced more discomfort when considering pushing Darren rather than simply flicking the lever. This extends our understanding of the role of personal involvement, indicating that although people are more likely to make utilitarian decisions, once their experience of discomfort for making such a decision is increased they are less likely to make utilitarian decisions. This is consistent with previous work suggesting the engagement of more rational processes when considering impersonal dilemmas, such as the trolley scenario, versus more emotional processes when considering personal dilemmas, such as the footbridge dilemma (Greene et al., 2001; 2004).

Interestingly, the role of personal involvement was also found to interact with the perspectives that people considered when they were required to justify why each course of action might be the right thing to do (utilitarian and non-utilitarian). Participants across groups were less likely to take the characters' perspectives into account and more likely to consider the agent when the personal involvement was high versus low. Increasing the personal involvement thus appears to focus more attention on the agent's role, presumably because it becomes more difficult to behave in a utilitarian manner and/or to feel justified in doing so.

In addition to examining the influence of personal involvement, a novel manipulation was also included, the type of harm: physical versus social. Previous studies of utilitarian decision-making have tended to involve either exclusively cases of physical harm, or to include cases of extreme emotional/illegal harm (such as selling pictures of your child to pornographers in order to make money) (e.g. Greene et al.,

2001; 2004). The present study aimed to explore utilitarian decision-making in real-life-type scenarios, to investigate whether the established principles of utilitarian decision-making hold true for everyday social situations. How do people balance risks and benefits in a utilitarian dilemma in a social context? Can social harm be compared and equated in the same manner as physical harm? Whilst utilitarian decisions were found to be more common than non-utilitarian decisions, this was influenced by the type of harm, whereby participants across groups were more likely to make utilitarian decisions for physical versus social harm. Interestingly, these utilitarian decisions involving physical harm were rated to be more uncomfortable, suggesting potential conflict between personal distress and knowing the 'right' thing to do.

This pattern of findings suggests that utilitarian decision-making might be more likely to occur in physical situations where the risks involved are greater and the consequences of the options are more black and white. Estimating the risks versus benefits involved for the various characters in physical harm scenarios might be said to be fairly clear-cut: should I kill Darren or let five people die? The harm incurred by letting Darren die or letting the friends die are directly equivalent (i.e. death), and thus the undisputed utilitarian option is to choose to sacrifice Darren in order to save the five friends. However, in social situations, such as the break-up dilemma presented above (see Figure 16), estimating the risks and benefits to the various parties might be considered to be less straightforward. It is more difficult to quantify the disappointment experienced by the group and weigh this against hurting Amir's feelings, in order to decide which course of action represents the lesser of the two evils. In contrast, the harm incurred by the two parties in the social scenarios (i.e. hurting the feelings of Amir or disappointing the friends) is not directly equivalent, although they are similar in nature. Thus, it is possible that participants viewed the harm incurred by hurting Amir's feelings as higher than the sum of the harm incurred by disappointing the group. Therefore, in social scenarios participants might have been less likely to view sacrificing the one (i.e. Amir) as the utilitarian option, and rather to have viewed protecting Amir at the expense of the group to incur incur less harm overall, even though more individuals were harmed.

We might have expected to see a greater consideration of the characters' and agents' perspectives in situations involving social rather than physical harm, due to the greater demands involved in making utilitarian decisions in the social versus physical scenarios. However, this pattern was not confirmed, since there was no main effect and an interaction in the opposite direction emerged, namely that people were more likely to reason about both the characters' and the agents' perspectives when considering situations where physical rather than social harm occurred. Therefore, although social scenarios might involve a less clear-cut balance of harm, this did not seem be the crucial factor driving consideration of the different perspectives; rather, it seems that physical harm was more likely to engender greater consideration of both the characters' and agents' perspectives. It is possible that this might reflect the extent of the harm caused; it might be said that the social harm described in the current scenarios is less severe than the physical harm,

since the consequences, although upsetting, are not life and death. Better matching of the severity of harm involved across the two scenarios sets, with more severe social harm or more minor physical harm, may have reduced or removed the effect of this manipulation.

11.4.3.2 Utilitarian decision-making and autistic traits

On the basis of a previous study examining utilitarian decision-making in people with ASD (Gleichgerrcht et al., 2013), it was expected that the high versus low AQ group would make more utilitarian choices in personal scenarios than the low AQ group. However, no differences between groups emerged for the courses of action they chose. What might account for this failure to replicate? The previous study conducted by Gleichgerrcht et al. included a clinical sample of participants diagnosed with ASD. In comparison the present study included healthy participants with sub-clinical levels of autistic traits. Thus, it might be argued that any group differences would be less marked on the present task than on the previous study. Whilst this is a reasonable possibility, there is nevertheless evidence to suggest that the high AQ group in the present study displayed executive and empathic weaknesses relative to the low AQ group, which might be expected to lead to more utilitarian decision-making. (Please note that the same sample of participants examined in the present study were described in Chapter 10, where tests of executive functioning and empathic processing are reported: see Chapter 10.3.3).

Alternatively, there were a number of differences between the task used by Gleichgerrcht et al. and that developed for the present study. As with the majority of previous literature examining utilitarian decision-making, Gleichgerrcht et al. only examined the trolley and footbridge dilemmas, and it is thus possible that the group differences detected reflect the idiosyncratic nature of these scenarios. As discussed above, physical versus social harm might be argued to be more clear-cut, and the harm posed to the various parties in trolley/footbridge scenarios is extremely salient: either one person or five people die. The ASD group examined by Gleichgerrcht et al. might have been more sensitive to this salient contrast between choices (i.e. kill one to save five, or let five die) and thus have been more likely to act in a "rational" manner even when the extent of personal involvement was increased. On the present study any possible group differences resulting from the trolley scenario might have been masked by the inclusion of a range of scenarios, which presented more realistic but potentially less clear-cut situations. In order to examine this possibility, the groups were compared on the personal and impersonal trolley/footbridge scenarios alone. However, still no group differences emerged, suggesting that the discrepancy between the study findings cannot simply be accounted for by the particular characteristics of the classic trolley/ footbridge dilemmas.

Did the present task unintentionally water down any influence of autistic traits on utilitarian decision-making? There is one further difference between the tasks: in the present study the character who

was the single victim (e.g. Darren or Amir in the examples shown) was always known to the participant and given a name, whereas in the classic scenario used by Gleichgerrcht et al. the characters were strangers and nameless. This amendment was made to the classic scenario in order to balance the physical and social items as well as possible. However, it might have impacted upon people's willingness to be utilitarian; once the victim had a name (e.g. 'Darren' or 'Amir') people may have felt more personal loyalty to them, and this might have weakened the effect of the agent's personal involvement by making both the personal and impersonal scenarios more personal in nature. Despite this, in the present study there were effects of personal involvement in the predicted direction, suggesting that this factor still played a part as expected. With respect to group differences, if anything it might be argued that those with high versus low AQ traits would be less influenced by any acquaintance with the characters, and therefore it is difficult to account for the lack of group differences on this basis.

Turning now to the discomfort ratings, as expected the high AQ group was found to experience less discomfort when considering both utilitarian and non-utilitarian courses of action. Interestingly, this difference was found across the board; it had been expected that the high versus low AQ group ratings of discomfort might differ selectively for scenarios involving social versus physical harm, or to interact with the influence of the personal involvement factor. Instead, it appears that the high AQ group were less aware of the impact of harm on both parties, despite making similar utilitarian versus non-utilitarian choices.

Did a lack of experienced discomfort reduce the complexity of the decisions for the high AQ group? People with high AQ traits might be expected to show less consideration of the characters' perspectives than the low AQ group, in view of their difficulties with 'putting themselves in others' shoes'. This in turn may well be linked to the reduced discomfort they experienced in making utilitarian and non-utilitarian decisions. Indeed, although the groups did not differ in their utilitarian versus non-utilitarian choices, there were differences in their verbal rationales relating to such choices. The high AQ group was found to consider the characters' perspectives more for the physical versus social scenarios, whereas the low AQ group displayed the opposite pattern. Although the physical scenarios have more severe consequences, the impact of these on the characters is clear-cut, i.e. death. Social scenarios might be said to be more demanding with respect to the need to consider others' perspectives: in comparison to the prospect of dying, the impact of harm on the character requires more detailed consideration. Thus, the high AQ participants may have failed to appreciate the more nuanced implications of social harm, and instead have focused on the severity of harm inflicted upon the various characters. In contrast, the high AQ group considered the agent's perspective more for social versus physical scenarios, whereas the low group showed the opposite pattern. This suggests that the high AQ group might have failed to appreciate

that inflicting physical harm onto others, such as killing someone, would be a more onerous decision for the agent than deciding to inflict social harm, such as excluding someone.

Taken together, the present findings show that whilst the high and low AQ groups did not differ in the choices they made about the most appropriate course of action, they did differ in their reasoning about such choices and experienced discomfort. This pattern might be said to be comparable with previous work showing an intact knowledge of what is 'right and wrong', but a more limited appreciation of the implication of violating such rules and reasoning about this in those with ASD (Grant et al., 2005; Zalla et al., 2009; please see section 11.4.2 for a brief discussion of this). Making utilitarian or non-utilitarian decisions probably lies somewhere in between the concept of knowing the rules versus understanding and reasoning about them, but could certainly be said to be less demanding than the verbal questions which tapped understanding and reasoning about moral dilemmas.

11.4.3.3 The possible contribution of executive and empathic processes

As discussed, both cognitive and emotional processes are thought to contribute to morality (please see Chapter 4.1 for a more detailed discussion). Haidt et al., (2001) suggested that intuitive emotional responses may lead to rapid moral judgments such as, "incest is wrong under any circumstances", but that cognitive processes are required to justify such judgments. With respect to utilitarian decision-making, the tendency to make 'rational' utilitarian judgments on the trolley scenario has been linked to brain regions involved with general reasoning skills, whereas people tend to make more 'instinctive' non-utilitarian judgments on the 'footbridge' dilemma which is reflected in their greater engagement of brain regions involved in emotional processing (Greene et al., 2001). It is thus possible that both executive functioning and empathic processes might be involved in performance on the 'Utilitarian Judgments' task, and potentially underpin the group differences detected.

Broadly speaking, both the findings from the moral questionnaire and the 'Utilitarian Judgments' task suggest that the high AQ group appeared to know what is 'right and wrong' and to apply this to the scenarios when they are asked to choose a course of action, but that they struggled to justify the legitimacy of different courses of action when asked to explain them. It is possible that the high AQ participants made broad decisions about the best course of action, without necessarily integrating fully their knowledge of moral rules into the scenario context to consider the pros and cons. This could be underpinned by an executive failure to apply their knowledge of moral rules, and/or to integrate this with empathic processes.

As mentioned, there is some evidence for executive weaknesses in the high AQ group, as described in Chapter 10.3.3. It is possible that executive processes played a role in weighing up different courses of action and reasoning about them. Group differences might not have manifested when participants were asked to make a choice between utilitarian versus non-utilitarian courses of action, since this could be

argued to be relatively low in executive demand and to rely more heavily upon social knowledge. However, group differences emerged when the participants were asked to justify the basis for different courses of action (as measured by the verbal responses), for which the executive demand could be argued to be higher. It is possible that weaker executive skills in the high AQ group meant that when making their verbal responses, they were less able to appreciate the complexities and/or longer-term consequences of a decision both for the characters and the agent, particularly for social scenarios where the outcomes were less clear-cut. It is less clear how weaker executive skills could explain the reduced experience of discomfort by the high AQ group, since this is an emotional reaction that is likely to be predominantly mediated by empathic processes.

Alternatively, the findings could be explained in terms of impaired empathic processes. In addition to weaker executive skills, there was also some evidence for differences in cognitive and emotional empathy in the high AQ group (please see Chapter 10.3.3). It is possible that cognitive and/or emotional empathy played a role in appreciating others' perspectives and resonating with their pain. For instance, failing to appreciate the impact of harm upon the characters could have resulted in less discomfort in the high AQ group for both social and physical scenarios. With respect to the verbal responses, given the well-established difficulties with cognitive empathy including perspective-taking, consideration of both the characters' and the agent's perspectives is likely to have been more difficult for the high AQ group. This could therefore account for the high AQ group's greater focus on implications of physical versus social harm for the characters, and the converse pattern for the agent.

11.4.4 Strengths and weaknesses of the 'Utilitarian Judgments' task

Most work examining utilitarian decision-making to date has focused on the influence of the agent's role via manipulating the extent of their personal involvement in the scenarios. In contrast, the findings from this novel task highlight the importance of examining a broader set of factors that might influence utilitarian decision-making. For instance, the inclusion of both a question relating to the choice of utilitarian or non-utilitarian course of action, and a question relating to the ability to reason about each of these courses of action, advances our understanding of how people weigh-up the different parties' perspectives when considering utilitarian dilemmas. It allowed for assessment of how different scenario factors may influence the likelihood of considering the various characters and/or the agent's perspectives. With respect to examining utilitarian decision-making in clinical or sub-clinical groups, assessing justifications as well as choices proved to be more sensitive to subtle differences.

Furthermore, most previous work has also only focused on physical harm or very extreme forms of emotional/illegal harm. This study included instances of both physical and social harm, in order to investigate whether the principles of utilitarian decision-making established from examining extreme harm

translate to more everyday situations. The findings suggest that utilitarian courses of action might also be preferred for cases of social harm, at least under certain circumstances, but that this effect is not as strong as for physical harm. This highlights the extreme and unrealistic nature of the trolley/ footbridge classic scenarios, which do not represent everyday situations and perhaps elicit more utilitarian decision-making than we are likely to see in other kinds of dilemmas.

As has been discussed (see section 11.4.2.1), it is not clear whether any performance differences on the social versus physical harm scenarios in this task were attributable to the social/ physical dimension or rather to the severity of the harm involved. The sensitivity of the task to the social/physical dimension could have been improved by making the scenarios involving social harm more severe, by including consequences such as 'social ostracism'. The present task included cases of people being excluded from groups or events, but these did not have such enduring or severe consequences. Alternatively, the physical harm scenarios could have been made less severe, by including cases such as physical injury rather than simply death. However, for either augmentation it would be extremely difficult to match the social/physical scenarios whilst maintaining the same overall scenario structure of the trolley/footbridge dilemmas. In this initial study of utilitarian decision-making for social harm it was important to include the trolley/footbridge scenarios and to use these as a basis for developing further scenarios, in order to allow for comparability of the task with previous literature. Future work might seek to explore utilitarian decision-making not only in different settings (i.e. physical or social harm/ hypothetical versus everyday), but also in relation to different scenario features (i.e. number of individuals involved in either party, or the role of the agent as a bystander or active participant in the scenario). This would, however, require departing from using the basic trolley/footbridge scenario structure in order to allow for a wider range of manipulations to be explored.

Despite attempts to match the new scenarios to the classic scenarios as closely as possible, there was also a need to match the physical and social scenarios to one another. As mentioned, there was one key difference between the present study and those used previously, whereby the single victim (e.g. Darren or Amir in the examples shown) was always known to the participant and given a name. This was a necessary adaptation to the trolley/footbridge scenarios, which allowed the physical and social scenarios to be matched as closely as possible. However, it might well have impacted upon people's willingness to be utilitarian, and have weakened the personal involvement effect, although this was still detected. It would be interesting to systematically explore this in future tasks by comparing known characters with strangers in the scenarios, although this again poses practical difficulties with respect to the matching of the materials.

11.4.5 Conclusion

The present study involved developing a novel scenario-based task to compare utilitarian decision making in physical versus social harm in people with high versus low autistic traits. The groups did not differ in their choice of course of action, but they differed in their consideration of the characters' and the agent's perspectives when reasoning about why utilitarian or non-utilitarian decisions might be appropriate. The high AQ group also experienced less discomfort when making such decisions. These findings are broadly consistent with previous literature suggesting that people with ASD might 'know' 'right from wrong', but have a poorer understanding of the implications of violating moral rules and are less able to reason about their contingencies.

This study also extends the literature on utilitarian decision-making in neurotypical individuals. The findings suggest that the focus on extreme types of harm explored in classic utilitarian dilemmas might over-estimate people's tendency to behave in a utilitarian manner when applied to everyday situations. This concludes the experimental work presented in this thesis.

PART 4

CHAPTER 12: GENERAL DISCUSSION

12.1 Overview

The present thesis explored how autistic traits influence everyday social functioning. The seven experimental chapters presented will now be discussed and their implications considered. Firstly, the methodological contribution of this body of work will be explored, including a consideration of its strengths and weaknesses. Secondly, the theoretical contribution of this work for the understanding of autistic traits, and the relationship to aetiological models of ASD, will be examined. Thirdly, the potential clinical implications of this work for both the assessment and management of everyday social difficulties will be explored. Finally, a number of outstanding questions and possible avenues for future research will be discussed.

12.2 Novel contribution to the literature

Successful social functioning is crucial for the development of relationships (e.g. Lopes, Salovey, Côté, Beers, & Petty, 2005), which are in turn thought to play a protective role in both psychological (Segrin & Taylor, 2007) and physical health (Cohen, Doyle, Skoner, Rabin & Gwaltney, 1997). In addition to enhancing an individual's general wellbeing, socially skilled behaviour has benefits for others within the social environment, such as considering the feelings of others and engaging in pro-social behaviour (e.g. Channon et al., 2012). There are a number of psychiatric and neurological conditions which are associated with disrupted social behaviour, including brain injury, neurodegenerative and neurodevelopmental disorders, in particular ASD.

The work presented in this thesis has made three key novel contributions to the existing literature. Firstly, these studies explored complex social behaviour in the broader autistic phenotype, examining a sub-clinical sample of people scoring high on a self-report measure of autistic traits, as opposed to a clinical sample of people diagnosed with ASD. Secondly, a range of more ecologically valid experimental tasks tapping social cognition was developed in order to elucidate real-life-type social behaviours. Finally, aspects of social functioning were addressed which have been relatively under-researched in relation to ASD: pro-social behaviour and moral reasoning. These three contributions will be detailed below in section 12.3, and the theoretical implications of the findings from these studies will be discussed in section 12.4. The links to psychological models of ASD, and the clinical implications will be discussed in sections 12.5 and 12.6.

12.3 Methodological contribution

12.3.1 The broader autistic phenotype

The work presented in this thesis is novel in that it adopted a trait-based approach to examine the influence of autistic traits on everyday social behaviour. Participants from four successive community samples were screened using a self-report measure of autistic traits, and subsets drawn from the highest and lowest deciles of each sample were compared in a series of studies. This has extended our understanding of the autism spectrum by identifying weaknesses in social functioning in those with high levels of autistic traits that are, qualitatively speaking, broadly consistent with the profile of people with a diagnosis of ASD. This supports the idea of autism as a spectrum condition which lies on one end of a continuum (e.g. Baron-Cohen et al., 2001), and is in line with the revisions to DSM-5 which subsumes different subtypes that previously had separate diagnostic criteria (see Chapter 1.1 for a discussion of this issue). The notion that people with sub-clinical levels of autistic traits might demonstrate similar types of impairment in social behaviour, but perhaps to a lesser extent, adds another strand to existing evidence indicating a continuous distribution for autistic traits (Skuse, Mandy & Scourfield, 2005), and with other work indicating similar genetic aetiologies for people with high autistic traits in the general population and those who met diagnostic criteria (Robinson et al., 2011).

12.3.2 Balancing experimental control with ecological validity

As highlighted, another novel contribution of this thesis is the exploration of 'real-life-type' situations. Whilst there is a great deal of experimental work examining both typical and atypical social behaviour, this tends to focus on abstract laboratory-based tasks. This existing body of work has proved instrumental in illuminating the fundamental processes that enable humans to interact with one another, and in characterising core deficits that might lead to problems with social cognition and behaviour. However, to date there is a paucity of experimental work which seeks to translate these findings into real-life-type social behaviours and to directly address difficulties with everyday functioning (Channon & Crawford, 1999); this was the aim of the present thesis.

A series of scenario-based tasks was developed which adhered to the fundamental principles of controlled experimental work, but also aimed to achieve greater ecological validity. Thus, the tasks were administered in a systematic manner and the crucial aspects of experimental control were preserved, such as the counterbalancing and matching of materials across different task conditions, the use of structured questions, and rigorous quantitative scoring of qualitative data to tap different aspects of performance. Moreover, the different scenarios included in the tasks were derived from real-life insights garnered from the experiences of both neurotypical individuals and people with ASD in various clinical settings.

The task demands of the various experimental studies were varied to include a range of more or less demanding measures, ranging from ratings scales, to forced-choice option questions and free verbal responses. This allowed for an examination of different levels of difficulty upon task performance. In addition, the tasks were designed to assess different aspects of performance, including behavioural (e.g. what would you do?), cognitive (e.g. what would you be thinking?) and emotional (e.g. what would you be feeling?) aspects, and also the reasoning behind people's judgments and decisions (e.g. why would you do/ think/ feel this?).

12.3.3 Exploring the 'social' in social cognition

The third main contribution of this thesis is the specific focus of the novel scenario-based tasks on different aspects of social functioning ranging from pro-social behaviour to moral judgment and reasoning. The first study (reported in Chapter 5) presented the 'Above and Beyond' task, which addressed the lengths that people are willing to go to help others. The scenarios in this task involved situations in which a character is in need of help (e.g. a man falls over heavily in front of you - it looks like he might need medical attention), but where there is a small inconvenience incurred in doing so (e.g. if you stop to help you will be late to an important work meeting). The primary question this was: "Are those high in autistic traits less pro-social?".

Participants high in autistic traits were found to behave less pro-socially, which was thought to possibly reflect a reduced understanding of the expectation to help others. This was followed up in the next study with the 'Social Expectations' task (reported in Chapter 6), which compared performance on scenarios with clear-cut versus ambiguous social rules underpinning the expectation to help. For instance, participants were asked whether they would give up their seat to an elderly person (clear-cut rule) versus someone carrying a heavy parcel (ambiguous rule). The question of whether people high in autistic traits understood the inherent expectations to help others in the scenarios was assessed by examining both their ability to estimate it and to reason about it. Whilst those high in autistic traits could identify the social expectations for both clear-cut and ambiguous social rules, they were found to rely more heavily upon rule-based explanations.

One factor that was potentially thought to contribute to performance was the extent to which people are deserving of help. Thus, a third study again assessed pro-social behaviour (reported in Chapter 7), examining whether those high in autistic traits were differentially influenced by how much the story character was deserving of their help. Characters with either a strong or a weak justification for asking for help were compared in the 'Social Favours' task. For instance, one scenario involved a character who wanted help carrying a large package up some stairs, either because they had a bad back (strong

justification), or because they did not want to get their shirt dirty (weak justification). Interestingly, contrary to predictions, the high autistic trait group behaved similarly to the low autistic trait group on this task.

The fourth study (reported in Chapter 8) sought to explore further the nature of 'deservingness'. This time the experiment shifted in focus, from active pro-social behaviour to help the individual, to moral judgments of their deservingness from a more impersonal standpoint. The positive and negative 'Deservingness' tasks explored whether participants thought characters were deserving of various positive and negative outcomes, such as passing their driving task either because they drove really well (deserved) or flirted with the instructor (underserved); or getting reprimanded in class either because the character did not do their homework (deserved) or because the instructor was in a bad mood (undeserved). As with the 'Social Favours' task, no differences emerged between groups, suggesting that people high in autistic traits understood the nature of 'deservingness' and could both accurately differentiate between situations where outcomes were or were not deserved, and respond appropriately.

The fifth study (reported in Chapter 9) continued the exploration of negative events, shifting the focus to thinking about how to avoid them, rather than judging if they were deserved. The ability to generate counterfactual thoughts in response to one's own and/or another's mistakes was examined. For instance, participants were told that either they or their colleague had accidentally left their laptop on a train; or that either they or their colleague had mistakenly laughed at a serious proposal in a work meeting. This study explored the roles of both the 'perpetrator' and the 'victim' in stimulating counterfactual thinking. The central question explored in this experiment was whether those high in autistic traits would generate fewer counterfactual thoughts in relation to their own and others' mistakes. As expected, the high AQ group produced fewer counterfactual thoughts than the low AQ group.

A sixth study (reported in Chapter 10) complemented this by addressing the content rather than the fluency of counterfactual thinking. Participants were shown scenarios where the character made a mistake that had a negative consequence for them. For instance, in one scenario the participant is driven to their graduation ceremony by their sibling, but their sibling forgets to fill up the car with fuel and they break down on the motorway and miss the ceremony. In another scenario, the participant's flatmate forgets it was their grandmother's funeral and invites friends over to their shared flat that evening. On the 'Counterfactual Judgments' task, those with high autistic traits were found to differ in the quality of their counterfactual thinking. Taken together, these two studies suggest that people high in autistic traits struggle to produce counterfactual thoughts, and that those produced are of a different quality.

The final study (reported in Chapter 11) built upon the previous experiments which explored the ramifications of doing the 'wrong thing' by instead examining ethical dilemmas where there are no clear right or wrong answers. This sought to replicate and extend the findings of a classic task, in order to examine if established principles about utilitarian decision-making held true for more everyday situations.

The 'Trolley' and 'Footbridge' problems (Foot, 1967; Thomson, 1976) present participants with a utilitarian dilemma, where they have to decide whether to intervene and save five railway workers by sacrificing one, either by flicking a switch (impersonal version) or by pushing someone onto the track (personal version). The 'Utilitarian Judgments' task developed for this study included the trolley/ footbridge scenarios and three further utilitarian dilemmas involving inflicting physical harm; four original social harm scenarios were also developed. Contrary to expectations, those high in autistic traits made similar decisions to the low autistic trait group, differentiating between situations involving high or low personal involvement (flick switch or push person) and the type of harm involved (physical or social). However, the high autistic trait group differed in their ability to reason about this and experienced less discomfort when making utilitarian decisions.

12.3.4 Extending the findings

The experimental chapters presented in this thesis have already made specific suggestions for possible ways in which the sensitivity of each task could be improved, or the line of work extended. However, more broadly, there are two key ways in which this body of work could be further advanced.

Firstly, whilst this thesis addressed a range of situations, these virtually all involved either averting negative outcomes (task examining pro-social behaviour and utilitarian decision-making), or responding to negative outcomes (tasks examining deservingness and counterfactual-thinking). One of the studies included positive as well as negative outcomes (the 'Deservingness' tasks presented in Chapter 8) and the pattern of group performance was similar, suggesting that this might not be confined to negative outcomes, but this needs further exploration. These situations also fell into two types of broad aspects of social behaviour, helping others (pro-social behaviour), or making moral judgments and reasoning about these. Thus, it is unclear to what extent the overall pattern of findings regarding the performance of the high versus low autistic trait groups might apply to a broader set of situations. Future work should seek to explore the ramifications of autistic traits for everyday social functioning to a number of other aspects of social behaviour. For instance, instead of focusing on helping others, it might be interesting to explore other types of pro-social behaviour, such as dealing with awkward situations, or "letting others down gently". Moreover, moral notions of 'fairness' or 'what is just' could be explored in other everyday contexts, such as situations requiring co-operation with others, or in line with principles of reciprocity.

Secondly, further work should also seek to further elucidate the aspects of preserved versus impaired performance identified in those with high levels of autistic traits. For instance, group differences emerged when participants were required to reason about social and moral rules, but not when they were required to identify them. It was postulated that this might reflect a more rudimentary knowledge of social and moral rules in those with high versus low levels of autistic traits. However, it is also possible that this

discrepancy reflects the complexity of task demands, or some other as yet unidentified factor. In order to identify the significance of this finding, it would be helpful to explore whether there are also islets of preservation in other aspects of everyday social functioning. For more discussion of preserved versus impaired aspects of performance, please see section 12.4.

12.3.5 Limitations of the methodological approach

12.3.5.1 Study design

The present work adopted a trait-based approach to assessing everyday functioning in the autistic spectrum. This involved screening community samples using a self-report measure of autistic traits and then inviting in groups of the highest and lowest scoring people to take part in the study. Whilst this approach proved sensitive to elucidating differences between those scoring high versus low on autistic traits, there are a number of potential limitations that should be considered, including the use of self-report measures and the selection of small groups of extreme autistic trait scorers.

Firstly, self-report measures have been criticised on the grounds that people may report inaccurately when assessing their own personality or behavioural tendencies, either due to lack of insight, or a desire to present oneself in a 'good light' (Lilienfeld & Fowler, 2006). Nevertheless, it would have been difficult to assess autistic traits in a population of young adults without relying on self-report measures. Whilst there are existing parental or teacher report measures that assess autistic traits (e.g. Social Responsiveness Scale; Constantino et al., 2003), these were developed for screening ASD in children and adolescents rather than for an adult population. Alternatively, diagnostic measures like the ADI (Lord, Rutte & Le Couteur, 1994) and ADOS (Rutter, DiLavore & Risi, 2002) could be used, but these were developed to detect clinical impairment and to diagnose ASD, rather than to assess the presence of autistic-like traits. Therefore, it would be difficult to discern how to apply such measures to assessing a sub-clinical population, and arguably their use would be inappropriate. The ADI and ADOS were also again developed primarily for assessing children and younger adolescents, and there are questions as to their ethical and methodological suitability to administer meaningfully to a sub-clinical adult population, and their sensitivity to more subtle difficulties with social behaviour (please see section 12.6.1 for a detailed discussion of these issues). Moreover, whilst self-report measures have limitations, studies have also found that they tend to cohere with observer report (Lilienfeld & Fowler, 2006). The AQ has also been shown to have good internal consistency and construct validity, strong test–retest reliability and robust self versus parental report reliability (Baron-Cohen et al., 2001). It has also been found to significantly differentiate those with ASD versus those without, with people with a diagnosis of ASD typically scoring above 32/50. Interestingly, the mean AQ scores of the high autistic trait groups used in the present thesis also tended to

be fairly high, ranging from a lowest score of 25 and a highest score of 46 on the AQ. This suggests that some individuals included in the high autistic trait groups might have met criteria for ASD had they been fully assessed, but to the authors knowledge none had previously received a diagnosis of ASD or any other neurodevelopmental or neurological condition.

Secondly, why choose the AQ specifically? There are a number of other self-report measures that tap some of the features associated with autistic traits. For instance, the Social Skill Inventory (SSI; Riggio, 1986) is a self-report measure that assesses multiple social and emotional aspects of behaviour, and provides an overarching index of social skill. The SSI has been used widely and has been found to correlate significantly with other measures of personality (Riggio & Carney, 2003), with social behavioural measures (Riggio, 1986), with social network size (Riggio & Carney 2003), and with socially awkward or uncomfortable behaviour. However, unlike the AQ, this measure was not developed to screen for autism or autistic traits. It therefore does not tap the full range of personality traits associated with the autistic spectrum, including the non-social features, such as enhanced attention to detail at the expense of processing contextual information. Although this thesis was predominantly focused on examining social behaviour, more broadly it aimed to extend understanding of the autistic spectrum, and thus using a measure that tapped all associated aspects was seen to be most appropriate.

Thirdly, the present thesis adopted an extreme scorers approach, selecting those from each sample who scored within the highest and lowest deciles of the overall range of AQ scores. In fact, the AQ contains five subscales that tap the range of deficits associated with ASD, including problems with social skill, communication, imagination, and an enhanced attention to detail and difficulties with attention switching. Whilst looking at the impact of the global features of the autism spectrum could be considered a strength it might also be viewed as a limitation, since this approach did not allow for exploration of the selective contribution of different deficits.

It seems most likely that ASD is the result of multiple deficits and that one single account cannot explain the range of social and non-social features associated with it (Happé, Ronald & Plomin, 2006). ASD is also a very heterogeneous condition, with individuals experiencing different degrees of severity in different domains (Happé, Ronald & Plomin, 2006). Analysis confirmed that across samples, the high and low AQ groups significantly differed for all five subscales, and so it is not possible to explore the contribution of different deficits in the present thesis. Examining the performance of people scoring high versus low for different AQ subscales would potentially elucidate the specific influence of each area of impairment for everyday functioning, including the potential contribution of non-social differences to social behaviour, and vice versa. Thus, future work could seek to explore the role of different deficits, as measured by those scoring high on one or several of the five subscales, rather than simply selecting the highest scorers overall.

Fourthly, using only samples of people scoring high versus low for autistic traits might have resulted in a lack of normative data for the novel experimental tasks presented. It is unclear how people who were not selected on the basis of their extreme autistic traits would perform on these tasks, and whether the performance of the high or the low AQ groups are more representative of the population mean.

It is also unclear how the findings might translate to a more extreme clinical population. As yet it is not clear whether the performance of a sub-clinical sample of people scoring high for autistic traits is similar to the performance of a clinical sample of people diagnosed with ASD. It remains unclear if there is a simple 'dose-response' relationship between levels of autistic traits and/or the clinical features of ASD and impairment, and where a clinical threshold should sit. Thus, in the future it would be important to administer these types of tasks to a clinical sample of people diagnosed with ASD in order to determine if they show a similar, although perhaps exaggerated, pattern of performance, or whether their social performance qualitatively differs from those high in autistic traits, but without a diagnosis. This has important implications for the clinical utility of these findings.

Finally, the sample sizes included in the present studies were relatively small, ranging from 19 to 27 participants per group. Cohen's (1992) table was used to establish that these samples sizes were sufficient to detect large effect sizes with a power set at 80% and alpha at 10%; most of the effect sizes detected in practice were indeed large. Nonetheless, future work should seek to enhance statistical power, as this might reveal additional or stronger group differences. An alternative approach would be to use a correlational rather than a group design, where larger samples of people scoring across the range for autistic traits could be compared for performance on various tasks. It should, however, be noted that testing much larger samples than those included in the present thesis would realistically result in the use of simpler tasks that could be administered much more quickly, or online. Although this type of study might have greater statistical power, it is likely to provide less rich information about the nature of social difficulties, and possibly to be less representative of real-life-type problem solving than the present body of work.

12.3.5.2 A note on gender

It is important to note that ASD is more commonly diagnosed in males rather than females (Fombonne, 2009), and that autistic trait scores tend to be higher for males than for females drawn from the general population (Baron-Cohen et al., 2001). In order to ensure that any group differences identified in the present thesis were more likely to be attributable to the role of autistic traits rather than gender, a specific recruitment strategy was adopted: matched numbers of males and females were recruited for the high and low autistic trait groups respectively.

Arguably this approach represents a strength of the present body of work, enabling one to disentangle to influence of autistic traits from the possibly confounding role of gender. However, it could be argued that the participants included in the experimental phase of the present studies were not entirely representative of the general population, since the groups' trait scores might have been more differentiated had gender not have been controlled for (i.e. if the highest and lowest scorers had been recruited, regardless of gender).

With respect to the clinical population of people diagnosed with ASD, it is unclear to what extent the diagnostic bias towards males reflects biological or sociological gender differences (Baron-Cohen, Knickmeyer & Belmonte, 2005; Halladay et al., 2015). Clinical reports have suggested that females with ASD might have better social skills than males with ASD, and have a greater ability to 'mask' or compensate for their everyday impairments (Halladay et al., 2015), and other work suggests that clinicians are less aware of the subtler profiles of differences exhibited by females on the spectrum. Ideally, future work should aim to explore the role of autistic traits, and/or the presence of ASD, within the context of gender differences. This would help to clarify the extent to which the profiles of males versus females differ across the spectrum.

12.3.5.3 Experimental materials

As outlined above, a suite of experimental tasks were developed to tap different aspects of social behaviour in a more naturalistic and fine-grained way than traditional laboratory-based tasks. However, there are two key limitations to this approach that should be addressed. Firstly, although the tasks could be claimed to be more ecologically valid than traditional measures of social cognition, the scenarios were still conducted 'offline'. The content of the tasks (i.e. the problems and situations explored in the scenarios) was representative of real-life-type social behaviours, but the scenarios could not be said to match the demands of real-life social interactions. In real-life social interactions people receive feedback from others, and experience both the short- and long-term consequences of their actions. They also might experience stronger emotional or physiological responses in real-life, such as anxiety or stress at handling sensitive situations. Facilitating this was not possible within the experimental constraints; however, interactive social scenarios could be explored in alternative paradigms, such as using virtual reality. Avatars could be programmed to give participants feedback on the basis of their decisions, and could also provide non-verbal cues to guide performance, which were not examined in the present body of work. Visual cues might make the experience more 'real' and thus enhance the ecological validity of the studies.

Secondly, complex, everyday scenarios of this nature are very likely to draw upon both cognitive and emotional processes known to be involved in social cognition, and hence disentangling their relative contributions to performance on the experimental tasks presented is likely to be extremely difficult.

Differences on task performance may well have involved impairment in several areas, possibly across both the cognitive and emotional domains. Whilst this could be argued to be a limitation, it might also represent a strength of the current approach, highlighting the difficulty of disentangling different processes and finding neat dissociations once we move away from simple laboratory-based tasks and try to tackle more complex behaviours and stimuli.

Taken together, the studies in this thesis have advanced our understanding of the psychological processes involved in social decision-making in sub-clinical populations; this in turn potentially informs the clinical assessment and management of high-functioning people with ASD by identifying particular factors that might facilitate or impinge upon successful social functioning.

As well as a number of key strengths and points of originality there are also some potential limitations to the present body of work, concerning both the design of the studies and the nature of the experimental materials used. These issues should be carefully considered when designing future studies, or interpreting the evidence from these studies.

12.4 Theoretical contribution

The key theoretical contribution of this body of work was the exploration of the influence of autistic traits upon complex and subtle aspects of social functioning, which has been neglected in previous work. Two themes were explored, pro-social behaviour (Chapters 5-7), and moral judgment and reasoning (Chapter 8-11). The findings from each of these strands of work will now be brought together and their implications discussed.

12.4.1 Pro-social behaviour

Pro-social behaviour refers to intentional acts that are designed to help others, such as sharing, donating, co-operating and volunteering (Brief and Motowidlo 1986). It is thought to be important for both society and the individual (Eisenberg & Miller, 1987; Eisenberg et al., 1998); behaving pro-socially has been found to aid social bonding, to have a positive impact on social adjustment, self-esteem, and to contribute towards psychological wellbeing and physical health (Coie et al. 1990; Eisenberg et al. 1998; Puffer, 1987). Pro-social behaviour is believed to be motivated by both cognitive and emotional empathic routes; for example, feeling more empathy has been linked to a greater concern for others' welfare and more helping behaviours (Batson, 1991).

People with ASD are thought to have difficulties with cognitive and possibly emotional empathy, and correspondingly they tend to display less socially sensitive behaviour (Blair, 2008). The work presented in this thesis aimed to explore pro-social behaviour within the autistic spectrum; surprisingly, this has not been explored in detail before. The postulated empathic deficit in people with ASD would be expected to

result in less pro-social behaviour, and thus, we might also expect to see reduced pro-social behaviour in those with high versus low levels of autistic traits.

The first three studies presented in this thesis examined two aspects of performance that are thought to guide pro-social behaviour; responsivity to others' needs and the ability to identify unwritten social rules. Responsivity to others' needs was assessed using measures that tapped both behavioural and affective responses. From a behavioural angle, participants were asked to describe what they would do in each situation ('Above and Beyond' task presented in Chapter 5), or more generally how likely they were to help others ('Social Expectations' and 'Social Favours' tasks presented in Chapters 6 and 7). From an affective angle, across the three tasks participants were asked questions that tapped their ability to identify the characters' needs or to resonate with them. The ability to identify and draw upon unwritten social rules was assessed in the 'Social Expectations' task, presented in Chapter 6. The findings from these three tasks will now be brought together and discussed in relation to these two aspects of performance.

12.4.1.1 Responsivity to others' needs in pro-social behaviour

As discussed, people with ASD are known to have problems with identifying others' needs, but are postulated to have a largely intact ability to share their emotional states and resonate with them (Blair, 2008). Thus, ASD is thought to be associated with impaired cognitive empathy and intact emotional empathy, although there is some conflict in the literature regarding whether emotional empathy is indeed preserved (Blair, 2008). Therefore, in the present body of work it was expected that those high in autistic traits would show difficulties with aspects of performance that relied upon identifying and interpreting others' mental states, and possibly on measures that examined the extent to which they resonated with others' needs.

The first two studies (the 'Above and Beyond' and 'Social Expectations' tasks) presented in Chapters 5 and 6 examined willingness to help others (behavioural responsivity), personal reward for helping others and sympathy for others in need (affective responsivity). With respect to their behavioural responsivity, participants scoring high versus low for autistic traits were, as expected, less pro-social on both the 'Above and Beyond' and the 'Social Expectations' tasks. Moreover, group differences in pro-social behaviour were not modified by a range of task demands; on the 'Above and Beyond' task participants scoring high versus low in autistic traits were less likely to help characters in need both when asked to generate a response, and when asked to select a response from a forced-choice option. On the 'Social Expectations' task, participants scoring high versus low in autistic traits were less likely to help characters even when the social rule guiding the expectation to help was clear-cut versus ambiguous (e.g. giving up your seat to elderly person walking with a stick versus a young person carrying a heavy parcel). With respect to their affective responsivity, participants scoring high versus low for autistic traits

experienced less personal reward for helping others on the 'Above and Beyond' task. Participants high versus low in autistic traits were also less sympathetic to the characters' needs on the 'Social Expectations' task.

However, this pattern of findings did not hold up when performance of those scoring high versus low in autistic traits was examined on the 'Social Favours' task, presented in Chapter 7. In contrast to the previous findings, those scoring high versus low on autistic traits did not differ either in their behavioural or their affective responsivity, since they were as likely to behave pro-socially towards characters and did not differ in their ratings of the personal sacrifice they experienced for helping others. As discussed in Chapter 7.4, differences between the task designs can probably best account for this discrepancy. On the previous tasks, the characters were always deserving of help, whereas on the 'Social Favours' task the characters were deliberately either more or less deserving of help, providing either a strong or weak justification for requesting help (e.g. carry a package for someone because they have a bad back or have a dirty shirt). The characters described in the scenarios of the 'Social Favours' task also directly requested help from the participants, whereas on the previous tasks this was not always the case. This combination of factors might have made the characters' needs more salient on the 'Social Favours' task than on previous tasks, and thus have inadvertently cued responsivity to their needs.

In summary, the lack of behavioural responsively observed on some but not all of the tasks may depend upon the extent to which the tasks had more salient cues available to support performance; those high in autistic traits might require more direct or deliberate signals to facilitate responsivity to others' needs at a similar level to that of the low autistic trait participants. The role of reliance on task cues and the use of social rules to navigate social performance will be further explored in the next section.

12.4.1.2 Identifying and reasoning about the expectations of pro-social behaviour

In addition to reduced affective responsivity, it is also possible that the reduced behavioural responsivity displayed by those high in autistic traits is underpinned by a more limited understanding of the expectation to help others. The 'Social Expectations' task presented in Chapter 6 directly investigated this by exploring the ability to identify expectations of help in situations with clear-cut versus ambiguous social rules, and also the ability to reason about them. Participants high in autistic traits were able to assess the strength of the social expectation to help that was inherent in the scenarios, both for clear-cut and ambiguous scenarios. However, they showed more limited ability to justify why they should help, relying on more 'rule-based' reasoning which reiterated the scenario (e.g. you should always give your seat to elders) rather than identifying with the character's plight (e.g. she might be in pain/ she needs it more than me). This suggests that the participants high in autistic traits could identify the unwritten rules guiding pro-

social behaviour, but that they had a more limited understanding of these than those low in autistic traits, and that in the absence of internal affective responses they may have relied more heavily upon these external cues to guide pro-social behaviour.

Consistent with idea of intact ability to identify social rules, no group differences were found on any of the measures of the 'Social Favours' task presented in Chapter 7. This examined the deservingness of the characters requesting help, suggesting that those with high levels of autistic traits might have an intact understanding of 'deservingness', or the basic principles guiding when someone is justified in receiving help. There were also no group differences found when participants were asked to estimate how satisfied the characters in need would be with different courses of pro-social behaviour on the 'Above and Beyond' task presented in Chapter 5. This is again consistent with the notion that people with high AQ traits can differentiate between courses of action on various bases.

However, the ability to reason about these scenario aspects (i.e. why characters might be more or less deserving of help, or satisfied with different types of help) was not assessed on either of these tasks, since only rating measures were used. Arguably, ratings assess participants' ability to identify aspects of the scenarios that might draw upon their knowledge of social rules. It would have been interesting to simultaneously explore both people's ability to identify these aspects of the scenarios, and to reason about them on the 'Above and Beyond' and 'Social Favours' tasks. Perhaps, like on the 'Social Expectations' task, a free verbal response measure might have proved a more sensitive measure of performance, and might have revealed poorer reasoning behind the intact ability to identify scenario aspects in the high autistic trait groups. Free verbal responses, such as those elicited on the 'Social Expectations' task, require participants to delve deeper and to justify their ratings. Arguably, this necessitates a more sophisticated understanding of the scenario and requires participants to apply their knowledge about social rules rather than to simply reiterate it.

Taken together it seems likely that those with high AQ traits "know" the social rules, but only at a surface level and that this might result in a failure to apply them appropriately. It also appears that less complex questions or the provision of task cues might mask the presence of more subtle differences between the groups. This is broadly consistent with some previous work indicating relative rather than absolute social deficits in ASD, where high-functioning individuals tend to show proficiency on simple structured tasks, but difficulties with more advanced and naturalistic tasks (Loveland et al., 2001; Channon et al, 2001; Channon et al., 2014).

Are reduced responsivity and a limited understanding of the social rules guiding behaviour confined to pro-social behaviour, or do these deficits also extend to broader aspects of social judgment and reasoning? Both responsivity and the use of rules to guide behaviour were further examined in the next four experimental chapters, which shifted in focus from pro-social behaviour to moral judgments. The

three studies already discussed focused on responding to others' needs by helping them. In these situations pro-social behaviour was always considered the 'right' thing to do, and the tasks were concerned predominantly with examining willingness to help on the basis of various task demands. By contrast, the next four studies focused on moral judgments and reasoning in situations where there were no 'right' or 'wrong' courses of action, but instead there were abstract principles guiding what was 'optimal' or 'fair'.

12.4.2 Moral judgment

Moral judgment refers to the processes by which people differentiate between 'right and wrong' (Fiske, 2004). As reviewed in Chapter 4.1, both cognitive and emotional factors are thought to contribute to moral judgment; thus people are thought to rely on a combination of both deliberate processes and more intuitive and affective influences to make moral judgments and reason about these (Haidt, 2001; Greene & Haidt, 2002). Cognitive aspects of morality refer to knowledge of moral rules, and the ability to apply these and to reason about them under different and novel conditions. Cognitive aspects of morality also encompass more abstract judgments of intentionality, cause and blame, but these are arguably coloured by emotional processes. For instance, you might blame someone for an outcome because they were seen to be causally responsible for it, but the extent to which you blame them might be influenced by the nature of your existing relationship with them. Emotional aspects of morality refer to self-conscious emotions such as the experience of regret, guilt, or shame in response to making a mistake or doing something 'bad'. Emotional aspects might also include the experience of distress at witnessing 'unjust' or 'immoral' circumstances, such as when outcomes are perceived to be unfair or when people are seen to behave in a morally reprehensible manner.

Surprisingly little work has examined moral judgment and reasoning in people with ASD, despite the putative role of empathic emotional processes in mediating this. Some work indicates that people with ASD are able to make 'black and white' moral judgments, such as to correctly differentiate between conventional versus moral transgressions (e.g. Blair, 1996; Leslie et al., 2006), but that differences emerge when more complex aspects of morality are examined (e.g. Channon et al., 2010; 2011). Therefore, it was expected that people with high versus low levels of autistic traits would differ in their performance on the various tasks examined in this thesis, but that they might show preservation of aspects of performance that relied more directly upon straightforward knowledge of what is 'right and wrong'.

In a series of four experiments, both cognitive and emotional contributions to moral judgment and reasoning were explored. The first study assessed judgments of deservingness for positive versus negative outcomes, the next two examined counterfactual reflections upon negative outcomes, and the final study investigated ethical decisions about courses of action that would inevitably lead to negative outcomes. With respect to emotional contributions to moral judgment, these were explored via ratings of self-conscious

emotions and of the experience of distress in response to the various negative events. With respect to cognitive contributions to moral judgment, these were explored via two main modes: firstly, the role of moral rules and principles in guiding behaviour, and secondly, the use of counterfactual thinking to reflect upon bad outcomes. The findings from these four studies will now be summarised with respect to both emotional and cognitive aspects of performance.

12.4.2.3 Emotional responses to moral judgments

Across three of the four tasks, emotional responses to the various negative outcomes were examined via ratings of self-conscious emotions and experienced discomfort. Whilst people with ASD have been found to display no difficulties with processing basic emotions like happiness or anger, some previous work has indicated difficulties with more complex and self-conscious emotions, such as regret or guilt (e.g. Capps et al., 1992). Self-conscious emotions are thought to differ from basic emotions in that they require an awareness of others, whereas basic emotions do not (e.g. Frith & Happé, 1999). On the two tasks exploring counterfactual reflections in response to negative events (the 'RCFT' task and the 'Counterfactual Judgments' task in Chapters 9 and 10), self-conscious emotions of regret and guilt were examined; in line with previous findings, the groups were expected to differ on these measures, with participants scoring high versus low for autistic traits displaying a reduced experience. However, in contrast with predictions, no group differences were identified on these measures.

When presented with ethical dilemmas on the 'Utilitarian Judgments' tasks presented in Chapter 11, participants were required to give ratings of discomfort for following different courses of action, where either one person or a group of individuals suffered physical or social harm. Although discomfort for harming others has not been directly examined in people with ASD before, it was expected that the high autistic trait group would experience less discomfort, on the basis of well-established difficulties with considering others' perspectives. In line with predictions, participants scoring high versus low in autistic traits displayed reduced discomfort for making these difficult decisions.

What does this pattern of findings tell us about the emotional responses of people high in autistic traits when making moral judgments? Initially, it might appear contradictory that participants scoring high versus low in autistic traits experienced less discomfort for making difficult decisions on the 'Utilitarian Judgments' task, but were not found to differ in their ratings of self-conscious emotions on the counterfactual thinking tasks. However, it should be noted that there is a key difference between the present tasks and those used previously that did identify differences in self-conscious emotions in those with ASD. Broadly speaking, previous tasks required participants to recognise emotions from static visual stimuli (i.e. pictures of faces), whereas the present tasks asked participants to estimate their own or another's emotional experience; for a detailed discussion of this please see Chapters 9.4 and 10.4. This

raises the possibility that people with ASD, or at least those with high levels of autistic traits, might experience self-conscious emotions in a typical fashion, although they might have difficulties in recognising these in static stimuli, particularly without any provided context.

There are also a number of differences between the counterfactual tasks that addressed self-conscious emotions, and the 'Utilitarian Judgments' task that addressed discomfort. These will now be considered and the extent to which they could account for the discrepancy in the pattern of findings between these types of tasks explored. The counterfactual tasks examined in Chapters 9 and 10 included negative events that were of a realistic nature, such as missing your graduation ceremony or being inappropriately laughed at within a work context. Whilst these could be considered to be very inconvenient or indeed distressing outcomes, they are fairly mild in terms of the impact of their severity, and their consequences relatively short-term. In contrast, the 'Utilitarian Judgments' task included a mixture of both physical and social negative events, which ranged from very severe and life-altering (e.g. death) to similarly minor and short-term (e.g. being socially excluded) outcomes. Could the more severe nature of the negative events included in the 'Utilitarian Judgments' task account for the group differences identified here? If this were the case, then we would expect to see that the high versus low autistic trait groups differed only for the physical but not the social harm items. In fact, an overall group difference was found, whereby the high versus low autistic trait group experience reduced discomfort across both types of negative events. Thus, it seems unlikely that it was the severity of the negative event driving group differences on the 'Utilitarian Judgments' task.

Is it possible that the counterfactual tasks were less sensitive to group differences in emotional responsivity in some other way? Differences in the salience of the scenario factors seems likely to account best for the mixed pattern of findings. The counterfactual tasks required participants to give ratings of regret in response to overtly 'bad' behaviour that led to a negative event, or to give ratings of guilt in response to thinking about different kinds of counterfactuals, ranging from practical to punitive in nature. It could be argued that for both of these sets of ratings, it was relatively clear that the agent of the negative outcome 'should' experience regret for their 'bad' actions, and guilt for actively wanting to punish another person for their mistakes, but that these emotions might be expected to be weaker when thinking practically about how to avoid the situation. Therefore, the salience of the agent's role as 'in the wrong' might have cued an appropriate response, potentially masking any group differences in emotional responding.

In contrast, the 'Utilitarian Judgments' task described situations where the agent's actions were not morally 'wrong', since participants were evaluating the choice to minimise harm by inflicting pain on another, versus or choosing not to interfere; both of these courses of action could be argued to be 'right', depending upon your moral standpoint. Rather, participants were required to evaluate how uncomfortable

they would feel with either course of action: choosing to sacrifice one for the greater good, or not acting. It could be argued that on the 'Utilitarian Judgments' task it was unclear whether an emotional response should be elicited at all, since the agent was not doing anything 'wrong', but rather managing a difficult ethical decision. If this is the case, then it follows that the high autistic trait participants would not have been 'cued' to consider their emotional experience, and instead would have focused primarily on the logical aspects of each course of action. This could therefore account for the group difference identified on the discomfort ratings on the 'Utilitarian Judgments' task, but not on the self-conscious emotion measures included in the two counterfactual thinking tasks.

In summary, there appears to be some evidence in participants with high versus low autistic traits of reduced affective responsivity to making decisions that negatively impact upon others. It seems likely that the 'Utilitarian Judgments' tasks was more sensitive to revealing group differences, since the events being judged were less clear-cut than those depicted in the 'RCFT' and the 'Counterfactual Judgments' tasks. It is possible that if more sensitive measures were used, such as asking people to describe their emotional responses or to justify them, group differences would have emerged across all three of the tasks. This once again highlights the limitations of rating measures, and the need both to consider the complexity of the demands involved in performance and to probe deeper in order to unmask group differences.

12.4.2.1 Identifying and reasoning about moral rules

Across the four studies included in part B of the experimental chapters, a number of different measures were used to explore both direct and indirect knowledge of moral rules. With respect to direct knowledge of moral rules, the best evidence comes from the two questionnaires that were included in Chapter 11 to assess knowledge of moral transgressions (MBI) and ethical standpoints (EPQ). No group differences were found on either the MBI or the EPQ, suggesting that people with high versus low levels of autistic traits were equally able to identify moral rules and to understand how these should be applied, at least in theory. However, it is important to note that questionnaires do not provide a deep probe of moral knowledge, and thus these types of 'direct' measures might not be sufficiently sensitive to reveal group differences. The ability to apply this knowledge and to reason about it was examined on the 'Deservingness' tasks and the 'Utilitarian' tasks, presented in Chapters 8 and 11.

The 'Deservingness' tasks explored judgments of people's deservingness for positive and negative events. The high autistic trait group was able to correctly differentiate between deserved and undeserved outcomes and give similar ratings to those of the low autistic trait group, both for positive and negative events (e.g. passing your driving test because you drove well or flirted with the instructor; or being reprimanded in class because you did not do your homework or because the instructor was in a bad mood). This extends the findings of the 'Social Favours' task presented in Chapter 7, and taken together these two

studies suggest that people high versus low in autistic traits were equally able to differentiate between scenarios where things were deserved or not, and to respond accordingly. These tasks could be argued to provide a more naturalistic measure of moral knowledge than the questionnaires (MBI or EPQ), which asked participants simply to demonstrate their moral knowledge rather than to draw upon it in real-life-type scenarios. However, both of the questionnaires and the scenario-based tasks included ratings only, and as discussed, ratings might still not go far enough in terms of demand complexity to be sufficiently sensitive to detect group differences. In future work it would be interesting to include free verbal responses, which it appears are more sensitive to revealing group differences.

The 'Utilitarian Judgments' task presented in Chapter 11 also explored putting knowledge of moral rules into practice. Participants were presented with ethical dilemmas and required to choose which party to save. In contrast with the findings of previous study (Gleichgerrcht et al., 2013), both the high and low autistic trait groups behaved similarly in terms of their choices, tending to make more utilitarian decisions. However, when participants were asked to reason about why following either a utilitarian or a non-utilitarian course of action might be appropriate, the groups differed with respect to considering different parties' perspectives. This measure arguably provided a more sensitive measure of moral knowledge, since participants were likely to draw on their knowledge of moral rules to justify these.

Overall, an intact ability to identify but a poorer capacity to reason about moral rules in the high AQ group was revealed in the? present studies. Looking across the range of evidence exploring moral judgment and decision-making, it appears that people with high AQ traits "know" what is wrong or right in an abstract sense. However, they may struggle to appreciate the complexities of such situations; this is a similar pattern of findings to that established in relation to pro-social behaviour.

12.4.2.2 Counterfactual thinking in response to negative events

Counterfactual thinking is an aspect of both social and moral behaviour that refers to mental representations of alternatives, mostly in response to negative events (McNamara et al., 2003) It is believed to play an important role in social learning via reflection upon mistakes, and consideration of how to solve future problems (Epstude & Roese, 2008). Counterfactual thinking is thought to involve both cognitive and emotional aspects of performance, including the ability to inhibit reality, to imagine alternative worlds, and to experience reflective emotions, such as regret. Some limited work has explored counterfactual thinking in people with ASD, finding evidence of impairment, at least under some conditions (e.g. Begeer et al., 2009). Thus, it was expected that participants scoring high versus low for autistic traits would experience difficulty with the counterfactual thinking tasks in this thesis.

The two studies presented in Chapter 9 and 10 explored counterfactual thinking in situations that were primarily concerned with dealing with the consequences of accidental transgressions for self and/or

others. These studies examined both the ability to produce counterfactual thoughts and the nature of counterfactual thoughts. On the 'RCFT' task presented in Chapter 9, people with high levels of autistic traits were found to produce fewer counterfactual thoughts in response to both their own and others' mistakes. This suggests that people with high levels of autistic traits were less fluent in counterfactual thinking. The 'Counterfactual Judgments' task presented in Chapter 10 examined whether the nature of their counterfactual thoughts also differed. Here, both groups showed a greater preference for practical over emotional or punitive counterfactuals, although those with high autistic traits endorsed punitive counterfactuals in relation to others' mistakes slightly more than those with low autistic traits. Taken together, this suggests that people with high levels of autistic traits both might struggle to produce counterfactual thoughts and might experience thoughts of a different quality.

The two counterfactual thinking tasks presented in Chapters 9 and 10 also included ratings of blame for self- and other-actions. The findings suggested that the high AQ group made harsher blame attributions, but only under some conditions. As discussed, blame might draw upon both cognitive and emotional factors and thus differences between groups could stem from impairment in either of these routes. When making attributions of blame people are thought to rely both on causal ascriptions of responsibility and to consider their own affective response (Cushman, 2008). Overall, these two studies suggest that when contemplating or judging negative outcomes, the high autistic trait group was slightly less forgiving of mistakes, with a tendency to blame and to punish people more for their transgressions. Thus, it could be argued that these two studies provide further evidence of altered ability to make moral judgments in people with high levels of autistic traits once they are required to go beyond 'black and white' decisions.

12.4.3 Understanding the influence of autistic traits on everyday social behaviour

The findings from the studies assessing pro-social behaviour and those assessing moral judgment and reasoning will now be brought together in order to address what this body of work has elucidated about the influence of autistic traits for everyday social behaviour. These seven experimental studies assessed situations where participants could choose to intervene in order to prevent or minimise a negative outcome for someone (tasks tapping pro-social behaviour), and also situations where participants had to decide how to deal with negative outcomes that had already occurred or were inevitable (tasks tapping moral judgment and reasoning). Regardless, across these two strands of work there were similarities in the patterns of behaviour of the high versus low autistic trait groups. Two broad aspects of performance were examined which cut across the range of tasks used: affective and cognitive. The conclusions from this body of work will now be summarised accordingly.

12.4.3.1 Affective aspects of performance

Overall, it emerged that people with high versus low levels of autistic traits were less responsive to others' needs. Firstly, on the pro-social tasks used in Chapters 5 and 6 people high in autistic traits were less able to recognise others' needs (lower sympathy ratings), experienced less personal responsivity to others' needs (lower personal reward ratings) and demonstrated less behavioural responsivity (less pro-social behaviour). Secondly, on the 'Utilitarian Judgments' task used in Chapter 11, people high in autistic traits experienced less discomfort at making difficult decisions involving sacrificing others. Although here the high autistic trait groups' responsivity was reduced overall, it should be noted that those high in autistic traits still showed a broadly similar pattern of behaviour on these tasks, and differentiated between the various task conditions in a similar manner. This suggests that there is an overall dampening of affective responsivity in those high versus low in autistic traits under certain conditions, rather than an absolute lack of it.

Furthermore, this pattern of reduced responsivity was not found to be ubiquitous, since no group differences emerged on a number of other tasks. Studies assessing deservingness were presented in Chapters 7 and 8 ('Social Favours' task and the 'Desvingness' tasks for positive and negative outcomes), and no group differences emerged on any of the measures included in these tasks. On the counterfactual tasks examined in Chapters 9 and 10, there were also no group differences identified for ratings of self-conscious emotions in response to 'bad' behaviours. This indicates that whilst there is some evidence for dampened affective responsivity in those with high versus low autistic traits, this is dependent upon the task conditions and/or complexity of the task demand involved. It is possible that these tasks ('Social Favours', 'Deservingness' and the counterfactual tasks) might have been less sensitive to detecting possible differences in performance. It was previously suggested (see Chapter 7.4 and Chapter 8.4) that intact performance on the tasks assessing understanding of 'deservingness' for help ('Social Favours' task), or of 'deservingness' for various outcomes (Deservingness' tasks) might indicate an intact ability of those high in autistic traits to recognise the social rules underpinning 'fairness'. Moreover, on the counterfactual tasks it was suggested that the ratings of self-conscious emotions might have been relatively undemanding, since it was fairly obvious that the agent of the 'bad' behaviour should feel 'regret' or 'guilt' for their actions. Thus, it seems likely that the factors underpinning performance on the 'Social Favours', 'Deservingness' and the counterfactual tasks were more salient than for other tasks included in the thesis, which might have operated to cue appropriate responding in the high autistic trait group. It is possible that group differences might have emerged had these tasks the 'Social Favours', 'Deservingness' and the counterfactual tasks included deeper probes of affective responsivity, such as asking people to describe their own or others' emotional states.

Nonetheless, the reduction in affective responsivity identified across several of the tasks included in this thesis is likely to have important implications for everyday social functioning. It is conceivable that reduced affective responsivity might play a causal role in the atypical judgments and the less sensitive social behaviour observed in the high versus low autistic trait groups. It has been postulated that empathic processes play a motivating role in pro-social behaviour, whereby people are more likely to help others if they feel good about themselves for doing so, or can alleviate their own experience of vicarious personal distress (e.g. Eisenberg et al., 1989; Eisenberg, 2007). Thus, if the high autistic trait group experienced less reward for helping others they might have been less motivated to help others, which might be reflected in their reduced pro-social behaviour on the 'Above and Beyond' and 'Social Expectations' tasks. Conversely, if the high autistic trait group experienced less discomfort at witnessing others in pain, they might have been less likely to behave in a manner driven by emotional factors and more likely to behave in a rational manner which deliberately considered the pros and cons of each possible course of action. In support of this, although no differences between group were identified in utilitarian choices on the 'Utilitarian Judgments' task, the high autistic trait group displayed altered decision-making, since the groups differed in how they weighed the involved parties' perspectives.

Overall, it appears that to enhance social sensitivity in those with high levels of autistic traits it might be necessary to raise external signals which support people's decisions to help others, in order to compensate for a lack of internal signals. The clinical implications of this will be further explored in section 12.6.3.

12.4.3.2 Cognitive aspects of performance

A second key finding concerns differences between the groups in their understanding of the rules that guide behaviour across both pro-social and moral situations. It appears that people with high levels of autistic traits are equally able to identify the social rules underpinning the expectation to behave in a certain way, but that they display a more limited understanding of these rules when asked to explain them. This is in keeping with previous literature suggesting that people with ASD 'know' what is 'right and wrong', but that their ability to reason about it is poorer (Grant et al., 2005; Zalla et al., 2011). If their understanding of social rules is more limited, and also is not integrated with intuitive processes that motivate socially sensitive behaviour, then simply 'knowing' the rules might not be sufficient. Thus, limited social knowledge could account for the less pro-social behaviour and poorer reasoning observed in situations where the rules underpinning behaviour were not clear (i.e. the 'Above and Beyond' and 'Social Expectations' task presented in Chapters 5 and 6). An intact ability to identify rules, but a poorer understanding of these, could also account for the high autistic trait group's intact 'black and white' moral judgments demonstrated, but difficulty with more complex moral judgments and/or the need to justify these.

A reduced ability to engage in counterfactual thinking, as demonstrated on the 'RCFT' task presented in Chapter 9, might lead to more limited social understanding in everyday life. Counterfactual thinking is thought to be important for learning from our mistakes, both in order to reflect upon the implications and to avoid future problems (REDS). It is thought to depend upon similar processes to those involved in false belief tasks, and thus to draw upon the same general skills as used in cognitive empathy. Whilst little work has explored counterfactual thinking in relation to the social performance of those with ASD, there is some evidence to suggest that counterfactual thinking is more limited in this group (e.g. Begeer et al., 2009). If people high in autistic traits have a more limited ability to engage in counterfactual thinking, then they might have had developed a less thorough understanding of the different situations presented in the tasks, and have failed to acquire a range of sophisticated social problem solving skills that can be flexibly deployed.

In summary, two key findings emerged with respect to the influence of autistic traits on everyday social behaviour. Firstly, people high in autistic traits appeared to be less responsive to others' needs, although this was not found to be universal. Secondly, they were found to show an intact knowledge of social and moral rules on direct measures of their knowledge, but a poorer ability to explain or justify these on more indirect measures of social or moral knowledge. Intact knowledge of rules might prove to be both a "blessing and a curse". People high in autistic traits might rely more heavily on social rules to guide behaviour, but fail to fully appreciate these or to integrate them with context specific sources of information, resulting in less socially sensitive behaviour. These two findings are not mutually exclusive, but rather demonstrate two sides of the same story: a greater reliance upon social rules is likely to occur in the absence of sufficient intuitive internal signals to guide behaviour.

This body of work has highlighted that it is essential to consider both the quality and the nature of social behaviour, not just its presence or absence. It has also emphasised the difficulty of dissociating cognitive from affective aspects of social behaviour in real-life-type tasks, and that in everyday life these two processes are likely to be used in concert despite being theoretically dissociable.

12.5 Links to psychological models of ASD

Psychological theories of ASD are useful in conceptualising the disorder and in drawing links between observed abnormal behaviours and postulated cognitive impairments. As discussed in Chapter 2.2, several major psychological theories have been proposed to account for the symptoms associated with ASD, some of which are more focused on explaining the non-social features, such as the theory of executive dysfunction (e.g. Hill, 2004), and some of which can better account for the social features of ASD, such as impaired cognitive empathy (e.g. Baron-Cohen et al., 1985; Baron-Cohen, 1995).

12.5.1 Summary of findings from executive and empathic measures

In this thesis, measures of both executive functioning and of empathic processing were included in several of the studies (Chapter 7 and 8; Chapter 10 and 11). Two measures of executive function that assessed inhibitory processing and verbal fluency were used, since these were seen as particularly important for performance on the social scenario-based tasks explored in this thesis. Both inhibition of inappropriate or insensitive behaviour and the ability to generate new ideas are likely to play a role in considering the best course of action and in formulating an appropriate response.

Analysis of these measures revealed that as expected, both of the high trait groups performed worse than their low autistic trait group counterparts for tests of inhibitory processing, involving naming the colour ink of a printed conflicting colour word (e.g. RED printed in green ink). For verbal fluency, group differences were revealed on only one of the two samples tested with this measure; on the sample included in Chapter 10 and 11 the high autistic trait group performed worse for letter but not category fluency, but no differences were detected between groups for the sample included in Chapter 7 and 8. It is important to note that the pattern of findings was somewhat patchy, and that the differences identified indicated a relative weakness rather than a clinically significant level of impairment, since the high autistic trait groups still performed within the average range.

With respect to empathic processes, two different measures were used in order to tap various aspects of empathy, since this is a multidimensional construct (e.g. Davis, 1983; Thoma & Bellebaum, 2012). Firstly, the MIT, a scenario-based test of cognitive empathy requiring interpretation of others' mental states via non-literal language or actions, was included. As expected, this revealed group differences, whereby the high autistic trait group performed worse than the low autistic trait group. However, again the findings were somewhat patchy; the high versus low autistic trait groups included in Chapters 7 and 8 differed in their interpretation of non-literal actions but not language, whereas the high versus low autistic trait groups included in Chapter 10 and 11 displayed the opposing pattern. This suggests that the differences detected between groups might have been subtle, and thus the findings were relatively unstable across different samples.

Secondly, the IRI was administered, which is a self-report questionnaire with four subscales assessing both cognitive and emotional empathic processes. Here, those high autistic traits were expected to perform worse for subscales tapping cognitive empathy, but not necessarily to differ on those tapping emotional empathy. In fact, a much more complex pattern of results was found, with the high versus low autistic trait groups differing for both cognitive and emotional empathy subscales, but with different subscales differentiating the samples; the high versus low autistic trait groups included in Chapters 7 and 8 differed for 'Fantasy' (cognitive) and 'Empathic Concern' (emotional) only, but the high versus low autistic trait groups included in Chapters 10 and 11 differed for 'Perspective-taking' (cognitive) and 'Personal

distress' (emotional) only. This again is likely to reflect the subtle nature of high versus low autistic traits group differences and the instability of such differences across small sample sizes.

12.5.2 The role of executive functioning and empathic processing in social performance

The mild differences identified on measures of executive functioning and empathic processing are consistent with the notion that those with high levels of autistic traits show deficits similar in nature to those seen in people with a diagnosis of ASD, although perhaps to a lesser extent (e.g. Best et al., 2008; Freeth et al., 2013). It should be noted that these data only pertain to the participants included in four of the seven experimental studies presented in this thesis (Chapters 7 and 8, and Chapter 10 and 11). However, given the similarities between the samples examined across all of the studies it seems reasonable to assume that a broadly similar pattern of differences between groups on both the executive and the empathic measures would apply to the high versus low autistic trait participants included in the three remaining studies (i.e. Chapter 5, 6 and 9).

It also seems likely that the findings of reduced affective responsivity and more rule-based behaviour in the high autistic trait group might stem from poorer executive and/or empathic processes, at least in part. With respect to executive functioning, poorer inhibitory control might have led to more impulsive responding and a lack of insight into the other characters' needs. This might be reflected in the reduced behavioural and affective responsivity displayed by the high autistic trait groups across a number of tasks. Reduced verbal fluency could also account for the more simplistic rule-based responses on free-verbal measures, and the reduced counterfactual abilities of the high autistic trait groups.

With respect to empathic processes, a reduced ability to appreciate the mental states of others via cognitive empathy could also account for the findings; reduced affective responsivity might reflect a lack of understanding of the characters' predicaments and their own emotional states. A greater reliance on social or moral rules to guide behaviour might reflect the use of compensatory mechanisms in lieu of more intuitive processes. It is more difficult to see how emotional empathic processes could account for the tendency towards rule-based responses, but a deficit in the ability to resonate with and share the emotional states of others could directly account for the high autistic trait groups' reduced affective responsivity; although this seems less likely in light of evidence indicating impaired cognitive but intact emotional empathy.

It is important to note that both executive and empathic processes are likely to have played a role in task performance, and thus their contributions should not be considered mutually exclusive. Indeed, some work suggests that ASD is likely to be the result of multiple cognitive deficits and that no single theory can explain the full range of associated impairments (Happé et al., 2006). The role of these different processes cannot be directly attributed to task performance, and it is also difficult to separate their

contributions to task performance. However, considering the role of executive and empathic processes provides a useful framework for interpreting the findings on the present studies.

12.5.3 Alternative accounts

It is also likely that compensatory social knowledge is used in everyday social performance (Channon et al., 2001). It is possible that this might interact with cognitive processes, such as executive functioning or empathic processing, and be relied upon to guide social behaviour to different extents by different people (see Chapter 6.4 for a detailed discussion of this issue).

The contribution of alternative theoretical accounts also cannot be ruled out, even though there is no direct data available. For instance, the theory of 'weak central coherence' (e.g. Happé, 1996) postulates that those with ASD might pay more attention to the details at the expense of the global picture. This is predominantly thought to have implications for non-social behaviour, but it might also contribute to less sophisticated social performance, via a lack of consideration of the contextual details. However, it is unclear how this could account for the particular pattern of results with respect to the finding of reduced affective and behavioural responsivity and more rule-based behaviour displayed by the high autistic trait group.

One further factor that was not been directly addressed in the present thesis is the possible role of co-occuring alexithymia. As discussed in Chapter 2.3.2, alexithymia is a sub-clinical condition characterised by difficulties with identifying and describing emotions of the self (Nemiah et al., 1976) and corresponding impairments in social cognition (e.g. Parker et al., 1993). These features may overlap with the social and communication deficits seen in ASD, which are also known to frequently co-occur with alexithymia (Bird & Cook, 2013). Some work has explored the extent to which co-occuring alexithymia can help to account for some of the emotional features of ASD, which are not necessarily seen in all individuals; co-occuring alexithymia in people with ASD has been found to predict emotional awareness (Silani et al., 2008), the extent of arousal experienced by witnessing another in pain (Bird et al., 2010), and also the ability to recognise emotional from both facial and vocal stimuli (Cook, Brewer, Shah & Bird, 2013; Heaton et al., 2012). It thus seems possible that the presence of alexithymia in a large percentage of the ASD population, up to 50%, (Hill, Berthoz & Frith, 2004; Berthoz & Hill, 2005) could account for the discrepant findings regarding whether people with ASD experience difficulties with processing emotions, at least to some extent (e.g. Blair, 2008). One study also suggested that those scoring higher on the AQ also tended to score higher on a self-report measure of alexithymia, the Toronto Alexithymia Scale (TAS-20; Lockwood et al., 2013).

With respect to the present body of work, co-occuring alexithymia could therefore potentially account for some of the differences observed between the high and low autistic trait groups, in particular

for the reduced affective responsivity seen in the high autistic trait groups across a number of tasks. More generally, the presence or absence of alexithymia in those with high versus low levels of autistic traits might be an important factor in explaining some of the inconsistencies seen across different samples of high scorers. In order to directly assess this, participants would need to be screened with both the AQ and the TAS-20. If a group comparison design was adopted, as with the present body of work, then four different groups of individuals would be required for each study; one group of participants scoring high for the AQ only, one group of participants scoring high for the TAS-20 only, one group of participants scoring high for both measures, and finally one group of participants scoring low for both measures. Another approach would be to use a correlational design that assesses people scoring high and/or low for both measures across the full range of scores.

The bigger question is whether alexithymia can account for differences beyond those on measures of emotional functioning, and hence what implications it might have for understanding ASD. For instance, if some but not all individuals with ASD have deficits in both emotional and cognitive aspects of empathy, this double deficit may be linked to greater impairment on other aspects of social cognition, rather than confined to emotional measures. However, there is little evidence examining this to date, and any such implications are merely speculative. Moreover, it is difficult to see how a central deficit in emotional functioning could explain any performance deficits on non-social tasks.

12.6 Clinical implications

12.6.1 Background and the relevance of the present work

The poorer social skills of people with ASD are thought to lead to dysfunction in a range of areas, including difficulties with emotional, social, academic and occupational functioning (e.g. Church, Alisanski, & Amanullah, 2000; Gillot & Standen, 2007). As discussed in Chapter 3.5, ASD is a neurodevelopmental disorder that is diagnosable from the early years of infancy and is thought to be present across the life-course. However, owing to the heterogeneous nature of the disorder and the differing profiles of individuals on the spectrum, the time of identification can vary widely and might be considerably later for some people. This is particularly true for higher-functioning people whose symptoms and difficulties might not become apparent or problematic until relatively late in life, when faced with increasingly complex environments and meeting the additional demands of adulthood (Howlin, 2006; Howlin, Goode, Hutton & Rutter, 2004).

Standardised diagnostic tools for ASD rely on either parental report or on play- or observation-based assessments. These tools can prove inappropriate or insensitive to detecting the difficulties of older

adolescents or adults, particularly those who are higher-functioning. For instance, parents might not be available or may struggle to recall sufficient information about the individual's early years. On the other hand, play or observation-based assessments which are pitched at children and younger adolescents might appear patronising to older and/or higher-functioning people on the spectrum, or might be too simplistic to detect the more subtle difficulties displayed by such individuals. Thus, assessing ASD in older adolescents and adults is a particularly complex challenge and there is a need for additional tools that are appropriate for direct assessments of an adult's social and cognitive abilities (Seltzer et al., 2004).

Higher-functioning adults with ASD possess average or above average intellectual capacities, allowing them to engage with more demanding work and social environments than those who are lower-functioning. However, they might also lack the social skills required to navigate these more complex social situations, such as working as part of a team or understanding the nuances of different types of relationships. Despite the difficulties of higher-functioning persons in managing independent life, most social skill programmes developed focus on children and/or younger adolescents with ASD, and thus are often unsuitable for adults, in particular higher-functioning persons (Howlin & Lounds-Taylor, 2015; Channon et al., 2012). These programmes tend to focus on training cognitive skills, such as taking other people's perspectives (e.g., Fisher & Happé, 2005; Turner-Brown, Perry, Dichter, Bodfish, & Penn, 2008), or on teaching basic social skills, such as eye contact (Rao et al., 2008). Whilst these have proven somewhat effective within a laboratory context, any gains seen do not typically generalise to real-life situations, as people might struggle to apply the targeted skills or behaviours to novel contexts (Howlin & Yates, 1999; Hadwin, Baron-Cohen, Howlin & Hill, 1997; Locke, Rotheram-Fuller, Xie, Harker & Mandell, 2014).

In summary, there is a need to improve both diagnostic and intervention tools for those with ASD to ensure they are suitable for the full range of people affected across the life-course. The present thesis primarily focused on extended understanding of clinical symptomatology for everyday functioning within the sub-clinical range of the autistic spectrum. This involved developing a range of tools to assess different aspects of everyday social functioning, and generated a number of insights regarding areas of perseveration and impairment. Whilst this body of work did not involve testing people with a diagnosis of ASD, the findings might have implications for informing the assessment and management of ASD in high-functioning people with clinical levels of impairment. A number of possible implications and suggestions for clinical practice will now be outlined.

12.6.2 Additional diagnostic tools

Real-life-type tasks such as those described in the thesis might prove to be useful tools for identifying both impaired and preserved social abilities in those with ASD, with respect to both the affective and cognitive aspects. They provide a means of profiling individuals' strengths and weaknesses, by

using a combination of response measures that probe both responsivity to others, and the ability to identify and use social knowledge. They could thus represent potentially useful tools for use in clinical settings, where they could contribute both to the identification and assessment of people's specific deficits. It would be important to view these types of tools as complementary to existing tools, as their role would be to supplement rather than to replace these, in order to gain a better picture of higher-functioning individuals' difficulties. It would remain essential to gain as clear a developmental picture as possible, but these could provide additional age- and ability-appropriate tools for assessing individuals' presentation.

The tools developed and used in the present thesis have also highlighted the importance of assessing the quality of people's social performance, and not simply the presence or absence of specific behaviours. This principle should be incorporated into clinical practice by ensuring that assessments are sensitive to teasing apart these different aspects of performance, and probe deep enough to reveal subtler difficulties in higher-functioning and/or older people.

12.6.3 Transforming interventions

The social scenario-based tasks presented in this thesis and the insights they have garnered regarding the influence of autistic traits for everyday behaviour might also help to guide the development of more successful intervention strategies. A number of findings have emerged from the studies that might have implications for interventions: i) using external cues to guide social behaviour; ii) targeting the discrepancy between identifying and reasoning about social rules; iii) enhancing social motivation. These will now be explored in detail below.

12.6.3.1 Using external cues to guide social performance

As discussed throughout this thesis, it appeared that the high autistic trait groups might have used compensatory strategies to successfully complete some aspects of task performance, relying more heavily upon external cues to guide their performance in the light of their reduced experience of internal signals that also motivate appropriate social behaviour. Whilst any compensatory strategy usage was not sufficient to fully disguise their difficulties, the high autistic trait groups performed similarly to those with low levels of autistic traits when the task demands were relatively simple, and/or when direct knowledge of social or moral rules was sufficient. This is consistent with previous work indicating that individuals with high-functioning ASD may rely on compensatory strategies (e.g. Hill & Frith 2003), such as the application of learned social rules to alleviate empathic deficits. Thus, the findings from the present thesis provided some tentative evidence that the provision of external cues might enhance the social performance of people in broader autistic phenotype, and possibly those with ASD.

Some previous work has examined the role of external cues in supporting the social performance of people with ASD, such as focusing on picking upon signals provided by others, like gestures (e.g. McClannahan & Krantz, 1999). However, this has exclusively concentrated on child and adolescent populations, and often the lower-functioning end of the spectrum. The tasks presented in this thesis could help to bridge this gap by using a similar approach to explore more subtle and sophisticated behaviours that are appropriate for supporting higher-functioning or older adolescents and adults. For instance, on the 'Above and Beyond' or 'Social Expectations' tasks, the scenarios could be adapted to ensure that the plights of the characters were made as salient as possible, in order to cue pro-social behaviour. Materials such as these could be used to hone the use of compensatory strategies, in combination with understanding of what others may be thinking or feeling.

Whilst there is some evidence that training people with ASD to notice others' signals is effective at improving social performance (e.g. Mclannahan & Krantz, 1999; Alberto & Troutman, 2005), the present findings suggest that people with high levels of autistic traits are naturally less responsive to such signals, and thus this approach might prove challenging. Instead, the present findings suggest that situational external cues might provide a useful alternative target for supporting social learning in those with ASD. Future interventions for people with ASD could focus on making the rules governing social situations as concrete as possible, ideally in real-life settings. For instance, when travelling on public transport, their attention could be drawn to signs indicating that priority should be given to elderly, pregnant or disabled persons.

12.6.3.2 Targeting the discrepancy between identifying and reasoning about social rules

The findings from the present thesis have also highlighted the presence of a discrepancy between the ability to identify social and moral rules, versus the ability to reason about these or apply them. Firstly, people with high versus low levels of autistic traits were equally able to identify both social and moral rules or norms, but provided less sophisticated or detailed reasoning when asked to explain these across a number of tasks. Secondly, people with high versus low levels of autistic traits were also less likely to act in accordance with social and moral rules, even when they could identify the expectation to behave in a certain way inherent in the situations presented. This pattern of findings is consistent with previous work in clinical populations, indicating that people diagnosed with high-functioning ASD "know" the unwritten rules, but only at a surface level (e.g. Zalla et al., 2011). It appears that direct knowledge of social and moral rules is not sufficient, and that a more sophisticated understanding of these is necessary for successful social performance. Therefore, whilst using cues to raise the external signals that are readily identified by

those with ASD could provide a helpful strategy as discussed, this alone is unlikely to be sufficient for facilitating successful social performance. Thus, a complementary approach that seeks to help people learn about the principles behind such rules might also be necessary.

In a clinical setting, the present tasks could be refined to identify such discrepancies in higher-functioning people and to guide intervention strategies. Improving both people's awareness of rules and also their understanding of these rules, such as possible caveats, might support the generation of more flexible and socially appropriate responses, especially in complex and unpredictable social situations (Channon et al., 2010; Howlin et al., 2004). Thus, social skill training could focus on expanding a person's awareness of the principles behind unwritten social and moral rules, and the contingencies for their application. With respect to pro-social behaviour, a person with ASD might identify that it is 'right' to help someone in need, but fail to appreciate the longer-term implications of this, or struggle to recognise that helping behaviours might be less appropriate or more important in different situations. With respect to moral dilemmas, a person with ASD might understand that "it is wrong to hit someone". However, they might fail to fully comprehend the reasons behind why such rules should be observed, or that such rules are guiding principles rather than entirely black and white. For instance, a person with ASD might fail to appreciate that there are certain occasions where it is appropriate to breach the rule, such as in situations involving self-defence.

Social skill training could therefore highlight the different aspects of such rules by developing a set of guiding principles to use in addition to directly cueing knowledge of such rules. This would seek to help people with ASD to appropriately judge a specific situation and to apply knowledge of social and moral rules more appropriately and flexibly.

12.6.3.3 Enhancing social motivation

That a lack of social motivation may be a central characteristic of ASD is a fairly recent concept, whereby those with ASD are thought to find social interactions less intrinsically rewarding than neurotypical individuals (Chevallier, Kohls, Troiaini, Brodkin & Schultz, 2012). The present findings provide some support for a lack of experience of social reward in the broader autistic phenotype. Firstly, those high versus low in autistic traits were found to experience less personal reward for helping others on the 'Above and Beyond'. Secondly, and more generally, they were found to exhibit reduced affective responsivity for others' needs across tasks. Thus, it seems possible that those with high levels of autistic traits might find social interactions inherently less rewarding, which might underpin their less pro-social behaviour and possibly other aspects of poorer social functioning.

Whilst this is an area that has as yet not received much attention, it seems possible that enhancing the motivation derived from social situations might be useful for guiding and improving current intervention strategies to support people with a diagnosis of ASD. This could potentially be achieved by making the personal gains for engaging in socially appropriate behaviour more salient. With respect to pro-social behaviour, treatment approaches might focus on minimising the personal costs involved in helping others, by making it less effortful to help others or by highlighting how easy it might be to provide help. Alternatively, emphasising the personal gains to be obtained from helping others might provide a more direct way of motivating pro-social behaviour. As well as the obvious benefits for others, a range of personal benefits might also be achieved by behaving pro-socially, including the opportunity for reciprocal help in the future, or a chance to enhance their personal reputation in the eyes of others. Such benefits could be explicitly spelt out to people with ASD in social skill training, and they could then be trained to apply these principles via the use of situational cues and/or rules of thumb. This approach could also potentially be applied to children and young adolescents with ASD, by using clear rewards for behaving in a pro-social manner towards siblings or friends (e.g. being rewarded additional time playing if they share their toys with others), and also setting clear disadvantages for refusing to co-operate with others (e.g. being docked time playing if they refuse to let another child have a go on the computer).

The implications of enhancing social rewards for moral judgment and reasoning are less clear-cut. However, the overall principles of highlighting the personal gains for behaving in a pro-social manner to others could also be used to target other inappropriate social behaviours. For instance, the high autistic trait groups displayed a greater tendency to blame others and to want to punish them for their mistakes on the counterfactual thinking tasks used in Chapters 9 and 10. This is also consistent with previous work indicating harsher judgments of blame in those with ASD (Channon et al., 2011). Social skill training could seek to target and potentially reduce such 'punitive' behaviours by highlighting that being more forgiving towards others for their mistakes might result in kinder personal treatment in the future, and that it will encourage others to have a favourable opinion of them.

It currently remains unclear whether any reward deficits are specific to social stimuli or reflect a general reward processing deficit in ASD (e.g. Dichter et al., 2012). It might be interesting for future work to study how people with ASD process social stimuli and rewards in the context of everyday situations, and the potential presence of differences in the extent to which they seek and take pleasure from different aspects of social versus non-social sources of reward. This would also be interesting to examine in sub-clinical populations, such as the high autistic trait groups explored in the present study.

The novel experimental tasks developed for each experiment, and the insights gained from this body of work, both have potential clinical implications for the assessment and for the management of

ASD. Firstly, the scenario-based tasks presented might prove to be useful additional tools for the assessment of ASD in higher-functioning and/or older people. These might complement the use of current diagnostic tools, which are mostly currently geared towards younger persons and assess more simplistic behaviours. Secondly, the tasks developed for this thesis might provide a useful basis for the improvement of current social skill training tools. These might be used to teach those with ASD to identify and to learn to reason about social and moral rules that guide behaviour, and to apply these more flexibly according to a range of different contexts. Finally, a number of themes emerged from the findings of the seven experiments presented, which might inform the approach of both the assessment and the management of ASD, particularly for higher-functioning and/or older people. With respect to the assessment of ASD, the present findings have highlighted the importance of assessing the quality and nature of social behaviours, and not just considering their presence or absence. With respect to the management of ASD, the present findings also have three clear implications for social skill training, including the use of external or situational cues to guide performance, targeting the discrepancy between identifying and reasoning and social rules, and the importance of enhancing the motivation to act in a socially sensitive manner.

12.7 Concluding comments

The present thesis aimed to investigate how autistic traits influenced everyday social functioning. A range of novel scenario-based tasks were developed that examined specific aspects of social behaviour, particularly pro-social behaviour and moral judgment and reasoning. Overall, the findings showed that people with high levels of autistic traits tended to be less behaviourally and emotionally responsive to others' needs, in similar fashion to those with ASD, consistent with the conceptualisation of a continuum of trait severity. The findings also suggested that people with high levels of autistic traits showed a relatively intact awareness of the largely unwritten social and moral rules that underpin everyday situations, but that their understanding of these appeared to be more limited. Broadly speaking, the findings highlighted that such scenario-based tasks are more sensitive than traditional social cognition tasks in identifying the nature and severity of impairments in everyday social interactions and functioning. These studies are the first to explore complex social behaviour in people with high versus low levels of autistic traits, and have expanded understanding of the nature of their deficits in everyday functioning. The body of work also has potential clinical implications for the assessment and management of higher-functioning and or older people with ASD.

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.

Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., Happé. F., Frith, C. & Frith, U. (1999). The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans. *Neuroreport*, *10*(8), 1647-1651.

Alberto, P. A. & Troutman, A. C. (2005). *Applied Behavior Analysis for teachers* (7th ed.). Upper Saddle River, NJ: Merrill/Prentice Hall.

Alexander, A. L., Lee, J. E., Lazar, M., Boudos, R., DuBray, M. B., Oakes, T. R., Miller. J., Lu, J., Jeong, E-K., McMahon, W., Bigler, D. & Lainhart, J. E. (2007). Diffusion tensor imaging of the corpus callosum in Autism. *Neuroimage*, *34*(1), 61-73.

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126*, 556.

Alicke, M. D., Mandel, D. R., Hilton, D. J., Gerstenberg, T. & Lagnado, D. A. (2015). Causal conceptions in social explanation and moral evaluation: a historical tour. *Perspectives on Psychological Science. 106*(6), 790-812.

Allik, H., Larsson, J. O., & Smedje, H. (2006). Health-related quality of life in parents of school-age children with Asperger Syndrome or high-functioning autism. *Health and Quality of Life Outcomes*, *4*(1), 1-8.

Almeida, R. A., Dickinson, J. E., Maybery, M. T., Badcock, J. C., & Badcock, D. R. (2012). Visual search targeting either local or global perceptual processes differs as a function of autistic-like traits in the typically developing population. *Journal of Autism and Developmental Disorders, 43*(6), 1–15.

American Psychiatry Association (2000). Diagnostic and Statistical Manual of Mental Disorders. (4th Ed.). DSM-IV TR (Text Revision). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). Diagnostic and Statistical Manual of Mental Disorders, (5th Ed.): Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). Ausitsm Spectrum Disorder Factsheet.

Anagnostou, E., Soorya, L., Brian, J., Dupuis, A., Mankad, D., Smile, S., & Jacob, S. (2014). Intranasal oxytocin in the treatment of autism spectrum disorders: A review of literature and early safety and efficacy data in youth. *Brain Research, 1580*, 188-198.

Andari, E., Duhamel, J. R., Zalla, T., Herbrecht, E., Leboyer, M., & Sirigu, A. (2010). Promoting social behavior with oxytocin in high-functioning autism spectrum disorders. *Proceedings of the National Academy of Sciences*, *107*(9), 4389-4394.

Anderson H.H. & Anderson G.L. (1951) *An introduction to projective techniques and other devices for understanding human behaviour*. NJ, USA: Prentice Hall, Englewood Cliffs.

Anderson, S. W., Barrash, J., Bechara, A. & Tranel, D. (2006). Impairments of emotion and real-world complex behavior following childhood- or adult-onset damage to ventromedial prefrontal cortex. *Journal of the International Neuropsychological Society, 12(*2), 224-35.

Anderson, S. W., Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience, 2*(11),1032-37.

Anscombe, E. (1958). "Modern Moral Philosophy". *Philosophy. 33*(24).

Auyeung, B., Baron-Cohen, S., Chapman, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Foetal testosterone and the child systemizing quotient. *European Journal of Endocrinology*, *155*(suppl 1), S123-S130.

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

Bacon, A. L., Fein, D., Morris, R., Waterhouse, L., & Allen, D. (1998). The responses of autistic children to the distress of others. *Journal of Autism and Developmental Disorders, 28*, 129-142.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology: Section A, 49(1),* 5-28.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological medicine*, *25*(01), 63-77.

Barberis, C., & Tribollet, E. (1996). Vasopressin and oxytocin receptors in the central nervous system. *Critical Reviews in Neurobiology, 10*, 119–154.

Barnes, J., Lombardo, M., Wheelwright, S. & Baron-Cohen, S. (2009). Moral dilemmas film task: a study of spontaneous narratives by individuals with autism spectrum conditions. *Autism Research, 2*(3), 148-156.

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

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in cognitive sciences*, *6*(6), 248-254.

Baron Cohen, S. (2008). *Autism and Asperger syndrome.* Oxford: Oxford University Press.

Baron-Cohen, S. (2009). Autism: the empathizing–systemizing (E-S) theory. *Annuals of the New York Academy of Sciences*, *1156*(1), 68-80.

Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T., & Chakrabarti, B. (2009). Talent in autism; hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philosophical Transactions of the Royal Society, 364*, 1377-1383.

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: implications for explaining autism. *Science*, *310*(5749), 819-823.

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., O'Riordan, M., Jones, R., Stone, V., & Plaisted, K. (1999). A new test of social sensitivity: Detection of faux pas in normal children and children with Asperger syndrome. *Journal of Autism and Developmental Disorders*, *29*(5), 407-418.

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

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

Bartels, D. M. & Pizarro, D. A. (2011). The mismeasure of morals: antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition, 121,* 154-161.

Bartz, J. A., Zaki, J., Bolger, N., & Ochsner, K. N. (2011). Social effects of oxytocin in humans: context and person matter. *Trends in Cognitive Sciences*, *15*(7), 301-309.

Batson, C. D. (1991). The altruism question. Hillsdale, NJ: Erlbaum.

Batson, C. (1998), Prosocial behavior and altruism. In: Lindzey G, Gilbert D, Fiske ST, editors. The Handbook of Social Psychology, Oxford University Press, 282– 316.

Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and empathy: Two qualitatively distinct vicarious emotions with different motivational consequences. *Journal of Personality*, *55*(1), 19-39.

Bayliss, A. P., & Kritikos, A. (2011). Brief report: Perceptual load and the autism spectrum in typically developed individuals. *Journal of Autism and Developmental Disorders, 41*(11), 1573–1578.

Beauchamp M. H., Anderson V. (2010). SOCIAL: an integrative framework for the development of social skills. *Psychological Bulltein, 136*, 39-64.

Beck, S. R., & Guthrie, C. (2011). Almost thinking counterfactually: Children's understanding of close counterfactuals. *Child Development*, *82*(4), 1189-1198.

Beck, S. R., Riggs, K. J., & Gorniak, S. L. (2009). Relating developments in children's counterfactual thinking and executive functions. *Thinking & Reasoning*, *15*(4), 337-354.

Begeer, S., Malle, B. F., Nieuwland, M. S., & Keysar, B. (2010). Using theory of mind to represent and take part in social interactions: Comparing individuals with high-functioning autism and typically developing controls. *European Journal of Developmental Psychology, 7(1),* 104-122.

Begeer, S., De Rosnay, M., Lunenburg, P., Stegge, H., & Terwogt, M. M. (2014). Understanding of emotions based on counterfactual reasoning in children with autism spectrum disorders. *Autism*, *18*(3), 301-310.

Begeer, S., Terwogt, M. M., Lunenburg, P., & Stegge, H. (2009). Brief report: additive and subtractive counterfactual reasoning of children with high-functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *39*(11), 1593-1597.

Beldarrain, M. G., Garcia-Monco, J. C., Astigarraga, E., Gonzalez, A., & Grafman, J. (2005). Only spontaneous counterfactual thinking is impaired in patients with prefrontal cortex lesions. *Cognitive Brain Research*, *24*(3), 723-726.

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2003). Motion perception in autism: a "complex" issue. *Journal of Cognitive Neuroscience, 15*(2), 218-225.

Berthoz, S. & Hill, E. L. (2005). Reliability of the Bermond-Vorst Alexithymia Questionnaire: data from adults with autism spectrum disorder, their relatives and normal controls. *European Journal of Psychiatry, 20,* 291–8

Berthoz, S., Armony, J. L., Blair, R. J. R. & Dolan, R. J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain, 125,* 1696–708.

Best, C. S., Moffat, V. J., Power, M. J., Owens, D. G., & Johnstone, E. C. (2008). The boundaries of the cognitive phenotype of autism: Theory of mind, central coherence and ambiguous figure perception in young people with autistic traits. *Journal of Autism and Developmental Disorders*, *38*(5), 840-847.

Bickman, L. (1972). Social influence and diffusion of responsibility in an emergency. *Journal of Experimental Social Psychology, 8*(5), 438-445.

Bird, G., & Cook, R. (2013). Mixed emotions: The contribution of alexithymia in the emotional symptoms of autism. *Translational Psychiatry, 3*.

Bird, G., Silani, G., Brindley, R., White, S., Frith, U., & Singer, T. (2010). Empathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, awq060.

Bishop-Fitzpatrick, L., Minshew, N. J., & Eack, S. M. (2014). A systematic review of psychosocial interventions for adults with autism spectrum disorders. In *Adolescents and adults with autism spectrum disorders* (pp. 315-327). Springer New York.

Blair, R. J. R. (1995). A cognitive developmental approach to morality: investigating the psychopath. *Cognition, 57,* 1-29.

Blair, R. J. R. (1996). Brief report: Morality in the autistic child. *Journal of Autism and Developmental Disorders*, *26*(5), 571-579.

Blair, R. J. R. (2008). Fine cuts of empathy and the amygdala: Dissociable deficits in psychopathy and autism. *The Quarterly Journal of Experimental Psychology, 61* (1), 157-170.

Blair, R. J. R. Jones, L., Clark, F., & Smith, M. (1997). The psychopathic individual: a lack of responsiveness to distress cues? *Psychophysiology*, *34*(2), 192-198.

Blakemore, S. J. (2008). The social brain in adolescence. *Nature Reviews Neuroscience*, *9*(4), 267-277.

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

Boninger, D. S., Gleicher, F., & Strathman, A. (1994). Counterfactual thinking: From what might have been to what may be. *Journal of Personality and Social Psychology, 67*(2), 297-307.

Booth, R., Charlton, R., Hughes, C., & Happé, F. (2003). Disentangling weak coherence and executive dysfunction: planning drawing in autism and attention–deficit/hyperactivity disorder. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1430), 387-392.

Bowler, D. M. (1992). "Theory of mind" in Asperger's syndrome. *Journal of Child Psychology and Psychiatry,33*(5), 877–893.

Boucher, J. (1988). Word fluency in high-functioning autistic children. *Journal of Autism and Developmental Disorders, 18*, 637–645.

Bunge, S. A. (2004). How we use rules to select actions: A review of evidence from cognitive neuroscience. *Cognitive, Affective, & Behavioral Neuroscience, 4*(4), 564–579.

Buon, M., Dupoux, E., Jacob, P., Chaste, P., Leboyer, M., & Zalla, T. (2013). The role of causal and intentional judgments in moral reasoning in individuals with high functioning autism. *Journal of Autism and Developmental Disorders*, *43*(2), 458-470.

Burgess, P. W. & Shallice, T. (1997). *The Hayling and Brixton Tests. Test Manual.* Bury St Edmunds: Thames Valley Test Company.

Burnett, H. G., & Jellema, T. (2013). (Re-) conceptualisation in Asperger's Syndrome and typical individuals with varying degrees of autistic-like traits. *Journal of Autism and Developmental Disorders, 43*(1), 211–223.

Branscombe, N. R., Wohl, M. J., Owen, S., Allison, J. A., & N'gbala, A. (2003). Counterfactual thinking, blame assignment, and well-being in rape victims. *Basic and Applied Social Psychology*, *25*(4), 265-273.

Brewer, R., Marsh, A. A., Catmur, C., Cardinale, E. M., Stoycos, S., Cook, R., & Bird, G. (2015). The impact of autism spectrum disorder and alexithymia on judgments of moral acceptability. *Journal of Abnormal Psychology*, *124*(3), 589.

Brief, A. P., & Motowidlo, S. J. (1986). Pro-social organizational behaviours. *Academy of Management Review, 11*(4), 710–725.

Callenmark, B., Kjellin, L., Ronnqvist, L. & Bolte, S. (2014). Explicit versus implicit social cognition testing in autism spectrum disorder. *Autism 18*, 684– 693.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J. R., & Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science*, *304*(5674), 1167-1170.

Capps, L., Kehres, J., & Sigman, M. (1998). Conversational abilities among children with autism and children with developmental delays. *Autism, 2* (4), 325-344.

Capps, L., Yirmiya, N., & Sigman, M. (1992). Understanding of simple and complex emotions in non-retarded children with autism. *Journal of Child Psychology and Psychiatry, 33*(7), 1169–1182.

Carrington, S., Templeton, E., & Papinczak, T. (2003). Adolescents with Asperger syndrome and perceptions of friendship. *Focus on Autism and Other Developmental Disabilities, 18(4),* 211-218.

Carlo, G., Allen, J. B., Buhman, D. C. (1999). Facilitating and disinhibiting prosocial behaviors: The nonlinear interaction of trait perspective taking and trait personal distress on volunteering. *Basic & Applied Social Psychology, 21,* 189–197.

Castelli, F., Frith, C., Happe, F., Frith, U. (2002). Autism, Asperger Syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain,* 125, 1839-49.

Cassidy, S., Ropar, D., Mitchell, P., & Chapman, P. (2014). Can adults with autism spectrum disorders infer what happened to someone from their emotional response? *Autism Research, 7(1),* 112-123.

Channon, S., Charmon, T., Heap, J., & Crawford, S. & Rios, P. (2001). Real life type problem solving in Asperger's syndrome. *Journal of Autism and Developmental Disorders, 31,* 461-469.

Channon, S., Collins, R., Swain, E., Young, M. B., & Fitzpatrick, S. (2012). The use of skilled strategies in social interactions by groups high and low in self-reported social skill. *Journal of Autism and Developmental Disorders, 42(7)*, 1425-1434.

Channon, S., & Crawford, S. (1999). Problem-solving in real-life-type situations: The effects of anterior and posterior lesions on performance. *Neuropsychologia, 37 (7)*, 757-770.

Channon, S., & Crawford, S. (2010). Mentalising and social problem-solving after brain injury. *Neuropsychological Rehabilitation, 20(5)*, 739-759.

Channon, S., Crawford, S., Orlowska, D., Parikh, N., & Thoma, P. (2014). Mentalising and social problem solving in adults with Asperger's syndrome. *Cognitive Neuropsychiatry*, *19*(2), 149-163.

Channon, S., Fitzpatrick, S., Drury, H., Taylor, I., & Lagnado, D. (2010). Punishment and Sympathy Judgments: Is the Quality of Mercy Strained in Asperger's Syndrome? *Journal of Autism and Developmental Disorders*, *40*(10), 1219-1226.

Channon, S., Lagnado, D., Fitzpatrick, S., Drury, H., & Taylor, I. (2011). Judgments of cause and blame: Sensitivity to intentionality in Asperger's Syndrome. *Journal of Autism and Developmental Disorders*, *41*(11), 1534-1542.

Channon, S. Pellijeff, A. & Rule, A. (2005). Social cognition after head injury: Sarcasm and theory of mind. *Brain and Language, 93(2),* 123-134.

Channon, S., Rule, A., Maudgil, D., Martinos, M., Pellijeff, A., Frankl, J. & Shieff, C. (2007). Interpretation of mentalistic actions and sarcastic remarks: Effects of frontal and posterior lesions on mentalising. *Neuropsychologia*,*45*(8), 1725-1734.

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: evidence from the empathy quotient (EQ) and the "reading the mind in the eyes" test. *Social Neuroscience, 1*(2), 135-148.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231-239.

Christ, S. E., Kanne, S. M., & Reiersen, A. M. (2010). Executive function in individuals with subthreshold autism traits. *Neuropsychology*, *24*(5), 590.

Christensen, J. F., & Gomila, A. (2012). Moral dilemmas in cognitive neuroscience of moral decision-making: a principled review. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1249-1264.

Church, C., Alisanski, S., & Amanullah, S. (2000). The social, behavioral, and academic experiences of children with Asperger syndrome. *Focus on Autism and other Developmental Disabilities*, *15*(1), 12-20.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Cohen, I. L. (1994). Artificial neural network analogue of learning in autism. *Biological Psychiatry, 36*, 5– 20.

Cohen, D. J., Young, J. G., Lowe, T. L., & Harcherik, D. (1980). Thyroid hormone in autistic children. *Journal of Autism and Developmental Disorders*,*10*(4), 445-450.

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M. (1997). Social ties and susceptibility to the common cold. *Jama*, *277*(24), 1940-1944.

Coie, J. D., Dodge, K. A., & Kupersmidt, J. B. (1990). Peer group behavior and social status. In S. R. Asher & J. D. Coie (Eds.), Peer rejection in childhood (pp. 17–59). New York: Cambridge University Press.

Coke, J. S., Batson, C. D., & McDavis, K. (1978). Empathic mediation of helping: A two-stage model. *Journal of Personality and Social Psychology*, *36*(7), 752.

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L. & Reich, W. (2003). Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders*, *33*(4), 427-433.

Constantino, J. N. (2011) The quantitative nature of autistic social impairment. *Paediatric Reviews, 69*, 55–62.

Cook, E. H. (1998). Genetics of autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *4*(2), 113-120.

Cook, R., Brewer, R., Shah, P., & Bird, G. (2013). Alexithymia, not autism, predicts poor recognition of emotional facial expressions. *Psychological Science*, 0956797612463582.

Cooper, N. R., Simpson, A., Till, A., Simmons, K., & Puzzo, I. (2013). Beta event-related desynchronization as an index of individual differences in processing human facial expression: Further investigations of autistic traits in typically developing adults. *Frontiers in Human Neuroscience, 7*, 159.

Coricelli, G., & Rustichini, A. (2010). Counterfactual thinking and emotions: regret and envy learning. *Philosophical Transactions of the Royal society B: Biological sciences*, *365*(1538), 241-247.

Crozier, S., & Tincani, M. (2007). Effects of social stories on prosocial behavior of preschool children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 37*(9), 1803–1814.

Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, *108*(2), 353-380.

Darley, J. M. & Latane, B. (1968). Bystander intervention in emergencies: diffusion of responsibility. *Journal of Personality and Social Psychology, 8,* 377-383.

Dawson, G. & Fernald, M. (1987). Perspective taking ability and its relationship to the social behaviour of autistic children. *Journal of Autism and Developmental Disorders, 17,* 487-498.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10,* 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality & Social Psychology, 44,* 113–126.

D'Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology*, *27*(2), 152.

Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, *41*(2), 127-138.

Decety, J., & Michalska, K. J. (2010). Neurodevelopmental changes in the circuits underlying empathy and sympathy from childhood to adulthood. *Developmental Science*, *13*(6), 886-899.

Dennett, D. C., (1978). Beliefs about beliefs. *Behavioural and Brain Sciences, 1,* 568-570.

de Gelder, B., Vroomen, J., & van der Heide, L. (1991). Face recognition and lip reading in Autism. *European Journal of Cognitive Psychology, 3*, 69-86.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. Psychological Corporation.

De Martino, B., Harrison, N. A., Knafo, S., Bird, G., & Dolan, R. J. (2008). Explaining enhanced logical consistency during decision making in autism. *Journal of Neuroscience, 28,* 10746-10750.

Department of Health. (2010). Fulfilling and rewarding lives: the strategy for adults with autism in England.

Department of Health. (2014). Think Autism Fulfilling and Rewarding Lives, the strategy for adults with autism in England: an update.

Dewey, M. (1991). Living with Asperger's syndrome. In U. Frith (Ed), *Autism and Asperger Syndrome* (pp. 184-206). New York, NY, US: Cambridge University Press.

Dichter, G. S., Felder, J. N., Green, S. R., Rittenberg, A. M., Sasson, N. J., & Bodfish, J. W. (2012). Reward circuitry function in autism spectrum disorders. *Social Cognitive and Affective Neuroscience*, *7*(2), 160-172.

Dolan, M. & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine, 34,* 1093-1102.

Domes, G., Heinrichs, M., Kumbier, E., Grossmann, A., Hauenstein, K., & Herpertz, S. C. (2013). Effects of intranasal oxytocin on the neural basis of face processing in autism spectrum disorder. *Biological Psychiatry*, *74*(3), 164-171.

Drayton, S., Turley-Ames, K. J., & Guajardo, N. R. (2011). Counterfactual thinking and false belief: The role of executive function. *Journal of Experimental Child Psychology*, *108*(3), 532-548.

Dziobek, I., Rogers, K., Fleck, S., Bahnemann, M., Heekeren, H. R., Wolf, O. T., & Convit, A. (2008). Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *Journal of Autism and Developmental Disorders*, *38*(3), 464-473.

Ehlers, S. & Gillberg, C. (1993). The Epidemiology of Asperger syndrome. A total population study. *Journal of Child Psychology and Psychiatry*, *34* (8), pp. 1327-1350

Eigisti, I-M. & Shapiro, T. (2003). A systems neuroscience approach to autism: biological, cognitive and clinical perspectives. *Mental Retardation and Developmental Disabilities Research Reviews, 9,* 205–215.

Eisenberg, N. (2007). Empathy-related responding and prosocial behaviour. *Novartis Foundation Symposium, 278*, 71–80.

Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., & Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behaviour: a multimethod study. *Journal of Personality and Social Psychology*, *57*(1), 55.

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (1998). Prosocial development. *Handbook of Child Psychology*.

Eisenberg, N., Guthrie, I. K., Murphy, B. C., Shepard, S. A., Cumberland, A., & Carlo, G. (1999). Consistency and development of prosocial dispositions: A longitudinal study. *Child Development*, 1360-1372.

Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviours. *Psychological Bulletin*, *101*(1), 91.

Eisenberg, N., McCreath, H., & Ahn, R. (1988). Vicarious Emotional Responsiveness and Prosocial Behavior Their Interrelations in Young Children. *Personality and Social Psychology Bulletin*, *14*(2), 298-311.

Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcín, C. et al. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research*, *5*(3), 160-179.

Eslinger, P. J., Grattan, L. M., Damasio, H., Damasio, A. R. (1992). Developmental consequences of childhood frontal lobe damage. *Archives of Neurology, 49*(7),764-9.

Epstude, K., & Roese, N. J. (2008). The functional theory of counterfactual thinking. *Personality and Social Psychology Review*, *12*(2), 168-192.

Eycke, K. D., & Müller, U. (2015). Brief report: new evidence for a social-specific imagination deficit in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(1), 213-220.

Falk, A., Fehr, E., & Fischbacher, U. (2008). Testing theories of fairness—Intentions matter. *Games and Economic Behavior*, *62*(1), 287-303.

Feather, N. T. (1999). Judgments of deservingness: Studies in the psychology of justice and achievement. *Personality and Social Psychology Review*, *3*(2), 86-107.

Feather, N.T. (2006). Deservingness and emotions: applying the structural model of deservingness to the analysis of affective reactions to outcomes. *European Review of Social Psychology, 17,* 38-73.

Feather, N.T. & Sherman, R. (2002). Envy, resentment, *schadenfreude* and sympathy: reactions to deserved and undeserved achievement and subsequent failure. *Personality and Social Psychology Bulletin, 28*(7), 953-961.

Fecteau, S., Pascual-Leone, A., & Théoret, H. (2008). Psychopathy and the mirror neuron system: preliminary findings from a non-psychiatric sample. *Psychiatry Research*, *160*(2), 137-144.

Fernell, E., Eriksson, M. A., & Gillberg, C. (2013). Early diagnosis of autism and impact on prognosis: a narrative review. *Clinical Epidemiology*, *5*, 33.

Fisher, N. & Happé, F. (2005). A training study of theory of mind and executive function in children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *35*(6), 757-771.

Fletcher, P. C. & Henson, R. N. A. (2001). Frontal lobes and human memory. *Brain, 124(5),* 849- 881.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research*, *65*(6), 591-598.

Foot, P. (1967). The problem of abortion and the doctrine of the double effect in virtues and vices. *Oxford Review, 5,* 5-15.

Forsyth, D.R. (1980). A taxonomy of ethical ideologies. *Journal of Personal & Social Psychology, 39,* 175-184.

Forsyth, D. R., Nye, J. L., & Kelley, K. (1988). Idealism, relativism, and the ethic of caring. *The Journal of Psychology*, *122*(3), 243-248.

Foxton, J. M., Stewart, M. E., Barnard, L., Rodgers, J., Young, A. H., O'Brien, G., & Griffiths, T. D. (2003). Absence of auditory 'global interference' in autism. *Brain*, *126*(12), 2703-2709.

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

Frey, B. S., & Meier, S. (2004). Social comparisons and pro-social behavior: Testing" conditional cooperation" in a field experiment. *American Economic Review*, 1717-1722.

Frith, U. (1989). *Autism and "Theory of Mind"*. In C. Gillberg (Ed.), Diagnosis and Treatment of Autism. (pp. 33-52). New York: Plenum Press.

Frith, U. (2003). *Autism, explaining the enigma* (2nd ed.). Oxford: Blackwell.

Frith, U., & de Vignemont, F. (2005). Egocentrism, allocentrism, and Asperger syndrome. *Consciousness and Cognition, 14*(1), 719–738.

Frith, U., & Happé, F. (1994). Autism: Beyond "theory of mind". *Cognition*, *50*(1), 115-132.

Frith, U. & Happé, F. (1999). Theory of mind and self-consciousness: what is it like to be autistic? *Mind & Language, 14(1),* 82-89.

Frith, U., Happe, F, & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development, 2,* 208 – 124.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, *10*(4), 483-527.

Furnham, A. (1998). *Measuring the beliefs in a just world.* In Montada, L. and Lerner, M. J. (eds.), Responses to Victimization and Belief in a Just World, Plenum, New York, 141-162.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*, 493–501.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences, 8*, 396–403.

Gamliel, I., Yirmiya, N., Jaffe, D. H., Manor, O., & Sigman, M. (2009). Developmental Trajectories in Siblings of Children with Autism: Cognition and Language from 4 Months to 7 Years. *Journal of Autism and Developmental Disorders*, *39*(8), 1131–1144.

García-Villamisar, D. A., & Dattilo, J. (2010). Effects of a leisure programme on quality of life and stress of individuals with ASD. *Journal of Intellectual Disability Research*, *54*(7), 611-619.

García-Villamisar, D., & Hughes, C. (2007). Supported employment improves cognitive performance in adults with autism. *Journal of Intellectual Disability Research*, *51*(2), 142-150.

Gardener, H., Spiegelman, D., & Buka, S. L. (2009). Prenatal risk factors for autism: comprehensive meta-analysis. *The British Journal of Psychiatry*, *195*(1), 7-14.

Gardener, H., Spiegelman, D., & Buka, S. L. (2011). Perinatal and neonatal risk factors for autism: a comprehensive meta-analysis. *Pediatrics*, *128*(2), 344-355.

Gavanski, I., & Wells, G. L. (1989). Counterfactual processing of normal and exceptional events. *Journal of Experimental Social Psychology*, *25*(4), 314-325.

Geschwind, D. H. (2011). Genetics of Autism Spectrum Disorders. *Trends in Cognitive Sciences*, *15*(9), 409–416.

Ghaziuddin, M. (2010). Brief report: Should the DSM V drop Asperger syndrome? *Journal of Autism and Developmental Disorders, 40*(9), 1146-1148.

Ghaziuddin, M. (1997). Autism in Down's syndrome: family history correlates. *Journal of Intellectual Disability Research*, *41*(1), 87-91.

Gillberg C. (1991). Outcome in autism and autistic-like conditions. *Journal of American Academy of Child and Adolescent Psychiatry, 30,* 375-82.

Gillott, A. & Standen, P. J. (2007). Levels of anxiety and sources of stress in adults with autism. *Journal of Intellectual Disabilities*, *11*(4), 359-370.

Gleichgerrcht, E., Torralva, T., Rattazzi, A., Marenco, V., Roca, M., & Manes, F. (2013). Selective impairment of cognitive empathy for moral judgment in adults with high functioning autism. *Social Cognitive and Affective Neuroscience*, *8*(7), 780-788.

Gleichgerrcht E., Torralva, T., Roca, M., Pose, M. & Manes, F. (2010). The role of social cognition in moral judgment in frontotemporal dementia. *Social Neuroscience, 6*(2), 113-22.

Gökçen, E., Petrides, K. V., Hudry, K., Frederickson, N., & Smillie, L. D. (2014). Sub-threshold autism traits: The role of trait emotional intelligence and cognitive flexibility. *British Journal of Psychology, 105*(2), 187–199.

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

Gould, J. & Ashton-Smith, J. (2011). Missed diagnosis or misdiagnosis: girls and women on the autism spectrum. *Good Autism Practice, 12* (1), pp. 34-41

Grant, C. M., Boucher, J., Riggs, K. J., & Grayson, A. (2005). Moral understanding in children with autism. *Autism*, *9*(3), 317-331.

Grant, A. M., & Gino, F. (2010). A little thanks goes a long way: Explaining why gratitude expressions motivate prosocial behavior. *Journal of Personality and Social Psychology*, *98*(6), 946.

Grant, C. M., Riggs, K. J., & Boucher, J. (2004). Counterfactual and mental state reasoning in children with autism. *Journal of Autism and Developmental Disorders*, *34*(2), 177-188.

Gray, C. A. (1998). *Social stories and comic strip conversations with students with Asperger syndrome and high-functioning autism,* In: Current Issues in Autism (pp. 167-198). Springer US.

Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work?. *Trends in Cognitive Sciences*, *6*(12), 517-523.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*(3), 1144-1154.

Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389-400.

Greene, J. D., Somerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science, 293,* 2105-2108.

Gregory, S. G., Connelly, J. J., Towers, A. J., Johnson, J., Biscocho, D., Markunas, C. A., Lintas, C., Abramson, R. K., Wright, H. H., Ellis. P., Langford, C. F., Worley, G., Delong, G. R., Murphy, S. K., Cuccaro, M. L., Persico A., Pericak-Vance, M. A. (2009). Genomic and epigenetic evidence for oxytocin receptor deficiency in autism. *BMC Medicine, 7*, 62.

Grinter, E. J., Maybery, M. T., Van Beek, P. L., Pellicano, E., Badcock, J. C., & Badcock, D. R. (2009). Global visual processing and self-rated autistic-like traits. *Journal of Autism and Developmental Disorders, 39*(9), 1278–1290

Guastella, A. J., Einfeld, S. L., Gray, K. M., Rinehart, N. J., Tonge, B. J., Lambert, T. J., & Hickie, I. B. (2010). Intranasal oxytocin improves emotion recognition for youth with autism spectrum disorders. *Biological Psychiatry*, *67*(7), 692-694.

Geurts, H. M., Verté, S., Oosterlaan, J., Roeyers, H., & Sergeant, J. A. (2004). How specific are executive functioning deficits in attention deficit hyperactivity disorder and autism?. *Journal of Child Psychology and Psychiatry*, *45*(4), 836-854.

Hadwin, J., Baron-Cohen, S., Howlin, P., & Hill, K. (1997). Does teaching theory of mind have an effect on the ability to develop conversation in children with autism? Jo*urnal of Autism and Developmental Disorders*, *27*(5), 519-537.

Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814.

Haidt, J., Bjorklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Unpublished manuscript*, University of Virginia.

Hale, C. M. & Tager-Flusberg, H. (2005). Social communication in children with autism: The relationship between theory of mind and discourse development. *Autism, 9* (2), 157 – 178.

Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., Messinger, D. & Szatmari, P. (2015). Sex and gender differences in autism spectrum disorder: summarizing evidence gaps and identifying emerging areas of priority. *Molecular Autism*, *6*, 36.

Hamilton, A. F., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience, 26*, 1133–1137.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Development Disorders, 24*, 129-154.

Happé, F. G. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development*, *66*(3), 843-855.

Happé, F. G. (1996). Studying weak central coherence at low levels: children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry*, *37*(7), 873-877.

Happé, F. (1999). Autism: Cognitive deficit or cognitive style? *Trends in Cognitive Sciences, 3* (6), 216-222.

Happé, F., Booth, R., Charlton, R., & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition, 61* (1), 25-39.

Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., Dolan, R., Frackowiak & Frith, C. (1996). 'Theory of mind'in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport*, *8*(1), 197-201.

Happé, F. & Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36 (1), 5-25.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience, 9* (10), 1218 – 1220.

Hare, R. D., Hart, S. D., & Harpur, T. J. (1991). Psychopathy and the DSM-IV criteria for antisocial personality disorder. *Journal of Abnormal Psychology*, *100*(3), 391.

Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., & Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind & Language*, *22*(1), 1-21.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual: Revised and expanded.* Odessa, FL: Psychological Assessment Resources.

Heaton, P., Reichenbacher, L., Sauter, D., Allen, R., Scott, S., & Hill, E. (2012). Measuring the effects of alexithymia on perception of emotional vocalizations in autistic spectrum disorder and typical development. *Psychological Medicine*, *42*(11), 2453-2459.

Heavey, L., Phillips, W., Baron-Cohen, S., Rutter, M. (2000). The Awkward Moments Test: A Naturalistic Measure of Social Understanding in Autism. *Journal of Autism and Developmental Disorders, 30,* 3.

Heerey, E. A., Keltner, D., & Capps, L. M. (2003). Making sense of self-conscious emotion: linking theory of mind and emotion in children with autism. *Emotion*, *3*(4), 394.

Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, *8*(1), 26-32.

Hill, E., Berthoz, S., & Frith, U. (2004). Brief report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. *Journal of Autism and Developmental Disorders, 34*(2), 229–235.

Hill, E. L. & Bird, C. M. (2006). Executive processes in Asperger syndrome: Patterns of performance in a multiple case series. *Neuropsychologia, 44(14),* 2822-2835.

Hill, E. L. & Frith, U. (2003). Understanding autism: insights from mind and brain. *Philosophical transactions of the Royal Society B: Biological sciences, 358(1430),* 281-289.

Hiller, A., & Allinson, L. (2002). Understanding embarrassment among those with autism: Breaking down the complex emotion of embarrassment among those with autism, *Journal of Autism and Developmental Disorders, 32,* (6), 583-592.

Hodges, J. R., Patterson, K., Ward, R., Garrard, P., Bak, T., Perry, R., & Gregory, C. (1999). The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal variants of frontotemporal dementia) from early Alzheimer's disease: a comparative neuropsychological study. *Neuropsychology*, *13*(1), 31.

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders*, *38*(8), 1555-1566.

Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice.* Cambridge, England: Cambridge University Press.

Hoffman, M. L. (2001). Empathy and moral development: Implications for caring and justice. Cambridge: Cambridge University Press.

Horder, J., Wilson, C. E., Mendez, M. A., & Murphy, D. G. (2014). Autistic traits and abnormal sensory experiences in adults. *Journal of Autism and Developmental Disorders*, *44*(6), 1461-1469.

Howard, M. A., Cowell, P. E., Boucher, J., Broks, P., Mayes, A., Farrant, A., & Roberts, N. (2000). Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. *Neuroreport*, *11*(13), 2931-2935.

Howlin, P. (2000). Outcome in adult life for more able individuals with autism or Asperger syndrome. *Autism*, *4*(1), 63-83.

Howlin, P., Goode, S., Hutton, J., & Rutter, M. (2004). Adult outcome for children with autism. *Journal of Child Psychology and Psychiatry*, *45*(2), 212-229.

Howlin, P., & Lounds-Taylor, J. L. (2015). Addressing the need for high quality research on autism in adulthood. *Autism, 19*(7), 771-773.

Howlin, P., & Yates, P. (1999). The potential effectiveness of social skills groups for adults with autism. *Autism*, *3*(3), 299-307.

Hudson, M., Nijboer, T. C., & Jellema, T. (2012). Implicit social learning in relation to autistic-like traits. *Journal of Autism and Developmental Disorders, 42*(12), 2534–2545.

Hughes, R. & Huby, M. (2002). The application of vignettes in social and nursing research. *Journal of Advanced Nursing*, *37*(4), 382-386.

Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*(4), 590-619.

Hughes, C. Leboyer, M., Bouvard, M. (1997) Executive function in parents of children with autism. *Psychological Medicine, 27,* 209–220.

Hughes, C., Plumet, M-H, Leboyer, M. (1999). Towards a cognitive phenotype for autism: increased prevalence of executive dysfunction and superior spatial span amongst siblings of children with autism. *Journal of Child Psychology and Psychiatry, 40,* 705–71.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia, 32,* 477–492.

Hunt A., & Shepherd C. (1993). A prevalence study of autism in tuberous sclerosis. *Journal of Autism and Developmental Disorders, 23*, 323–339.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience, 7*(12), 942–951.

Ingudomnukul, E., Baron-Cohen, S., Wheelwright, S., & Knickmeyer, R. (2007). Elevated rates of testosterone-related disorders in women with autism spectrum conditions. *Hormones and Behavior*, *51*(5), 597-604.

Izuma, K., Matsumoto, K., Camerer, C., & Adolphs, R. (2011). Insensitivity to social reputation in autism. *Proceedings of the National Academy of Sciences, 108*(42), 17302–17307.

Jacob, S., Brune, C. W., Carter, C. S., Leventhal, B. L., Lord, C., & Cook, E. H. (2007). Association of the oxytocin receptor gene (OXTR) in Caucasian children and adolescents with autism. *Neuroscience Letters*, *417*(1), 6-9.

Jameel, L., Vyas, K., Bellesi, G., Roberts, V., & Channon, S. (2014). Going 'Above and Beyond': Are those high in autistic traits less pro-social?. *Journal of Autism and Developmental Disorders*, *44*(8), 1846-1858.

Jameel, L., Vyas, K., Bellesi, G., Cassell, D., & Channon, S. (2015). Great Expectations: The Role of Rules in Guiding Pro-social Behaviour in Groups with High Versus Low Autistic Traits. *Journal of Autism and Developmental Disorders*, *45*(8), 2311-2322.

Jankowski, K. F., & Takahashi, H. (2014). Cognitive neuroscience of social emotions and implications for psychopathology: examining embarrassment, guilt, envy, and schadenfreude. *Psychiatry and Clinical Neurosciences*, *68*(5), 319-336.

Joliffe, T., & Baron-Cohen, S. (1999). A test of central coherence theory: linguistic processing in high-functioning adults with autism or Asperger syndrome: is local coherence impaired? *Cognition, 71* (2), 149-185.

Jones, R. S., Zahl, A., & Huws, J. C. (2001). First-hand accounts of emotional experiences in autism: A qualitative analysis. *Disability & Society*, *16*(3), 393-401.

Jorde, L. B., Hasstedt, S. J., Ritvo, E. R., Mason-Brothers, A., Freeman, B. J., Pingree, C., Mo, A. (1991). Complex segregation analysis of autism. *American Journal of Human Genetics, 49*(5), 932-938.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review*, *93*(2), 136.

Kaland, N. (2011). Brief report: Should Asperger syndrome be excluded from the forthcoming DSM-V? *Research in Autism Spectrum Disorders*, *5*(3), 984-989.

Kasari, C., Chamberlain, B., & Bauminger, N. (2001). Social emotions and social relationships in autism: can children with autism compensate? In J. Burack, T. Charman, N. Yirmiya, & P. Zelazo (Eds.), Development and autism: Perspectives from theory and research. Hillsdale, NJ: Erlbaum Press.

Klin, A., Lang, J., Cicchetti, D.V., Volmar, F.R. (2000). Inter-rater Reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorder, 30*(2), 163-167.

Klin, A., & Volkmar, F. R. (2000). Treatment and intervention guidelines for individuals with Asperger syndrome. *Asperger Syndrome*, 340-366.

Knobe, J. (2003). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, *16*(2), 309-324.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, *446*(7138), 908-911.

Kohlberg, L. (1958). *The development of modes of moral thinking and choice in the years 10 to 16*. University of Chicago.

Kohlberg, L. (1971). "Stages in moral development as a basis for moral education." In C.M. Beck, B.S. Crittenden, and E.V. Sullivan, Eds. *Moral Education: Interdisciplinary Approaches*. Toronto: Toronto University Press.

Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*(3), 754-770.

Lagnado, D. A., Gerstenberg, T., & Zultan, R. I. (2013). Causal responsibility and counterfactuals. *Cognitive Science*, *37*(6), 1036-1073.

Lai, M. C., Lombardo, M. V., Pasco, G., Ruigrok, A. N., Wheelwright, S. J., Sadek, S. A., Chakrabarti, B. & Baron-Cohen, S. (2011). A behavioral comparison of male and female adults with high functioning autism spectrum conditions. *PloSOne*,*6*(6), e20835.

Lai, M. C., Lombardo, M. V., Suckling, J., Ruigrok, A. N., Chakrabarti, B., Ecker, C., Deoni, S. C. L. (2013). Biological sex affects the neurobiology of utism. *Brain*, *136*(9), 2799-2815.

Lainhart, J. E., Piven, J., Wzorek, M., Landa, R., Santangelo, S. L., Coon, H., & Folstein, S. E. (1997). Macrocephaly in children and adults with autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(2), 282-290.

Lam, K. S., Aman, M. G., & Arnold, L. E. (2006). Neurochemical correlates of autistic disorder: a review of the literature. *Research in developmental disabilities*, *27*(3), 254-289.

Latané, B., & Darley, J. M. (1970a). The unresponsive bystander: Why doesn't he help?

Latané, B., & Darley, J. M. (1970b). Social determinants of bystander intervention in emergencies. *Altruism and Helping Behavior*, 13-27.

Lawson, J. Baron-Cohen, S. & Wheelwright, S. (2004). Empathising and systemising in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders, 34* (3), 301-310.

Leaf, J. B., Taubman, M., Bloomfield, S., Palos-Rafuse, L., Leaf, R., McEachin, J., & Oppenheim, M. L. (2009). Increasing social skills and pro-social behavior for three children diagnosed with autism through the use of a teaching package. *Research in Autism Spectrum Disorders, 3*(1), 275–289.

Leslie, A. M., Mallon, R., & DiCorcia, J. A. (2006). Transgressors, victims, and cry babies: Is basic moral judgment spared in autism? *Social Neuroscience*, *1*(3-4), 270-283.

Levine, M., Crowther, S. (2008). "The Responsive Bystander: How Social Group Membership and Group Size Can Encourage as Well as Inhibit Bystander Intervention." *Journal of Personality and Social Psychology, 95*(6): 1429–1439.

Levine, M., Prosser, A., Evans, D., & Reicher, S. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin*, *31*(4), 443-453.

Lilienfeld, S.O. & Fowler, K.A. (2006). The self-report assessment of psychopathy: Problems, pitfalls and promises. In Patrick, C.J. (Ed.), *Handbook of Psychopathy*. Guildford Press.

Lin, A., Tsai, K., Rangel, A., & Adolphs, R. (2012). Reduced social preferences in autism: Evidence from charitable donations. *Journal of Neurodevelopmental Disorders, 4,* 8.

Liu, X., Kawashima, M., Miyagawa, T., Otowa, T., Latt, K. Z., Thiri, M., Nishida, H & Sasaki, T. (2015). Novel rare variations of the oxytocin receptor (OXTR) gene in autism spectrum disorder individuals. *Human Genome Variation*, *2*.

Locke, J., Rotheram-Fuller, E., Xie, M., Harker, C., & Mandell, D. (2014). Correlation of cognitive and social outcomes among children with autism spectrum disorder in a randomized trial of behavioral intervention. *Autism*, *18*(4), 370-375.

Lockwood, P. L., Bird, G., Bridge, M., & Viding, E. (2013). Dissecting empathy: high levels of psychopathic and autistic traits are characterized by difficulties in different social information processing domains. *Frontiers in Human Neuroscience*, *7:*760.

Long, A. A. & Sedley, D. N. (1987). The Hellenistic Philsophers: Transaction of the Principal Sources with Philosophical Commentary. *Cambridge: Cambridge University Press.* 366–367.

Lopes, P. N., Salovey, P., Côté, S., Beers, M., & Petty, R. E. (2005). Emotion regulation abilities and the quality of social interaction. *Emotion*, *5*(1), 113.

Lord, C., Rutter, M. & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24* (5), 659-685.

Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., & Piven, J. (2009). The neuropsychological profile of autism and the broad autism phenotype. *Archives of General Psychiatry*, *66*(5), 518–526.

Lord, C., Petkova, E., Hus, V., Gan, W., Lu, F., Martin, D. M. et al. (2012). A multisite study of the clinical diagnosis of different Autism Spectrum Disorders. *Archives of General Psychiatry*, *69*(3), 306-313.

Lounds-Taylor, J., Dove, D., Veenstra-Vander-Weele, J., Sathe, N. A., McPheeters, M. L., Jerome, R. N., & Warren, Z. (2012). Interventions for adolescents and young adults with autism spectrum disorders. *Comparative Effectiveness Reviews, No. 65*

Loveland, K. (2009). Toward an ecological theory of autism. In J. A. Burack, T. Charman, N. Yirmiya, & P. R. Zelazo (Eds.), The development of autism: Perspectives from theory and research. Hillsdale, NJ: Erlbaum.

Loveland, K. A., Pearson, D. A., Tunali-Kotoski, B., Ortegon, J., & Gibbs, M. C. (2001). Judgments of social appropriateness by children and adolescents with autism. *Journal of Autism and Developmental disorders*, *31*(4), 367-376.

Low, J., Goddard, E., & Melser, J. (2009). Generativity and imagination in autism spectrum disorder: Evidence from individual differences in children's impossible entity drawings. *British Journal of Developmental Psychology*,*27*(2), 425-444.

Lupfer, M.B & Gingrich, B.E. (1999). When bad (good) things happen to good (bad) people: the impact of character appraisal and perceived controllability on judgments of deservingness. *Social Justice Research, 12(3),* 165-188.

Lykken, D. T. (1957). A study of anxiety in the sociopathic personality. *The Journal of Abnormal and Social Psychology*, *55*(1), 6.

Macrae, C. N., & Milne, A. B. (1992). A curry for your thoughts: Empathic effects on counterfactual thinking. *Personality and Social Psychology Bulletin*, *18*(5), 625-630.

MacDonald, K. & Feifel, D. (2013). Helping oxytocin deliver: considerations in the development of oxytocin-based therapeutic for brain disorders. *Frontiers in Neuroscience, 7,* 35. *Intelligence Quotient: Testing, Role of Genetics and the Environment and Social Outcomes*, pp. 35-54.

MacDonald, K., & MacDonald, T. M. (2010). The peptide that binds: a systematic review of oxytocin and its prosocial effects in humans. *Harvard review of psychiatry*, *18*(1), 1-21.

Mandel, D. R., & Dhami, M. K. (2005). "What I did" versus "what I might have done": Effect of factual versus counterfactual thinking on blame, guilt, and shame in prisoners. *Journal of Experimental Social Psychology*, *41*(6), 627-635.

Manning, R., Levine, M., & Collins, A. (2007). The Kitty Genovese murder and the social psychology of helping: the parable of the 38 witnesses. *American Psychologist*, *62*(6), 555.

Martin, I., & MacDonald, S. (2004). An exploration of causes of non-literal language problems in individuals with Asperger syndrome. *Journal of Autism and Developmental Disorders, 34* (3), 311-328.

Mazurek, M. O. (2013). Loneliness, friendship, and well-being in adults with autism spectrum disorders. *Autism*, *0*(0) 1-10.

McClannahan, L. E. & Krantz, P. J. (1999). *Activity schedules for children with autism*. Bethesda, MD: Woodbine House.

McDonald, S., Flanagan, S., & Rollins, J. (2011). The awareness of social inference test (revised). Sydney, Australia: Pearson Assessment.

McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. Journal of Head Trauma Rehabilitation, 18, 219–238

McNamara, P., Durso, R., Brown, A., & Lynch, A. (2003). Counterfactual cognitive deficit in persons with Parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(8), 1065-1070.

McPartland, J. C., Reichow, B., & Volkmar, F. R. (2012). Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder.*Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(4), 368-383.

Mendez, M. F. (2006). What frontotemporal dementia reveals about the neurobiological basis of morality. *Medical Hypotheses, 67*(2), 411-8.

Mendez, M. F., Anderson, E., & Shapira, J. S. (2005). An investigation of moral judgement in frontotemporal dementia. *Cognitive and Behavioral Neurology*, *18*(4), 193-197.

Mestre, M. V., Samper, P., Frías, M. D., & Tur, A. M. (2009). Are women more empathetic than men? A longitudinal study in adolescence. *The Spanish Journal of Psychology*, *12*(01), 76-83.

Meyer, J. A., Mundy, P. C., Van Hecke, A. V., & Durocher, J. S. (2006). Social attribution processes and comorbid psychiatric symptoms in children with Asperger syndrome. *Autism, 10*(4), 383–402.

Mill, J. S. (1863). Utilitarianism. Chapter 3: Of the ultimate sanction of the principle of utility.

Miller, M. B., Sinnott-Armstrong, W., Young, L., King, D., Paggi, A., Fabri, M., Polonara, G. & Gazzaniga, M. S. (2010). Abnormal moral reasoning in complete and partial callosotomy patients. *Neuropsychologia*, *48*(7), 2215-2220.

Minio-Paluello, I., Baron-Cohen, S., Avenanti, A., Walsh, V., & Aglioti, S. M. (2009). Absence of embodied empathy during pain observation in Asperger syndrome. *Biological Psychiatry, 65*(1), 55–62.

Mitchell, J. P., Banaji, M. R., & MacRae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*(8), 1306-1315.

Modahl, C., Green, L. A., Fein, D., Morris, M., Waterhouse, L., Feinstein, C., & Levin, H. (1998). Plasma oxytocin levels in autistic children. *Biological Psychiatry*, *43*(4), 270-277.

Moll, J., de Oliveira-Souza, R., & Eslinger, P. J. (2003). Morals and the human brain: a working model. *Neuroreport*, *14*(3), 299-305.

Moll, J., de Oliveira-Souza, R., Eslinger, P. J., Bramati, I. E., Mourão-Miranda, J., Andreiuolo, P. A., & Pessoa, L. (2002). The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *The Journal of Neuroscience*, *22*(7), 2730-2736.

Moll, J., Oliveira-Souza, D., & Zahn, R. (2008). The neural basis of moral cognition. *Annals of the New York Academy of Sciences*, *1124*(1), 161-180.

Moll, J., Oliveira-Souza, R. D., Zahn, R., & Grafman, J. (2008). The cognitive neuroscience of moral emotions.

Monsch, A. U., Bondi, M. W., Butters, N., Paulsen, J. S., Salmon, D. P., Brugger, P., & Swenson, M. R. (1994). A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. *Neuropsychology*,*8*(1), 25.

Monsch, A. U., Bondi, M. W., Butters, N., Salmon, D. P., Katzman, R., & Thal, L. J. (1992). Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*, *49*(12), 1253-1258.

Moran, J. M., Young, L. L., Saxe, R., Lee, S. M., O'Young, D., Mavros, P. L., & Gabrieli, J. D. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences*, *108*(7), 2688-2692.

Moriguchi, Y., Ohnishi, T., Decety, J., Hirakata, M., Maeda, M., Matsuda, H., & Komaki, G. (2009). The human mirror neuron system in a population with deficient self-awareness: An fMRI study in alexithymia. *Human Brain Mapping*, *30*(7), 2063-2076.

Moriguchi, Y., Ohnishi, T., Lane, R. D., Maeda, M., Mori, T., Nemoto, K., Matsuda, H. & Komaki, G. (2006). Impaired self-awareness and theory of mind: an fMRI study of mentalizing in alexithymia. *Neuroimage*, *32*(3), 1472-1482.

Mueser, K. T., & Bellack, A. S. (2007). Social skills training: Alive and well? *Journal of Mental Health, 16*(5), 549–552.

Müller, E., Schuler, A., & Yates, G. B. (2008). Social challenges and supports from the perspective of individuals with Asperger syndrome and other autism spectrum disabilities. *Autism, 12(2),* 173- 190.

National Autistic Society. (2004). A place in society.

National Autistic Society. (2011). Great expectations.

Nelson, E. E., & Guyer, A. E. (2011). The development of the ventral prefrontal cortex and social flexibility. *Developmental Cognitive Neuroscience, 1(*3), 233–245.

Nemiah, J. C., Freyberger, H. E., & Sifneos, P. E. (1976). Alexithymia: a view of the psychosomatic process. *Modern Trends in Psychosomatic Medicine*, *3*, 430-439.

N'gbala, A., & Branscombe, N. R. (1995). Mental simulation and causal attribution: When simulating an event does not affect fault assignment. *Journal of Experimental Social Psychology*, *31*(2), 139-162.

Nicolle, A., Ropar, D., & Beck, S. R. (2014). Regret and disappointment in autism spectrum disorders: The matter of thinking versus feeling. *Cortex*, *30*(1), e3.

Niedenthal, P. M., Tangney, J. P., & Gavanski, I. (1994). " If only I weren't" versus" If only I hadn't": Distinguishing shame and guilt in conterfactual thinking. *Journal of Personality and Social Psychology*, *67*(4), 585.

Nucci, L. P. (1985). Children's conceptions of morality, societal convention, and religious prescription. *Moral dilemmas: Philosophical and Psychological Issues in the Development of Moral Reasoning*, 137-174.

Nucci, L. & Nucci, M. (1982). Children's social interactions in the context of moral and conventional transgressions. *Child Development, 53*, 403–412.

Nucci, L . & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child Development, 49*, 400 –407.

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin, 133*(2), 310–327.

Osguthorpe, R. T., & Scruggs, T. E. (1986). Special education students as tutors: a review and analysis. *Remedial & Special Education, 7*(4), 15–25

Ozonoff, S. (1998). Assessment and remediation of executive dysfunction in autism and Asperger syndrome. In *Asperger Syndrome or High-functioning Autism?* (pp. 263-289). Springer US.

Ozonoff, S., Pennington, B. F. & Rogers, S.J. (1991). Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32,* 1081 – 1105.

Ozonoff, S., South, M., & Provencal, S. (2005). Executive functions. In F.R. Volkmar, R. Paul, A. Klin & D. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders* (pp. 606-627).

Parker, J. D., Taylor, G. J., Bagby, R. M. (1993). Alexithymia and the recognition of facial expressions of emotion. *Psychotherapy and Psychosomatics, 59*, 197–202.

Patil, I., & Silani, G. (2014). Reduced empathic concern leads to utilitarian moral judgments in trait alexithymia. *Frontiers in psychology*, *5*.

Patrick, C. J., Cuthbert, B. N., & Lang, P. J. (1994). Emotion in the criminal psychopath: fear image processing. *Journal of Abnormal Psychology*, *103*(3), 523.

Peterson, D. M., & Bowler, D. M. (2000). Counterfactual reasoning and false belief understanding in children with autism. *Autism*, *4*(4), 391-405.

Petrinovich, L., O'Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology, 64*, 467–478.

Philip, R. C., Dauvermann, M. R., Whalley, H. C., Baynham, K., Lawrie, S. M., & Stanfield, A. C. (2012). A systematic review and meta-analysis of the fMRI investigation of autism spectrum disorders. *Neuroscience & Biobehavioral Reviews*, *36*(2), 901-942.

Piaget, J. (1932). The moral judgment of children. *London: Routledge & Kegan-Paul.*

Piven, J., Arndt, S., Bailey, J., Havercamp, S., Andreasen, N. C., & Palmer, P. (1995). An MRI study of brain size in autism. *American Journal of Psychiatry*,*152*(8), 1145-1149.

Piven, J. & Palmer, P. (1997) Cognitive deficits in parents from multiple-incidence autism families. *Journal of Child Psychology and Psychiatry, 38,* 1011–1021.

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*, *39*(05), 777-783.

Pearl, J. (2000). Causal inference without counterfactuals: Comment. *Journal of the American Statistical Association*, 428-431.

Perner, J. (1998). 13 The meta-intentional nature of executive functions and theory of mind. *Language and Thought: Interdisciplinary Themes*, 270.

Perner, J., Leekham & Wimmer, H. (1987). 2-year-olds difficulty with false belief – the case for a conceptual deficit. British Journal of Developmental Psychology, 5, 125-137.

Poljac, E., Poljac, E., & Wagemans, J. (2012). Reduced accuracy and sensitivity in the perception of emotional facial expressions in individuals with high autism spectrum traits. *Autism, 17*(6), 668–680.

Ponnet, K., Buysse, A., Roeyers, H., & De Clercq, A. (2008). Mind-reading in young adults with ASD: Does structure matter?. *Journal of Autism and Developmental Disorders*, *38*(5), 905-918.

Ponnet, K., Buysse, A., Roeyers, H., & De Corte, K. (2005). Empathic accuracy in adults with a pervasive developmental disorder during an unstructured conversation with a typically developing stranger. *Journal of Autism and Developmental Disorders, 35(5),* 585-600.

Ponnet, K. S., Roeyers, H., Buysse, A., De Clercq, A., & Van Der Heyden, E. (2004). Advanced mind-reading in adults with Asperger syndrome. *Autism, 8(3),* 249-266.

Premack, D., Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences, 1,* 515-526.

Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations*, *9*(1), 29-43.

Puffer, S. M. (1987). Pro-social behavior, non-compliant behavior and work performance among commission salespeople. *Journal of Applied Psychology, 72*(4), 615–621.

Rao, P. A., Beidel, D. C., & Murray, M. J. (2008). Social skills interventions for children with Asperger's syndrome or high-functioning autism: A review and recommendations. *Journal of Autism and Developmental Disorders, 38*(2), 353–361.

Reichow, B., Steiner, A. M., & Volkmar, F. (2013). Cochrane review: Social skills groups for people aged 6 to 21 with autism spectrum disorders (ASD). *Evidence-Based Child Health: A Cochrane Review Journal, 8*(2), 266–315.

Reiss, A. L., Freund, L., Behavioral phenotype of fragile X syndrome: DSM-III-R autistic behavior in male children. (1992). *American Journal of Medical Genetics, 43*, 35–46.

Rende, B., Ramsberger, G., & Miyake, A. (2002). Commonalities and differences in the working memory components underlying letter and category fluency tasks: a dual-task investigation. *Neuropsychology, 16(3),* 309.

Richer, J. (1976). The social avoidance of autistic children. *Animal Behavior, 24*(4), 898–906.

Riggio, R. E. (1986). Assessment of basic social skills. *Journal of Personality and social Psychology*, *51*(3), 649.

Riggio, R. E., & Carney, D. R. (2003). *Social Skills Inventory Manual.* Mind Garden.

Riggio, R. E., Tucker, J., & Coffaro, D. (1989). Social skills and empathy. *Personality and Individual Differences, 10(1),* 93 – 99.

Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, *gta*, 73–90.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review Neuroscience*, *27*, 169-192.

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happé, F., & Ronald, A. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry*, *68*(11), 1113-1121.

Roese, N.J. (1997). Counterfactual thinking. *Psychological Bulletin, 121*(1), 133-148.

Roeyers, H., Buysse, A., Ponnet, K., & Pichal, B. (2001). Advancing advanced mind-reading tests: Empathic accuracy in adults with a pervasive developmental disorder. *Journal of Child Psychology and Psychiatry, 42(02),* 271-278.

Rogers, K., Dziobek, I., Hassenstab, J., Wolf, O. T., & Convit, A. (2007). Who cares? Revisiting empathy in Asperger syndrome. *Journal of Autism and Developmental Disorders, 37* (4), 709-715.

Rosen, Frederick (2003) Classical Utilitarianism from Hume to Mill. Routledge, page 132

Rosenblatt, M. (2008). *I Exist: the message from adults with autism in England*. London: The National Autistic Society.

Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9(2),* 331-349.

Rutkowski, G. K., Gruder, C. L., & Romer, D. (1983). Group cohesiveness, social norms, and bystander intervention. *Journal of Personality and Social Psychology*, *44*(3), 545.

Rutter, M., DiLavore, P. C., & Risi, S. (2002). *Autism Diagnostic Observation Schedule: ADOS*. Los Angeles, CA: Western Psychological Services.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1255.

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain, 128*(5), 1102–1111.

Scott, F. J., Baron-Cohen, S., & Leslie, A. (1999). 'If pigs could y': A test of counterfactual reasoning and pretence in children with autism. *British Journal of Developmental Psychology*, *17*, 349-362.

Schaller, M., & Cialdini, R. B. (1988). The economics of empathic helping: Support for a mood management motive. *Journal of Experimental Social Psychology, 24*(2), 163–181

Scheeren, A. M., de Rosnay, M., Koot, H. M., & Begeer, S. (2013). Rethinking theory of mind in high-functioning autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 54(6),* 628-635.

Schnall, S., Haidt, J., Clore, G. L., & Jordan, A. H. (2008). Disgust as embodied moral judgment. *Personality and Social Psychology Bulletin. 34*(8), 1096-1109.

Schreiber, C. (2011). Social skills interventions for children with high-functioning autism spectrum disorders. *Journal of Positive Behaviour Interventions, 13*(1), 149–162.

Schwartz, S. H. (2010). Basic values: How they motivate and inhibit prosocial behavior. *Prosocial motives, emotions, and behavior: The better angels of our nature*, *14*, 221-241.

Schwarz, E., Guest, P. C., Rahmoune, H., Wang, L., Levin, Y., Ingudomnukul, E., Ruta, L. & Bahn, S. (2011). Sex-specific serum biomarker patterns in adults with Asperger's syndrome. *Molecular Psychiatry*, *16*(12), 1213-1220.

Segrin, C., & Taylor, M. (2007). Positive interpersonal relationships mediate the association between social skills and psychological well-being. *Personality and Individual Differences*, *43*(4), 637-646.

Seltzer, M. M., Shattuck, P., Abbeduto, L., Greenberg, J.S. (2004). Trajectory of development in adolescents and adults with autism. *Mental Retardation & Development Disabilities Research Reviews, 10*, 234-47.

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry, 34* (8), 1351-1364.

Shallice, T. & Burgess, P.W. (1996). The domain of supervisory processes and temporal organisation of behaviour. *Philosophical Transactions of the Royal Society of London, B351,* 1405–1412

Shaver, K. G. (1985). The attribution of blame: Causality, responsibility, and blameworthiness. New York: Springer-Verlag. Siegrist, M., Earle, T., & Gutscher, H. (20

Shilling, V. M., Chetwynd, A., & Rabbitt, P. M. A. (2002). Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. *Neuropsychologia*, *40*(6), 605-619.

Sigman, M., Spence, S.J. & Ting, W. A. (2006). Autism from developmental and neuropsychological perspectives. *Annual Reviews of Clinical Psychology, 2,* 327-55.

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study. *Social Neuroscience*, *3*(2), 97-112.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science, 303* (5661), 1157–1162.

Skuse, D. H., James, R. S. Bishop, D. M. V, Coppin, B., Dalton, P., Aamodt-Leeper, G., Bacarese-Hamilton, M., Creswell, C., McGurk, R., & Jacobs P.A. (1997). "Evidence from Turner's syndrome of an imprinted X-linked locus affecting cognitive function." *Nature*, 387, 705-708.

Skuse, D. H., Mandy, W. P., & Scourfield, J. (2005). Measuring autistic traits: Heritability, reliability and validity of the social and communication disorders checklist. *The British Journal of Psychiatry, 187*(6), 568–572.

Smalley, S. L., Tanguay, P. E., Smith, M., & Gutierrez, G. (1992). Autism and tuberous sclerosis. *Journal of Autism and Developmental Disorders*, *22*(3), 339-355.

Smetana, J. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52*, 1333–1336.

Smetana, J. G., & Braeges, J. L. (1990). The development of toddlers' moral and conventional judgments. *Merrill-Palmer Quarterly (1982-)*, 329-346.

Sofroniew, M. V. & Weindl, A. (1981). Central nervous system distribution of vasopressin, oxytosin, and neurophysin. In J. L. Martinez, R. A. Jensen, R. B. Mesing, H. Rigter, & J. L. McGaugh (Eds.), Endogenous peptides and learning and memory processes (pp. 327–369). New York: Academic Press.

Solca, F., Poletti, B., Zago, S., Crespi, C., Sassone, F., Lafronza, A. & Ciammola, A. (2015). Counterfactual Thinking Deficit in Huntington's Disease. *PloS one*, *10*(6), e0126773.

Southgate, V., & de C Hamilton, A. F. (2008). Unbroken mirrors: Challenging a theory of autism. *Trends in Cognitive Sciences, 12*(6), 225–229.

Spek, A., Schatorjé, T., Scholte, E., & Van Berckelaer-Onnes, I. (2009). Verbal fluency in adults with high functioning autism or Asperger syndrome. *Neuropsychologia*, *47*(3), 652-656.

Spek, A. A., Scholte, E. M., Van Berckelaer–Onnes, I. A. (2010). Theory of mind in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders, 40* (3), 280-289.

Spreckley, M., & Boyd, R. (2009). Efficacy of applied behavioral intervention in preschool children with autism for improving cognitive, language, and adaptive behavior: a systematic review and meta-analysis. *The Journal of Pediatrics*,*154*(3), 338-344.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10,* 640-656.

Stuss, D. T., & Knight, R. T. (2013). *Principles of frontal lobe function*. Oxford University Press.

Stuss, D. T., & Levine, B. (2002). Adult clinical neuropsychology: lessons from studies of the frontal lobes. *Annual Review of Psychology, 53(1),* 401-433.

Sze, J. A., Gyurak, A., Goodkind, M. S., & Levenson, R. W. (2011). Greater emotional empathy and prosocial behavior in late life. *Emotion, 12*(5), 1129.

Thoma, P., & Bellebaum, C. (2012). Your error's got me feeling–how empathy relates to the electrophysiological correlates of performance monitoring. *Frontiers in Human Neuroscience*, *6*:135.

Thomson, J. J. (1976). Killing, letting die, and the trolley problem. *The Monist*, *59*(2), 204-217.

Travis, L. L., & Sigman, M. S. (1998). Social deficits and interpersonal relationships in autism. *Mental Retardation and Developmental Deficits, 4*(2), 65–72.

Tsuchida, A. & Fellows, L. K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex, 49(7),* 1790-1800.

Turner, M. (1999). Annotation: Repetitive behaviour in autism: A review of psychological research. *Journal of Child Psychology and Psychiatry*, *40*(06), 839-849.

Turner-Brown, L. M., Perry, T. D., Dichter, G. S., Bodfish, J. W., & Penn, D. L. (2008). Brief report: Feasibility of social cognition and interaction training for adults with high functioning autism. *Journal of Autism and Developmental Disorders*, *38*(9), 1777-1784.

Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.

Van Boxtel, J. J., & Lu, H. (2013). Impaired global, and compensatory local, biological motion processing in people with high levels of autistic traits. *Frontiers in Psychology, 4*, 209.

Van Hoeck, N., Ma, N., Ampe, L., Baetens, K., Vandekerckhove, M., & Van Overwalle, F. (2013). Counterfactual thinking: an fMRI study on changing the past for a better future. *Social Cognitive and Affective Neuroscience*, *8*(5), 556-564.

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose–response meta-analysis of multiple outcomes. *Clinical Psychology Review*, *30*(4), 387-399.

Vyas, K. (2015). Psychopathic traits and everyday social behaviour. A thesis submitted for the degree of Doctor of Philosophy by Karishma Vyas, Department of Experimental Psychology, University College London, September 2015.

Vyas, K., Jameel, L., & Channon, S. (in preparation a). The 'Social Favours' task: pro-social behavior in groups high and low in psychopathic traits.

Vyas, K., Jameel, L., Channon, S. (in preparation b). Judgments of deservingness for positive and negative outcomes in groups high versus low in psychopathic traits.

Wainryb, C., & Turiel, E. (1994). Dominance, subordination, and concepts of personal entitlements in cultural contexts. *Child Development*, *65*(6), 1701-1722.

Waldmann, M. R. & Weigmann, A. (2010). A double causal contrast theory of moral intuitions in trolley dilemmas. In S. Ohlsson & R. Catrambone. (Ed.,). *Proceedings of the 32nd Annual Conference of the Cognitive Science Society.* (2589-2584). Autisn, TX: Cognitive Science Society.

Warren, Z. E., Foss-Feig, J. H., Malesa, E. E., Lee, E. B. & Lounds Taylor, J. (2011). Neurocognitive and behavioral outcomes of younger siblings of children with autism spectrum disorder at age five. *Journal of Autism and Developmental Disorders, 42*, 3, 409-418.

Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology*, *56*(2), 161.

Wheatley, T., & Haidt, J. (2005). Hypnotic disgust makes moral judgments more severe. *Psychological Science*, *16*(10), 780-784.

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., Weil, L. & Wakabayashi, A. (2006). Predicting autism spectrum quotient (AQ) from the systemizing quotient-revised (SQ-R) and empathy quotient (EQ). *Brain research*, *1079*(1), 47-56.

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "open-ended" executive function tests in autism. *Autism Research*, *2*(3), 138-147.

White, S. J., Frith, U., Rellecke, J., Al-Noor, Z., & Gilbert, S. J. (2014). Autistic adolescents show atypical activation of the brain′ s mentalizing system even without a prior history of mentalizing problems. *Neuropsychologia*, *56*, 17-25.

White, S. W., Ollendick, T. H., & Bray, B. C. (2011). College students on the autism spectrum: Prevalence and associated problems. *Autism, 15*(6), 683–701.

Wimmer, H. & Perner, J. (1983). Beliefs about beliefs – representation and constraining fucntiong of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 103-128.

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine, 11,* 115-129.

Witkin, H. A., Oltman, P. K., Raskin, E. & Karp, S. (1971). A manual for the embedded figures test. Palo Alto, CA: Consulting Psychologists Press.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.

Wu, S., Jia, M., Ruan, Y., Liu, J., Guo, Y., Shuang, M., Gong, X., Zhang, Y., Yang, X. & Zhang, D. (2005). Positive association of the oxytocin receptor gene (OXTR) with autism in the Chinese Han population. *Biological Psychiatry*, *58*(1), 74-77.

Young, L., Bechara, A., Tranel, D., Damasio, H., Hauser, M., & Damasio, A. (2010). Damage to ventromedial prefrontal cortex impairs judgment of harmful intent. *Neuron*, *65*(6), 845-851.

Yogev, A. and Ronen, R. (1982), "Cross-age tutoring: effects on tutors' attributes", *Journal of Educational Research*, Vol. 75 No. 5, pp. 261-268.

Zago, S., Delli Ponti, A., Mastroianni, S., Solca, F., Tomasini, E., Poletti, B. & Porta, M. (2014). Counterfactual Thinking in Tourette's Syndrome: A Study Using Three Measures. *Behavioural Neurology*.

Zalla, T., Barlassina, L., Buon, M., & Leboyer, M. (2011). Moral judgment in adults with autism spectrum disorders. *Cognition*, *121*(1), 115-126.

Zalla, T., & Leboyer, M. (2011). Judgment of intentionality and moral evaluation in individuals with high functioning autism. *Review of Philosophy and Psychology*, *2*(4), 681-698.

Zalla, T., Sav, A. M., Stopin, A., Ahade, S., & Leboyer, M. (2009). Faux pas detection and intentional action in Asperger Syndrome. A replication on a French sample. *Journal of Autism and Developmental Disorders*, *39*(2), 373-382.

Zeeland, S.-V., Ashley, A., Dapretto, M., Ghahremani, D. G., Poldrack, R. A., & Bookheimer, S. Y. (2010). Reward processing in autism. *Autism Research, 3*(2), 53–67.

Zultan, R., Gerstenberg, T., & Lagnado, D. A. (2012). Finding fault: Counterfactuals and causality in group attributions. *Cognition*, *125*(3), 429-440.

Zürcher, N. R., Bhanot, A., McDougle, C. J., & Hooker, J. M. (2015). A systematic review of molecular imaging (PET and SPECT) in autism spectrum disorder: Current state and future research opportunities. *Neuroscience & Biobehavioral Reviews*, *52*, 56-73.

Zwaigenbaum, L., Bryson, S., & Garon, N. (2013). Early identification of autism spectrum disorders. *Behavioural Brain Research*, *251*, 133-146.