Systemising Emotions:
Teaching Emotion Recognition to People with Autism Using Interactive Multimedia

Ofer Golan

Wolfson College

Department of Psychiatry
Autism Research Centre

Supervisor: Professor Simon Baron-Cohen

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Cambridge

Table of Contents

Preface

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Cambridge. The research described herein was conducted under the supervision of Professor Simon Baron-Cohen in the Department of Psychiatry, University of Cambridge.

Except where acknowledgement and reference are made to previous work, this work is, to the best of my knowledge, original. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text. Neither this, nor any substantially similar dissertation has been, or is being, submitted for any other degree, diploma, or other qualification at any other university. This dissertation does not exceed 60,000 words in length.

Ofer Golan
September 6th, 2006

Abstract

Recognition of emotions and mental states (ER) in others is a core difficulty for individuals with autism spectrum conditions (ASC). In contrast, they show good skills in ‘systemizing’- understanding non-agentive systems. This thesis investigated whether improvement of ER in ASC is possible when systemising strengths are harnessed. A computer based training program was evaluated, as the computerised environment is well liked amongst individuals with ASC for its predictability and controllability. Previous interventions into ER in ASC were narrow in scope, focused on basic emotions and on still facial expressions only, all of which might have limited the ability to generalise from these training programs.

This thesis evaluated the effectiveness of *Mind Reading*, a computer program teaching ER from a wide range of facial expression videos and recorded speech segments, systematically presented. Three different experiments tested the effectiveness of a minimum of 10 hours of software use over a period of 10-15 weeks among individuals with ASC. Experiments included evaluation of independent use of the software by adults and by 8-11 year olds with ASC, and tutor and group supported use of the software in adults with ASC. ER skills were assessed on four levels of generalisation before and after the training period, and compared to matched ASC and typically developing control groups.

Results showed improved ER for software users from faces and voices, compared to the ASC control groups. Improvement was mostly limited to faces and voices which were included in the software. Generalisation to stimuli not included in the software was found in the children experiment, in the vocal and visual channels separately. Follow up assessment after a year showed greater improvement on general socio-emotional functioning measures among child and adult software users, compared to ASC controls.

These results suggest that individuals with ASC can improve their ability to recognise emotions using systematic computer-based training with long term effects, but may need further tutoring to prevent hyper-systemising, and to enhance generalisation to other situations and stimuli. The reasons behind generalisation difficulties and the study’s limitations are discussed, and suggestions for future work are offered.

Papers submitted for publication

Parts of the work in Chapters 6 and 7 and in appendices 1,2 and 3 has been submitted for publication in the following papers:

Chapters 6 and 7

Golan, O., & Baron-Cohen, S. (2006). Systemizing empathy: Teaching adults with Asperger Syndrome and High Functioning Autism to recognize complex emotions using interactive multimedia. *Development and Psychopathology, 18* (2), 591-617.

Appendix 1

Golan, O., Baron-Cohen, S., & Hill, J. J. (2006). The Cambridge Mindreading (CAM) Face-Voice Battery: testing complex emotion recognition in adults with and without Asperger Syndrome. *Journal of Autism and Developmental Disorders, 36* (2), 169-183.

Golan, O., & Baron-Cohen, S. (submitted). The Cambridge Mindreading Face-Voice Battery for children (CAM-C): testing basic and complex emotion recognition in children with and without Autism Spectrum Conditions.

Appendix 2

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (in press). The 'Reading the Mind in the Voice' test - Revised: A study of complex emotion recognition in adults with and without Autism Spectrum Conditions. *Journal of Autism and Developmental Disorders*.

Appendix 3

Golan, O., Baron-Cohen, S., Hill, J. J., & Golan, Y. (2006). 'Reading the Mind in Films' - testing recognition of complex emotions and mental states in adults with and without autism spectrum conditions. *Social Neuroscience,* 1(2), 111-123.

Golan, O., Baron-Cohen, S., & Golan, Y. (submitted). 'Reading the Mind in Films -C' - testing recognition of complex emotions and mental states in children with and without autism spectrum conditions.

Table of Tables

Table of Figures

Abbreviations

AQ	The Autism spectrum Quotient
AS	Asperger Syndrome
ASC	Autism Spectrum Conditions
CAM-A	The Cambridge Mindreading Face-Voice Battery – Adult version
CAM-C	The Cambridge Mindreading Face-Voice Battery – Child version
CAST	the Childhood Asperger Syndrome Test
C-FAT	Child Feature-based Auditory Task
ED	Executive Dysfunction
EF	Executive Function
ER	Emotion Recognition
FQ	The Friendship and relationship Questionnaire
HFA	High-Functioning Autism
RME-A	The Reading the Mind in the Eyes test – Adult version
RME-C	The Reading the Mind in the Eyes test – Child version
RMF-A	The Reading the Mind in Films test – Adult Version
RMF-C	The Reading the Mind in Films test – Child Version
RMV-R	The Reading the Mind in the Voice test – Revised
ToM	Theory of Mind
VABS-S	Supplement to the Vinelands Adaptive Behavior Scales
WCC	Weak Central Coherence

Acknowledgements

As much as I love my work, there was no chapter I enjoyed writing more than this one. During my five year journey that ends in this thesis, I have had the privilege of meeting and working with many great men, women and children, who supported, inspired and enriched me and my work in many ways. This is the place to thank them for their contribution, with apologies to anyone I might have mistakenly left out.

First and foremost, I would like to thank Professor Simon Baron-Cohen, my supervisor and mentor. Since you first replied to my email from Israel, you never ceased to amaze me with your elegant combination of professionalism, modesty, empathy, generosity, and great sense of humour. Thank you, Simon, for being my role model.

I am also grateful to my Israeli mentor Dr Raya Ariel, for sparking my interest in autism spectrum conditions, for encouraging me to embark on this journey and for believing in my ability to complete it long before I thought I could start it. Thank you, Raya, for your long lasting support, friendship, and good advice.

This project would have not begun without the help of Elizabeth and Sidney Corob, CBE, who agreed to support a young Israeli psychologist, and made it possible for me to come to Cambridge and start this work. Thank you for your kindness and generosity towards our little family. You are always in our hearts.

Along this period, my work was supported by the National Alliance for Autism Research, the Cambridge Overseas Trust, B'nai B'rith Leo Baeck scholarships, and the Wingate Foundation. I am grateful to all these organisations and hope they found their investment in this project and in me fruitful.

Conducting a clinical study in such a scope required lots of enthusiasm and goodwill from many people and organisations, who helped in recruiting participants, organising groups and facilitating them, hosting our research team, and providing us with all we needed. Many thanks go to Robin Wallach, Susan Owen, and Barbara Bendon at

Umbrella Autism Cambridgeshire, Andrea Macleod and Rob Whiskens at Autism West Midlands, Jill Howard and Gill Wilson at Autism London, Lynne Moxon at the European Services for People with Autism, Harriet Fisher and James Graham at the Interact Centre, Moira Gommon and Steve Owen at the Wirral Autistic Society, and Jane Taylor at Brookside Family Consultation Clinic. Thanks to the staff and pupils of Mayfield Primary School, and Kings Hedges Primary School in Cambridge for their help with recruiting and assessing child participants. Special thanks to Pamela Yates, who arranged and hosted the social skills groups, and to Sally Wheelwright and Jenny Hannah for assistance in recruiting adult participants.

Throughout this project, I was very lucky to share the burden with several research assistants, who were as kind and welcoming to participants as they were professional and hard-working. I am indebted to Yael Golan, Sarah Johnson, Emma Chapman, Ilaria Minio-Paluello, and Yael Granader. This project would have never been completed without your help.

The Autism Research Centre in Cambridge was a great place to work in. Besides the international mix of keen and creative minds, it was simply fun coming to the lab, knowing my fellow ARC agents will be there, willing to endure my loud presence. I would like to thank all of my colleagues at the ARC past and present, but especially Chris Ashwin, Johnny Lawson, Bhisma Chakrabarti, Jacqueline Hill, Matthew Belmonte and Marie Gomot for your friendship and for hours of brain storming together, which surely enriched the work presented in these pages. Thanks also to Paul LaCava at Kansas University, for all these hours of online research discussions, which helped me understand my own work better. Special thanks to Jenny Hannah and Paula Naimi for patiently teaching me the ways of the English culture, the English language, and the English etiquette and for leading me through the mysteries of English bureaucracy.

Wolfson College provided me not only with kind hospitality by friendly staff, but also with the opportunity to meet friends from all over the world. These friends have provided me with priceless support through the long journey towards this thesis. I am grateful to all of them, but in particular to Jennifer and John Williams, Gil Klein, and

Merav Cohen, whose friendship I cherish and intend to continue enjoying for years to come.

For me, the most difficult aspect of spending five years away from my home in Israel was the need to interact with my dear family through the little green eye of a webcam. I would like to thank my brothers, my brother and sisters in law, my parents in law, and especially my grandparents, for their patience and support from afar and for always being there for me when I needed them. My love and eternal gratitude goes to my parents, for bringing me up to ask questions and to pursue knowledge, and for teaching me the sense of empathy and care for other people's emotions. I hope this work manages to combine these two gifts you have granted me.

Although words cannot describe it well, I would like to thank Yael and Leor, my loved ones. Yael - thank you for embarking on this adventure with me, and for making these five years something to look back to with a smile. Thank you for pushing me to finish, often at the cost of your own work. I would have never managed it without your love and support, for which I am forever indebted. Leor - I often called you 'my anti-thesis', since it was impossible to focus on writing in your charming presence. You are too young to know how much light you brought into my life, but I hope you will be reading this one day and know how grateful I was to have you with me for the last months of this journey.

Lastly, I would like to thank the adults, the children and the families who took part in this study who taught me about autism spectrum conditions more than any book could ever do.

Introduction

1 Autism

1.1 Description and diagnosis

Autism is a pervasive neuro-developmental condition, which is estimated to affect 0.1% of the population (Medical Research Council, 2001). It was first described by the American Psychiatrist Leo Kanner in 1943, as 'autistic disturbances of affective contact'. Kanner described a group of children, who from birth were socially and emotionally withdrawn, and who showed stereotyped and obsessive behaviours. These children had delayed speech development, were echolalic, and had very literal understanding of language. They interacted better with objects than with people, and were most satisfied when left alone (Kanner, 1943).

From the research conducted on autism ever since Kanner's description, it has become clear that autism has a genetic basis (Folstein & Rosen-Sheidley, 2001) and that individuals with autism show unique patterns of brain development and activity (Frith & Hill, 2004). Nevertheless, despite the shift in understanding autism - from the product of 'cold mothering' (Bettelheim, 1967) to a genetic based neuro-developmental condition - more than fifty years after Kanner's original paper the diagnosis of autism still relies on behavioural observation and developmental history. The symptoms used to diagnose autism today are very similar to Kanner's original description. The 4th edition of the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM-IV) and the 10th edition of the International Classification of Diseases of the World Health Organisation (ICD-10) set this triad of clusters of symptoms as a basis for diagnosis of autism (American Psychiatric Association, 1994; World Health Organisation, 1994):

A. Qualitative impairments in social interaction

This cluster deals with the interest, the understanding, and the emotional and behavioural manifestation of social interaction. Individuals with autism show less interest in social interaction. They lack spontaneous seeking to share enjoyment, interests or achievements with others (e.g.: by showing, bringing, or pointing out objects of interest), and rarely look for comfort from others, when in distress. They also have poor empathy, and fail to offer comfort to others. Their interaction style lacks social or emotional reciprocity, and they have difficulties using non verbal behaviours (e.g. eye gaze, facial expression, body postures, gestures) to regulate social interaction. Consequently, they fail to develop age appropriate peer relationships.

B. Qualitative impairments in communication

This cluster focuses on the verbal and non verbal communication difficulties manifested in autism. Individuals with autism show a delay in, or a total lack of the development of spoken language, without attempts to compensate through alternative modes of communication. Those who develop language show difficulties in the ability to initiate or sustain a conversation with others, and their verbal comprehension is very literal and rigid. Their language problems are mainly in the pragmatics of language – the use of language sensitive to the social context. Individuals with autism also have difficulties in the expression and perception of non verbal aspects of communication, including facial expressions, intonation and gestures. They show poor flexibility in language expression and a relative lack of creativity and fantasy in thought processes. Children with autism lack varied, spontaneous, age appropriate, make-believe play or social imitative play.

C. Restricted, repetitive & stereotyped behaviour patterns, interests & activities

This cluster deals with the rigidity and the stereotypic manner of thought and behaviour in autism. Individuals with autism tend to impose rigidity and routine on a wide range of aspects of day-to-day functioning. They have difficulties adjusting to change and dealing with novelty (e.g. going to a new school) and would adhere inflexibly to specific routines or rituals even on known everyday activities (e.g. changing the regular route to school). Play patterns of children with autism are also repetitive, and as shown above, lack imagination. They also tend to have encompassing preoccupation with one or more stereotyped and restricted patterns of interest (e.g. dates, routes, timetables), and have persistent preoccupation with parts of objects (e.g. car wheels rather than the whole car). Stereotyped behaviour is also manifested through characteristic stereotyped and repetitive motor mannerisms (e.g. hand flapping or twisting), and specific sensory interests (e.g. smell or feel of objects).

The diagnosis of autism also requires an onset of this picture of abnormal functioning prior to age 3 years, and an exclusion of criteria of other pervasive developmental disorders (such as Rett's Syndrome or Childhood Disintegrative Disorder).

In addition to the cluster triad, ICD-10 also includes several other features that are characteristic of autism, such as high levels of anxiety, sleeping and eating disturbances, auditory and tactile hypersensitivity, lack of spontaneity, initiative, and creativity in organising one's daily life, and decision-making difficulties. The specific manifestation of deficits characteristic of autism changes throughout the life span, but the deficits continue with a broadly similar pattern of problems.

In about 75% of individuals with autism the pervasive presentation of deficits described above is accompanied by mental retardation. Co-morbid occurrence of other conditions with autism is also common. Among these are neuro-developmental conditions such as epilepsy (Steffenburg, Steffenburg, & Gillberg, 2003), Tourette syndrome (Baron-Cohen, Scahill, Izaguirre, Hornsey, & Robertson, 1999), tuberous sclerosis (Bolton & Griffiths, 1997), and attention deficit and hyperactivity disorder. Other co-morbidities include psychiatric conditions such as schizophrenia, affective disorders, eating disorders, and obsessive-compulsive disorder (Gillberg & Billstedt, 2000). The genetic and neurological foundations of these co-morbidities are yet to be fully understood.

1.2 Asperger Syndrome and the autistic spectrum

Almost in parallel to Kanner, Hans Asperger, an Austrian pediatrician, described in 1944 a group of children who presented with what he called 'autistic psychopathy'

(Asperger, 1944). These children presented with similar difficulties to the ones Kanner had described: failing to socially integrate, lacking eye contact, speaking in flat or inappropriate intonation, failing to emotionally interact, using stereotypic movements, and presenting with circumscribed interests. However, unlike the children Kanner had described, Asperger's children were verbal, and had normal and even superior intelligence. Asperger named them 'little professors', as he argued that these children compensate for their poor socialising skills with original thought, which could lead them to great achievements.

Asperger's paper did not receive the same level of attention as Kanner's, mostly since it was written in German during the 2nd world war. However, in the late 1970s, the British psychiatrist Lorna Wing discovered Asperger's writings, and coined the term 'Asperger Syndrome' (AS; Wing, 1981). The diagnosis criteria of AS are similar to those of autism, but they require a lack of any clinically significant general delay in language or cognitive development. These features differentiate between the two diagnoses.

The existence of a different form of autism in children who show the cluster triad symptoms but do not necessarily have language delay or mental retardation was further supported in a pioneering autism epidemiological study conducted in the former London borough of Camberwell (Wing & Gould, 1979). In addition to children with 'Kanner's autism', the researchers identified a group of children who had impairments of social interaction, communication and imagination, together with a repetitive stereotyped pattern of activities. These children did not fit into the full picture of autism as described by Kanner, suggesting that autism is not a single condition, but rather a spectrum of conditions. The autistic spectrum has 'Kanner's autism' at the most severely affected and lower functioning side and AS higher on the spectrum, towards typical functioning. Autism Spectrum Conditions (ASC) are believed to share the underlying neurobiological bases (Volkmar & Klin, 2005).

In some cases, children with ASC show an initial language delay, prior to age 3 years, but as they grow up, they develop language and have normal levels of IQ (i.e. no mental retardation). These children are usually diagnosed with High Functioning

Autism (HFA). Though some studies argue that individuals with AS and HFA differ neuro-anatomically (Lotspeich et al., 2004) or neuro-psychologically (Rinehart, Bradshaw, Brereton, & Tonge, 2002), they usually present a similar clinical picture, and are often tested as similar high functioning ASC (Frith, 2003). This work will adopt this approach.

The prevalence of ASC is currently estimated as 0.6% of the population (Chakrabarti & Fombonne, 2001, 2005). Current estimates of the prevalence of AS are around 0.4-0.5% of the population (Kadesjo, Gillberg, & Hagberg, 1999; Scott, Baron-Cohen, Bolton, & Brayne, 2002). These estimates show that autism and ASC are much more common than previously believed (Frith, 2003). Both autism and ASC are more common in males than in females. The male to female ratio in autism is about 4:1. This male dominance is even more distinct in AS and HFA, where male to female ratio goes up to 9:1 (Baron-Cohen, 2003).

1.3 Cognitive Models of Autism

As mentioned above, research efforts aim to explain the underlying mechanisms of autistic symptomatology. Neuro-developmental research has revealed unique patterns of activation in the autistic brain (Belmonte et al., 2004; Frith & Hill, 2004). However in the absence of a comprehensive neurological explanation of autism, the existing explaining models come from the field of cognitive psychology. Supported by behavioural, developmental and neuro-imaging research studies, the cognitive models provide an extensive explanation for the autistic triad of symptom clusters. During the last two decades, three cognitive models have attempted to explain the cognitive, emotional, and behavioural deficits in autism: Impaired Theory of Mind, Executive Dysfunction, and Weak Central Coherence. These three cognitive models are reviewed below, followed by a fourth, more recent model, which examines strengths, as well as deficits in ASC.

1.3.1 Impaired theory of mind – Mindblindness

Since the middle of the 1980s, the Mindblindness hypothesis has been the most acceptable, or at least the most studied cognitive model of autism. First, the typical development of Theory of Mind (ToM) will be reviewed, followed by a description of the ToM deficit in autism.

1.3.1.1 Typical development of Theory of Mind

‘Theory of Mind’, a term coined by Premack & Woodruff, (1978), refers to the ability to predict the mental states of another, or to form meta-representations of others’ intentions, beliefs, desires, and emotions and to distinguish them from oneself’s. ToM is also referred to as ‘mentalising’ (Frith, 2003), or as ‘mindreading’ (Wellman, 1992).

This ability to ‘read’ others’ minds, and predict their thoughts, emotions, and future behaviour, requires an understanding of the difference between the mental and the physical. The mental is much more flexible and abstract than the physical: it can be true or false (as in the case of beliefs); it exists only in the mental world; and it can be completely independent of reality (as in the case of fantasies). Another quality of the mental is that it is private and individual – my mental state is invisible to others. We may share some thoughts or ideas, but we may also have completely different intentions, beliefs, or desires with regards to the same events (Wellman, 1992).

The cognitive capacity to ‘mindread’ starts developing from birth, though newborns are already equipped with an innate special interest in humans, and especially in human faces (Johnson & Morton, 1991; Mondloch et al., 1999). During their first year of life, infants come to construe people as what Flavell calls ‘compliant agents’ (Flavell, 1999), i.e. entities that are self-propelled and capable of independent movement (agents) but also influenceable at a distance by communicative signals (compliant). During the second year of life, infants develop an understanding of intentionality and desire. They learn that agents may relate to objects in different ways, such as wanting to get them, or wanting to avoid them. They also learn to associate these with the direction of the agent's attention, and with the agent's emotional expression. Infants at their second year of life also start engaging in a triadic interaction with agents (Tomasello, Carpenter, Call, Behne, & Moll, 2005). They develop the ability for joint attention, i.e. the ability to attend to an object together with another agent. This allows them to learn about the intentions and emotional attitudes of agents towards objects. They also begin to use mental state words to describe what they see, desire, or feel towards objects or agents (Wellman, 1992). These early development stages will be further reviewed in chapter 2, with regards to emotion recognition skills.

One of the major developments of the second year of life is the emergence of the ability to pretend. Between 18-24 months of age, infants start playing with objects 'as if' they were something else, e.g. using a broom as a guitar. This form of pretend play is the first opportunity for infants to create 'meta-representations' of objects. Pretend play has been proposed as a building block for mentalising. Pretend play also includes role playing, in which the child could be the parent, with the doll as its child. Such imaginative play allows the child to create meta-representations of others' mental states, and to distinguish them from her own, thus building her ToM skills (Leslie, 1987).

Between the ages of 3-4 years, children understand the link between attending to an object and knowing about it. An agent that does not *attend* to an object or an event will not *know* about it. The most common channel of attention is visual attention. When 3-4 year old children were shown a picture of two girls, one looking into a box, and the other laying her hand on the edge of the box, they could reliably indicate which girl *knows* what is in the box. Therefore children at this age are already aware that *seeing* leads to *knowing* (Pratt & Bryant, 1990).

The next central milestone in the development of ToM appears around the age of 4-5. At this age children are able to separate the mental and the physical, and understand that people's subjective beliefs can be different to objective reality. A four year old child can grasp that different agents have different beliefs, which can differ from her own, and which may bring the agents to different behavioural outcomes (Wellman & Lagattuta, 2000).

To experimentally test the child's ability to mentalise at this level, Wimmer and Perner created a new experimental paradigm, assessing the child's understanding of another agent's false belief (Wimmer & Perner, 1983). The children watched a puppet play scenario in which a protagonist puts an object into a location x and then witnessed that in the absence of the protagonist the object is transferred from x to location y. Since this transfer occurred when the protagonist was not present to see it, they had to assume that the protagonist still believed that the object was in x. Participants had to indicate where the protagonist would look for the object on his return. Below the age of 4, children said that the protagonist would look for the object in its new location (y), failing to distinguish between the objective reality and the protagonist's false belief, or between what the child knows and what the protagonist knows about the object's location. However, from the age of 4, children could recognise the protagonist's *false* belief and said she would look for the object where she left it. This ability to understand the difference between the mental and the physical, and to distinguish between their own mental state and others', marks the development of the child's ToM ability.

ToM carries on developing and refining after the age of 4 (some aspects of it will be discussed in chapter 2 with regards to the development of complex emotion recognition skills). It underlies our ability to understand human behaviour and to predict it, to co-operate and communicate, to engage in social relationships, and to function in society (Wellman & Lagattuta, 2000). Without ToM, one is bound to constantly feel confused by other people's behaviour, failing to understand the motives that underlie human action. As we will see in the next section, this is how the mindblindness model pictures the world through the eyes of a person with autism.

1.3.1.2 Theory of Mind in Autism

From the descriptive symptomatology of autism given at the beginning of this chapter, it is easy to conclude that individuals with autism may have underlying ToM difficulties. Their poor eye contact and lack of imaginative play, their lack of empathy and difficulties forming relationships deviate from the typical development of ToM described above. Adding to these are the reports of adults with autism, who struggle to understand the intentions and to predict the behaviour of people around them, and who often feel like ‘an anthropologist on Mars’ (Sacks, 1995). However, the first attempt to assess the ToM deficit *experimentally* was conducted in 1985, using the false belief paradigm described above. Baron-Cohen, Leslie, and Frith, presented children with autism, children with Down syndrome, and typically developing controls with the ‘Sally-Anne task’, shown in figure 1 below. Using puppets, the experimenter showed the child Sally, who has a basket and Anne, who has a box. Sally puts her marble in her basket and leaves. Anne then moves Sally’s marble from the basket to her box. Sally returns, and the child is asked where she will look for her marble.

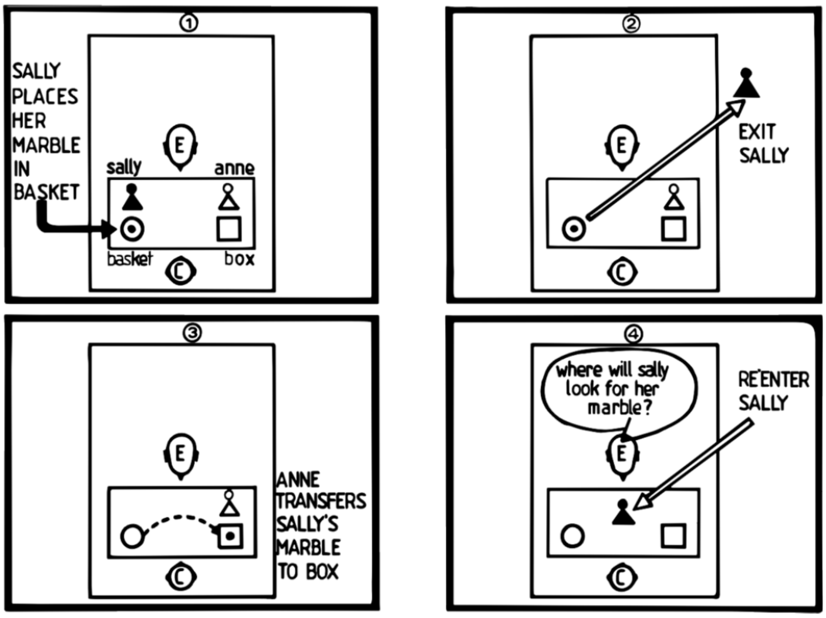


Figure 1.1: The Sally-Anne false belief task (adapted from Baron-Cohen, Leslie, & Frith, 1985)

Results showed that children with autism, despite having an average mental age that was higher than 5 years, consistently said that Sally would look for her marble in Anne’s box, thus failing to acknowledge Sally’s false belief. In comparison, children with Down syndrome, who were younger and had a lower mental age, answered this question correctly. The authors concluded that the failure of children with autism to understand the mental state of belief in others suggests they have an underlying deficit in ToM (Baron-Cohen, Leslie, & Frith, 1985). This finding was later replicated with different variations on the false belief task. For example, to rule out the possibility that children with autism failed to relate to puppets as humans, the experiment was replicated with the experimenters in place of the puppets, coming to the same conclusions (Leslie & Frith, 1988).

In order to rule out a verbal understanding confound in the Sally-Anne task, Baron-Cohen and colleagues administered to the same participants a non-verbal picture sequencing task. The children were asked to put the cards in the right order to create a story. Three kinds of stories were included- mechanical, behavioural and intentional. The mechanical stories depicted interaction between objects (e.g. a balloon flying in the wind, hitting a tree and bursting); the behavioural stories depicted people engaging in behaviours that do not require mental attribution (e.g. buying sweets in a shop), and the intentional stories depicted people engaging in behaviours that require mental state attribution (e.g. A boy who bought sweets goes home without *noticing* there is a hole in his bag of sweets. When he gets home he is *surprised* to find the sweets are not where he *expected* them to be). The children with autism managed to arrange and to explain the behavioural stories as well as, and the mechanical stories better than, the Down syndrome or typically developing controls. However, they performed significantly worse than the controls on the intentional stories, supporting the ToM deficit hypothesis (Baron-Cohen, Leslie, & Frith, 1986).

Another study used a different false belief paradigm, to check if the child with autism simply has different assumptions about what other people normally expect to happen. Hence, this paradigm involved the child herself experiencing a false belief: A tube of ‘Smarties’ sweets was shown to the child, who was asked what it contains. When the child answered it contains Smarties, the experimenter opened the tube and showed the child it actually contains a pencil. The child was then asked what another child coming in would say the box contains. Children with autism, having just experienced the disappointment of finding a pencil instead of sweets, said the new child would say the box contains a pencil. When questioned about their own false belief, some of the children with autism now insisted that they had all along believed the tube contained a pencil. Others recalled they originally thought the tube contains sweets, but still could not attribute their own false belief to others (Perner, Frith, Leslie, & Leekam, 1989).

Children with autism also fail to link attending to something with knowing about it (Leslie & Frith, 1988). In a study using Pratt & Bryant's 'seeing leads to knowing' task, described above, which is passed by typically developing 3 year olds, Baron-Cohen & Goodhart compared performance of young adolescents with autism with matched controls with mental retardation. The adolescents with autism performed at chance and significantly worse than the controls (Baron-Cohen & Goodhart, 1994). Another study using schematic pictures of faces surrounded by four different kinds of snacks, asked children to tell which snack the agent wants, or is going to get. Typically developing children and learning disabled children used the direction of the agent's gaze to tell what it wants and likes, and which snack it is going to pick. However, children with autism failed to make a mentalistic interpretation of what or where the agent was looking, in order to tell what it wants, likes or thinks (Baron-Cohen, Campbell, Karmiloff-Smith, Grant, & Walker, 1995). This was further replicated in several studies of children with autism (Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997; Phillips, Baron-Cohen, & Rutter, 1998).

Further support to the ToM impairment in autism comes from studies showing difficulties in emotion recognition in autism, especially with cognition based emotions (Baron-Cohen, Spitz, & Cross, 1993). This area will be reviewed in detail in chapter 2.

Children with autism do come to pass false belief tasks at a later stage in development. An analysis of several false belief studies revealed that typically developing children pass false belief tasks by the age of 5, and learning disabled children at the mental age of 5. However, children with autism were only able to pass these tasks at the average mental age of 9:2 (Happe, 1995). This finding, as well as findings of high functioning children with ASC who pass the false belief tasks mentioned above indicate that their ToM understanding does improve with age. Such improvement could result from compensation strategies individuals with autism employ in order to bypass their earlier difficulties. Such compensation strategies include using logic rules (Frith, 2003), relying on good verbal skills (in high functioning individuals with ASC, e.g. Grossman, Klin, Carter, & Volkmar, 2000), or relying on gained life experience (Grandin, 1995). In order to test the ToM deficit in autism in older or with higher functioning children with autism, more advanced tasks are required.

One example of such a task is the ‘second order’ false belief paradigm, which requires the child to relate the mental state of one agent to that of another. In the case of the Sally-Anne task, a second-order question would be: ‘If while Anne was moving the marble into the box, Sally was peeking through the keyhole, where would Anne think Sally will look for the marble?’. This question requires separating our knowledge of what Sally knows and Anne doesn’t know, which will then lead to the conclusion that Anne now has a false belief, and she would expect Sally to look for the marble in the basket, whereas in fact Sally will look for it in the box. In a study using another story following similar lines, children with autism who pass first order false belief tasks, failed the second order false belief question (Baron-Cohen, 1989).

More advanced ToM tasks were also designed for the ability to look for clues about mental states in people’s eyes. In a series of studies, high functioning adults and children with ASC and matched controls were presented with a task called ‘Reading the mind in the eyes’. Using strips of photos of the eye region only, participants were asked to pick a label that best describes the mental state of the person in the picture (e.g. scheming, preoccupied, decisive. The words in the children’s task were simplified). Figure 1.2 shows one item from this task. In all three studies, participants with ASC were significantly impaired in reading a person’s mental/emotional state from the eye region (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001). These findings show that the inability to use cues from the eyes for mentalising is impaired throughout the lifespan and throughout the autistic spectrum.

Figure 1.2: An item from the adult version of the ‘Reading the Mind in the Eye’ test

Another study tested the use of contextual cues for ToM in social situations, using a task called ‘the Strange Stories Test’ (Happe, 1994a). This assesses the ability to provide context-appropriate mental state explanations for non-literal statements made by story characters (e.g., white lies, sarcastic statements). Happe’s study with adolescents found specific deficits on this task, a result that was later replicated with adults with AS or HFA (Jolliffe & Baron-Cohen, 1999a).

In his book ‘Mindblindness’, Baron-Cohen summarised a decade of research into ToM in autism. Further to the findings presented above, he suggested that three of the major symptom groups in autism – the abnormalities in social development, communication development and pretend play – are a result of mindblindness - degrees of disability to mindread, to create theories about others’ minds (Baron-Cohen, 1995). If individuals with autism have some impairment in the ability to attribute mental states to others, to separate their knowledge from others’ or to use cues such as eye direction to interpret intentions, it is not surprising that they have difficulties engaging in conversations, that their communication style is very concrete, that they struggle with imaginary play and pretending, and that they fail to form social relationships.

The ToM deficit in autism received further support from several neuro-imaging studies, showing that individual with autism have lower activation than controls in brain areas that belong to the ToM network. This network, also known as ‘the social brain’, was first proposed by Brothers & Ring (1992). It includes the medial prefrontal cortex, which allows the representation of the mental as separate from the real, the anterior singulate and orbito-frontal cortices which, along with the superior temporal cortices and the amygdala, form a network of brain regions that implement computations relevant to social processes. The superior temporal sulcus (STS) and the temporal poles bilaterally are involved in processing explicit behavioural information such as the perception of eye direction and intentional behaviour (the STS) and the retrieval from memory of personal experiences (the temporal poles), both are required functions for successful mentalising. Other areas that are involved in ToM are the amygdala and the fusiform gyrus. The former is related to processing of emotional content and biological movement, and the latter to face perception (Baron-Cohen, 1995; Frith & Frith, 2003). Areas of the social brain are illustrated in figure 1.3.

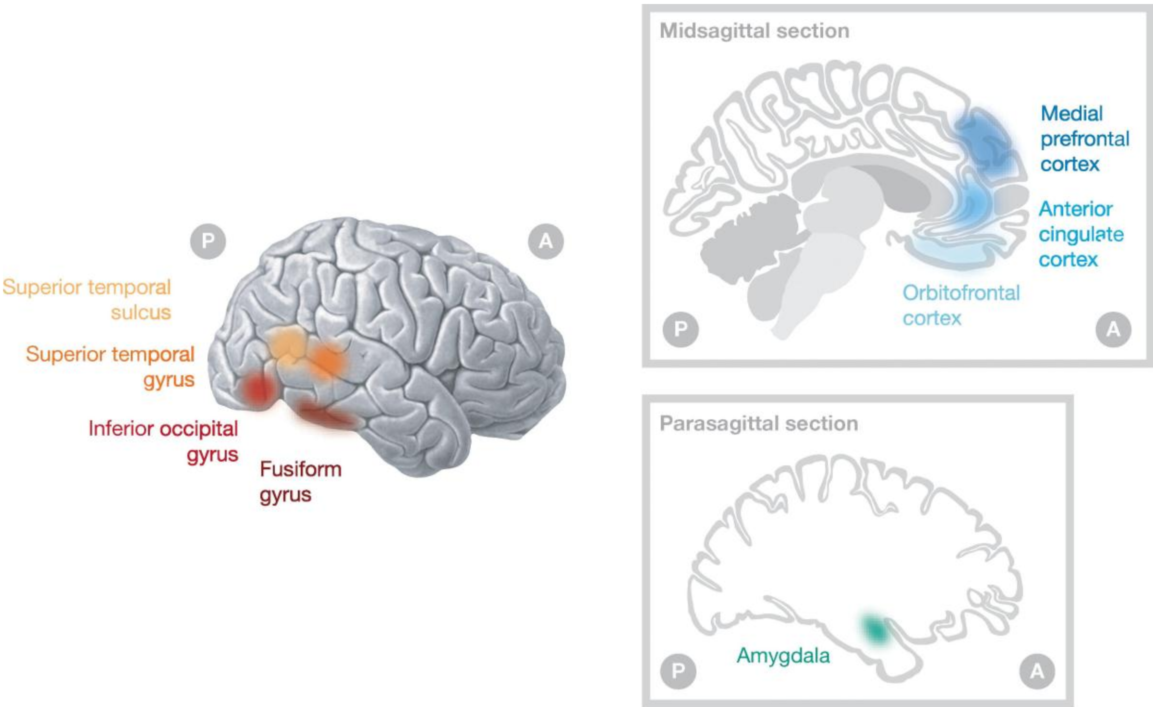


Figure 1.3: The Social Brain Network (adapted from Baron-Cohen & Belmonte, 2005)

Happe and colleagues conducted a PET study with participants with AS and matched controlls. Participants were asked to read stories which required mentalising and control stories which did not. When reading the mentalising stories, participants with AS showed no activation of the left medial prefrontal cortex, which was active among the controls (Happe et al., 1996). A similar study used PET to measure brain activity of participants with ASC and matched controls whilst listening to theory of mind stories. Both groups showed similar patterns of brain activity, but activation in the medial frontal area of the brain was less intensive and extensive in the ASC group, compared to controls (Nieminen-von Wendt et al., 2003). These findings support the localisation of ToM in the brain, with the left prefrontal medial cortex as a key region in a network employed during mentalising.

However, to rule out a verbal confound involved in these task, Castelli and colleagues conducted a neuro-imaging study using a non-verbal ToM task. Participants were PET scanned while watching animated sequences, depicting 2 triangles moving about on a screen in 3 different conditions: moving randomly, moving in a goal-directed fashion, and moving interactively with implied intentions. The last condition frequently elicited descriptions in terms of mental states that viewers attributed to the triangles (e.g. ‘the small triangle is pursuing the large one, but the large one is not interested’). The autism group gave fewer and less accurate descriptions of the mentalistic animations, but equally accurate descriptions of the other animations compared with controls. When watching the mentalistic animation, typical participants showed increased activation in the medial prefrontal cortex, the superior temporal sulcus and the temporal poles, whereas the autism group showed less activation than the control group in all these regions (Castelli, Frith, Happe, & Frith, 2002).

Specific mentalising brain regions were also found active when the ‘Reading the mind in the eyes’ test was used in a neuro-imaging study: Participants with ASC had less extensive frontal activation than controls, and no activation of the amygdala. The control group had a significantly stronger response in the left amygdala, the right insula and the left inferior frontal gyrus (Baron-Cohen et al., 1999). A related study using mental state vs. non-mental state words in an auditory paradigm involving SPECT found the orbito-frontal cortex (OFC) is another important region in the social brain, which is less active in autism, compared to controls (Baron-Cohen et al., 1994).

The literature reviewed above gives extensive support to the impaired ToM deficit of autism, both behaviourally and neuropsychologically. Nevertheless, the ToM hypothesis cannot explain all the characteristic phenomena in autism. In particular, it does not explain asocial aspects of autism. One example of this is the different visuospatial processing seen in autism. Individuals with autism tend to focus on particular objects of interest which engage them (e.g. spinning or mechanical objects, or even local features of these objects), while disregarding other stimuli in the vicinity. Another example is the superior attention to detail and the deficit in global attention (Plaisted, 2000). Finally, the ToM hypothesis has nothing to say about the repetitive behaviour patterns in autism (Happe, 1994b). The cognitive models of autism presented next aim to address these aspects.

1.3.2 Executive Dysfunction

'Executive function' (EF) is a broad term for cognitive functions underlying flexible goal oriented behaviour, such as planning, working memory, impulse control, inhibition and set shifting as well as the initiation and monitoring of action. Problems with EF, or Executive Dysfunction (ED) may cause marked difficulties in novel or ambiguous situations, while performance on routine, well-learnt tasks remains intact (Hughes, 2001). People with ED show rigidity and perseveration, being explained by poverty of initiation of new non-routine actions and the tendency to be stuck in a given task set. At the same time the ability to carry out routine actions can be excellent and is manifested in a strong liking for repetitive behaviour and sometimes elaborate rituals (Hill, 2004a). Neurologically, these functions are governed by frontal areas of the brain. Hence, ED can be found in patients with frontal brain damage, as well as in disorders with frontal lobe abnormalities such as ADHD, Tourette Syndrome, OCD and schizophrenia (Hill, 2004b). Several characteristics of autism spectrum conditions, such as repetitive behaviours, difficulties with flexible adaptation to novelty, and difficulties with planning, can all be related to ED. Individuals with autism usually require prompts and externally provided structure to help them switch to another activity, but perform routine activities very well. Hence, an executive dysfunction theory of autism was proposed (Ozonoff, 1995a; Russell, 1997). Although it is unclear whether ED is a primary cause of autism, or whether it is merely a result of associated neurological damage, it is nevertheless an important model, since it aims to explain non-social features of autism, which were not explained by the ToM model.

Experimental investigation of the ED theory of autism has explored various EF skills, often with conflicting results. Here I will review the functions in which deficits have been found in autism spectrum conditions.

Cognitive flexibility (also known as *set shifting*) is the ability to shift to a different thought or action according to changes in a situation. A lack of cognitive flexibility would result in perseverative, stereotyped behaviour. A task used in studies of cognitive flexibility is the Wisconsin Card Sorting Task (WCST). This task requires sorting of cards according to one of three dimensions – colour, shape or number. The rule, according to which the cards must be sorted is not revealed to the examinee, and changes occasionally. After each attempt, the examiner tells the examinee whether s/he has placed the card correctly (i.e., followed the correct rule), but does not give the examinee the rule explicitly. To succeed in this task, the examinee needs to use the examiner’s feedback in order to infer that the rule has changed and then to shift the response set. Perseveration on this task is viewed as a failure to shift set to the new sorting criterion. When compared to matched typical controls, high functioning adults with autism showed high levels of perseveration in their response to the WCST (Rumsey, 1985). They had difficulty in shifting to sorting according to another rule, and instead continued to sort using the first rule. Similar results were found with children with HFA, matched with children with learning disabilities (Ozonoff, 1995a; Ozonoff, Pennington, & Rogers, 1991), or with other neuro-developmental conditions (Ozonoff & Jensen, 1999). This perseverative pattern on the WCST was still manifested when the children with autism were assessed 3 years later (Ozonoff & McEvoy, 1994). Another study of lower functioning children with autism used a modified version of the WCST, in which all ambiguous cards were removed from the deck, and participants were explicitly told when to shift set. Despite the simplification and scaffolding, the autistic group made significantly more errors and perseverative responses than controls (Prior & Hoffmann, 1990). Though these responses on the WCST of individuals with autism have not been related to their perseverative tendencies in daily life, the findings suggest this may be the case.

Hughes and colleagues have used another test of set shifting with children and adolescents with autism, whose intelligence level varied from moderate learning disabilities to normal levels. They were compared to controls with learning disabilities and typically developing controls. The task, called the Intradimensional–extradimensional shift task, taken from the Cambridge Neuropsychological Test Automated Battery (CANTAB), uses two dimensions – shape and colour. It is presented in several stages, providing a more precise identification of the locus of difficulty on a set-shifting task than is possible using the WCST. Participants were asked to discriminate between one of two (or in later stages, one of four) stimuli, with shifts in discrimination required within-set or between-sets. Children and adolescents with autism were impaired in comparison to both of the comparison groups only when an extradimensional shift was required. This study suggests that it is not that autistic individuals perseverate generally, but rather that they experience an autism-specific pattern of perseveration (Hughes, Russell, & Robbins, 1994).

Contrary to these findings, several other studies had found no difficulties with set shifting in individuals with autism: Minshew and colleagues reported no difference between the number of perseverative errors produced by a group of high functioning adults and adolescents with autism compared to matched controls (Minshew, Goldstein, Muenz, & Payton, 1992). In addition, it was found that when the effects of full scale and verbal IQ are removed from the analysis, significant group differences on level of perseveration disappear (Liss et al., 2001; Rumsey, 1985). This suggests that perseveration tendencies may be related to verbal IQ or to understanding of the instructions and directions. In addition, Ozonoff found that children with autism performed significantly better on the computerised version of the WCST, than on the traditional version administered by an examiner. Her findings suggest that some of the difficulties seen in set shifting tasks, and indeed other tasks which require interpersonal interaction, may actually depend on the social impairment in autism (Ozonoff, 1995b).

Another skill included under EF which has been studied in autism is *planning*. Planning is a complex, dynamic operation in which a sequence of actions must be constantly monitored, re-evaluated, and updated. This requires the conceptualisation of changes from the current to the planned situation, looking ahead to identify alternatives, making choices, and then implementing the plan and amending it if required (Hill, 2004a). Planning ability is commonly studied using ‘the Tower of

Hanoi' or 'the Tower of London' tasks. On both tasks examinees are asked to move disks from a predefined sequence on three different pegs to match a goal state determined by the examiner. This should be performed in as few moves as possible and following a number of specific rules. The task requires comparing the current sequence with the desired one, planning the moves that the examinee wants to conduct, and monitoring the success of the solution s/he came up with.

As with set shifting tasks, when children and adolescents with autism were assessed, they perform significantly worse on these tasks, compared to matched typical controls (Ozonoff & Jensen, 1999), and worse than matched controls with other neuro-developmental diagnoses such as ADHD, Tourette Syndrome and dyslexia (Ozonoff & Jensen, 1999; Ozonoff, Pennington, & Rogers, 1991). In addition, a longitudinal study had revealed that this impairment was maintained 3 years after the original assessment (Ozonoff & McEvoy, 1994).

Hughes et al, in their study described above, included a computerised version of the tower task, called 'the Stockings of Cambridge'. The children and adolescents with autism performed significantly worse than the age matched learning disabled control group, and worse than the younger typically developing control group matched on mental age. However, similar to their findings on set shifting, the researchers found that the group with autism had a planning deficit only on the more complex puzzles, which required a large number of steps (4-5 steps), but not on the easier puzzles (2-3 steps; Hughes, Russell, & Robbins, 1994). This finding supports a planning deficit in autism, and since every day life requires quite complex planning (e.g. in schoolwork), this may explain why individuals with autism have difficulties coping with these life skills.

Interestingly, contrasting findings to the research described above were found in studies using other planning tasks. One example is the mazes task from the Wechsler Intelligence Scale for Children (WISC-R). Examinees are asked to draw a way out of a maze without crossing lines or going into dead ends, hence requiring pre-planning of the route. A study using this task with children with HFA found no difference between their performance and the performance of matched controls with language delay (Liss et al., 2001). In another study, using a kinematic reach-to-grasp task, which requires planning of motor movements, low IQ children with autism performed significantly worse than age matched typically developing controls, but also worse than high functioning children with autism, suggesting the planning deficit may be related to IQ rather than to autism (Mari, Castiello, Marks, Marraffa, & Prior, 2003). Studies conducted with pre-school children with autism, using simpler non-verbal EF tasks found no difference between their performance and the performance of matched developmentally delayed controls (Dawson et al., 2002; Griffith, Pennington, Wehner, & Rogers, 1999).

The results of these studies raise a question mark related to the tower tasks reviewed above. Tower tasks are quite complex and involve a number of other processes (e.g., working memory, inhibition of non-goal directed moves) rather then merely planning. Future studies should use more fine-grained tasks to understand planning abilities in autism beyond the effect of having low IQ. This point is further illustrated by the findings of a study of high functioning children with ASC, which tested the association of different ED and ToM tasks in autism. No significant correlations between the EF tasks, including a tower task, were found with the social and the repetitive behaviour clusters of autistic symptoms. Correlations with the communication cluster became non-significant when verbal IQ was controlled. In addition, when controlling for verbal IQ, no correlations between EF and ToM measures were found either (Joseph & Tager-Flusberg, 2004). These findings suggest that the ED model may be unable to explain autistic symptoms, or underlie ToM difficulties, when verbal IQ is partialled out. A study of very high functioning adults with AS (selected for their talent in mathematics and physics) using the Tower of Hanoi also showed no impairment on this task (Baron-Cohen, Wheelwright, Stone, & Rutherford, 1999), suggesting that an ED is not a core or universal impairment in ASC.

Unfortunately, there are hardly any neuro-imaging studies of executive functions in autism to help clarify the picture. One example of such a study used fMRI to investigate the performance of high-functioning adults with autism, compared to typically developed controls on an occulo-motor spatial working memory task and a visually guided saccade task. Compared to controls, the group with autism showed significantly less activity in dorsolateral prefrontal cortex and posterior cingulate cortex, suggesting abnormalities in neocortical circuitry in autism (Luna et al., 2002). Interestingly, Dawson et al found reduced performance on EF tasks that relate to the dorsolateral PFC among pre-schoolers with autism, suggesting difficulties employing this area for EF in autism (Dawson, Meltzoff, Osterling, & Rinaldi, 1998).

Another study investigated patterns of brain activation of adults with HFA and matched controls during a very simple visually paced finger movement task that requires planning, response inhibition, and execution of a prompted action sequence. Overall, the two groups used the same brain regions during the task. However, participants with autism showed less activation in perirolandic cortex and supplementary motor areas and greater activation in posterior and prefrontal cortices, compared to controls. They also showed greater than normal individual variability of the functional organisation for motor control in the brain. These findings suggest there is less distinct regional activation/deactivation patterns in autism, which may be related to the motor impairments in autism and to the disturbances of functional differentiation in the autistic brain (Muller, Pierce, Ambrose, Allen, & Courchesne, 2001). Difficulties with these basic functions, the authors argue, may underlie higher complex cognitive functioning demonstrated in the cognitive studies above.

More research of the ED model of autism would help determine the relevant brain circuits for the different functions pertained in this general term and amalgamate findings of low level difficulties with findings of complex executive functioning deficits found in autism. The model of reduced connectivity between brain regions in autism (Belmonte et al., 2004; Courchesne & Pierce, 2005) may help in bridging this gap. This model will be reviewed in the next section.

By pointing to potential areas of difficulty, especially in complex tasks, the executive dysfunction model of autism raises important questions about the methodology used in autism studies. These potential difficulties should be taken into account when testing individuals with autism, so that planning and mental flexibility requirements are set to minimum to avoid confounds. The ED model also helps to explain the repetitive rigid behaviour aspects of autism. However, it does not necessarily explain the social and communication deficits in autism, which are central. Since by definition the ED stresses *dysfunction*, i.e. the difficulties observed in autism, it also fails to explain the superior performance and isles of ability seen in autism spectrum conditions. The next two models will also aim to provide an explanation for these.

1.3.3 Weak central coherence

The weak central coherence (WCC) theory of autism was suggested by Uta Frith in her 1989 book ‘Autism: explaining the enigma’. She suggested that individuals with autism have a unique cognitive style, which focuses on the details and misses their context and the holistic picture. According to Frith, the natural perceptual-cognitive tendency to group details together and look for a higher holistic meaning, or a central coherence, is lacking in autism. Without the ability to group details in context and give them higher meaning, individuals with autism experience a fragmented world (Frith, 1989). Interestingly, the increased focus on details on expense of the whole was already noted by Kanner (1943), in his original description of autism.

Some experimental support for the WCC in autism has come from studies of reading: Frith and Snowling used homographs (words with one spelling, two meanings and two pronunciations) to examine the use of preceding-sentence context to derive meaning and determine pronunciation (e.g. ‘In her eye there was a big tear’; ‘In her dress there was a big tear’). If people with autism have WCC, then reading a sentence will be the same as reading a list of unconnected words, and sentence context will not help in clarifying the meaning of the ambiguous word. Results showed that children with autism failed to use the context of the preceding sentence to determine the pronunciation of homographs (Frith & Snowling, 1983). This finding was later replicated with high functioning adolescents and adults with ASC (Happe, 1997; Jolliffe & Baron-Cohen, 1999b). Such an inability to associate words with context is likely to lead to communication difficulties.

In a series of experiments conducted with adults with autism and AS, Jolliffe and Baron-Cohen found further support for the central coherence hypothesis. They presented participants with a sentence giving some contextual background for a situation (e.g. ‘Albert said he wouldn’t return to the restaurant’) and another sentence stating the outcome of the situation (e.g. ‘He left without giving a tip). Participants were then asked to choose from three optional explanations the one that explains why the outcome had occurred (e.g.' he was dissatisfied with the service' vs. 'the restaurant was closed when he arrived'). Typical control participants used the contextual information given in the first sentence to identify the most coherent inference to explain the outcome. However, participants with AS/HFA failed to make use of the context to draw a coherent inference.

A similar paradigm which used context sentences (e.g. 'John went to art class') followed by sentences that include words that have more than one meaning (e.g. 'he *drew* a gun') showed that participants with AS/HFA fail to make use of the contextual information to judge the appropriate meaning of the ambiguous word (Jolliffe & Baron-Cohen, 1999b). In another experiment, adults with AS/HFA performed significantly worse than matched controls on a task which required arranging sentences which include contextual cues in the right order to form a story. In comparison, no such group difference was found when the sentences included temporal, rather than contextual, cues (Jolliffe & Baron-Cohen, 2000). These findings show how difficult it is for individuals with ASC to draw together contextually related information and to construct overviews. This could explain some of the verbal communication problems that are characteristic to autism.

WCC in autism is not limited to the verbal domain. Individuals with ASC have difficulties processing detail into a whole in the perceptual domain too. In a study using a visuo-conceptual integration task, line drawings depicting simple objects had been cut into pieces and arranged in a puzzle-like fashion. Participants were required to mentally integrate the fragments in order to identify what each object would be if the pieces were correctly put together. High functioning adults with AS/HFA performed significantly worse than typically developed controls on this task (Jolliffe & Baron-Cohen, 2001a). In another task designed to examine visuo-conceptual integration, participants were required to visually integrate several objects so that they form the most coherent scene. To do this they were asked to identify the scene and to select which object is incongruent or odd for the established context (e.g. a nursery room scene with children and toys in it had a kitchen knife as the odd object). Adults with AS/HFA were significantly impaired in their ability to identify the types of scene and the odd objects. They were slower to respond and less accurate than controls

(Jolliffe & Baron-Cohen, 2001b). This demonstrates that individuals with an ASC struggle with integrating elements into a whole and to come up with higher order conceptualisation both visually and verbally.

The implications of such difficulties to social and communicational functioning in ASC are evident. Failing to pass a false belief task could be explained by an inability to integrate all the details of the social scene presented, to come up with the right prediction. Similarly, an inability to integrate relevant face features into a holistic concept could explain the difficulties on tasks which require reading cues from the face. Indeed, a recent study found that the inability of individuals with autism to process faces holistically was related to their configural style when processing objects (Behrmann et al., 2006).

The term *weak* central coherence stresses the negative aspect of this unique cognition style. However, it has some benefits, and could equally be labelled ‘strong attention to detail’. Indeed, studies that have assessed individuals with autism on tasks that require local rather than global attention had revealed that they actually perform better than typical matched controls. For example, ‘the embedded figures task’ (EFT) requires the examinee to find a target shape in a larger complex picture, that is, to disregard the overall gestalt of the picture and focus on its details. Children with ASC are more accurate on the EFT than IQ-matched controls (Shah & Frith, 1983) and adults with ASC have a shorter reaction time than typical controls on this task (Jolliffe & Baron-Cohen, 1997). Another example is the Block Design subtest of the Wechsler intelligence battery. This test of spatial ability requires copying an abstract pattern shown on paper, using coloured blocks. That is, it requires the ability to visualise a geometric gestalt as a combination of small constituent shapes, and to match those with the shapes on the blocks. Children with autism are faster and more accurate on this task, compared to typically developed controls (Shah & Frith, 1993).

O’Riordan and colleagues tested children with autism and matched typical controls on a variety of visual search tasks, which required finding a target amongst shapes varying in colour and orientation. Children with autism performed significantly better than matched controls. They showed an enhanced ability to discriminate between display items, which seemed to underlie their superior visual search (O'Riordan & Plaisted, 2001; O'Riordan, Plaisted, Driver, & Baron-Cohen, 2001).

Another superior aspect of local processing in autism is found with regards to visual illusions. On illusions such as the Müller-Lyer illusion, the Ebbinghaus illusion or the Ponzo illusion, participants are asked to compare the size of two features in two geometric drawings. Whereas the two features are actually identical, they appear to be different because of other features surrounding them. For example, in the Ebbinghaus illusion, two identical circles appear different in size since one of them is surrounded by larger circles (which make it appear relatively small), whereas the other is surrounded by smaller circles (which make it appear relatively big). When tested with visual illusions, children with autism were not biased by the surrounding features (i.e. by context) or by the need for global processing. They were able to focus only on the relevant details, and made accurate estimations with regards to the relative size of the shapes. In other words, they were not deluded by visual illusions (Happe, 1996). Similarly, Brosnan and colleagues, who studied the ability to process perceptual relationships between features, which could be perceived as a higher level gestalt, found that children with autism performed significantly worse on this task than controls. The authors concluded that children with autism fail to process inter-element relationships that would allow for the appreciation of larger perceptually coherent units that comprise of multiple elements and, consequently, context (Brosnan, Scott, Fox, & Pye, 2004).

Interestingly, several recent studies failed to replicate findings of perceptual WCC in ASC. In an experiment requiring participants to resize the compared features in four visual illusions, individuals with AS were found to have the same biases and make the same mistakes as controls, suggesting that this condition is not affected by WCC in the visual tract (Ropar & Mitchell, 1999). The researchers went on to use a battery of visuospatial tasks such as the embedded figures test, the Wechsler's block design, and the Rey complex figure test, but found no ASC superiority on these tasks and no correlation of the performance in these tasks with performance on visual illusions. They suggested that perception of illusions and performance on visuo-spatial tasks may rely on different mechanisms (Ropar & Mitchell, 2001). These findings show that the WCC model of autism still needs further clarifications, particularly in relation to reliability.

Some support for the WCC model may come from neuro-imaging studies. Models of brain functioning in autism suggest that WCC is the cognitive manifestation of altered connectivity between local ‘low-level’ perceptual brain systems and frontal brain regions in charge of integration or ‘coherence’. If brain regions that control integration are cut off from their perceptual inputs, the resulting cognitive-behavioural outcome would be enhanced local, on account of reduced global, processing (Baron-Cohen & Belmonte, 2005; Belmonte et al., 2004). In a neuro-imaging study using a visuo-spatial task which demands sustained, covert attention to lateral locations in the visual field brain functioning of adults with ASC was compared to matched controls. Individuals with ASC showed heightened ventral occipital brain activation and lowered pre-frontal, parietal, and temporal activation, compared to controls. This increased activation of early sensory visual areas in the autism group with reduced activation of pre-frontal areas, supports the above mentioned brain model of WCC (Belmonte & Yurgelun-Todd, 2003).

Another neuro-imaging study assessed the brain activity during verbal comprehension of active and passive sentences in high functioning adults with ASC and matched controls. Participants with autism engaged in more extensive processing of the meanings of individual words that comprise a sentence, as manifested in greater activation in Wernicke’s area. At the same time, the autistic participants showed less activation in Broca’s area, which is associated with semantic, syntactic and working memory processes, all of which serve to integrate the meanings of individual words into a coherent conceptual and syntactic structure. In addition, the group with ASC showed lower functional connectivity throughout the cortical language system, compared to controls (Just, Cherkassky, Keller, & Minshew, 2004). These findings of increased processing of units and decreased processing of integrative ‘high-level’ information, and the decreased connectivity between these areas provide further support to the WCC theory.

In a neuro-imaging study which examined the brain activity behind the superior performance of individuals with autism on the embedded figures test, Ring et al found a similar pattern to the one described in the studies above: Participants with AS/HFA showed lower activation than controls in dorsolateral prefrontal and parietal cortical areas of the brain, involved in working memory for objects and spatial relations. They also showed increased activation of ventral occipitotemporal regions, involved in visual object perception. These differences suggest that the normal strategy invokes a greater contribution from working memory systems when taking the task, while the strategy in ASC depends to an abnormally large extent on visual systems for object feature analysis, again stressing the local on account of the global (Ring et al., 1999).

The WCC model, though somewhat conceptually vague, has been very influential in explaining some of the difficulties in communication and in social functioning. It has also helped to highlight the strengths individuals with ASC possess and to provide an explanation to them. Next, I will review a model which takes this one step further, by integrating findings from all three models, proposing a distinction between strengths and difficulties in autism.

1.3.4 Empathising-Systemising

The Empathising-Systemising (E-S) model of autism emerged from the concepts of ‘folk psychology’ (the universal ability to understand the behaviour of other people in terms of their intentional states) and ‘folk physics’ (the universal ability to understand physical objects in terms of their causal/mechanical properties). The term ‘folk’ is used to suggest that the understanding of these domains is intuitive and develops without being formally taught (Pinker, 1998). The E-S model, and its extension in ‘the extreme male brain’ theory, makes use of these two factors and their distribution in the population to explain autistic phenomena, as well as sex differences in the general population (Baron-Cohen, 2002, 2003). I will start by defining each factor, followed by experimental evidence regarding its manifestation in ASC.

1.3.4.1 Empathising

Empathising is defined as the drive to identify emotions and mental states in others and to respond to them with an appropriate emotion (Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002). It includes two components – cognitive attribution and an affective reaction (see also Davis, 1983). The cognitive component of empathising is parallel to what has already been described above as ToM or mentalising, i.e. the ability to attribute mental states to oneself and others, as a natural way to understand the motives and behaviour of agents. Empathising expands beyond ToM as it also involves the ability and motivation to respond to the other's emotional or mental state with an *appropriate* spontaneous emotional reaction. The reaction is emotional (e.g. feeling sorry for a suffering person). It may involve an appropriate behaviour (e.g. helping or comforting the suffering person), but may also remain an internal affective reaction that is not externalised behaviourally. The appropriateness of the response is central to this part of empathising as it puts the response in context. Watching a person crying, for example, may lead us to interpret that this person is sad and respond with comforting. However, if we know that this person is only pretending to be crying in an attempt to trick another person, our affective response may be amusement (if we feel the trick was performed to entertain) or dismay (if we feel it was unacceptable). This example shows how flexible empathising is, and how both its cognitive and affective components depend on context (Baron-Cohen, 2003). Hence, empathising integrates components of ToM as well as central coherence (through context dependency) and a good empathiser would need to be able to use both. From that we can predict how difficult empathising is for individuals with autism spectrum conditions.

The centrality of an empathy deficit in autism spectrum conditions has been proposed before (Gillberg, 1992; Hobson, 1993; Wing, 1981). Baron-Cohen argues this deficit underlies symptoms in the three clusters of autism, accounting for the social deficits, socio-emotional communication difficulties, and limited imagination and pretence (due to inability to imagine others' minds). The empathising deficit in autism includes both the cognitive and the affective components, i.e. – both the understanding of others' intentions, emotions and beliefs, and the ability to come up with appropriate response (Baron-Cohen & Belmonte, 2005; Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002).

The development of the cognitive component of empathising in typical development and its impaired manifestation in autism was reviewed earlier under the ToM model, and will be further reviewed in chapter 2 with specific emphasis on emotion recognition. The affective component is somewhat harder to study, as it relates to a spontaneous internal reaction, which does not necessarily have a behavioural expression. However, its development can be observed from infancy through various behaviours, including affective reciprocity of facial and vocal emotional signals between the caregiver and the child, e.g. eye contact, social smiles (Bakeman & Adamson, 1984), the establishment of joint attention (especially social referencing; Tomasello, Carpenter, Call, Behne, & Moll, 2005), various aspects of pro-social behaviour (e.g. sharing, helping, offering comfort and affection), imitation, and imaginative play (Denham, 1998).

As with the cognitive component, the development of the affective component of empathising is impaired in autism. In a longitudinal study of high risk infants, who have siblings with a diagnosis of autism and of low risk infant controls, several socio-emotional features were lacking only in the high risk group, including eye contact, orienting to name, imitation, social smiling, reactivity, social interest and affect (Zwaigenbaum et al., 2005). Similar findings were found in a retrospective analysis of infancy videos of children who were later diagnosed with autism (Osterling, Dawson, & Munson, 2002).

Studies conducted with children with autism, matched controls with Down syndrome, and matched typically developing controls showed that compared to controls, children with autism were more ‘flat’ in their facial expressions (Yirmiya, Kasari, Sigman, & Mundy, 1989), and were impaired at sharing positive emotion in joint attention interactions (Kasari, Sigman, Mundy, & Yirmiya, 1990). Similar difficulties were found in the vocal expression of emotion among children and adults with AS/HFA (Paul, Augustyn, Klin, & Volkmar, 2005; Shriberg et al., 2001).

However, the lack of behavioural manifestation of the affective component of empathising could still mean that individuals with ASC have the affective response without externalising it. In order to bypass the expressed emotion deficit limitation, Yirmiya and colleagues attempted to measure affective empathy through self report. Adapting a paradigm previously used with typically developing children (Feshbach, 1982), the researchers played short video clips showing children experiencing different emotions to children with HFA and matched controls. Participants were asked to label the way the child in the film felt, and then to tell how they are feeling. Reports of own feeling that matched the label the child gave to the film character's feeling were considered evidence for affective empathy (even if they were mislabelled, i.e. failed to show cognitive empathy). Children with HFA gave fewer empathic responses compared to the matched controls (Yirmiya, Sigman, Kasari, & Mundy, 1992).

With high functioning adults with ASC, self report of affective (and cognitive) empathising can be collected using questionnaires. Adults with HFA rated themselves as experiencing lower levels of empathy, compared to matched controls from the general population. This was found on two different instruments: the Interpersonal Reactivity Index (Shamay-Tsoory, Tomer, Yaniv, & Aharon-Peretz, 2002) and the Empathy Quotient (Baron-Cohen & Wheelwright, 2004).

In addition to the social brain abnormalities reviewed in the section describing ToM, an interesting direction for studying the affective aspect of empathy comes from the neuro-imaging studies of 'mirror neurons'. These neurons were first discovered in the premotor frontal cortex of the macaque brain (Gallese, Fadiga, Fogassi, & Rizzolatti, 1996; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). They show activity in relation to specific actions performed by self as well as matching actions performed by others. Functional neuro-imaging experiments in humans, demonstrated that the neural circuit involved in action execution overlaps with that activated when actions are observed. This line of research studied the role of the mirror system in imitation of action, including facial expressions (Leslie, Johnson-Frey, & Grafton, 2004; Meltzoff, 2002), and in understanding intentionality (Blakemore & Decety, 2001; Williams, Whiten, Suddendorf, & Perrett, 2001). These findings suggest that mirror neurons also underlie representation of mental and emotional states in others, or in other words –

underlie empathy (Decety & Jackson, 2004; Gallese, 2003; Meltzoff, 2002). Since the activation of mirror neurons is automatic (i.e. does not require conscious decision), and since they exist in primates without imitative and ‘theory of mind’ abilities, mirror neurons are likely to be involved in the more automatic ‘emotionally contagious’ aspect of empathy, what Baron-Cohen and Wheelwright referred to as the affective component of empathising (Baron-Cohen & Wheelwright, 2004).

If mirror neurons have such a central role in imitation, understanding of intentions, and empathy, one can expect deviant functioning of this system in individuals with autism, who have difficulties with all of these. Indeed, several studies held so far with individuals with autism revealed differences in structure and functioning of the mirror system: Hadjikhani and colleagues conducted a structural analysis of brains of high functioning adults with ASC and matched controls. They found local decreases of gray matter in the ASC group in the inferior frontal cortex, the inferior parietal lobule, and the superior temporal sulcus. These areas belong to the mirror neuron system. The decreases were correlated with ASC symptom severity (Hadjikhani, Joseph, Snyder, & Tager-Flusberg, 2005). In an EEG study of the mirror neuron system, adults with ASC and matched controls watched a video of a hand in motion and performed the same movement themselves with their own hands. Whereas controls’ mirror neuron system responded similarly to the two conditions, the ASC group only showed activation in the self movement condition, supporting malfunctioning of the mirror neuron system in representing others’ movements (Oberman et al., 2005).

So far, only one imaging study tested the functioning of the mirror system in autism during a socio-emotional task: In this study, using fMRI, high functioning children with ASC and matched controls were shown 5 different emotional expressions and were asked to watch them and to imitate them. Unlike the typically developing children, the ASC group showed no activity in the mirror area in the pars opercularis of the inferior frontal gyrus, on either the imitation or observation conditions. Significant correlations were found between activity in this area and autistic symptom severity measures, in particular on the social cluster. Activity in the anterior component of the mirror neuron system was also significantly lower in the ASC group. In addition, the control group showed reliably greater activity in insular and periamygdaloid regions as well as in the ventral striatum and thalamus. In contrast, children with ASC showed greater activity in left anterior parietal and right visual association areas. The authors concluded that typically developing children can rely on a right hemisphere-mirroring neural mechanism interfacing with social brain areas such as the amygdala and the insula. This allows the meaning of the emotion to be directly felt in an empathic way and hence understood. In contrast, this mirroring mechanism appears to be lacking in children with ASC, who must then adopt an alternative strategy of increased visual and motor attention, since the internally felt emotional significance of the facial expression is probably not experienced (Dapretto et al., 2006).

Whereas this line of research has only just begun exploring empathy in typical development and in ASC, the preliminary findings seem promising and may help in better explaining both the affective and cognitive aspects of empathising. As shown on this last study, the findings of mirror system malfunctioning associate well with previous findings of social brain activation deficits in ToM tasks.

1.3.4.2 Systemising

Though the E-S model adds to our understanding of the socio-emotional deficit in autism by presenting the affective component of empathising, which was missing from the ToM model, the most novel contribution of this model is probably the formulation of the second factor it explores: systemising.

Systemising is defined as the drive to analyse and build systems, in order to understand and predict the behaviour of non-agentive events in terms of underlying rules and regularities. A systemiser seeks to analyse the system down to its lowest level of detail and identify parameters that may play a *causal* role in its behaviour. Systems can be of different nature: technical (e.g. the workings of a machine), natural (e.g. ocean tides), abstract (e.g. a mathematical model), taxonomic (e.g. criteria for arranging birds), musical (e.g. the structure of a symphony), etc (Baron-Cohen, 2003; Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002). Like empathising, systemising involves both the interest/motivation to understand systems, and the skill of understanding their rules and behaviours. However, the two vary on the object of interest: Whereas empathising revolves around agentive entities, systemising is more often (though not exclusively) focused on non-agents. Hence, whereas empathising is context dependent and unpredictable (as it depends on the intentions of agents), systemising is potentially accurate, rule-based, and predictable. Another way to describe the difference between the two is in terms of open and closed systems (Lawson, 2003): Systemising focuses on closed systems, so that when a systemiser becomes familiar with the rules that govern the system, s/he is able to predict its behaviour accurately. Implementing the rules of the system will always lead to the same outcome. Empathising, in contrast, deals with open 'systems', which means that no matter how well we know them (e.g. the ways humans behave in our society), there is always some ambiguity left, and other factors/contexts we cannot control or predict. This makes empathising less precise and more flexible.

Individuals with ASC tend to show good systemising skills. They are very attentive to details and prefer predictable, rule-based environments, features that are intrinsic to systemising. Anecdotal examples come from descriptions of savants with autism, which also revolve around their expertise in systems. Examples include individuals with autism who can give the day of the week to any given date within seconds, others who have learned to speak several languages in a short period of time, musicians with perfect pitch, who can play a musical piece perfectly after hearing it once, and painters who can draw exact copies of images they have watched in 15 minutes (Frith, 2003; Sacks, 1995).

But examples for enhanced systemising in autism are not limited to savants: Clinical descriptions of children with ASC denote their obsessive interest in systems from a very early age. A few examples include spinning objects (e.g. fans, washing machines), mechanical objects (e.g. trains), patterns (e.g. on pavement tiles, or on curtains), meteorology (obsession with the weather), geography (e.g. formation of mountains), astronomy (e.g. orbits of planets), and electronics (e.g. electronic circuits) (Attwood, 2003). Baron-Cohen and Wheelwright conducted a survey of circumscribed interests in children with ASC and in controls with Tourette syndrome. Compared to the controls, children with ASC were significantly more interested in areas relating to folk physics (i.e. to systemising) and significantly less interested in areas relating to folk psychology (i.e. empathising) (Baron-Cohen & Wheelwright, 1999).

Later in life, high-functioning individuals with ASC may use these good systemising skills in professional fields such as mathematics, physics, engineering and computers. (Baron-Cohen, 2003; Baron-Cohen, Wheelwright, Stone, & Rutherford, 1999). When the occupations of family members of individuals with ASC were compared to those of controls in the general population, a significantly larger proportion of fathers and grandfathers of children with autism were engineers (Baron-Cohen, Wheelwright, Stott, Bolton, & Goodyer, 1997). Similarly, the proportion of family members diagnosed with ASC was significantly higher among university students studying maths, physics and engineering, compared to humanities students (Baron-Cohen et al., 1998). Because of the genetic basis of autism, family members may share a 'broader phenotype' of autism, and may present with milder versions of ASC characteristics. Hence, in the absence of vocational studies of adults with ASC, these findings may offer an occupational direction high functioning children with ASC may be aiming to when they grow.

Experimental support for superior systemising in autism can be drawn in part from the experiments reviewed in the section on central coherence. In these experiments, individuals with ASC were found to be superior to typically developing controls on various tasks that involve scanning for details, analysing and manipulating systems. The examples, which were described in detail above include the Embedded Figures Test (Jolliffe & Baron-Cohen, 1997), the Wechsler Block Design subtest (Shah & Frith, 1993), and visual scanning tasks (O'Riordan, Plaisted, Driver, & Baron-Cohen, 2001; Plaisted, O'Riordan, & Baron-Cohen, 1998b). The WCC model used these examples to show the increased attention to detail in autism. The E-S model takes it one step further by arguing that great attention to detail is only one feature of the larger picture of systemising.

Experimental support for the rule-based system analysis feature of systemising, beyond the attention to detail level, was initially provided by the picture sequencing task described in the ToM section above. On the mechanical stories, depicting events caused by physics rules (e.g. a rock rolling down a mountain, hitting a tree and breaking it), children with autism performed significantly better than matched typically developed controls and matched controls with Down syndrome (Baron-Cohen, Leslie, & Frith, 1986). These findings suggest that even learning disabled children with autism systemise better than their non-autistic peers.

In addition, children with AS and matched typically developing controls were tested on a task of physics problems involving judgement of trajectories, forces, volume transformations, etc. These problems could be solved from everyday real world experience of the physical-causal world (e.g. on a picture of a well, with a rope and bucket attached to a lever on top of it, participants were asked which direction the handle needs to be turned in order for the bucket to go up). Children with AS performed significantly better then controls on this task (Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001). A similar study conducted with adult males with AS and with matched male and female controls used 'The Physical Prediction Questionnaire', which involves understanding physical systems. The task included mechanical diagrams and participants were asked to predict the movement of two levers or bobs in response to the movement of a connected lever. The AS and male control group performed equally well on this task, and significantly better than the female control group (Lawson, Baron-Cohen, & Wheelwright, 2004). Another study used a self report questionnaire to learn about systemising related behaviours and interests of adults with and without ASC. The questionnaire includes examples from everyday life in which systemising could be used to varying degrees, such as: 'When I listen to a piece of music, I always notice the way it's structured' or 'If I had a collection (e.g. CDs, stamps), it would be highly organised'. Adults with AS/HFA scored significantly higher on this questionnaire, compared to controls from the general population. (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003). Finally, on a physical equivalent of the false belief task (the false photo task), children with autism outperformed controls in predicting how a mechanical device (a Polaroid camera) would work (Leslie & Thaiss, 1992).

The E-S model of impaired empathising and enhanced systemising in autism offers an integrative explanation for many of the symptoms of ASC, some of which were not sufficiently explained by the cognitive models presented above. The socio-emotional deficits, which were explained by the ToM model, are explained in a similar way by the impaired empathising model, though this model manages to explain the lack of expressive emotional behaviours in ASC through the affective component of empathising. The need for accuracy in systemising can explain some of the pedantic language high functioning individuals with ASC possess, as well as the difficulty in understanding metaphors (which are not precise, e.g. when someone says they are ‘dying to go to sleep’ they are not really dying). Furthermore, systemising takes the executive dysfunction explanation of cognitive rigidity and repetitive behaviours in autism one step forward by looking at the *content* of these obsessional behaviours and suggesting their common denominator is an unusually strong focus on systems. Systemising also addresses the ‘need for sameness’ symptom of autism by suggesting this is caused by an attempt to hold the environment constant and to employ the known rules. For example, the difficulties in set shifting on the WCST may be related to insistence on the current rule-based system, which results in difficulties adapting to an unannounced change of rules. In the social world, individuals with autism may try unsuccessfully to employ strict rules like those characteristic of non-agentive exact systems with agentive, context-dependent, ambiguous social phenomena. For example, a child with ASC who learned that s/he should eat at the table, would have difficulties coping with a change of this rule on a picnic, failing to appreciate the different context, which justifies bending the rules.

The systemising model does not rule out executive dysfunction in autism, but places them in a secondary place, and relates them either to learning disabilities or to tasks which are hard to systemise (Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002).

Like the WCC model, the systemising model emphasises the great attention to detail in ASC. However, unlike the WCC, systemising does not see this as the main mechanism behind autistic cognition. Instead, it views attention to detail as the first stage of scanning for causality and rules in the environment in an attempt to come up with a complex systematic explanation to it. This is where these two models make different predictions: whereas the WCC model predicts individuals with ASC will fail to integrate details into a meaningful whole and are constrained to remain at the local level, the systemising model predicts that this integrative rule formation operation is not only possible, but is the core mechanism of the systemising process that guides the cognition of individuals with ASC. Hence, as long as a system can be defined according to exact rules, individuals with ASC will be able to comprehend it. A formulation that solves these conflicting predictions views great attention to detail with poor global coherence (the WCC stand) as a first stage in the development of the autistic mind. This stage is followed, when cognitive systems mature, by the urge to look for rules and create systems, including complex ones (the systemising stand). More developmental studies of systemising in the typically developing and autistic minds are needed in order to test the validity of this formulation (Baron-Cohen & Belmonte, 2005; Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002).

Though the research reviewed above supports the existence of impaired empathising and superior systemising in autism, further research is needed to support this model, in particular the systemising part, across different ages and levels of ability. More experimental support for superior systemising in systems other than physics would further test the E-S model. Finally, neuro-imaging studies suggesting a neural circuit for systemising are still required. Though findings of the superior local abilities presented in the WCC model section above support the systemising idea, there is still a need for neuro-imaging data of higher systemising functions such as rule-creation and system analysis.

Nevertheless, the E-S model presents a promising comprehensive explanation for autistic phenomena. In this thesis I make use of this model to formulate my questions and hypotheses. In particular, I test whether the good systemising skills in ASC can be harnessed to enhance one important aspect of cognitive empathising – the ability to recognise emotions and mental states in others. Would high functioning individuals with ASC be able to compensate for their poor empathising through systemising? This might be hard to implement, because the socio-emotional world is a context-related open system (Lawson, 2003), often unpredictable and difficult to analyse with strict rules. However, if provided with a *system* of emotions, it is plausible that systemising skills could be utilised to help them learn to recognise emotions. In the next chapters, I review the development of emotion recognition skills in typical development and in autism, and present the system of emotions which was evaluated in this thesis.

2 Emotion recognition

The ability to understand people's emotional and other mental states, what might be termed 'cognitive empathy', underlies social and communication skills (Baron-Cohen, 1995; Damasio, 1995; Dennett, 1987). Humans are skilled at integrating information from facial expressions, vocal intonation, context, and body language into a coherent picture of others' emotional states. During social interaction, we constantly monitor these states and adapt our social behaviour accordingly. This ability to detect, discriminate and respond to emotions, starts during the first years of life (Walker-Andrews, 1997) and continues developing through childhood, adolescence and adulthood (Harris, 1989). This chapter reviews the development of Emotion Recognition (ER) in typical development and in autism, beginning with the distinction emotion researchers make between 'basic' and 'complex' emotions. This distinction has affected much of the research conducted around emotion and the recognition of emotion. In addition, it has some interesting implications on studies of ER in ASC.

2.1 Basic and complex emotions

In an attempt to define what emotions are, philosophers, psychologists, and neuro-scientists have put forward a variety of explanations. These explanations refer to the neurological activity associated with the experience of emotion, the evolutionary account for why different emotions exist, the universality of emotions, the cognitive or information-processing level, and the structuring of emotions in different cultures (Griffiths, 1997).

2.1.1 Basic emotions

Darwin, who explored the expression of emotions in humans and animals, argued different species share the ability to experience and express certain emotions. This ability, he argued, is innate and can therefore be inherited (Darwin, 1872). According to Darwin, emotions were shaped through natural selection as they increased survival and hence reproductive fitness of the individual, enabling their genes to be passed on.

For example, the sight of another human/animal expressing anger or threat evokes the feeling of fear, followed by a fight or flight reaction. In this first example, perception of anger in others and experience of fear in the observer would have conferred an evolutionary survival advantage to the individual with this capacity. Equally, when a conspecific views the expression of fear in another, it evokes a similar feeling in the observer, which causes the observer to escape a potential approaching danger via vicarious exposure to the danger. This brief and effective means of 'communication' helps individual survival. Darwin's approach suggested that certain emotions are shared not only between humans, but also between other animals. Furthermore, these emotions are hardwired in activity patterns of the autonomic nervous system, and are thus evoked involuntarily. For example, when we see a snake approaching, our body starts reacting immediately and involuntarily. The series of actions includes increased heart rate, shallower breathing, sweating and 'butterflies' feeling in the stomach, all symptoms of fear, and shared by all humans and many animal species. The emotions shared between different species that have this clear neurological pattern, were named by the followers of the evolutionary approach 'basic' or 'primary emotions' (Damasio, 1995; LeDoux, 1998).

Darwin's argument about the existence of basic emotions which are shared by humans universally is supported by a seminal series of cross-cultural studies by Paul Ekman. Using images of posed Caucasian faces, Ekman and colleagues conducted multiple experiments (in Japan, New Guinea, Borneo, Brazil, and the United States) that sought to determine whether or not participants attribute similar emotional labels to the same images. Despite widely differing levels of literacy, exposure to media, and geography, it was found that between 70-80 percent agreement was achieved for six emotional expressions: anger, disgust, fear, happiness, sadness, and surprise. These emotions were also interpreted as being caused by highly similar situations, irrespective of culture (Ekman & Friesen, 1971; Ekman, Sorensen, & Friesen, 1969). This brought Ekman to label these emotions 'basic' (Ekman, 1999). All other emotions were argued to be subject to culturally specific "display rules", and hence were not universal or 'basic'.

Further support for the existence of discrete basic emotions comes from animal studies, neuropsychological, and neuro-imaging studies. These reveal brain areas associated with experiencing and recognising basic emotions. Recognition of basic emotions draws on a distributed set of structures that include the occipito-temporal neocortex, amygdala, orbito-frontal cortex and right fronto-parietal cortices. Recognition of fear relates especially to the amygdala (Adolphs et al., 2005; Calder, Lawrence, & Young, 2001; LeDoux, 1998), and the detection of disgust may rely on the insula and basal ganglia (Calder, 2003; Calder, Keane, Manes, Antoun, & Young, 2000; Phillips et al., 1998). Recognition of sadness was found to relate to the left amygdala and right temporal pole (Blair, Morris, Frith, Perrett, & Dolan, 1999), as well as the rostral supracallosal anterior Cingulate cortex (ACC) and the dorso-medial prefrontal cortex. Similar areas are associated with happiness, as well as the ventral tegmental area, which is part of the brain reward system (Lane, Reiman, Ahern, Schwartz, & Davidson, 1997; Murphy, Nimmo-Smith, & Lawrence, 2003; Phan, Wager, Taylor, & Liberzon, 2004). Anger recognition is associated with activity in the lateral orbito-frontal cortex and the ACC (Adolphs, 2002; Blair, Morris, Frith, Perrett, & Dolan, 1999). Perception of surprise has been found to be associated with activation in the medial temporal lobes, namely the right parahippocampal gyrus, an area associated with novelty detection (Schroeder et al., 2004). Although findings of neural networks for some of these emotions are still preliminary, these studies support the existence of discrete neural 'affect programs' for basic emotions (Griffiths, 1997).

Even if basic emotions are rooted in biology and evolution, they may not represent the full human emotional experience, which is much broader. However, non-basic emotions may lack clear neurological substrates. They may be less automatic, and may have a greater cognitive component. Because of their greater dependence on cognitive processes, these emotions, labelled 'complex' (Ben-Ze'ev, 2000; Griffiths, 2003) may be unique to humans.

2.1.2 Complex emotions

Various models have been created to systematically describe complex emotions. Damasio views complex emotions as subtle variations of the basic ones, e.g. *panic* and *shyness* are variations of *fear*. He argues that unlike the innate basic emotions, the complex variations are tuned by experience and would therefore vary according to the personal and cultural background of individuals (Damasio, 1995). Others described complex emotions in terms of basic emotion *blends* (Plutchik, 1980; Shaver, Schwartz, Kirson, & O'Connor, 1987). Using the analogy of colour mixing, Plutchik created a circle of basic emotions, in which mixing basic (or elementary) emotions create complex ones. The further away basic emotions are on the circle, the less likely they are to mix, and the more 'conflicting' is the emotion blend. For example, a blend of fear and surprise, which are close basic emotions, results in the complex emotion *alarm*. A blend of two distant emotions like joy and fear create a complex emotion akin to *guilt*. Blending basic emotions is argued to be a cognitive process (Plutchik, 1980). This idea of basic emotions underlying all complex emotions was used differently by Oatley and Johnson-Laird, who argued that basic emotions can be identified with a distinctive mode of information processing, built into our cognitive system. Complex emotions are cognitive elaborations of these modes, constructed from basic emotions plus appraisal judgements, i.e. cognitive representations. Because of their cognitive component, complex emotions are argued to be more amenable to cultural influence (Oatley & Johnson-Laird, 1987).

The psychologist Paul Harris suggested a distinction between situational (i.e. basic) and cognitive (i.e. complex) emotions. Situational emotions are capable of being triggered by situations alone (e.g. when I am watching my favourite TV show, I feel *happy*, when somebody changes the channel I feel *angry*), whilst cognitive emotions are primarily triggered by epistemic states (e.g. when I *expect* to see my favourite show and find out there is a special news programme instead, I feel *disappointed*). Hence, the difference between basic and complex emotions is in the involvement of cognitive components such as beliefs, intentions, or expectations (Harris, 1989). Griffiths, who extended Damasio's ideas, differentiated the rapid, automatic and stereotyped basic emotions, which he calls 'affect programs', from the 'higher cognitive emotions'. These emotions do not necessarily have stereotypic physiological activity associated with them. They are mental representations, which are much more associated with long term goals, desires and beliefs than the basic ones (Griffiths, 1997). Ben-Ze'ev extended this view by arguing that two properties distinguish complex emotions from basic ones: The attribution of intentionality, beliefs, desires and wishes to others, and the centrality of social emotions, which are both unique to complex emotions (Ben-Ze'ev, 2000).

Several attempts to contrast complex and basic emotions have emphasised the linguistic aspects. Ortony and colleagues explained this approach: “Emotions, of course, are not linguistic things. However, the most convenient access we have to them is through language” (Ortony, Clore & Foss, 1987, p. 342). Using this route, Johnson-Laird and Oately proposed a semantic analysis of the emotional corpora. Beyond the basic ones, emotions were construed as being either *relational emotions* [that are formed in relation to an object, e.g. love and hate], *caused emotions* [that are experienced for a known reason, e.g. enjoyment], *emotional goals* [that relate to a goal or desire, e.g. longing, disappointment], or *complex emotions* [that are experienced as a result of high-level evaluations of the self (e.g. shame, remorse) or the relation of the self to others (e.g. embarrassment, jealousy)]. Using these categories, Jonhson-Laird and Oatley analysed 590 emotions and mental states, and presented them in a taxonomy of emotional concepts in the English language (Johnson-Laird & Oatley, 1989). Ortony, Clore & Foss proposed their own general taxonomy of emotional words. Their componential analysis of the emotion corpora distinguished external conditions of affective words, which produce emotion labels according to an external description of a person (e.g. attractive) from internal conditions, which are directly related to the person’s experience. Internal conditions were subdivided into mental and non-mental emotion labels. Non-mental labels refer more to purely physical states (e.g. thirst), and are distinguished from actual emotion words. Within the mental domain, they allocated emotion words into five categories: *affective* words (e.g. sad), *cognitive* words (e.g. bored), *affective-behavioural* words, which have a behavioural component in addition to the affect (e.g. gloomy), *cognitive-behavioural* words (e.g. unfriendly), and *cognitive-affective* words (e.g. smug). Figure 2.1 presents Ortnoy et al’s model. Using these criteria, the authors created an emotion taxonomy of more than 500 words (Ortony, Clore, & Foss, 1987).

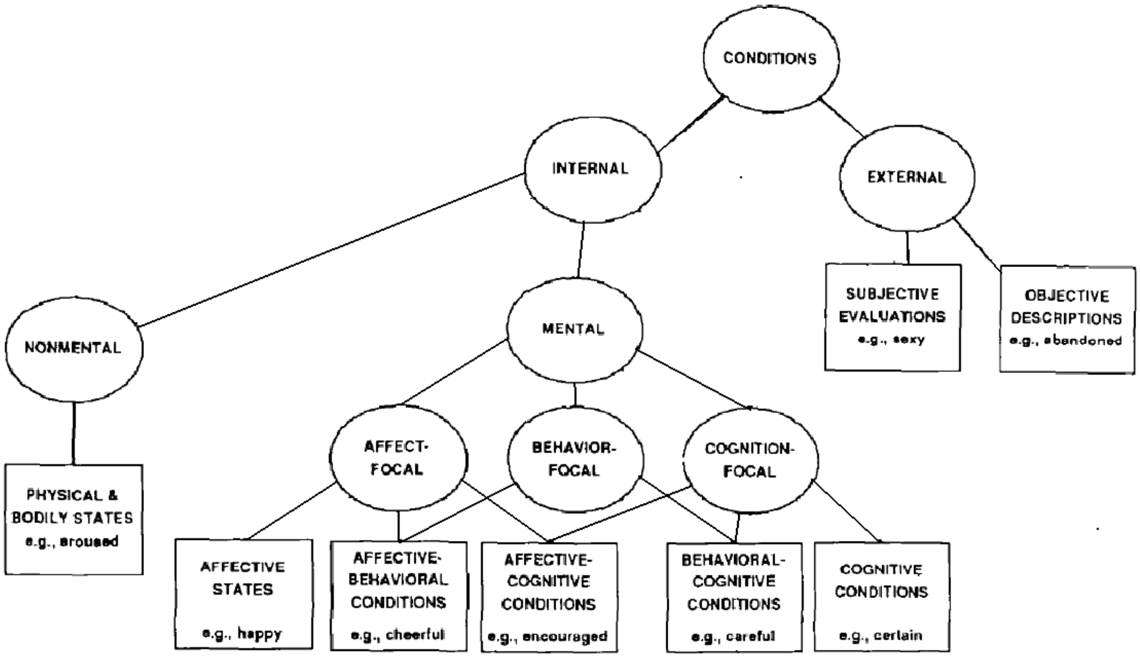


Figure 2.1: Structure of the affective lexicon (adapted from Ortony, Clore, & Foss, 1987)

2.1.3 Alternative views to the basic-complex classification

Not all schools accept the distinction between basic and complex emotions. One different view was suggested by James Russell and colleagues. Russell criticised the basic emotion model, and argued it lacked robust neurological support for distinct neural networks for each basic emotion (Posner, Russell, & Peterson, 2005; Russell, 2003). In his *circumplex* model of affect, Russell suggests that all emotional states arise from cognitive interpretations of core neural sensations from two fundamental neurophysiological systems: one is related to valence (a pleasure–displeasure continuum), and the other to the level of arousal, or alertness (Russell, 1980). Every emotion can be understood as a linear combination of these two dimensions, or as varying degrees of both valence and arousal. Happiness, for example, is conceptualised as an emotional state that is the product of strong activation in the neural systems associated with positive valence or pleasure, together with moderate activation in the neural systems associated with arousal. All other emotions arise from the same two neurophysiological systems but differ in the degree or extent of activation. Specific emotions arise out of patterns of activation within these two neurophysiological systems, together with cognitive interpretations and labelling of these core physiological experiences (Posner, Russell, & Peterson, 2005; Russell, 2003).

Another critical view of the basic-complex emotion idea comes from the social constructivist school, which argues that emotions are not automatic and involuntary, but are rather constructed of cognitive appraisals nested in behavioural scripts. These appraisals and scripts reflect the values of specific cultures and are unique to each culture. To refute the idea of biologically rooted basic emotions, social constructivists offer examples of emotions that are not associated with body states (e.g. guilt), and of the dependence of emotions which are considered 'basic' on cultural norms. For example, when anger is caused as a result of taking an offence, it requires a complex cognitive attribution (and so may be less automatic) and it differs between cultures, which may vary in whether an event is considered offensive (Prinz, 2004).

Whether they accept the distinction between basic and complex emotions or not, the centrality of cognitive processes for understanding of emotions is acknowledged by all schools. This, as well as the need for understanding others' mental states in complex emotions connects this field to 'theory of mind' (ToM) and cognitive empathy, described in chapter 1. It appears that successful recognition of complex emotions in others requires developed ToM abilities, whereas recognition of basic emotions may be possible even before the ability to mentalise. I will examine this in the next section, which reviews the literature about the development of ER abilities.

2.2 The development of emotion recognition skills

Although infants possess only rudimentary capacities to detect, discriminate, and recognise others' emotional expressions, they are born prepared to rapidly develop these skills during the first year (Walker-Andrews, 1997). Developmental studies of ER in infancy typically focus on the basic emotions of happiness, anger, fear and sadness. Studies have explored ER in visual and auditory channels separately, as well as using multimodal stimuli.

Soon after birth, neonates show a preference for human face-like stimuli (Johnson & Morton, 1991). During the first months of life, they distinguish between different emotional states of their caretaker, and respond to them differentially. In a study by Haviland and Lelwica, interactions of mothers with their 10 week old infants were videotaped. Mothers acted three facial and vocal expressions (happiness, sadness and anger) and their infants' facial behaviours were coded. By 10 weeks of age, infants respond differently to joy, anger, and sadness when the presentations are both facial and vocal. The infants could also mirror expressions of joy and anger (Haviland & Lelwica, 1987). Three to four month old infants already show sensitivity to adults' varying expressions, especially when they are dynamic, multimodal, interactive events (Walker-Andrews, 1997). Early sensitivity to the valence and intensity of emotional expressions was demonstrated when three month olds showed a preference for a smiling face over a neutral one, especially when the smile became more pronounced (Kuchuk, Vibbert, & Bornstein, 1986). Discrimination of emotional expression in the voice was shown by Walker-Andrews and Grolnick, who habituated 3- and 5-month old infants to either a woman's sad or happy voice, along with a slide of her face expressing the same affect. After the infants visually habituated, only the voice was changed either to happy or sad. The 3 month olds increased fixation to the picture when the sound changed from sad to happy, but not from happy to sad. The 5-month-olds dishabituated to both orders. This suggests that by 5 months, and possibly by 3 months of age, infants can discriminate between sad and happy vocal expressions (Walker-Andrews & Grolnick, 1983). Similar results were found using dynamic stimuli: Infants were shown several films of actresses facially and vocally depicting happy, sad, or angry expressions. Four month olds dishabituated to the novel expressions presented when changing from happy to sad. Seven month olds responded to differences between happy and angry expressions too (Caron, Caron, & MacLean, 1988).

By 7 months of age, infants can recognise that different examples of the same expression belong to the same category. When shown fear and happiness on faces of four different models using a paired-comparison procedure, infants were able to generalise their discrimination of these 2 expressions across the faces of the 4 models, if they were first presented with the set of happy faces. When presented with the fearful and happy faces in a discrimination task, infants preferred to look at the fearful faces, possibly because of the evolutionary importance of fear detection for survival (Nelson & Dolgin, 1985). At this age, infants can also detect incongruence between facial and vocal expressions of emotions: When presented simultaneously with two filmed dynamic facial expressions (from the set of happy, sad, neutral, and angry), accompanied by a single vocal expression that matched one facial expression, 7 month olds increased fixation to any of the facial expressions that were sound matched (Walker-Andrews, 1986; Walker, 1982).

At the end of their first year of life, infants become capable of social referencing, i.e. connect others' emotional expressions to environmental events (Walker-Andrews, 1997). For example, twelve month olds were presented with a set of novel toys and the infants' mothers were directed to pose happy, fearful, or neutral facial expressions. Infants remained closer to their mothers when they posed fear, stayed at a middle distance for neutral, and moved towards the toys when they expressed happiness (Klinnert, 1984).

As the studies above show there is evidence for the detection and discrimination of facial and vocal expressions of emotion during the first year of life. Infants' skills in categorising facial expressions and making inter-modal matches lead to their ability to use others' expressions to judge events. However, there are methodological problems in studies of infants: due to the absence of language, they are forced to use indirect dependent measures to assess ER (e.g., habituation and preference). These methodological discrepancies have led researchers to question whether the same construct of emotion expression recognition is being measured over development (when language is present or absent), and make it difficult to discuss continuity of these functions over development (McClure, 2000).

During the 2nd and 3rd years of their life, children become aware of the causes and consequences of emotions (Denham, 1998). They learn to recognise the distress of others and its causes, and engage in comforting behaviour on one hand, and in teasing and spiteful behaviour on the other (Harris, 1989). Developmental studies of ER in early childhood tested the children's ability to match verbal labels with facial and vocal emotional expressions, to match the expressions with each other, and to associate them with situations. For example, Widen and Russell tested 3-5 year olds on free labelling of basic emotion pictures. Children's emotional expression labelling increased with age in a systematic order: 3 year olds could label *happy*, *angry*, and *sad* expressions, 4 year olds were also able to label *fear*, and 5 year olds labelled *surprise* in addition to the former four. Recognition of *disgust* emerged later and was not consistently labelled by any of these age groups (Widen & Russell, 2003). A similar progression pattern has been found in other studies for basic ER from faces and from voices (Gross & Ballif, 1991; Stifter & Fox, 1987). When a non verbal paradigm was used, children as young as 2 years of age were able to sort photographs of facial expressions depicting emotions (e.g., happiness, sadness, anger) and physical states (e.g., sleepiness) into groups according to Russell's model's dimensions of pleasure and arousal (Russell & Bullock, 1985, 1986).

These findings suggest that even when verbal labelling is not fully developed, young children are able to discriminate facial and vocal expressions of emotion. In addition, it has been shown that younger children rely on facial expressions for information on another's emotional state to a greater extent than situational cues. A study exploring facial expressions and situational cues of emotion demonstrated that children's reliance on situational cues increased with age. Three to five year olds focused almost exclusively on facial expressions, whereas by eight or nine years of age, children relied additionally upon situational cues (Hoffner & Badzinski, 1989).

In the vocal channel, when 4-10 year olds were presented with conflicting emotional information in the linguistic and para-linguistic domains (e.g. a sad sentence uttered in a happy intonation), pre-schoolers relied almost exclusively on the linguistic content, whereas older children relied more on the speaker's intonation. However, pre-schoolers were able to judge the emotion from the intonation when the sentences were played in a foreign language or when the verbal content was masked (Morton & Trehub, 2001). This suggests that young children have difficulties in dealing with conflicting emotional information in the different channels. This ability emerges later, with the understanding of mixed emotions, display rules, and deception (Denham, 1998).

As described in Chapter 1, between 4-5 years of age, children become capable of understanding others' true and false beliefs (Flavell, 1999; Wellman, 1992). Harris argued that with this ability to imagine what others believe, want, know or feel, children become capable of understanding more complex, cognition based, emotions and develop cognitive empathy (Harris, 1989, 1994). In a series of cross-cultural studies in The United Kingdom, the Netherlands and Nepal, Harris and colleagues found that by 7-8 years of age, children become capable of understanding complex emotions that do not necessarily have clear facial expressions associated with them, and can give examples of situations that evoke complex emotions such as *pride, jealousy, gratefulness, worry,* and *guilt* (Harris, Olthof, Meerum-Terwogt, & Hardman, 1987). In addition to the development of ToM, the understanding of such social emotions require moral development, the comprehension of normative behaviour standards and the notion of self and other's responsibility for outcomes. For example, deliberately taking a younger child's toy evokes *guilt*, as this behaviour was intentional and contradicted social norms (Harris, 1989). Emotions such as *embarrassment*, *guilt*, and *shame*, sometimes referred to as 'self-conscious' or 'social' emotions actually require the ability for second order mental state reasoning, i.e. consideration of others' views of one's behaviour (Bennett & Matthews, 2000). The association of these emotions with ToM has also been demonstrated in brain imaging studies: reading descriptions of embarrassing and guilt provoking situations activated 'social brain' areas, which are involved in ToM such as the superior temporal sulcus, and medial prefrontal and lateral orbito-frontal cortices (Berthoz, Armony, Blair, & Dolan, 2002; Takahashi et al., 2004).

Throughout childhood, the accuracy and speed of ER improves (De Sonneville et al., 2002) children's emotional vocabulary expands, and they are able to recognise more subtle emotions and mental states (Vicari, Reilly, Pasqualetti, Vizzotto, & Caltagirone, 2000). However, empirical support for the development of complex emotion recognition has so far been limited to childhood (Herba & Phillips, 2004). The next section describes findings of a survey into the emotional vocabulary beyond early childhood.

2.2.1 Developmental changes in the emotion lexicon

Children's abilities to verbally label emotional expressions start developing from 2 years of age, with recognition of the basic emotions achieved by age 5 (Denham, 1998; Izard & Harris, 1995). An emotional vocabulary is available from quite an early stage: Bretherton and Beeghly interviewed mothers of 28 month old children about their vocabulary, and found that over 60% of the children were familiar with the emotional labels *happy*, *scared* and *mad* (i.e. angry), and were able to use them in their language. *Sad* was known to more than 50% of the children at this age, but more than 80% of them knew and used emotion words such as *like* and *love* (Bretherton & Beeghly, 1982). Another study reported that more than 75% of 3 year olds use emotion words for feeling *good*, *happy*, *sad*, *angry*, *loving*, *mean*, and *surprised* (Ridgeway, Waters, & Kuczaj, 1985). Toddlers first use emotion labels to refer to their own emotional state (at about 20–24 months). Later, such labels are used in reference to others' emotions, and finally, at about 3 to 3.5 years of age, children will use emotion terms for imaginary characters or in reference to those in a story (Reilly, McIntire, & Bellugi, 1990).

Whereas the development of basic emotion vocabulary has been extensively studied, studies describing the use of complex emotional and mental state language terms are scarce. There could be several reasons for the scarcity of evidence for development of complex emotion recognition. As shown above, complex emotions are considered to be culture and language dependent, and are not necessarily associated with unique facial expressions or with specific brain regions. Since classifications of complex emotions often make use of language based taxonomies, a possible route for studying them is the verbal route. However, studies describing complex emotions corpora have not tested the development of these emotions in the vocabulary of children or adults. One study had tested the development of complex (and basic) emotion labels in toddlers and pre-schoolers aged 18-71 months, using a list of 125 emotions (Ridgeway, Waters, & Kuczaj, 1985). However, this was not extended into primary school aged children and adolescents.

As part of a preparatory study towards the intervention project described in this thesis, a developmental survey of the emotional lexicon in school aged children and adolescents was conducted at the Autism Research Centre in Cambridge University (Baron-Cohen, Golan, Hill, & Wheelwright, submitted). This survey aimed to examine a much broader range of emotions and mental states in terms of when school children and adolescents understand these emotion words.

The emotion and mental state list used was defined through a search for all emotion related terms in the English language, using the electronic thesaurus in Microsoft Word. This revealed 1150 emotion words. An emotion was defined as a mental state with an emotional dimension that could be preceded by the phrase I feel x, or he/she looks x, or he/she sounds x (These criteria were adopted so that actors could portray these emotions for the *Mind Reading* intervention program on video and audio). Mental states that were epistemic with an emotional dimension were included (e.g. doubting). Mental states that were excluded were those that could be a purely bodily state (e.g. thirsty); slang words (e.g. chuffed); swear words (e.g. pissed off); and epistemic states with no emotional dimension (e.g. reasoning).

845 words identified in this way were then listed in a vocabulary checklist and groups of volunteers, stratified by age from the general population, were asked to state if they knew the meaning of each word. For individuals aged 11-18, self-report was accepted. For individuals aged 5-10 years old, parental or teacher report was collected. This questionnaire survey method allowed for 3 response options: *Clearly Understood*, *Not Understood*, and *Possibly Understood*. This method has been used in the Communication Development Inventory (Fenson et al., 1994) and found to be reliably correlated with comprehension tested in the lab. Only those items endorsed as *Clearly Understood* were judged to be within the comprehension of an individual. Data was collected by two school year bands (i.e. year 1+2, year 3+4, year 5+6, year 7+8, year 9+10, year 11+12 and 6 form college). These bands represent 7 age groups (5-6, 7-8, 9-10, 11-12, 13-14, 15-16, and 17-18 respectively). Sample size on the lower 3 bands (on which ratings were provided by parents) varied between 16-34 per band. On the other 4 bands (on which ratings were given by children), sample size varied between 42-134 per band. When applying a cutoff point of 70% of the age band for an emotion word to be satisfactorily recognised by an age band, the number of emotions

recognised was 51 for the year 1+2 band, 109 for year 3+4 band, 225 for year 5+6 band, 356 for year 7+8 band, 488 for year 9+10 band, 583 for year 11+12 band, and 702 emotion words for the 6 form college band, leaving 143 words that less than 70% of 17-18 year olds were familiar with. These findings show the gradual increase in the ability to label emotions and mental states along school years among English speakers in the UK. The full list of emotions and percentages of children recognised them is available elsewhere (Baron-Cohen, Golan, Hill, & Wheelwright, submitted). Findings of this survey were taken into consideration when creating the difficulty levels in the *Mind Reading* intervention program. Data from this survey was also used throughout the work described in this thesis, when selecting concepts and distracters for the various tasks used.

2.3 Emotion recognition in autism spectrum conditions

Emotion and mental state recognition are core difficulties for individuals with ASC (Baron-Cohen, 1995; Hobson, 1994). Such difficulties have been identified through cognitive, behavioural and neuro-imaging studies, and across different sensory modalities (Frith & Hill, 2004). Most ER studies carried out with individuals with ASC have focused on the recognition of the basic 6 emotion, either because of the universality of these emotions, or because of their clearer association with neurological routes found both in humans and animals. Studies assessing the recognition of these emotions report inconclusive findings in children and adults with ASC. Findings of the ER deficit in ASC become more robust when testing the recognition of complex emotions and mental states. Next, I review research findings of ER in ASC according to modality, and will discuss the possible reasons for this difference between basic and complex ER findings.

2.3.1 Emotion recognition in the face

Since the human face is so central in the expression and communication of emotion, the majority of ER studies have focused on the face. Several studies have shown that individuals with autism have face processing impairments, including impaired face discrimination and recognition, as well as emotional expression recognition difficulties. ASC are also characterised by atypical strategies for processing faces, characterised by reduced attention to the eyes and piecemeal rather than configural processing strategies.

In most ER studies of facial expressions, children and adults are shown photos or videos of the basic emotion facial expressions, and are asked to match them with a verbal label, other pictures of faces, or a matching vocal expression. In a series of studies conducted during the 1980', Peter Hobson and colleagues assessed the ER abilities of learning disabled adolescents with autism on 4 basic emotions (happy, unhappy, angry and scared). Performance of the autism group was compared to learning disabled controls without autism, and to typically developing controls matched on mental age. The group with autism was markedly impaired in matching drawn and photographed facial expressions with videotaped expressions and contexts (Hobson, 1986a), had difficulties matching facial expressions with emotional gestures (Hobson, 1986b), with vocal expressions on the six basic emotions (Hobson, Ouston, & Lee, 1988a), and with similar expressions in other faces (Hobson, Ouston, & Lee, 1988b), but performed as well as the controls on similar tasks involving objects instead of faces. This was also found with adult participants with autism (Macdonald et al., 1989).

Tantam and colleagues found that children with autism were significantly worse than controls at finding an odd facial expression of emotion out, and at labelling facial expressions of emotion (Tantam, Monaghan, Nicholson, & Stirling, 1989). Using a sorting-by-preference paradigm, Celani and colleagues tested basic emotional expression matching and identity matching among children with autism. To prevent participants from piecemeal processing of facial features in their answers, pictures were presented very briefly one after the other. Participants with autism performed significantly worse than matched typically developing and Down syndrome controls on matching of facial emotional expressions, but not on identity matching, suggesting the face processing difficulties in autism are specific to the emotional domain (Celani, Battacchi, & Arcidiacono, 1999). Similarly, Deruelle and colleagues, who tested recognition of emotion (happiness, disgust, and surprise), identity, gaze direction, and lip reading in children with autism, found poor performance in the autistic group on all measures except for identity matching (Deruelle, Rondan, Gepner, & Tardif, 2004).

Loveland and colleagues tested inter-modal ER on four basic emotions (happy, sad, angry, and surprised). They played videos of two facial expressions on a split screen and asked children with autism and matched controls with Down syndrome to choose the face that matches a vocal emotional expression played in parallel. The autism group performed more poorly than controls in detecting inter-modal correspondence of faces and voice (Loveland, Tunali Kotoski, Chen, & Brelsford, 1995).

Bormann-Kischkel and colleagues used Russel's circumplex model of emotions to test emotion recognition from facial expressions in children and adults with autism. *Excited*, *calm* and *sleepy* facial expressions were added to the basic 6 emotions. Participants were significantly less accurate than matched controls in choosing a facial expression to go with an emotional verbal label, but were as accurate as controls on a similar task involving colours rather than faces. In addition, there was no substantial difference between groups in their sorting of emotions according to pleasantness and arousal, though the judgments made by the autism group relied mostly on the lower part of the face (the mouth area). This has mostly affected misjudgements made on surprised faces (Bormann-Kischkel, Vilsmeier, & Baude, 1995). Despite being cross culturally recognised (and hence included in the basic 6 emotions), surprise has been argued to be more of a complex emotion, as it relies on false beliefs, and therefore requires ToM. Indeed, Baron-Cohen and colleagues found that children with autism, who can recognise happy and sad facial expressions, perform poorly on recognition of surprise (Baron-Cohen, Spitz, & Cross, 1993). However, Bormann-Kischkel et al.'s findings also suggest that the processing style of faces in autism is different to the normative style. This unique face processing style has received lots of research attention in behavioural, neuro-imaging, and gaze tracking studies.

Various studies showed individuals with autism process faces differently: In his early studies, Hobson described a lack of an 'inversion effect' in autism. Since the human brain is hard wired for perception of the face gestalt from birth, speed and accuracy of performance are hampered when faces are presented upside-down. However, such an effect was not found in the autism group, suggesting that the autistic brain does not treat (upright) faces as special (Hobson, Ouston, & Lee, 1988b). In addition, individuals with ASC tend to process faces in a feature-based approach, whereas controls from the general population process faces configurally (Young, 1998). For example, when tested on a task that required matching facial features in the context of a complete face, participants with autism were found to make more errors compared to a task in which the features were presented in isolation (Teunisse & De Gelder, 1994). When response time is not limited, individuals with ASC are significantly slower than controls in telling whether two face images are identical or not. This is related to their feature-based processing style, which is more time consuming (Behrmann et al., 2006).

In addition to piecemeal face processing, individuals with ASC appear to variably attend to different face parts: Joseph and Tanaka asked high functioning children with ASC and matched controls to match a target face with face parts. Whereas typically developing children performed better when relying on the eyes area of the face, the ASC group evidenced a whole-test advantage for mouths only, and was markedly deficient when face recognition depended on the eyes (Joseph & Tanaka, 2003). Deficient attention to the eyes in ASC were also reported in gaze tracking ER studies, using pictures of basic emotional expressions (Pelphrey et al., 2002), or ecologic social situations (Klin, Jones, Schultz, Volkmar, & Cohen, 2002a, 2002b), the latter also reporting extra attention to the mouth region. Lacking attention to the eyes in ASC has a clear association with poor ER.

The behavioural findings of a different face processing style in ASC are associated with similar reports from neuro-imaging studies of face processing. These report that participants with ASC show less activation in brain regions central to facial processing, such as the fusiform gyrus, and its area, also referred to as the Fusiform Face Area (FFA). Whereas the FFA was found to be extensively active when typically developing participants process emotions, individuals with ASC seem to use alternative neural sites to process faces such as the frontal cortex or the primary visual cortex (Critchley et al., 2000; Pierce, Muller, Ambrose, Allen, & Courchesne, 2001). Schultz and colleagues found that individuals with ASC use the inferior temporal gyri to process facial expressions, an area used by the typically developing controls to process objects. In addition, individuals with ASC demonstrated a pattern of brain activity during face discrimination that is consistent with feature-based strategies that are more typical of object perception (Schultz et al., 2000). Another study reported a strong positive correlation between the time spent fixating on the eyes in a face processing task, and activation of the FFA and the amygdala among individuals with ASC, suggesting that diminished gaze fixation may account for the fusiform hypoactivation to faces commonly reported in autism (Dalton et al., 2005).

The centrality of the amygdala in ER from faces has been widely discussed (Adolphs, 2002; Adolphs et al., 2005; Haxby, Hoffman, & Gobbini, 2002; Narumoto et al., 2000; Phillips, Drevets, Rauch, & Lane, 2003). The amygdala has been shown to play a critical role in the early stage processing of facial expression. It is a structure that quickly reacts to emotionally relevant stimuli and events (LeDoux, 1998). Due to its critical role in emotional arousal, it mediates the formation of emotional learning (Schultz, 2005). Studies of individuals with amygdala lesions show their ability to recognise emotions from facial expressions has been seriously hampered (Adolphs, Baron-Cohen, & Tranel, 2002; Adolphs & Tranel, 2003), with fear recognition suffering the greatest deficit (Calder et al., 1996). Similarities in social deficits and ER difficulties between individuals with amygdala lesions and individuals with ASC have given rise to the amygdala theory of autism, stressing the centrality of the amygdala in the socio-emotional deficit in ASC (Baron-Cohen et al., 2000; Howard et al., 2000). As described in Chapter 1, Baron-Cohen and colleagues found no amygdala activation in adults with ASC during the 'Reading the mind in the eyes' task, compared to significant activation in matched controls. These findings, using a task that relies on attention to the eye region, stress the importance of the amygdala for ER and mentalising, and the impact of its hypoactivation on socio-emotional functioning in ASC (Baron-Cohen et al., 1999). Lack of left amygdala activation among individuals with ASC was also found in the study of Critchley et al, who presented participants with photos of basic emotion facial expressions (Critchley et al., 2000). Howard and colleagues found recognition difficulties of fear (but not other basic emotions), which was associated with enlarged amygdalae in the ASC group, suggesting that the functional deficit is related to structural abnormalities (Howard et al., 2000). Other studies showed no variation in amygdala activation in ASC, in response to different intensities of fearful facial expressions which occurred with the typically developing controls (Ashwin, Baron-Cohen, Wheelwright, O'Riordan, & Bullmore, in press), and abnormal functional connectivity of the amygdala with other medial temporal lobe structures in people with AS during fearful face processing. (Welchew et al., 2005). The clear association between fear recognition and amygdala activation caused brain imaging studies of autism which involve the amygdala to focus on the detection of fear. Association between amygdala dysfunction and other emotions in ASC requires further investigation.

In contrast to the findings reported above, there are a number of studies conducted with children and adults with ASC that have failed to find ER difficulties in facial expressions of basic emotions. For example, Grossman and colleagues found no difference between children and adolescents with AS and matched controls on emotion labelling of facial expression photographs of five basic emotions (happy, sad, angry, afraid and surprised) (Grossman, Klin, Carter, & Volkmar, 2000). Gepner and colleagues found no difference in performance of children with autism and controls who were asked to match still and dynamic facial expressions of four basic emotions (happy, sad, surprise and disgust) presented on video, with photographs of emotional expressions (Gepner, Deruelle, & Grynfeltt, 2001). Castelli used facial expression images of the 6 basic emotions, which were 'morphed' to represent various levels of emotional intensities. She asked children with autism and matched controls to match these images with prototypical photos of the basic 6 emotions, and to verbally label the emotions. Results revealed that children with autism were as able as controls to recognise all six emotions with different intensity levels, and that they made the same type of errors (Castelli, 2005). A gaze tracking study using facial expressions of three basic emotions (angry, happy, surprised) and a neutral expression, found that children with autism had no difficulties labelling the emotions, and had the same fixation patterns as typically developing children (van der Geest, Kemner, Verbaten, & van Engeland, 2002). A lack of ER deficit on photos of facial expressions of the basic 6 emotions was also reported in adults with ASC (Adolphs, 2001; Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997).

A neuro-imaging study of high functioning adolescents with ASC found no differences on labelling or expression matching of fearful and angry photos in the behavioural level. The ASC group showed significantly less activity than the typically developing group in the FFA, but amygdala activation was not related to task performance in the ASC group (Wang, Dapretto, Hariri, Sigman, & Bookheimer, 2004). These findings show that high functioning individuals with ASC may be relatively unimpaired in ER of basic emotions, yet still show differences in the automatic processing of facial expressions. Similarly, in the mirror neuron study by Dapretto et al, described in Chapter 1, no ER differences were found between the children with ASC and their matched controls at the behavioural level, though clear group differences were found in activation of brain areas (Dapretto et al., 2006).

These findings with regards to ER in ASC can be explained in different ways. One possible explanation, following the ToM deficit theory, is that individuals with ASC can recognise situation-based emotions (such as happiness or fear) but fail to recognise belief-based emotions, due to their deficit in attributing mental states to others. As mentioned above, Baron-Cohen and colleagues compared recognition of situation-based happiness and sadness with belief-based surprise and found that compared to learning disabled and typically developing controls, matched on mental age, children with autism had no difficulty in recognising sadness or happiness, but found it harder to recognise surprise (Baron-Cohen, Spitz, & Cross, 1993). However, other studies (e.g. Castelli, 2005) found no ER deficit in recognition of surprise as well as the other basic emotions. Another explanation could be that individuals with ASC struggle to recognise social emotions specifically, such as pride or embarrassment (Hillier & Allinson, 2002; Kasari, Chamberlain, & Bauminger, 2001). The relative success with basic ER in ASC may also reflect that such individuals (especially those with normal intelligence) may be using compensation strategies to bypass their difficulties. For example, Grossman and colleagues showed children with AS pictures of 5 basic emotions, with matching or mismatching labels. The children with AS had no problem recognising these emotions or identifying the emotions when labelled with matching words, but (unlike the controls) had difficulties in recognising the emotions in the mismatching labels condition. This result suggests that, instead of recognising the emotions in the face, the children with AS were using the written label to answer the question. The authors concluded that individuals with AS use verbal mediation as a compensatory strategy, which may mask their deficits under certain circumstances (Grossman, Klin, Carter, & Volkmar, 2000). The use of compensatory strategies was also found in the imaging studies mentioned above, where activation in atypical, and lack of activation in typical brain regions was found, even when no behavioural deficit was reported. Since recognition of basic emotions from facial expressions can be seen in 5 year old typically developing children, it is plausible (as discussed in Chapter 1 with regards to false belief tasks) that older children and adults with ASC have gained enough experience to develop compensatory strategies for recognition of these emotions. However, such strategies may not suffice when recognition of more complex emotions is required.

Findings regarding complex ER in ASC are more consistent. Capps and colleagues asked children with high functioning autism (HFA) to label emotions from photographs and to provide examples of situations that made them feel this way. Compared to matched controls, children with HFA found it harder to recognise and explain pride and embarrassment. No group difference was found on expressions of sadness or happiness (Capps, Yirmiya, & Sigman, 1992). Adolphs and colleagues reported that high functioning adults with ASC, who had intact basic ER, experienced difficulties when asked to judge trustworthiness and approachability of people from photographs of their faces (Adolphs, Sears, & Piven, 2001). The social impact of difficulties understanding such a fundamental aspect of non-verbal communication is obvious.

Baron-Cohen and colleagues presented high functioning adults with ASC with photographs of basic and complex emotions and found the ASC group did not differ from controls from the general population on recognition of the basic emotions, but scored significantly lower on recognition of complex emotions (Baron-Cohen, Wheelwright, & Jolliffe, 1997). The group difference became more prominent when only the eye region was presented. As described in Chapter 1, children and adults with ASC had difficulties recognising complex emotions from pictures of the eyes only, when tested on the ‘Reading the Mind in the Eyes’ task (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001).

Most tasks of ER from faces, use still pictures rather than motion. The result is a relatively narrow range of complex emotions which can be studied, as the distinction between many emotions and mental states requires motion (e.g. relief). In addition, this makes the tests less naturalistic, and therefore may not assess an individual's actual ability to identify emotions from faces in motion. Appendix 1 reports tasks created for this study, to assess basic and complex ER from videos of full faces in motion, in children and in adults.

2.3.2 Emotion recognition in the voice

Emotion recognition from voices has been studied less frequently. Here too there are contradictory findings in relation to recognition of basic emotions: some studies found ER difficulties in children and adults with ASC when matching emotional voices with faces (Hobson, 1986a; Hobson, Ouston, & Lee, 1988a; Loveland, Tunali Kotoski, Chen, & Brelsford, 1995), whereas others showed no difference in face-voice matching of emotional expressions between participants with ASC and matched clinical groups (Loveland et al., 1997). Hall and colleagues tried to increase the salience of emotional faces of 4 basic emotions (happy, sad, surprised, and angry) using supporting prosodic information, by concurrently presenting participants with sounds of prosodic voices and pairs of facial expression photos. Participants were asked to match the emotion in the voice with the corresponding facial expression. Their findings showed that not only did the prosodic information fail to facilitate performance of participants with ASC, their performance actually decreased due to the presentation of stimuli in both channels (Hall, Szechtman, & Nahmias, 2003). However, with multi-modal studies, it is difficult to tell whether the deficit lies in vocal channel, in the visual channel, or in the integration of the two. Boucher and colleagues assessed children with autism on vocal emotion labelling and vocal-facial emotion matching of basic emotions. They found the children with autism were impaired on face-voice matching, but not on labelling of basic emotions from voices, suggesting the integration of facial and vocal stimuli hampers the performance of individuals with ASC (Boucher, Lewis, & Collis, 2000).

Very few of the studies investigating emotion recognition in ASC have studied recognition of emotions and mental states solely from vocal stimuli. Studies conducted with participants from the general population have evaluated use of the linguistic content of vocalisations, versus paralinguistic features when recognising emotions. These two features of emotional speech are processed separately in the brain, verbal content being processed in the left hemisphere and intonation being processed in the right hemisphere (McNeely & Parlow, 2001; Wildgruber et al., 2005). As described above, perception of emotional content in the voice is evident around age 4-5 months of age, with greater proficiency in relying on intonation for emotion recognition the older the child is (Morton & Trehub, 2001). In the general population, the processing of these two aspects of emotional speech and the use of both to label the speaker's mental state develop fairly early.

In individuals with ASC, however, difficulties using intonation and pragmatic/emotional stress in speech to make socio-emotional judgments have been reported, such as telling whether the speaker is calm or excited, or whether s/he is talking to a child or to an adult (Paul, Augustyn, Klin, & Volkmar, 2005). It is possible that in an emotion recognition task which includes verbal content and intonation, individuals with ASC will focus only on the linguistic content.

The studies assessing emotion recognition from voices described above focused only on basic emotions. Only two studies have been conducted assessing complex emotion recognition from voices in ASC: Kleinman and colleagues created an advanced ToM task that was based on the 'Reading the Mind in the Eyes' task but in the vocal domain. Their vocal task tested the recognition of 6 basic and 6 complex emotions, using a neutral sentence read with different intonations and found difficulties at both levels among individuals with ASC, compared to matched controls (Kleinman, Marciano, & Ault, 2001). Rutherford and colleagues created the 'Reading the Mind in the Voice' task, in which short segments of speech, taken from different BBC dramas, were played to participants, who were asked to choose one out of two possible words, each describing the speaker's possible mental state. Compared to controls from the general population, high functioning adults with ASC were deficient in their ability to recognise complex emotions and mental states (Rutherford, Baron-Cohen, & Wheelwright, 2002). Rutherford et al's task included only two possible answers per question and suffered sensitivity and ceiling effect problems. Appendix 2 presents a revised and improved version of the task, created for this study. In addition, Appendix 1 describes two new vocal tasks, assessing basic and complex ER in children and adults with ASC. These tasks allow to test recognition of individual emotional concepts, and are matched with the facial expression ER tasks mentioned above.

Imaging studies of vocal perception in ASC are also scarce, perhaps because the sound of the scanner itself (such as the echo-planar system) confounds auditory brain scanning. In the study described in section 1.3.1.2 (Nieminen-von Wendt et al., 2003), participants listened to ToM studies in the scanner. The design of this study does not enable us to tell if the differences in brain activation found between the ASC and the control group stem from difficulties attributing mental states, or if they could be due to the perceptual (auditory) channel selected. Indeed, Gervais and colleagues found that the areas in the superior temporal sulcus (STS), which are considered the auditory equivalent of the FFA and respond to sounds of voices in the general population, show no response to vocal sounds among adults with ASC. A normal activation pattern was found in response to non-vocal sounds (Gervais et al., 2004). These findings suggest that, as with the human face, the autistic brain does not specialise in processing the human voice and does not prioritise it as more salient than other sounds. In their research review of the vocal processing in ASC, McCann and Peppe found that research into prosody in ASC is an under researched area, and that existing studies have covered mostly prosodic expression, rather than comprehension (McCann & Peppe, 2003). More behaviour and imaging research into the recognition of emotion from voices in ASC is required.

2.3.3 Emotion recognition from context

As described in Chapter 1, emotions and mental states are felt and expressed in context. Context often gives information about the causes of emotions, or additional information needed to understand a basic expression in terms of a complex emotion. For example, a sad facial expression would be recognised as disappointed, if the context reveals that a failed expectation preceded it. Other examples include the use of sarcasm, deception, mixed emotions, or politeness (Ben-Ze'ev, 2000). Hence, decontextualised facial and vocal expressions are often insufficient for effective ER.

Studies assessing the ability of individuals with ASC to identify emotions and mental states from the context in which they are evoked have also shown deficits relative to typically developing controls or other clinical groups. As with ER from faces and voices, some studies reported difficulties with understanding of the causes and context of basic emotions (Fein, Lucci, Braverman, & Waterhouse, 1992; Hobson, 1986a), though findings of complex emotions and mental states studies were more robust. For example, in Baron-Cohen et al's picture sequencing study, described in section 1.3.1.2, the children with autism were able to sequence and explain the behavioural stories, even when situation-based basic emotions were included (e.g. a child running, falling and hurting her knee will be sad), but they had difficulties sequencing the mentalistic stories depicting surprise, which required understanding of false belief (Baron-Cohen, Leslie, & Frith, 1986). Similarly, when individuals with autism were tested on comprehension of situations, desires, and beliefs as causes of emotion, they were deficient compared to typical and learning disabled controls only in comprehension of belief-based emotions (Baron-Cohen, 1991). Another example is the study by Capps et al, described above, which showed that high functioning children with ASC were as able as controls to report situations that evoke happiness and sadness, but had difficulties describing situations that evoke pride and embarrassment (Capps, Yirmiya, & Sigman, 1992).

Deception usually involves concealing or altering one's expressions of emotion. Indeed, the understanding of deception in autism has been found to be deficient, compared to controls matched on mental age (Baron-Cohen, 1992; Sodian & Frith, 1992). Dennis and colleagues tested the understanding of deception in children with autism and AS and with matched controls using 'The Real and Deceptive Emotion Task'. In this task, children were presented with short narratives and with a scale of schematic facial expressions from happy to sad. The narratives described situations that involved or did not involve deception (An example of a situation that involved deception: 'Terry has a tummy ache, but he knows that if he told his mother about it, she wouldn't let him go out to play'). The children were asked to use the expression scale to tell how the protagonist is feeling inside, and how would his face look like, and explain the reason for their choices. Children with ASC could label the standard facial expressions but were less able than the controls to indicate the real emotions story characters feel, the deceptive emotions they express in the face, or the social reasons prompting a deceptive facial expression (Dennis, Lockyer, & Lazenby, 2000). This study demonstrates how the basic emotions could become too complex for individuals with ASC to understand when a context that requires mentalising is introduced. Similarly, Happe's Strange Stories Test (described in section 1.3.1.2) revealed a deficit in children and adults with ASC at providing context-appropriate mental state explanations for non-literal utterances made by story characters (Happe, 1994a; Jolliffe & Baron-Cohen, 1999a).

Baron-Cohen and colleagues created the Faux Pas test, which comprises short stories during which one character unintentionally says something they should have not said (e.g. ask a person standing in the restaurant to wipe the table, mistaking him for a waiter). Based on the context, participants were asked to tell what the character believed/knew when saying what s/he said (in the example above – 'Did the speaker think that the other person was a customer?'). Children with AS/HFA performed significantly poorer than control on this task, but were equal to controls on answering factual (rather than mental) questions on control stories (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999). Stone and colleagues used the Faux pas test in a neuropsychological study with patients with bilateral orbito-frontal lesions, who performed similarly to individuals with ASC, suggesting this area of the social brain is involved in this complex analysis of mental states (Stone, Baron-Cohen, & Knight, 1998). A similar instrument, using more subtle examples was created by Lawson and colleagues for adults. 'The Social Stories Questionnaire' contains 10 short stories and involves utterances made by one character that could upset another character in the story. Participants had to judge whether the section contained a potentially upsetting utterance and to judge whether this utterance (if present) would have upset the character concerned. Adults with AS performed significantly poorer than males and females from the general population on this task (Lawson, Baron-Cohen, & Wheelwright, 2004).

The findings presented above show that difficulties understanding complex emotions and mental states appear across all modalities, when presented separately. However, in reality, we are usually provided with multi-modal information on the different channels. The integration of this information eases the successful recognition of emotions and mental states. In the next section, I review studies using more ecologic, life-like stimuli, assessing whether the availability of multimodal emotional information can be utilised by in individuals with ASC in their ER.

2.3.4 Multi-modal emotion recognition

Judging complex emotions requires integration of multimodal information, including contextual information, prosody, and nonverbal visual cues (body postures and facial expressions) into a coherent holistic picture (Herba & Phillips, 2004). The Weak Central Coherence theory of autism would predict that individuals with ASC would experience difficulties utilising this breadth of information, and may actually find it harder to process than emotional information presented on a single perceptual channel. According to the empathising-systemising model, it is the imprecise nature of emotional information that hampers people with ASC as imprecision means such information is less easily systemised. As described in Chapter 1, altered connectivity between brain areas in ASC may affect the ability to integrate this multimodal socio-emotional information into a coherent whole (Baron-Cohen & Belmonte, 2005; Belmonte et al., 2004). Not many studies have assessed ER in ASC using ecologic stimuli. Those which have reported ER deficits among children and adults with ASC, compared to matched controls.

Pierce and colleagues assessed integrative socio-emotional understanding and the effect the number of social cues, played in different perceptual channels, had on it among children with autism, children with learning disabilities, and typically developing children. Participants were shown videotaped vignettes of child-child interactions in which the number of cues leading to the correct interpretation of the story varied from one to four (i.e., prosody, verbal content, nonverbal, or nonverbal with object). The children were asked whether the behaviour shown (e.g. one child snatches another child’s bag) is a good way to make friends, whether the child was mean or nice, how the recipient of the behaviour is feeling, and why. Results indicated that children with autism performed as well as both control groups on scenes containing one cue, but performed more poorly on scenes containing multiple cues. The authors related these findings to attentional problems in autism (Pierce, Glad, & Schreibman, 1997). Findings of this study suggest that due to their detail focused style, individuals with ASC fail to benefit from multimodal socio-emotional information.

In the study by Yirmiya and colleagues described in section 1.3.4.1, which used video clips from the Feshbach and Powell audiovisual test for empathy (Feshbach, 1982), children with HFA performed significantly lower than matched typically developing controls on emotion labelling. Four basic (happy, angry, sad, afraid) and one complex emotion (proud) were included (Yirmiya, Sigman, Kasari, & Mundy, 1992).

Studies assessing the ability of adults with ASC to recognise emotions and mental states from multimodal ecological stimuli are not widely available. ‘The Awkward Moments Test’ (Heavey, Phillips, Baron-Cohen, & Rutter, 2000) presented participants with seven short social situations, taken from television advertisements. The task items included facial expressions, body language, verbal content and prosody. Participants with AS/HFA and matched controls were asked to judge the protagonist’s mental state at the end of each scene. Participants with ASC performed at a significantly lower level compared to general population controls. In another study (Klin, Jones, Schultz, Volkmar, & Cohen, 2002a, 2002b), a single social scene from a feature film was presented to adults with ASC, while tracking their gaze. Comparing to typically developed controls, participants with ASC looked less at the eyes of characters, and more at characters’ mouths and surrounding objects, thus missing socio-emotional information pertinent for the understanding of the social situation.

In a recent study introducing a task called MASC (Movie for The Assessment of Social Cognition), Dziobek and colleagues asked participants with ASC and controls to watch a short film about an evening get together of four adults. The film was stopped at various points and participants were asked about the characters’ mental states. Individuals with ASC exhibited marked and selective difficulties in socio-emotional understanding (Dziobek et al., 2006). Interestingly, performance on the

MASC was not significantly correlated with performance on basic facial ER task or with performance on the Reading the Mind in the Eyes task, but was significantly correlated with participants' performance on the Strange Stories Test. These findings suggest that individuals with ASC relied on the (verbal) narrative rather than on facial expressions or eye direction when answering MASC questions. Whereas these findings could reflect on the tendency of individuals with ASC to use language to compensate for poor ER abilities, they could also stem from the fact that the original MASC was recorded in German and dubbed to English. Hence, it is possible that participants with ASC focused on the dialogue alone, due to incongruence between the visual and the auditory channels.

These studies suggest that individuals with ASC are impaired on multimodal ecological tests of social understanding, and that the 'breadth' of emotional information in the different channels challenges their ER skills. Appendix 3 reports a child and an adult version of a multimodal ecological ER task prepared for this study, which uses short scenes from feature films to test complex emotion and mental state recognition in individuals with ASC.

From the findings presented in this chapter, it appears that although the ER deficit in ASC is lifelong, some higher-functioning individuals develop compensatory strategies that allow them to recognise basic emotions. However, when recognition of more complex emotions and mental states is required, many find them hard to interpret in faces, voices, context and multimodal stimuli. This deficit (along with others) has considerable implications for the ability of children and adults with ASC to function socially. In the next chapter, I review interventions that have been used to train individuals with ASC to overcome this deficit, and present the new computer-based intervention evaluated in this thesis.

3 Emotion recognition interventions in Autism

In the previous chapter, I reviewed studies related to emotion and mental state recognition deficits in autism spectrum conditions. The skills of emotion recognition (ER) and mentalising are intuitive and automatic for most people, making it difficult to even imagine life without them. However, individuals with ASC have to be taught these skills, and work hard to bridge this empathising gap. Given the centrality of ER to socio-emotional functioning, there have been different attempts to train children and adults with ASC on recognition of emotions and mental states. In this chapter, I describe some of the interventions that have been evaluated. I then focus on computer based training methods in particular, and their advantages for individuals with ASC. Finally, I describe the computer-based intervention evaluated in this thesis.

Many interventions, from various disciplines, are available for individuals with ASC and their families, including medical, therapeutic, educational, and dietary programmes. Unfortunately not many of them have been evaluated by research (Jordan, 1999). Examples of some that have been evaluated include the intensive early intervention programme of Lovaas and his colleagues using Applied Behaviour Analysis (Lovaas, 1987); and the structured, visually based TEACCH educational programme (Schopler, Mesibov, & Hearsey, 1995; see Howlin, 1998; Jordan & Jones, 1999 for reviews). However, in this chapter, I discuss interventions that specifically focus on understanding and recognition of emotions and mental states.

3.1 Traditional interventions

3.1.1 Training lower functioning individuals with autism

Past attempts to teach emotion and mental state recognition to individuals with ASC were based on the ToM deficit model, and focused on the understanding of beliefs and desires, and their impact on basic emotions. Several studies reported attempts to teach children with autism to pass false belief tasks and to differentiate the mental from the physical. Bowler and Strom tried to teach children with autism and typically developing toddlers to pass the Sally-Ann first order false belief task. Participants were given five replications of the standard task with the inclusion of enhanced behavioural and emotional cues to the protagonist's false belief. This was done through enacting the protagonist’s search for the item where she left it, and her reaction when she could not find it. The results showed that children with autism, as well as some of the typically developing children, benefited to a significant extent from the enhanced cues, but not when the task was repeated without enhancement (Bowler & Strom, 1998).

Other attempts to teach ToM used a metaphor of mental states being like ‘pictures in the head’. Since children with autism can understand that photographs represent reality, but can also differ from it (Leslie & Thaiss, 1992), Swettenham and colleagues tried to teach children with autism the analogy that ‘people have photos in their head’, i.e. that mental states represent reality and that, like photographs, these can differ from reality, e.g. if they become outdated. Using a manikin with a slot for pictures in her head, children were taught (in a group) how *seeing* something creates a picture in the head (i.e. leads to *knowing*). Using this analogy, children practised false belief tasks to acquire the concept of false belief. All children with autism trained were able to use this strategy to predict behaviour of protagonists, but not their mental states. They managed to pass the Sally-Ann task, and the Seeing Leads to Knowing task (see Chapter 1 for descriptions), but not other false belief tasks, suggesting generalisation was limited (Swettenham, Baron-Cohen, Gomez, & Walsh, 1996).

McGregor and colleagues' training programme included the 'picture in the head' method and highlighted the protagonist's *intention* in the false belief task (e.g. what was Anne's intention when she hid the marble?). Training was given to adults with autism and to typically developing 3 year olds. Two matched control groups received no intervention. At the end of training, both intervention groups could pass the Sally-Ann task. However, whereas the typically developing 3 year olds managed to pass other false belief tasks they were not trained on, generalisation was limited in the autism group (McGregor, Whiten, & Blackburn, 1998a). Generalisation to videotaped false belief scenarios, rather than acted ones used during training, was also limited (McGregor, Whiten, & Blackburn, 1998b).

Fisher and Happe tried to improve generalisation by using Swettenham et al's protocol with children with ASC individually rather than in a group. Another group of children with ASC was trained on executive function tasks, and a third group received no intervention. Results showed an improvement on false belief task performance in both ToM and executive function training groups, including on variations the children were not trained on. The children's performance did not improve on the 'reading the mind in the eyes' task, or on teacher ratings of their understanding of mental states in everyday life (Fisher & Happe, 2005).

A similar method for teaching ToM used cartoon 'thought bubbles' to represent mental states. Parsons and Mitchell found that children with autism with an average verbal age of 7 years successfully interpreted thought bubbles as representational devices that could be used to infer an unknown reality and to inform about the content of people's beliefs. When they tested children with autism on false belief tasks with and without thought bubbles, which depicted the content of the protagonist's belief, performance was improved in the tasks which included bubbles (Parsons & Mitchell, 1999). Wellman and colleagues tested the effectiveness of thought bubbles as an instrument for individuals with autism to predict behaviours and mental states, to pass false belief tasks, and to generalise to other tasks, not included in training. Children with autism with an average mental age of 5 years went through a six stage training programme, using the thought bubble method to teach about knowledge of location of seen and unseen objects, behaviour, and false belief. Following training, children were able to predict behaviour and thoughts of protagonists in false belief tasks they were not trained on, and on the seeing leads to knowing task, which was not included in the training program. The authors concluded that the thought bubble concept is a good ‘prosthetic’ to help children with autism picture mental states (Wellman et al., 2002).

Hadwin and colleagues extended their training programme beyond understanding of false beliefs. Three groups of children with autism (with an average verbal mental age of 5 years) were included in the study: One group was trained to understand and recognise situation-based, desire-based, and belief-based emotions from various contexts. The second group was trained to understand beliefs and false beliefs. The third group’s curriculum was aimed to increase the level and quality of participants’ pretend play. The emotion group was trained to recognise happiness, sadness, anger, and fear from schematic drawings of facial expressions and from photographs of facial expressions. They were then taught to recognise situation-based emotions from drawings of emotion-eliciting situations (e.g. a big dog chasing a child, as a fear eliciting example). Teaching recognition of desire-based emotional context was limited to happiness and sadness (e.g. a child who gets the cupcakes she likes is happy), and so was the teaching of belief-based emotions. Belief-based emotions were taught with a combination of desire and belief (e.g. *believing* one will not get what one *wants* elicits sadness, even if the desire is fulfilled eventually, and vice versa), though the concept of surprise was not included, perhaps due to the children’s low mental age. The belief group was taught that different people can see the same thing differently, that seeing leads to knowing, that behaviour is affected by what one knows and believes, and that people could have false beliefs which will affect their behaviour. Variations of the Sally-Ann and Smarties false belief tasks were used for training. After 8 daily training sessions, of half an hour each, children in all 3 groups were tested on ER, understanding of belief and false belief, and pretend play. The ER and the belief groups improved in their performance on the kind of tasks they were being trained on, but not on the other domains. The pretend play group’s performance did not improve on any domain. A follow up assessment two months later yielded the same results (Hadwin, Baron-Cohen, Howlin, & Hill, 1996). The intervention created for this study was reported in detail in a separate reference (Howlin, Baron-Cohen, & Hadwin, 1999), used in several intervention studies (see below).

The interesting attempt to teach recognition of some of the basic emotions in the latter study suggests that with acquisition of defined and clear principles, situation and desire based basic emotions could be understood by lower functioning children with autism (though generalisation is still limited). However, no similar attempt to teach more complex emotions or social situations was reported. This may be due to the participants' young mental age, not enabling comprehension of more complex material. In addition, material taught was in many cases quite different to the socio-emotional functioning required of the children in everyday life. Association of ER with such variants of social functioning is described in the next section.

3.1.2 Training higher functioning individuals with autism

Teaching ER to individuals with Asperger Syndrome or High Functioning Autism is often undertaken as part of social skills training, in a group with other children with ASC or in a mixed group of children with ASC and typically developing peers (Spence, 2003; Strain & Hoyson, 2000). These training programmes typically include themes of basic interaction, conversation, play and friendship, emotion-processing, and social problem solving (Krasny, Williams, Provencal, & Ozonoff, 2003), as well as reducing socially inappropriate behaviour, personal hygiene, and others. Training is usually held in small groups, which limits scientific investigation of these programmes. For example, Barry and colleagues reported of a social skills group for children with HFA, implemented in an outpatient clinic setting. Training was effective in improving greeting and play skills, with less clear improvements noted in conversation skills (Barry et al., 2003). However, since only four children took part in this group, it is difficult to tell whether the lack of significant improvement is genuine, or whether it simply results from the small sample.

Ozonoff and Miller added training of theory of mind and false belief understanding to their social skills curriculum, which included conversation skills, expression of non-verbal cues and emotional expressions, and recognition of those in others. Role modelling by facilitators, and feedback on children's videotaped role playing, were used throughout the training. The 5 children with ASC who participated in the group and 4 matched controls who received no intervention were tested before and after the intervention using a battery which included first, second, and third order false belief tasks. In addition, parents and teachers rated the children's social skills before and after the intervention. After 14 weekly sessions, the intervention group significantly improved its performance on false belief tasks, whereas the control group's performance remained unchanged. However, the assessment was done on the tasks the intervention group was trained on. In addition, there were no changes on parents' and teachers' ratings of social skills, suggesting poor generalisation of learnt material to everyday social functioning (Ozonoff & Miller, 1995).

Positive results on a 7 month school-based socio-emotional and social interaction curriculum for 15 children with HFA aged 8–17 years were reported by Bauminger. Intervention focused on teaching interpersonal problem solving, affective knowledge, and social interaction, and participants showed improvement on all three areas. To assess affective knowledge children were asked to define 10 basic and complex emotions and to give examples of times they felt these emotions. At the end of the training programme, children were able to provide more examples of complex emotions, supplied more specific rather then general examples, and included an audience more often in the different emotions (Bauminger, 2002). However, in the absence of a control group, and since parents, siblings and peers were involved in the programme, it is difficult to tell which part of this improvement is related to the training itself, and which to extracurricular activities or simply time passing.

Another study of social skills training was reported by Solomon and colleagues, who ran 20 week social skills groups for 8-12 year old children with AS/HFA. Curricula included awareness of emotions (both basic and complex), face processing, theory of mind, conversational skills, and problem solving. Compared to a waiting list control group with AS/HFA, children in the intervention group showed significant improvement in basic emotion recognition from photos of faces and on problem solving, but not on complex ToM tasks such as the Strange Stories Test or the Faux Pas Test. Complex ER in faces was not assessed (Solomon, Goodlin-Jones, & Anders, 2004).

There are hardly any descriptions of social skills training for adults with ASC. Howlin and Yates reported a social skills group conducted with 10 adults with ASC that met over a year on a monthly basis. The principal function of the group was to provide the participants with a better understanding of their social difficulties and to foster more appropriate ways of dealing with these. There was a particular emphasis on improving conversational skills, including understanding of body language and others' emotions. Effects of the group were only assessed retrospectively, and relied on feedback from participants and families. Improvements in communication skills, understanding of others' body language and the ability to interpret other people's emotions were reported by group members, as well as improved ability to relate to people at home and outside (Howlin & Yates, 1999). Whereas ER was included in the curriculum of the group, it was not assessed beyond self report.

An adult group, which dealt specifically with ER, was reported by McKenzie and colleagues, who worked with adults with learning disabilities (but without ASC). Their training programme of 10 weekly sessions focused on recognition of emotions from facial expressions, teaching about the configuration of facial features in different expressions, the association of facial expressions with context, and recognition of emotion in line drawings, photographs and videotaped scenes. There was a significant overall increase in accuracy in identifying emotions following group training. The assessment included the emotions *happy, sad, angry, worried, bored* and *afraid*. In addition, a significant increase was found in the ability to label emotions depicted by line drawings typically used in symbol-based communication systems (McKenzie, Matheson, McKaskie, Hamilton, & Murray, 2000). Similar results were reported in a controlled trial which included a non-intervention control group of adults with learning-disabilities (Rydin-Orwin, Drake, & Bratt, 1999).

The latter examples demonstrate the merit of ER group training, which is a more economic and therefore a more feasible service. Unfortunately, social skills or emotion recognition groups are not widely available (Rogers, 2000), especially for adults with ASC (Howlin & Yates, 1999). Group interventions also require trained staff to facilitate the groups, which is not always available. Furthermore, unlike individuals with learning disabilities without ASC, group based interventions might actually be too socially demanding for people with ASC, and might therefore deter more socially anxious participants (Tantam, 2000). Finally, in such groups it is difficult to target the individual's specific level and pace of learning, potentially leaving some participants over-challenged, and others bored. The next section discusses a medium which addresses these needs.

3.2 Computer-based interventions

In the last two decades, increasing attempts to teach individuals with ASC social skills have used computer-based training. The use of computer software for this group has several advantages: Individuals with ASC favour the computerised environment since it is rule-based, predictable, and consistent. This fits well with their good systemising skills (Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002). The computer is also free from social demands, which they may find stressful. On the computer, information can be presented in a way that reduces the potentially confusing and anxiety-inducing, multimodal inputs that characterise 'real world' social situations (e.g. the visual channel can be played separately from the auditory channel, gestures can be separated from facial expressions, etc (Moore, McGrath, & Thorpe, 2000). Computer users can work at their own pace and level of understanding, immediate feedback is provided, and lessons can be repeated over and over again, until mastery is achieved. Finally, interest and motivation can be maintained through different and individually selected computerised rewards (Bishop, 2003; Moore, McGrath, & Thorpe, 2000; Parsons & Mitchell, 2002). In addition to that, computer based training is easily available commercially, and can therefore serve a wider audience. High-functioning children and adults can use it on their own, with little support required from professionals.

Amongst the potential limitations of computer-based training for teaching social skills and ER is the difference between the computerised medium and real life. In a way, the same reasons that make the computer environment more convenient for individuals with ASC might make it too distant from real social phenomena to be able to generalise. A fine balance needs to be struck when computer based training programs are used, so that the computerised environment structures and simplifies real life social phenomena, but does not distance itself too much from it, so that it can serve as a good scaffold from which to transfer learned knowledge back to real life.

Another possible flaw of using only computerised environment for training individuals with ASC is the risk that it becomes too appealing, thus strengthening their avoidance of social contact, as their needs are catered for by a non-human agent. To avoid that, Howlin suggested that computer-based training should always be combined with work with human facilitators (Howlin, 1998). However, social interaction has over the last decade become available through computers too. Using the internet to facilitate and learn social skills is now possible. Although little research supports its use currently, the use of email, instant messaging, chat rooms, media groups and other means allow individuals with autism to communicate in a secure and controlled social atmosphere. The increasing number of websites, online forums and communities set up by and for individuals with autism and their families provide, in addition to information and support, more opportunities for social interaction. Access to the internet via home computers, as well as via mobile communication interfaces such as cellular phones and hand held computers offer new opportunities for social interaction for individuals with autism (Bishop, 2003; el Kaliouby & Robinson, 2005).

Previous studies have found that the use of computers can help individuals with autism learn a variety of skills. Bosseler and Massaro used a computer-animated tutor to teach vocabulary and grammar for children with autism. After using the software several times a week for a period of 6 months, students with autism learned a significant number of new words and grammar. The program also led to learning and generalisation of new words (Bosseler & Massaro, 2003). Hetzroni and Tannous investigated the use of a program developed for enhancing communication functions of children with autism. Based on daily-life activities, the program presented children with situations that provide opportunities to use questions and answers formulated by parents (e.g. ‘What would you like to eat?’). Three different settings simulated daily activities around play, food, and hygiene. Following 18 sessions of software use, the 5 children assessed produced fewer sentences with delayed and irrelevant speech, engaged in fewer sentences involving immediate echolalia, and increased the number of communication intentions and the amount of relevant speech they produced. The children were also able to transfer their knowledge to their natural classroom environment (Hetzroni & Tannous, 2004).

Compared to these positive findings of computer-based training efficacy in the area of verbal communication, interventions into socio-emotional functioning showed more limited results, especially in the children's ability to generalise from the computerised environment into real life settings. Swettenham used a computer program to train children with autism, children with Down syndrome and typically developing 3 year olds on the Sally-Ann false belief task. The children were shown the scenario on the computer and instructed on the actions, the thoughts and beliefs of the characters. Following four days of training, with two sessions in each day, the children were tested on a close generalisation measure - solving the Sally-Ann task using dolls rather than on the computer. They were also tested on distant generalisation measures which included the variations on the Smarties task, and a non-verbal picture arrangement task depicting false belief scenarios. Children in all three groups were able to solve the close generalisation task. However, children in the autism group performed significantly lower than both the Down syndrome and the typically developing control groups on the distant generalisation measures. A follow-up assessment held after 3 months replicated the post intervention findings for the autism and Down syndrome group, though the typically developing group's performance significantly improved (Swettenham, 1996).

Rajendran & Mitchell evaluated the 'bubble dialogue' program in which two users can type in thought and speech content for story protagonists, using the bubble thought method. Story situations depicted different ToM and social themes such as simple and complex perspective-taking, false belief, lies and white lies, and making a friend. Dialogues were held between each of the participants and the experimenter. Two adults with AS used the program for one hour a week over a period of six weeks. Participants found the dialogues were useful in helping them consider the implications of thought and speech in social situations. However, there were no indications that the intervention improved their interpersonal understanding in the natural settings, as assessed by structured interviews with their carers, focused on social skills and understanding mental states (Rajendran & Mitchell, 2000).

Several studies reported evaluations of computer programs designed to teach emotion recognition and social skills: Bernard-Opitz and colleagues created a computer training program that presented users with eight distinct social problems. The social problems include themes of turn taking, requesting help and objects, giving in, and negotiating. Four of the tasks were labelled ‘easy’, i.e. requiring simple social problem-solving (e.g. not being able to reach an item), whereas the other four required more complex social problem-solving (e.g. being scolded for breaking an object) and were labelled ‘difficult’. Problem scenarios are animated on the computer, and followed by a choice of 4 possible solutions (2 correct and 2 incorrect), and an option to produce alternative solutions. Correct solutions, as well as appropriate innovative solutions are reinforced using animation. The user is then prompted to provide more solutions. When no more solutions are offered, the next problem is presented. This program was evaluated with 8 children with autism (with an average verbal mental age of 4 years and 9 months), and with 8 typically developing children (with an average chronological age of 4 and a half). Children in both groups used the program for 10 individual sessions interleaved with 6 assessment sessions. Tutor support was available through all sessions, but unlike the assessment sessions, in the training sessions, problem solutions were first explained thoroughly by the trainer. Overall, children with autism produced significantly fewer innovative solutions, compared to their typically developing peers. The improvement in the typically developing group was also more consistent. However, a steady increase across assessment sessions was observed for the autism group. Since the problems presented in the assessment sessions were not included in the training sessions, improvement on those provides supporting evidence for generalisation of social problem solving strategies among children with autism. However, generalisation into real life settings was not assessed (Bernard-Opitz, Sriram, & Nakhoda-Sapuan, 2001).

Silver and Oakes reported an ER training program named ‘the Emotion Trainer’. This program teaches ER of happiness, sadness, fear and anger from still photos of faces, situation-based context (e.g. the child saw a spider coming into the room), and belief-based context (e.g. the child thought there’s a spider in the room). The program also uses the methods developed for Hadwin et al’s study (Howlin, Baron-Cohen, & Hadwin, 1999), mentioned above, to teach the association between desire, belief and emotion. Children with ASC aged 10-18, with an average verbal mental age of 11 years were randomly assigned into one of two groups - computer intervention and no-intervention controls. Children in both groups were assessed before and after the intervention period on ER, using photos of facial expressions, the Strange Stories Test, and emotion eliciting context cartoons (taken from Hadwin et al's study). Children in the intervention group used the software for 10 daily sessions over a period of 2-3 weeks. Compared to the control group, the intervention group significantly improved on ER from situation-based and belief-based context, and on the Strange Stories Test. These results suggest the Emotion Trainer successfully teaches ER from context, including generalisation of knowledge to other tasks not included in the training (such as the Strange Stories Test). However, no significant improvement was found on ER from facial expressions. This could be related to the children being old and high functioning enough to successfully recognise basic emotions from facial expressions. Indeed, the authors reported that children in both groups performed at ceiling before the intervention on recognition of happy and fearful faces (Silver & Oakes, 2001).

Although this study reported promising results, it failed to teach and assess ER from facial expression at a level that is appropriate for the children's level of functioning. Teaching and assessment of more complex emotions might have resulted in more meaningful results. In addition, ER from vocal expression was not taught nor assessed, and neither was ecologic multimodal ER. Finally, no follow up measure of the children's socio-emotional behaviour 'in the real world' was taken, which leaves open the question of successful generalisation into real life.

Bölte and colleagues created a computer based training program called 'Frankfurt test and training of facial affect recognition' (FEFA). The program teaches recognition of the basic 6 emotions, and also includes photographs of 'neutral' facial expressions. Five hundred photographs of facial expressions are included in FEFA's training module. The test module includes fifty whole face items and forty eye region only items. Users are presented with a still photo of the whole face or the eye region and are asked to choose the correct emotion label out of the 7 possible labels (6 basic emotions or 'neutral'). Positive feedback is given using 'smiley' symbols. Feedback for wrong answers includes highlighting of the correct answer, written explanation and an example of a situational cartoon taken from Howlin et al (1999). In an evaluation of the program, twenty adults with AS/HFA were randomly allocated into an intervention and a control group. The intervention group used FEFA over a period of five weeks, consisting of two hours training a week. Both groups were assessed before and after the intervention period, using the face task and the eyes task taken from the training program, as well as external ER task not included in the training, to assess generalisation. The intervention group improved on ER from faces and eyes photos included in the training program but not on ER from the generalisation task (Bölte et al., 2002). A neuro-imaging study, assessing the effect of using FEFA for 5 weeks (with the support of a clinician) on brain activity replicated the behavioural results reported above, but failed to find increased activation in the fusiform gyrus. However, improved ER performance in the intervention group was accompanied by higher activation in the right medial occipital gyrus, a region involved in object and face recognition, and in the right superior parietal lobule, which is involved in visuo-spatial processing and in visual attention. Both areas have been assumed to be part of a compensatory facial processing network (Bölte et al., 2006). Activation of other social brain areas, such as the amygdala or the superior temporal sulcus, was not assessed. These findings suggest that individuals with ASC may improve on their ER abilities using alternative strategies to those used by the general population.

The computer-based interventions described above suffer from several limitations: first, they used drawings or photographs for training, rather than more life-like stimuli. This might have made generalisation harder than if more ecologically valid stimuli were used. Second, the programs teaching ER focused on basic emotions, which individuals with ASC can often recognise as well as the general population, as described in Chapter 2. Third, only facial expressions and context are used in these training programs. No reported program to date has systematically trained complex emotion recognition in the voice, nor in both visual and auditory channels. Fourth, the curriculum in the programs described by Silver and Oaks (2001) and by Bernard-Opitz et al (2001) is quite limited and hence limits the ability to analyse socio-emotional material systematically with a large variety of exercises. Indeed, results of the neuro-imaging study by Bölte et al suggest that with a large enough range of stimuli available, changes can be found even in brain functioning. Furthermore, the findings reported by Bölte et al (2006) of changes in activation of visual areas involved in visuo-spatial attention suggests individuals with ASC may be systematically scanning face features as part of their ER process. Hence, providing individuals with ASC with the opportunity to scan pre-arranged socio-emotional stimuli systematically, with a large enough selection of examples, could result in improvement at the behavioral level and changes in brain functioning. This thesis asked the particular question: can individuals with ASC, if they are provided with an organised system of socio-emotional material, harness their good systemising skills to overcome their empathising difficulties?

To test this question, this thesis evaluates an intervention program which provides users with the option to systematically and comparatively scan emotional stimuli to find common features that are shared between different protagonists expressing the same emotion. It is the first documented computer-based intervention program that enables this systematic study of both basic and complex emotions in both the visual and the auditory channels, using motion in visual stimuli (i.e. videos of faces rather than stills). This intervention program, called *Mind Reading* is described below.

3.3 Mind Reading: A systematic guide to emotions

Mind Reading (Baron-Cohen, Golan, Wheelwright, & Hill, 2004) is an interactive guide to emotions and mental states. It is based on a taxonomic system of 412 emotions and mental states, related to the emotional lexicon study described in Chapter 2. The inclusion criteria for words to be added to the list were those described in Section 2.2.1. This broad definition of emotion included concepts that fall in the 'internal-mental', but also the 'external' categories in the classification offered by Ortony et al (1987). The reason for this was that *Mind Reading* was designed to teach recognition of emotions and mental states from facial and vocal expressions. Hence, it includes concepts that traditionally would not fall under the definition of an 'emotion' (e.g. bored), as long as they have distinctive expressions in the face and the voice, and as long as they were deemed beneficial for social understanding. The 412 emotions and mental states included were defined by a lexicographer as having a sufficiently unique semantic definition. Other emotion and mental state labels were included in *Mind Reading* as 'similar emotions' (or synonyms) to the 412 concepts.

The 412 emotional concepts are organised in 24 thematic emotion groups (see Figure 3.1 for a list of the groups). Two of *Mind Reading*'s authors independently assigned each of the 412 emotion concepts to one of the 24 emotion groups. Where there was disagreement on assignment, a third author was given the same task, and independently categorised the emotion term. Assignment of these remaining emotion terms was judged to be valid if at least two of the 3 judges showed agreement. Every emotion concept was also coded for *valence* (or impact) as being either positive, negative, or neutral; and coded for *intensity* as being either strong, or not.

The 412 concepts were also organised according to 6 developmental levels, from age 5 to adulthood, based on data from the developmental emotion lexicon study (see Section 2.2.1). This allows users to start working with the software in the developmental level that matches their age, or level of functioning.

Afraid	Excited	Liked	Surprised
Angry	Fond	Romantic	Thinking
Bored	Happy	Sad	Touched
Bothered	Hurt	Sneaky	Unfriendly
Disbelieving	Interested	Sorry	Unsure
Disgusted	Kind	Sure	Wanting

Figure 3.1: The 24 groups included in *Mind Reading*

Each emotional concept is defined and demonstrated in six silent films of facial expressions, six voice recordings, and six written examples of situations that evoke this emotion. Facial expression clips are 3-5 seconds long. They show the face and shoulder area; hence the expressions sometimes include some shoulder movement (e.g. shrugging) as well as facial expression and head movement. Actors wear single colour shirts, and have no masking of any kind (i.e. no beard, glasses, hat, jewellery, etc). Voice recordings are 2-5 seconds long, and include short segments of speech comprising both prosody and verbal content. Where possible, neutral verbal content was used, so that ER is based mostly on prosody. Often, the same sentences were used for different emotions, with prosody being the only changed factor (e.g. 'You've done it again' could be used for the concept *congratulatory* as well as for *disappointed*, depending on prosody). In addition to the emotion specific media, each emotion group is introduced and demonstrated by a short video clip giving some clues for later analysis of the emotions in the group. The resulting library of emotional ‘assets’ (video clips, audio clips, or brief context examples) comprise 412 x 18 = 7416 units of emotion information to systematically analyse and learn. The face videos and voice recordings were performed by professional actors of both sexes, various age groups and ethnicities, to facilitate generalisation. All face video clips and voice recordings were validated by a panel of 10 independent judges, and were included in *Mind Reading* if at least 8 judges agreed the emotional label given described the face/voice. In order to ease the processing of expressions, to avoid over burdening the user and to encourage analysis of the emotion in each modality, faces and voices are presented separately for each emotion (i.e., silent face films and faceless voice recordings). Three different applications are available to study these expressions:

(1) *The emotion library* allows users to browse freely through the emotional media. Users can choose from the different emotion groups (see Figure 3.2a), play the scenarios giving examples of the emotions (Figure 3.2b), and go into individual emotion pages (Figure 3.2c). In each emotion page, faces, voices and context examples appear in separate tables. Users can play all 6 facial expressions to analyse the common features in them, listen to all 6 vocal expressions and find vocal patterns in them, read stories, and add their own notes and examples. To compare different emotional expressions in the face and the voice, a scrapbook, which can contain up to 60 different assets, is available. Users can also run searches for particular emotions they are interested in.

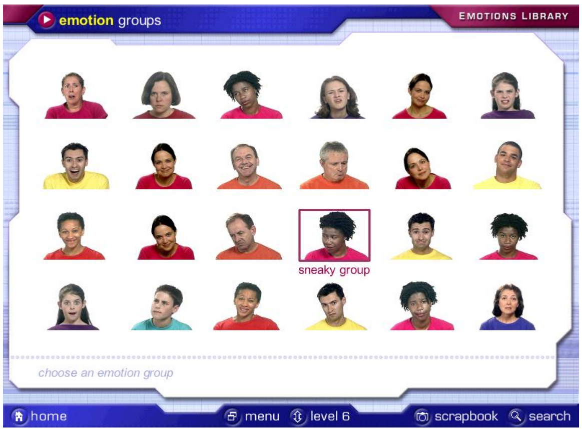
(2) *The learning centre* For children, less able adults and more structured learning, the learning centre offers lessons, quizzes and several reward collections to teach about emotions in a more directive way. In addition to teaching about the 24 emotion groups, it also includes lessons and quizzes about the 20 and 100 most commonly used emotions, as well as a ‘build your own lesson/quiz’ option. Questions require matching of stimuli within the same modality (i.e. faces with faces - see Figure 3.2d, voices with voices), as well as between modalities (face with voices, face and voices with verbal labels). Feedback is given when questions are answered, and the user cannot continue to the next question before the current one has been answered correctly. Rewards were included to make lessons and quizzes more appealing for

users. They include reward collections which were chosen for their potential appeal to users with ASC and were arranged systematically (e.g., pictures and information about space elements, clips of birds arranged by families, different types of trains to collect, flags of all the countries in the world etc. see Figure 3.2e). A reward is given when a quiz question is answered correctly on the first attempt, though this setting, as well as the number of possible answers to choose from can be modulated.

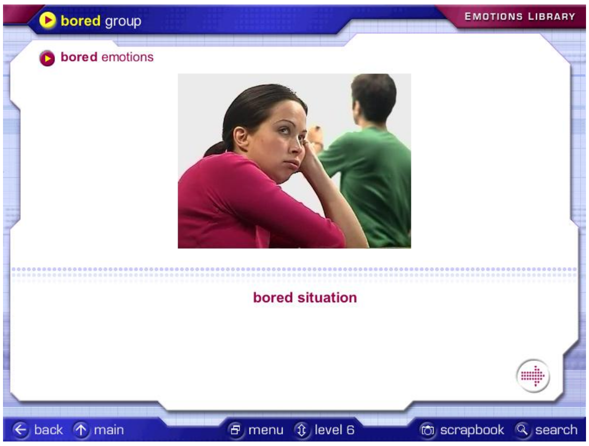
(3) *The game zone* comprises 5 educational games, allowing users to study about emotions in a playful, enjoyable environment. The games require different skills, such as guessing an emotion from parts of the face (the 'Hidden face' game, see Figure 3.2f), matching emotions using 'thought bubbles' to characters in social situations, and understanding the idea of different levels of intensity of emotion by sorting them on a continuum.

Vocal and animated helpers give instructions on every screen (Figure 3.2f shows one of the helpers). An external application – *the Mind Reading Manager* allows adult users/carers/teachers to change different settings (e.g. quiz difficulty level) and to monitor usage time for each user. The software was created for the use of children and adults of various levels of functioning, with the assumption that more able users could work with it independently, while lower functioning ones may require additional support. Figure 3.2 shows screen shot examples of the different sections of the programme.

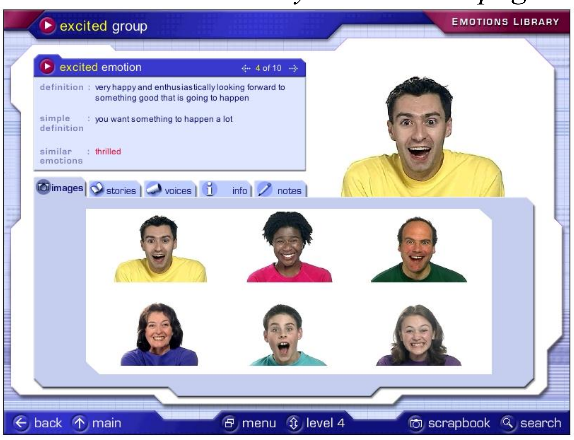
a. *The 24 emotion groups' main menu*



b. *The 'bored' group demo situation*

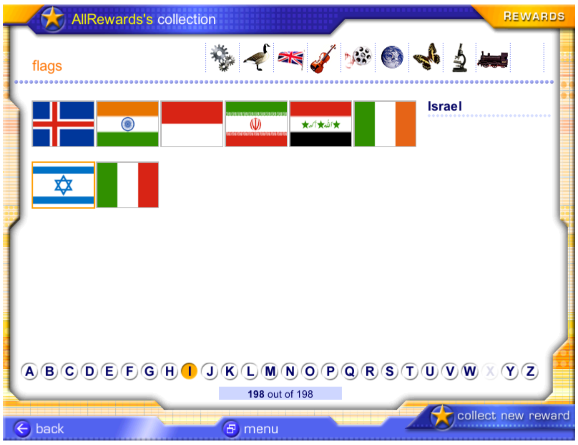


c. *The emotion library: an emotion page*



d. *The learning centre: a quiz question*

e. *The reward section - flags*



f. *The game zone: 'Hidden face'*

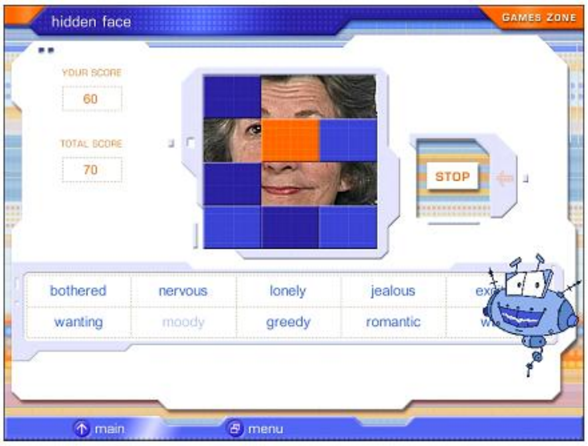


Figure 3.2: Screenshots from *Mind Reading*, the interactive guide to emotions (Baron-Cohen, Golan, Wheelwright, & Hill, 2004)

To summarise, *Mind Reading* was designed with the needs, difficulties, and strengths of individuals with ASC in mind. It provides them with a structured and guided system to analyse, enables them to separately focus on different perceptual modalities, and uses appealing rewards for them to enjoy. At the same time it presents the users with ecologic emotional stimuli in the face and the voice, and a breadth of emotions and mental states, both previously not provided by any other intervention.

As *Mind Reading* addresses a wide age range, the evaluation presented in this thesis included experiments with school aged children on one hand and with adults on the other. No previous attempts to teach ER to adults with ASC using computers have been reported, and there is a question of whether this can be done effectively at this relatively late stage. This study used the same design to assess the effectiveness of *Mind Reading* with children and adults, to test if similar effects would be found in both age groups. This design will be described in the next chapter.

Method

4 Design

To evaluate *Mind Reading*'s effectiveness in teaching individuals with ASC to recognise emotions, three controlled trials were conducted and participants were assessed on a variety of visual, vocal and multimodal ER tasks comprising both familiar and novel stimuli. This chapter describes the experiments conducted, the between and within group factors tested, and the study's hypotheses.

4.1 Experiments

The evaluation of *Mind Reading* included three experiments, two conducted with adults and one with 8-11 year old children:

(1) The first adult experiment (Expt 1) evaluated the effectiveness of *Mind Reading* in helping adults with ASC improve their ER skills when used on an individual basis at home, with no further assistance.

(2) The second adult experiment (Expt 2) evaluated the effectiveness of *Mind Reading* when used individually at home in conjunction with weekly tutor supported group meetings. Using group discussion, role-play, worksheets and analysis of emotions in newspapers and television programmes, this group aimed to consolidate participants' computer-acquired knowledge, in order to improve generalisation.

These two experiments enabled to test whether *Mind Reading* has an effect independent of any human tutor or social group effects, or whether the addition of a human tutor and social group enhances any effects.

(3) The evaluation conducted with children (Expt 3) was based on home use of the software by the children. Parents were asked to support their children's learning by suggesting emotions that may be challenging for them, or by discussing examples from their everyday life, to help the children generalise. Since it was assumed that all children in the intervention group will be supported by their parents when using the

software at home, no distinction between individual and tutor supported use of the software was made in the children evaluation study.

Since no emotion recognition training study was conducted with adults with ASC before, testing the effectiveness of *Mind Reading* with both adult and children groups allowed to investigate whether improvement of ER skills using computer-based intervention could be beneficial for adults as well as children.

4.2 Intervention

Despite variations in the developmental level of training undertaken by adults and children, and the addition of tutor and group support in Experiment 2, the intervention offered to the groups that used *Mind Reading* followed similar lines. In order to make the best of the advantages of computer based training, participants were asked to use *Mind Reading* according to their own needs and pace. The aim was to provide the participants with this systematic guide to emotions and to allow them to freely explore it with minimum direction. Participants were introduced to the software, its structure and organising principles, including the emotion taxonomy, the developmental levels and the different applications included in the software, the lessons, quizzes, rewards and games. A demonstration of a systematic analysis of an emotion was conducted with them. This included comparing faces expressing the same emotion to identify the unique facial features and motion that is common to all of them. The facial expressions were then compared to expressions of other emotions to identify the features and motion that distinguish the expression of these two emotions. A similar analysis was held with the vocal examples. Participants were encouraged to systematically analyse the facial and vocal stimuli in a similar manner. They were not instructed to study any particular emotion, or to take any particular lesson or quiz.

In all three experiments participants were asked to use the software for a minimum of 2 hours a week, for a period of 10 weeks, to assure a meaningful period for training, recognising that a longer duration might lead to individuals dropping out. Participants were asked to use the emotions library and learning centre as they please, but not to use the game zone for more than a third of the usage time.

4.3 Control groups

Each of the three experiments included three groups: A computer-intervention group of individuals with ASC that used *Mind Reading*, a matched no-computer-intervention group of individuals with ASC that did not use *Mind Reading*, and a third no-intervention matched group of controls from the general population that did not use *Mind Reading*, and were tested only once, for baseline measures.

(1) For the adult individual use study, the ASC control group had no other intervention during the whole duration of the experiment. This group only came in for assessment meetings, thereby simply controlling for the passage of time and for assessment-related improvement.
(2) In the tutor and group supported adult study, the ASC control group participants attended a social skills group, but without using *Mind Reading*, thereby controlling for the specific use of the software.
(3) In the children evaluation study, the ASC control group undertook no intervention during the entire duration of the experiment, with the exception of everyday school input, provided to both intervention and control groups.

4.4 Three rounds of assessment

In all 3 experiments, participants in the ASC intervention and the ASC control groups were assessed 3 times: a Pre-intervention assessment (time 1); a post intervention-period assessment, which was conducted 10-15 weeks after the first assessment (time 2); and a follow-up assessment, conducted a year after the second assessment took place (time 3). The measures used in time 1 included ER measures, which were then used again at time 2 to check for improvement (with the exception of one task, which was only used at time 2, see below for details), and more general socio-emotional functioning questionnaires which were then used again at time 3 to assess long term improvement. As described above, control groups from the general population were tested only once, to obtain baseline measures.

4.5 Assessing generalisation

As discussed in Chapter 3, generalisation is a great challenge for interventions in autism, especially in the socio-emotional domain. An effective intervention programme should enable its users to implement the knowledge and principles they have acquired into other contexts beyond the immediate training environment. Previous intervention evaluation studies (reviewed in Chapter 3) have included different levels of generalisation, getting further and further away from the training situation. For example, Swettenham (1996) tested the children, who were trained on the computer version of the Sally-Ann false belief task, with the same task using dolls rather than animated figures. This change of medium served as the close generalisation level. The distant generalisation level included changes both in medium and in content using other false belief tasks. Other studies (e.g. Rajendran & Mitchell, 2000) included another level of generalisation, assessing whether there was a change in real life socio-emotional skills through self and parental feedback.

In the case of *Mind Reading*, since facial and vocal expressions of emotion are presented separately, the ability to generalise was assessed within each modality, as well as in the multimodal holistic level. In addition, since the ToM and E-S models argue that the ability to recognise emotions and mental states in others underlies social functioning (Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002; Frith, 2003), one can predict that effective ER training should lead, in the long run, to improvement in real-life social functioning. Hence, the association of *Mind reading* use with real-life socio-emotional functioning was assessed. The study included four generalisation levels:

(1) *Close generalisation*: This was tested by playing face clips and voice recordings that were included in *Mind Reading* on a different computer

program, with more answers to choose from and with no feedback or support as that given by *Mind Reading*. The creation and validation of these tasks are described in Appendix 1.

(2) *Feature-based distant generalisation*: This level of generalisation assessed ER from faces and voices not included in *Mind Reading*, but still played separately (as it was during training). Due to the unique media included in *Mind Reading*, finding tasks that test complex emotion facial expression recognition in motion was not possible. Hence the 'Reading the Mind in the Eyes' test, which is a visual test of complex emotion and mental state recognition was used as a generalisation measure from faces. In the auditory channel, the 'Reading the Mind in the Voice' task for adults was revised for this study, and a similar task was created for children. The creation and validation of these tasks are described in Appendix 2.

(3) *Holistic distant generalisation*: since in reality the social environment requires the integration of both visual and auditory channels in context, an integrative ER assessment was conducted. Using scenes from feature films, this level comprised holistic socio-emotional stimuli, including faces, voices, body language and context. Participants were asked to identify the emotion of one of the characters in the scene. The creation and validation of these tasks are described in Appendix 3. To avoid improvement due to learning of the task items, and in order for the participants to be tested on a completely novel task, this task was only used at time 2.

(4) *'Real life' functioning*: This level aimed to assess whether using *Mind Reading* was associated with improvement in socio-emotional functioning in the long run. Measures included self report questionnaires for adults regarding friendship skills, and parental feedback on children's socialising and mentalising abilities. Questionnaires were administered before the intervention and one year after the second assessment.

The design of the three evaluation experiments is illustrated in Figure 4.1.

Experiment	Groups	Time 1 assessment	Intervention	Time 2 assessment	Follow up
Adult individual use	ASC MR users	Close: CAM-A Feature-distant: RME-A, RMV-R Holistic distant: N/A Real-life: FQ	Individual use of MR at home	Close: CAM-A Feature-distant:RME-A, RMV-R Holistic-distant: RMF-A	Real-life: FQ
	ASC controls		None		
	Typical controls				
Adult tutor and group supported use	ASC MR users	Close: CAM-A Feature-distant: RME-A, RMV-R Holistic distant: N/A Real-life: FQ	Individual use of MR at home, with weekly group meetings supported by tutor	Close: CAM-A Feature-distant:RME-A, RMV-R Holistic-distant: RMF-A	Real-life: FQ
	ASC controls		Weekly social skill group meetings supported by facilitator		
	Typical controls				
Child individual use	ASC MR users	Close: CAM-C Feature-distant: RME-C, C-FAT Real-life: VABS-S	Individual use of MR at home	Close: CAM-C Feature-distant: RME-C, C-FAT Holistic-distant: RMF-C	Real-life: VABS-S
	ASC controls		Regular school curriculum		
	Typical controls				

10-15 weeks

1 year

Note:
ASC= Autism Spectrum Conditions; MR= Mind Reading;
CAM= Cambridge Mindreading Face-Voice Battery (A-adults, C-children)
RME= Reading the Mind in the Eyes (A-adults, C-children); RMV-R= Reading the Mind in the Voice- Revised;
C-FAT= Children Feature-based Auditory Task; RMF= Reading the Mind in Films (A-adults, C-children);
FQ= Friendship Questionnaire; VABS-S= Supplement to the Vineland Adaptive Behaviours Scale;

Figure 4.1: Design of the intervention evaluation experiments

4.6 Hypotheses

For all three intervention experiments, it was predicted that:

1. Performance of participants with ASC would be lower on all emotion recognition tasks at time 1, compared to the typical control group.
2. ASC groups using *Mind Reading* would perform better on all emotion recognition tasks at time 2, compared to time 1, on both close and distant generalisation levels. Similarly, *Mind Reading* users would perform better on real-life measures at follow-up, compared to time 1.
3. Performance of ASC groups using *Mind Reading* would improve more than the ASC control groups on all measures. On the holistic distant generalisation task, taken only at time 2, *Mind Reading* users will perform better than ASC controls.

Among participants using *Mind Reading*, it was predicted that there would be a positive correlation between improvement measures and software usage time. The effects of IQ, age, and level of reported autistic symptoms on improvement measures was also calculated.

5 Participants

This chapter describes the participants who took part in the different studies, the selection and exclusion criteria for participants in the different groups, the parameters used for group matching, and the limitations of the selected sample. Since specific data about the characteristics of the groups differs between the different experiments, this is reported separately in the chapters describing the studies.

The ethics committee of Cambridge University Experimental Psychology Department approved all studies reported in this thesis. Written consent was given by all adult participants, and by parents of all child participants. In addition to parental consent, no child participated against his/her will (one child who was brought in by his mother, but found the assessment situation too stressful, was excused from the study).

Many studies into ASC have included only male participants. This is probably due to the lower proportion of females with ASC, compared to males, or to avoid possible confounds related to sex differences. However, in this thesis, since it presents individuals with ASC with a new intervention, which may be beneficial for males and females on the autistic spectrum, and since there were no hypotheses about differences in the ability of males and females to benefit from such an intervention, recruitment was not limited to males only.

5.1 Recruitment

5.1.1 ASC groups

For Experiments 1 and 3, the adult and child individual software use experiments, participants were recruited from a variety of sources, including the volunteer database of the Autism Research Centre, and a local clinic for adults with AS/HFA (CLASS - The Cambridge Lifespan Asperger Syndrome Service), a local clinic for children with AS/HFA (The Asperger Outreach Service in Brookside Family Consultation Clinic, Cambridge). In the two clinics, participants were contacted by the clinic administrator with information about the study, and only those who wished to hear more about the study were contacted by the research team. Participants were also recruited via adverts in the National Autistic Society magazine *Communication*, and in newsletters of regional autistic societies in Cambridgeshire, Hertfordshire, the Wirral, and London. In each experiment, participants were randomly allocated into the two ASC groups.

The call for volunteers specified that the study aims to evaluate a new computer based program for ER. Therefore, it is likely that people who do not usually volunteer for research studies may have been interested in this study due to the potential personal gain they might obtain from using the software. However, this may also mean that participants who have specific ER difficulties volunteered to the study, whereas other individuals with ASC, who do not have such difficulties, did not. It was assumed that the multiple recruitment methods and the random allocation of participants into the ASC intervention and control groups assisted in diminishing this potential bias.

For Experiment 2, the adult tutor and group supported use study, participants for the *Mind Reading* user group and the social skills control group were recruited via two support organisations and two colleges for individuals with ASC, where group meetings were held and where the first and second assessments took place. Since participants were recruited and group meetings were held by the organisations that had volunteered to help with the study, participants were not randomly allocated to the groups, but instead were assigned by their recruiting organisation. However, facilitators were blind to the study's design and to the role of their groups in it. From the organisations' perspective, the groups were offered as free interventions made available for them. Hence, it is unlikely that their choice of participants biased the groups' structure. Furthermore, since the participants were brought to the intervention groups by their local support workers, it is likely that this experiment included participants that would not normally volunteer to participate in research studies (indeed, for many of them it was the first time they took part in a study). Furthermore, it is possible that some of these participants did not necessarily feel they have particular difficulties in ER or in social skills, but joined the groups after being persuaded to do so by their support workers. For these reasons, participants recruited for this experiment may be quite different to those recruited for the other two experiments, and allow a better coverage of the potential range of high functioning people with ASC.

Neither adult nor child participants in the clinical groups were paid for their participation. Instead, participants in the ASC control groups received a copy of *Mind Reading* (which normally costs £80) free of charge after the second assessment, and participants in the ASC intervention group were allowed to keep the copy of the software they were already using. In addition, participants' travel and subsistence expenses were refunded. The lack of payment for their time may have deterred participants from a lower socio-economic background, who may have needed this incentive to take part. This was another potential bias in recruitment, though participants who joined the study came from quite heterogeneous socio-economic backgrounds.

5.1.2 Typical control groups

Adult participants were recruited for the typical control groups from a local employment agency, and from adverts placed around Cambridge. An attempt was made to recruit male and female participants from a wide range of ages, educational and occupational backgrounds. Participants were paid for their time.

Child participants were recruited through a mainstream primary state school in Cambridge. Parents were sent letters through the school, to ask if their children would take part in a validation study of a new computerised guide to emotions, designed to help children with ASC learn to recognise emotions. Children were tested at school, and were awarded £10 book vouchers to thank them for participating. Testing typically developing children in school, rather than at the Autism Research Centre like the children with ASC, was done for practical reasons. However, it suggests a potential bias as children may have concentrated less due to the testing environment. They may have also been less motivated to take the tasks, as they were taken out of lessons. To control for that, we tried to avoid taking the children out of lessons they particularly liked. In addition, no testing was done during break time. This potential bias may have reduced the performance of children in this group, compared to the

ASC groups that were tested out of school. This may actually suggest that performance differences found between the ASC and the typical control groups were potentially bigger if all children were tested in the lab.

Another factor that differentiates the adult and child participants in the typical control groups from children in the ASC groups is the payment. The difference between the intrinsic motivation many of the participants with ASC may have had for taking part in the study, and the extrinsic motivation elicited by payment in the typical controls is inherent to this kind of clinical studies (Brewer, 2002).

5.2 Inclusion and exclusion criteria

To be included in the ASC intervention or ASC control groups, participants had to have a diagnosis of Asperger Syndrome (AS) or High Functioning Autism (HFA) from a specialist centre, using DSM IV or ICD 10 criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). Although *Mind Reading* was designed to train individuals on the whole autistic spectrum, high-functioning participants with ASC were chosen for this study, to avoid the need for individual support in using the software. Since *Mind Reading* is based on a verbal taxonomy of emotions, verbal intelligence above 70 (i.e. 2 standard deviations below the average) was set as an inclusion criterion. In fact, in experiments 1 and 3, all the participants had a verbal IQ of at least 80 (i.e. within the average range). Setting such an IQ threshold is a limitation of the study, as it is unclear to what extent having learning disabilities may affect the ability to gain from using *Mind Reading*.

Adult participants were accepted to the study only if they had not participated in any intervention that is related to ER during the last three months and had no plans for engaging in another intervention while the study was ongoing. Social clubs were allowed, as were courses that were not directly related to ER (e.g. debating group). This inclusion criterion was set to avoid confounds of other interventions conducted in parallel to *Mind Reading*. Though it may potentially cause a sampling bias due to individuals in need of ER input being excluded, this has in fact not happened with any participant.

As children with ASC often get support related to social skills through school, taking part in such activities was not set as an exclusion criterion in the child study. Instead, the ASC intervention and control groups were matched on the proportion of children who were getting social skills support at school during the study. Children in the ASC intervention group were asked to use the software at home only, and to keep it separate from their social skills training, until the study is completed.

Lack of psychiatric or neurological diagnoses, and no direct family members with ASC served as inclusion criteria for participants in the typical groups. These were set since studies of family members of individuals with ASC reported that they perform worse on socio-emotional tasks, compared to controls from the general population or to relatives of individuals with other conditions, such as Down syndrome (Dorris, Espie, Knott, & Salt, 2004; Yirmiya, Shaked, & Erel, 2001). In addition, typical adult participants were screened using the Autism Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), and typical child participants were screened using the Childhood Asperger Syndrome Test (CAST; Scott, Baron-Cohen, Bolton, & Brayne, 2002) filled in by their parents. Participants were excluded if they scored above the cut-off for ASC on these questionnaires. Four adults (but no children) were excluded for this reason. If we view autism as a continuous spectrum that extends into typical functioning, then the exclusion of these typical controls, who were screened out due to their high scores on the AQ might have artificially 'dichotomised' this continuum. However, since the diagnosis of AS and HFA has only become common since the 1990's (Baron-Cohen, Wheelwright, Robinson, & Woodbury-Smith, 2005), there is a chance that these participants were undiagnosed adults with ASC. It was therefore preferred to take this precaution and exclude them from the study.

Participants with ASC were included even if they had additional diagnoses (such as depression, ADHD, or epilepsy), as such co-morbidities are quite common in ASC (Gillberg & Billstedt, 2000). However, we made sure that the proportion of participants with co-morbidities is balanced between the ASC intervention and control groups.

Another inclusion criterion in the *Mind Reading* user groups was free access to a personal computer with the required specifications for the software to run. Though the majority of participants had no problems with this, some of the participants had to upgrade their computers' memory or hard drive space. This requirement might have limited the range of participants taking part, in terms of socio-economic status, to those who can afford a computer or an upgrade, if needed. In order to limit this to minimum, we allowed the participants to use the software in friends' or relatives' houses, or in a local community centre, as long as they use it on their own. It was felt that loosening the intervention protocol this way was important so that a wider variety of participants could take part.

5.3 Group matching

In order to limit the influence of confounding variables as much as possible, the groups were matched on several factors that were deemed relevant: verbal and performance IQ, chronological age, sex ratio, and Socio-Economic Status (SES). In addition, the ASC groups were matched on the level of autistic traits participants have, on the proportion of participants with co-morbidities, on type of school they attend (in the children study), and on the proportion of participants attending social skills training (in the children study), as described above.

Matching on verbal IQ was required due to the centrality of verbal content in the training program, as seen both in emotional labels to faces and voices, and in the verbal content of vocal expressions. Matching on performance IQ was important as it includes visuo-spatial abilities, and has an effect on face processing (Hobson, Ouston, & Lee, 1988a, 1988b). Performance IQ was also found to predict ER and ToM in children with ASC (Buitelaar, van der Wees, Swaab-Barneveld, & van der Gaag, 1999). It was therefore important to control for it.

Due to time constraints, participants' IQ was tested using the Wechsler Abbreviated Scale of Intelligence (WASI), comprising the vocabulary, similarities, block design and matrix reasoning tests. The WASI produces verbal, performance and full scale IQ scores, with correlations of .88, .84 and .92 with the full Wechsler scales (Wechsler, 1999).

Matching the groups on chronological age was required, as ER and the emotional vocabulary continue to develop through the lifespan (see Chapter 2). They were matched on sex ratio due to findings of improved ER abilities in females compared to males (Baron-Cohen, 2002, 2003). Although these sex differences were not thoroughly investigated in males and females with ASC, it was suggested this pattern is found amongst them too (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003).

Matching the groups on SES is central in intervention studies (Brewer, 2002) and was quite important on this study due to the technological requirements involved. High spec computers with improved performance may be more easily available and accessible for adults and families with higher SES, as are the experience and knowledge of using computers. SES was measured through questions about occupation and educational level in the adult studies, and about parents' education and occupation in the child study. These questions were included in the general background questionnaire administered to participants at time 1 of the study.

In order to assure the two ASC groups in each experiment do not differ on the level of their autistic traits, participants were matched on their self reported level of these traits (in adults, using the AQ), or parental report of them (in children, using the CAST). This was done to assure that none of the ASC groups has participants with more severe manifestation of ASC traits, which may confound their performance or their ability to improve.

5.4 Limitations

Beyond the limitations and the potential biases mentioned above, there are a few additional limitations of the studies in this thesis:

The ASC groups in these studies comprised participants diagnosed with AS or HFA. As mentioned in chapter 1, there is some controversy about whether AS differs from HFA beyond the early differences in language development (Lotspeich et al., 2004; Rinehart, Bradshaw, Brereton, & Tonge, 2002). As these diagnoses were not treated as separate in this thesis, potential differences between the two, if they affect ER abilities, systemising, or the capacity to learn from computers, could not be tracked. However, since both conditions involve difficulties in socio-emotional functioning and in ER, and since in both the language and intelligence are at least within the normal range in the age groups tested, we allowed participants with these two kinds of diagnoses to be grouped together.

The design of the study includes groups with ASC and other groups from the general population with no psychiatric or neurological diagnoses. Having another clinical group (e.g. Tourette syndrome, conduct disorder or schizophrenia) might have helped to clarify whether ER difficulties, systemising strengths, or improvement due to the use of *Mind Reading* are specific to ASC, whether they can be found in other clinical conditions, or whether they simply represent the effect of having a psychiatric diagnosis of any kind. Due to the complex design, including two ASC groups in each experiment, no other clinical groups were recruited.

Another limitation of the design stems from the assessment of typical controls only at time 1, unlike the two clinical groups tested in each experiment. A one-time assessment does not allow to compare the post intervention measures of the ASC groups to the typical groups, since these may have improved their performance at time 2 due to taking the tasks for the second time or due to time passing. The assessment of the typical groups only once was done for practical reasons, related to the way these groups were recruited.

Relying merely on clinician's diagnoses may be a limitation of these studies too. As ASC are diagnosed behaviourally by clinicians, there is a chance of wrong or inaccurate diagnosis. Whilst the studies did not use measures such as the Autism Diagnostic Interview (ADI; Lord, Rutter, & Le Couteur, 1994) or the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) which are often used to validate diagnoses in ASC studies, the AQ and CAST, mentioned above, were used to verify the level of autistic traits participants possessed. Whereas the CAST (like the ADI) relies on parental report, the AQ is a self report instrument. This is a possible limitation, since the scores of participants who lack awareness to their condition may misrepresent the actual severity of their symptoms. To correct this, when participants scored below cut-off on the AQ, their parents or support workers were asked to fill in a parental version of the AQ (Baron-Cohen, Hoekstra, Knickmeyer, & Wheelwright, 2006). This occurred only in Experiment 2, and will be discussed in more detail there.

One of the major limitations of this study is the relatively loose control over the intervention period, since in all 3 experiments *Mind Reading* users used it at home and out of the research team's control. Whereas *Mind Reading* monitored their usage time, as a control for the participants' own report of the time they spent using the software, there was no way to verify that the participants actually used the software themselves, that they were not helped by anyone, or that they spent the required time in each area of the software. On all of these, it was necessary to depend on participants' reports. It was assumed that participants would not deliberately sabotage the experiment, or try to cheat, especially as this is an area of difficulty in ASC (Frith, 2003; Sodian & Frith, 1992). This limitation is inherent to clinical intervention studies that extend beyond the research lab (Brewer, 2002). However, since no systematic bias was expected by the participants, it was assumed that their reports are reliable.

Intervention Studies

6 Adult intervention study: Independent use of *Mind Reading*

This experiment compared the effect of an intervention group using *Mind Reading* at home to a no-intervention control group who took the assessment twice with no intervention. The groups were assessed at three levels of generalisation, and at follow-up one year later. The need for a no-intervention AS/HFA group was to assess whether any improvement was related to the intervention or simply to the passage of time and to taking the tasks twice. The experiment's design and hypotheses are detailed in Chapter 4.

6.1 Method

6.1.1 Participants

Three groups took part in this experiment: one AS/HFA intervention group, one AS/HFA control group and one typical control group. Participants in the clinical groups had all been diagnosed with AS/HFA in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). Participants filled in the Autism Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; see description below), to assess their self reported level of autistic traits. Eighty percent of the participants scored above a cut-off point of 32, which exactly matches the percentage originally reported by Baron-Cohen et al (2001). 88% of the participants scored above 26, which has recently been suggested in two separate studies as a more sensitive cut-off point for the AQ (Kurita, Koyama, & Osada, 2005; Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005). Participants were recruited from several sources, including the Autism Research Centre's volunteer database, a local clinic for adults with AS/HFA, and an advert in the National Autistic Society magazine *Communication.* Participants were accepted to the study only if they had not participated in any related intervention during the last three months and had no plans for engaging in another intervention during the study. Participants were randomly allocated into two groups:

(1) Software home-users: 19 participants (14 males and 5 females) were asked to use the software (provided free of charge) at home, by themselves, for two hours a week over a period of ten weeks (a total of 20 hours). Participants were included in the study if they completed a minimum of ten hours of work with the software. If they did not complete this minimum, participants were given an extension of up to 4 weeks to do more work with the software. Out of 24 participants originally recruited to this group, 3 withdrew during the 10 week period and 2 others were excluded at the end, as they failed to reach the 10 hour minimum. No specific pattern was found for these participants: They varied in their age range (21-43), education (three had carried on studying beyond compulsory education, two had not), and employment status (2 were unemployed, 3 were employed). Their IQ, AQ and Time 1 assessment task scores ranged within one standard deviation of their group means. All of them related dropping out/not completing their work to being too busy.

(2) AS/HFA control group: 22 participants (17 males and 5 females) attended the assessment meetings with a 10-15 week period between them, during which they did not take part in any intervention related to emotion recognition.

(3) Typical control group: 28 participants were recruited for this group from a local employment agency. Participants reported no psychiatric history and no occurrence of autism spectrum conditions in their families. After screening for autistic spectrum conditions using the AQ (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), 4 participants were excluded for scoring above the more sensitive cut-off of 26. The remaining 24 participants (19 males and 5 females) attended one assessment meeting.

All participants were given the Wechsler Abbreviated Scale of Intelligence (WASI), (Wechsler, 1999) and scored above 70 on both verbal and performance scales.

One way analysis of variance of participants' AQ scores was significant ($F[2,62]=81.01$, $p<.001$). Tukey post-hoc comparisons revealed that the two clinical groups scored significantly higher on the AQ compared to the typical control group, but did not differ from each other on it. Participants with AS/HFA were asked to report any psychiatric comorbid diagnoses. 5 participants in each clinical group had another psychiatric diagnosis, such as depression or ADHD. No difference was found between the AS/HFA groups in the proportion of participants with psychiatric comorbidity (χ^2[1]=0.07, n.s.). The groups' background data is reported in Table 6.1.

	Software home users (N=19)	AS/HFA controls (N=22)	Typical controls (N=24)	F (2,62)
Age	30.49 (10.27) 17.5-48	30.95 (11.24) 17.5-52	25.33 (9.08) 17.5-51	2.14
Verbal IQ	108.26 (13.29) 80-127	109.68 (10.0) 93-129	115.79 (13.66) 86-138	2.31
Performance IQ	111.95 (12.63) 94-134	115.27 (12.32) 97-140	112.54 (8.93) 92-129	0.53
AQ	37.16 (8.43) 20-47	38.23 (7.45) 16-49	14.04 (5.93) 6-26	81.02[1]
				χ^2 (2)
% Females	26.3%	22.7%	20.8%	0.18
% Left handed	21.1%	9.1%	8.3%	1.91
% Employed	47.4%	40.9%	45.8%	0.20
% A levels or above[2]	68.4%	63.6%	58.3%	0.47
% with comorbid diagnoses	26.3%	22.7%	N/A	0.07[3]

Table 6.1: Means (standard deviations), ranges, and proportions of background variables for the three groups of Experiment 1.

The three groups were matched on age, verbal and performance IQ, handedness and sex. They spanned an equivalent range of employment and educational levels. As shown in Table 6.1, no significant differences were found between the groups for age

[1] p<.001. All other test results are not significant (p>.05).
[2] A levels are the first component of non compulsory education in the UK.
[3] df=1 for this test.

(F[2,62]=2.14, n.s.), verbal IQ (F[2,62]=2.31, n.s.), performance IQ (F[2,62]=0.53, n.s.), sex (χ^2[2]=0.18, n.s.), handedness (χ^2[2]=1.91, n.s.), education (χ^2[2]=0.47, n.s.) and employment status (χ^2[2]=0.20, n.s.). In addition, no difference was found between the two clinical groups in the length of time between the two assessment meetings (t[27.7]=1.57, n.s.).

6.1.2 Instruments

6.1.2.1 Close generalisation: The Cambridge Mindreading Face-Voice Battery, Adult version (CAM-A)

This battery tests complex ER from short silent clips of faces and from voice recordings. 50 faces and 50 voices, taken from *Mind Reading*, test recognition of 20 different complex emotions and mental states (e.g. intimate, insincere). The face task comprises silent clips of adult actors, both male and female, of different ethnicities, expressing the emotions in the face. The voice task comprises recordings of short sentences expressing various emotional intonations. In both tasks, 4 adjectives are presented after each stimulus is played and participants are asked which adjective best describes how the person feels. Items were presented on a computer screen in random order, using DMDX experimental software (Forster & Forster, 2003). A handout of definitions of all the adjectives used in the battery was available for the participants at the beginning and through the assessment. There was no time limit for answering. The CAM-A took about 45 minutes to complete. It provides an overall facial and an overall vocal emotion recognition score (max=50 for each of them), as well as individual scores for each of the 20 emotions assessed (pass/fail, i.e. recognised above chance or not) and an overall number of the emotions correctly recognised (max=20). Test-retest correlations, calculated for the ASC control group in the current experiment were r=0.94 for the face scale, r=0.81 for the voice scale, and r=0.97 for the number of concepts recognised (p<.001 for all). Creation and validation of the battery is described in detail in Appendix 1.

6.1.2.2 Distant generalisation visual task – Reading the Mind in the Eyes, Adult version (RME-A)

The most recent version of the test (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) was used. This is an improvement over the original version (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997) in that it increased the number of items used and the number of answer options per item (from 2 to 4), to improve its power. The revised task has 36 items, in which participants are presented with a photograph of the eyes region of the face and they are asked to choose which one of four adjectives or phrases best describes the mental state of the person pictured (see Figure 1.2 for an example). In the adult study, the pictures and adjectives were played on a computer screen in random order (using DMDX software). A definition handout was provided at the beginning of the task and could be consulted throughout the assessment. Therefore, there was no time limit for answering. The task took about 15 minutes to complete. Test–retest correlation, calculated for the AS/HFA control group in the current experiment was $r=0.86$ ($p<.001$).

6.1.2.3 Distant generalisation auditory task – Reading the Mind in the Voice, Revised (RMV-R)

This task includes 25 items and is a revised version of the original task (Rutherford, Baron-Cohen, & Wheelwright, 2002). Each item includes a short segment of speech, followed by 4 emotion and mental state adjectives (adding 2 to those of the original task). Participants are asked to tell how the speaker is feeling by choosing one of the 4 available answers. Foils were selected to match the content of the verbalisations but not the intonation, thus making the task harder to answer. This avoided ceiling effects that the original version of the test was prone to. The test items were played on the computer in random order, preceded by an instruction slide and two practice items. Participants were given a definition handout before the beginning of the task. There was no time limit for answering. The RMV-R took about 20 minutes to complete. Test–retest correlation, calculated for the AS/HFA control group in the current experiment was r=0.8 (p<.001).The creation and validation of the revised task is described in detail in Appendix 2.

6.1.2.4 Holistic distant generalisation task – Reading the Mind in Films, Adult version (RMF-A)

This task comprises 22 short social scenes taken from feature films. Each scene includes visual, vocal and some contextual information. Scenes are presented on the computer screen (using DMDX software). Participants are presented with 4 adjectives and are asked to choose the one that best describes the way a target character feels at the end of the scene. Foils were selected to match some aspects of the scene (e.g. content of speech) but not all of them (e.g. facial expression, intonation, etc), thus making the task harder to answer. A handout of definitions of all the adjectives used in the task was available for the participants at the beginning and through the assessment. There was no time limit for answering. The RMF-A took about 20 minutes to complete. Task score is the number of correctly recognised emotions (max=22). Creation and validation of the task are described in detail in Appendix 3.

6.1.2.5 Follow-up measure: The Friendship and Relationship Questionnaire (FQ)

This questionnaire assesses the degree to which respondents have a need or enjoy close, empathic supportive and caring friendships, that are important to them, their interest in people, their ability to enjoy interaction with others for its own sake (e.g. not as part of work) and their perception of friendship as important. It is scored such that if a person's relationships are more based around a shared activity they get a lower score, and if a person's relationships are more based around close, confiding, intimacy and emotional reciprocity as an end in itself, they get a higher score. The FQ comprises 27 multiple choice questions (and 8 foils). Score for each item ranges between 0-5. Hence, the maximum score on the FQ is 135. Examples of items include: "How easy do you find it to make new friends?" or "I don't have anybody who I would call a best friend". Adults with ASC tested on the FQ scored significantly lower than male and female controls from the general population. Internal consistency calculated for the questionnaire revealed Cronbach's α of 0.75 (Baron-Cohen & Wheelwright, 2003).

The FQ was sent to all adult participants prior to their time 1 assessment. The follow-up measure of the adult intervention experiments included the FQ, which was preceded by a short questionnaire about the use of *Mind Reading* since the second assessment meeting. Participants were asked to estimate how much time they spent using *Mind Reading* since the second assessment meeting, and to rate how helpful *Mind Reading* was for their ability to recognise emotions in every day life, their ability to understand social situations, and their confidence in social situations on a 1 (not helpful at all) to 4 (very helpful) likert scale. Participants were also asked to describe how relevant and useful they found *Mind Reading* in their everyday life and in which areas. These questions provided some more qualitative information about the experience the participants had with *Mind Reading* and its relevance for their socio-emotional life in the long run. The follow up questionnaire, including the FQ, appears in Appendix 4.

6.1.2.6 The Autism Spectrum Quotient (AQ)

The AQ is a self-report questionnaire, which measures the degree to which any (adult) individual of normal IQ possesses traits related to the autistic spectrum. The AQ has subscales for communication (e.g. 'I frequently find that I don't know how to keep a conversation going'), social skill (e.g.' I find it difficult to work out people's Intentions'), imagination (e.g. 'I don't particularly enjoy reading fiction'), attention to detail (e.g. 'I tend to notice details that others do not'), and attention switching (e.g.' I frequently get so strongly absorbed in one thing that I lose sight of other things'). Scores range from 0-50, and the higher the score, the more autistic traits a person possesses. Subscales' internal consistency ranges between 0.63-0.77 and test-retest reliability for the AQ is r=0.7. The AQ was recently reported to have good discriminative validity when validated against diagnostic interviews in a clinical setting (Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005).

6.1.3 Procedure

Participants were individually tested at the Autism Research Centre in Cambridge. Four trained experimenters individually helped the participants through the assessments. Three experimenters were blind as to which group the participants belonged. Participants in the intervention group were asked to help in the evaluation of a new piece of software. It was explained they would need to commit to using *Mind Reading* for 2 hours a week over a period of 10 weeks and to be assessed before and after this training period. Participants of both control groups were asked to take part in an emotion recognition study, helping to validate new tasks. For this reason, participants in the AS/HFA control group were asked to come for two assessments, separated by a 10-15 week period. Participants' written consent was obtained. All participants were told they were free to withdraw from the study at any time.

In the first assessment background information was collected, as well as the AQ and FQ that participants filled out in advance. Participants were then seated in front of IBM compatible computers with 15 inch monitors to take the various ER tasks. They were given headphones for the voice tasks. The CAM-A, RME-A, and RMV-R tasks were presented in a random order. Controls from the general population also took the RMF-A (as this was their only assessment meeting). Two breaks were given during the CAM-A battery, and one during each of the other tasks. Participants were also allowed a break to freshen up between tasks. In between the tasks, two subtests of the WASI were administered.

After the assessment was completed, participants in the ASC and typical control groups were thanked and told their involvement was over. The participants of the intervention group were introduced to *Mind Reading* in detail. This included a presentation of the emotion groups and levels, the emotions library, learning centre, and game zone, and the reward section. Participants watched a demonstration of a systematic analysis of an emotion, comparing different faces and voices to identify the unique facial/intonation features of this emotion. They were encouraged to analyse the facial and vocal stimuli systematically, using the emotion group structure and the feature-based analysis of the facial and vocal examples of emotion. Participants were asked to use the emotions library and learning centre as they wished, but not to use the game zone for more than a third of the usage time (to ensure they do not spend too much time playing instead of systematically working through the emotions). Participants were also told *Mind Reading* logged their work and they were asked to bring the log file to the second assessment meeting, for usage time verification. The whole assessment meeting took about 3 hours. The research team provided technical support for installation and use of *Mind Reading* during the time between the two assessments. In addition, participants of the intervention group were approached by telephone at least once, to check they were still committed to the study and to working with the software.

In the second assessment meeting participants with ASC from the intervention and control groups took the same ER tasks they took in the first assessment, including the RMF-A and the other two WASI subtests. Task administration order was randomised. The log files of participants in the intervention group were checked to verify they had used the software for the required amount of time. Average usage time was 17.5 hours, (SD=6.7, range: 10-36). These participants were then asked for their feedback about the program, their experience with it and comments about its usefulness. Participants of the AS/HFA control group were asked about any possible effects of the first assessment meeting on their interest in emotions during the period between the assessments. All participants were then debriefed about the aims and design of the study and were rewarded with a complimentary copy of *Mind Reading* (or were allowed to keep the copy they used). This assessment meeting took about 3 hours.

6.2 Results

One sample Kolmogorov-Smirnov tests were conducted for all task scores in the three groups. Distributions of all scores in both groups did not differ from normal. Hence, parametric analysis was used.

First, performance of the three groups on the emotion recognition tasks at Time 1 was explored. Five one way ANOVAs were conducted, testing group differences on the emotion recognition tasks used at Time 1. Using Holm’s Sequential Rejective

Bonferroni Procedure[4], significant differences were found between the groups on the CAM-A face task ($F[2,62]=13.82$, $p<.001$), CAM-A voice task ($F[2,60]=11.53$, $p<.001$), CAM-A number of emotional concepts recognised ($F[2,60]=12.77$, $p<.001$), the RME-A ($F[2,62]=6.10$, $p<.01$), and the RMV-R task ($F[2,62]=4.92$, $p<.02$). Pre-planned comparisons with Bonferroni corrections revealed no significant differences between the two clinical groups on any of the task scores, and significantly higher scores of the typical control group on all tasks, comparing to the two AS/HFA groups. These findings support Hypothesis 1. Table 6.2 shows the means and standard deviations of the groups' emotion recognition scores at Time 1 and Time 2.

	Software home users[5]		AS/HFA controls		Typical controls
	Time 1	Time 2	Time 1	Time 2	
CAM-A face task (Max score=50)	31.32 (8.76)	37.50 (7.78)	32.55 (8.38)	34.77 (8.19)	42.04 (5.18)
CAM-A voice task (Max score=50)	33.76 (6.59)	38.89 (6.16)	35.18 (7.36)	36.64 (7.87)	42.08 (4.19)
CAM-A no. of concepts recognised (Max score=20)	9.82 (5.16)	13.61 (4.82)	10.55 (5.23)	11.27 (5.3)	16.08 (2.99)
Reading the Mind in the Eyes-A (Max score=36)	23.05 (6.75)	23.84 (4.72)	23.86 (6.75)	23.05 (7.30)	28.54 (3.12)
Reading the Mind in the Voice-R (Max score=25)	16.05 (2.93)	16.68 (3.89)	16.09 (3.87)	17.41 (3.51)	18.58 (2.39)
Reading the Mind in Films-A (Max score=22)		11.79 (3.77)		12.86 (3.39)	15.46 (2.43)

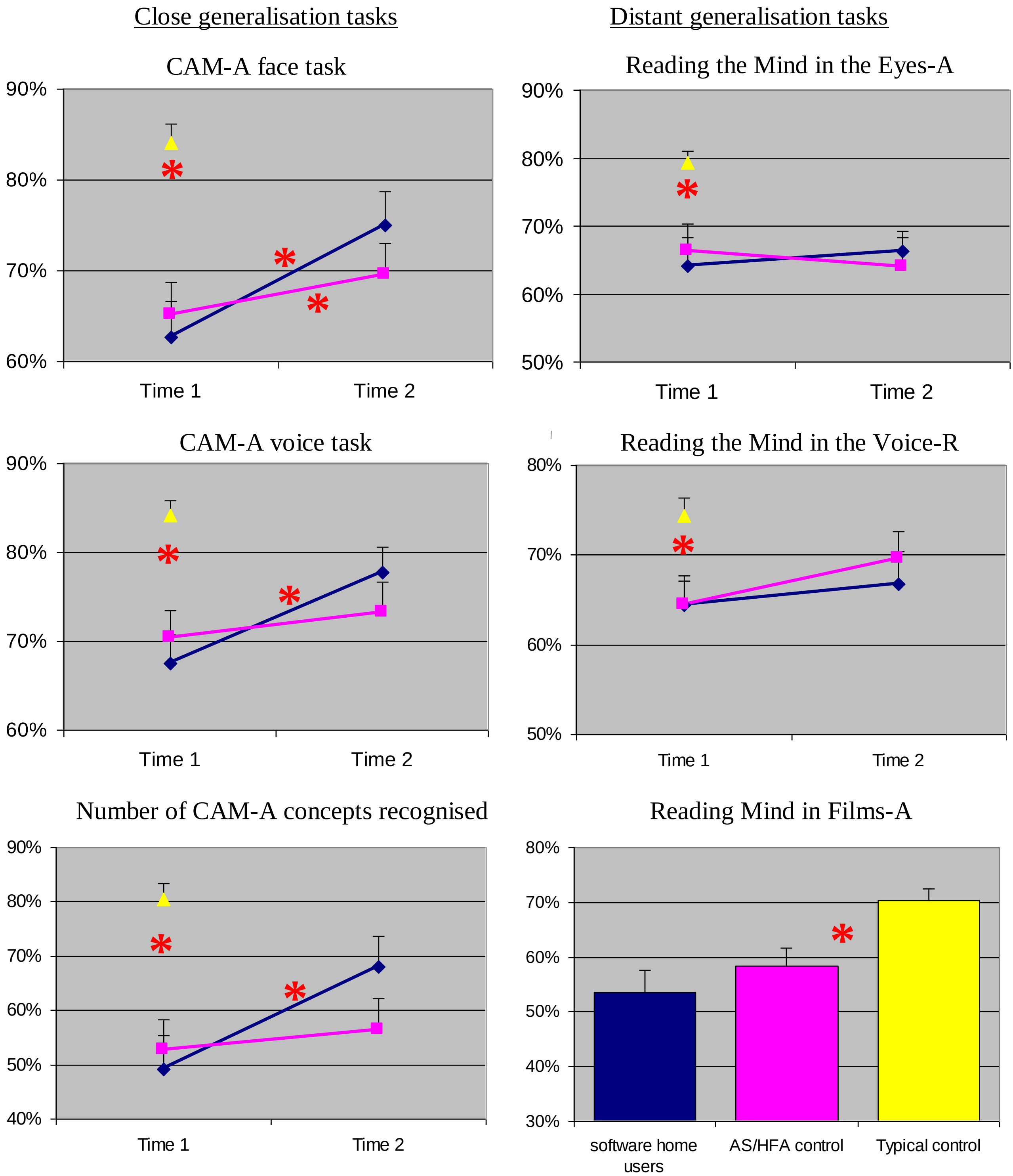
Table 6.2: Means (and standard deviations) of the emotion recognition measures in the 3 groups of Experiment 1 at Time 1 and Time 2

[4] Holm's Procedure is held to minimise the chance of type 1 error due to multiple comparisons, while keeping the power of the analysis (i.e. avoiding type 2 errors). The procedure is conducted stepwise, comparing successively higher p-values with increasingly greater significance levels. The different comparisons are ordered in increasing order according to their p-values. The smallest p-value is then compared against the most conservative α level (in this case, $0.05/5=0.01$). If p is smaller than α, H0 is rejected and the p-value of the next comparison is compared against a less conservative α level (in this case $0.05/4=0.0125$). The procedure is continued stepwise (i.e. $\alpha=0.05/3$, $0.05/2$ and eventually 0.05), with each successive test conducted at progressively higher significance levels. When one of the effects fails to reach significance, the procedure is stopped and all remaining effects are deemed non-significant. (Holm, 1979; see also Zhang, Quan, Ng, & Stepanavage, 1997).

[5] Two participants of this group did not complete the CAM-A voice task at Time 1 and one participant did not complete the CAM-A faces at Time 2. These participants were excluded from the analysis of the three CAM scores. Other than that, groups' sizes are identical to Table 6.1.

Next, five multivariate analyses of covariance with repeated measures were conducted, to examine the differences between the intervention and AS/HFA control group on the various tasks at Time 1 and Time 2. Age, verbal and performance IQ were used as covariates. Using Holm's Sequential Rejective Bonferroni procedure, significant time by group interactions were found for all CAM-A measures: faces ($F_{wilks}[1,35]=11.82$, $p<.002$), voices ($F_{wilks}[1,34]=7.51$, $p<.01$), and number of concepts recognised ($F_{wilks}[1,33]=8.38$, $p<.01$), but not for Reading the Mind in the Eyes ($F_{wilks}[1,36]=1.46$, n.s.) or Reading the Mind in the Voice-R ($F_{wilks}[1,36]=0.47$, n.s.). No main effects were found significant, but the effect of the covariate verbal IQ was significant in relation to CAM-A voice scores beyond time or group ($F[1,34]=5.11$, $p<.05$).

Simple main effect analyses for the three CAM-A scores with Bonferroni corrections revealed that in accordance with Hypotheses 2 and 3, the intervention group improved significantly from Time 1 to Time 2 on all three scores (for faces: $t[17]=5.37$, $p<.001$; for voices: $t[16]=5.24$, $p<.001$; for concepts recognised: $t[15]=3.96$, $p<.005$). The AS/HFA control group scores did not change significantly from Time 1 to Time 2 on the CAM-A voices ($t[21]=1.43$, n.s.) and the number of concepts recognised ($t[21]=1.25$, n.s.), but did so on the CAM-A faces task ($t[21]=3.51$, $p<.005$). However, when a t-test was conducted on Time 2 minus Time 1 score differences, the improvement of the intervention group was significantly greater than that of the AS/HFA control group ($t[38]=3.38$, $p<.005$). Table 6.2 shows the mean scores of all tasks at Time 1 and Time 2. Figure 6.1 shows these results in terms of proportion of correct answers. Proportions were used in the graphs instead of raw scores, in order to keep the scale uniform for all tasks.



Legend: ◆ Software Users ■ AS/HFA Controls ▲ Typical Controls * p<.01

Figure 6.1: Mean proportions of correct responses (with standard error bars) for the 3 groups on the two levels of generalisation – Experiment 1

Next, score differences for the 20 CAM-A concepts were computed (Time 2 score minus Time 1 score). A multivariate analysis of variance was conducted for the 20 difference scores, with group as the independent variable. The MANOVA did not yield a significant group difference beyond emotional concept ($F_{wilks}[20,19]=1.36$, n.s.) but yielded significant individual between group effects for the concepts: *grave* ($F[1,38]=5.81$, $p<.05$), *uneasy* ($F[1,38]=5.2$, $p<.05$), *lured* ($F[1,38]=6.98$, $p<.05$), *intimate* ($F[1,38]=5.70$, $p<.05$), *insincere* ($F[1,38]=6.79$, $p<.05$), and *nostalgic* ($F[1,38]=18.48$, $p<.001$). These findings suggest that the intervention group's recognition of these emotional concepts improved more than it did in the AS/HFA control group. However, care should be taken when interpreting these findings, due to the lack of an overall effect for the MANOVA, and the multiple tests conducted under it.

An analysis of the RMF-A scores was conducted next, to test for holistic distant generalisation. This task was only taken at Time 2 by the clinical groups, so differences could only be measured between groups. Therefore, a one way ANOVA was conducted on the three groups' RMF-A scores and was significant ($F[2,62]=7.68$, $p<.01$). Pre-planned contrasts with Bonferroni correction revealed no difference between the intervention group ($M=11.79$, $SD=3.77$) and the AS/HFA control group ($M=12.86$, $SD=3.39$) for this task ($t[62]=1.07$, n.s.). However, the typical control group ($M=15.46$, $SD=2.43$) scored significantly higher than both groups ($t[62]=3.81$, $p<.001$). These results suggest the intervention group did not perform better than the AS/HFA control group on this level of generalisation.

Lastly, a correlation analysis was conducted between the time participants had used the software and the improvement scores of each task. Significant Spearman Rho correlations were found for software usage time with improvement on the CAM-A voice task ($r_{spearman}=0.59$, $p<.01$), and with RMF-A task scores ($r_{spearman}=0.45$, $p<.05$). No other correlations were significant.

In their feedback at the end of the Time 2 assessment meeting, participants who used *Mind Reading* reported that the software was both useful (reported by n=9 participants) and enjoyable to work with (n=5). Several participants referred to the systematic nature of the software and its merit for them, e.g.: "*I am rather detailed* *and pedantic in the way that I learn things…I have to work through things systematically. This often involves playing around with details again and again until I feel the information has 'locked' itself into my brain. Mind Reading allowed me to concentrate on the bits I found most difficult and study them closely without having to do anything else. For me this was the best bit*". More specifically, participants said that comparing different examples of the same emotion, and distinguishing different emotions were helpful (n=4). And that *Mind Reading* was useful for understanding of subtle differences between expressions (n=3). In addition, participants said the software could be worked with independently, with no need for external support.

With regards to *Mind Reading*'s effect on their social functioning in everyday life, participants suggested they became more interested in looking at people's faces and engaging in eye contact (n=9), e.g. "*I did an experiment on the bus today, watching the face of a woman who was talking to someone. I could detect smiley type reactions, interest and empathy. Then I watched a child react to his parents, and noticed the automatic eye contact there. I never used to look at these before"*. Besides their interest in emotions in everyday life, participants reported that using the software made social situations easier for them (n=2), e.g "*I find I can now tell friendly smiles from polite or false ones*" or "*I can recognise how other people feel at my work place*". However, other participants argued that training using *Mind Reading* was too distant from the required functioning in real life, and that they found it difficult to transfer their knowledge (n=2), e.g. "*I feel it is not too hard to distinguish the faces according to the program with all the cues and limited choices, but it is not the same in real life when you are faced with a greater number of choices and the fact that one's attention is engaged in other matters*". Another participant commented that the separation between faces and voices had made learning harder, and different to real-life requirements.

6.3 Follow-up

6.3.1 Procedure

About a year (12-15 months) after the Time 2 assessment, participants in the ASC groups were sent the follow-up questionnaire shown in Appendix 4. The questionnaire was emailed to participants who had access to email, or posted to those who did not. A second letter was sent a month later to participants who did not respond to the first one.

6.3.2 Participants

Of the 41 participants with ASC who took part in the study at Time 1 and Time 2, eight could not be located for the follow up, and 18 completed and returned the follow-up questionnaires, nine who originally belonged to the AS/HFA intervention group (4 females, 5 males), and nine from the AS/HFA control group (3 females, 6 males). The two groups did not differ on age, verbal IQ, performance IQ, and the AQ and FQ scores from Time 1. In addition, there was no difference on the average estimated number of hours using *Mind Reading* since Time 2 between participants from the intervention group (M=6.11, SD=11.52) and participants from the AS/HFA control group (M=4.44, SD=3.91 ; t[16]=.41, n.s.). Background information for these participants appears in Table 6.3

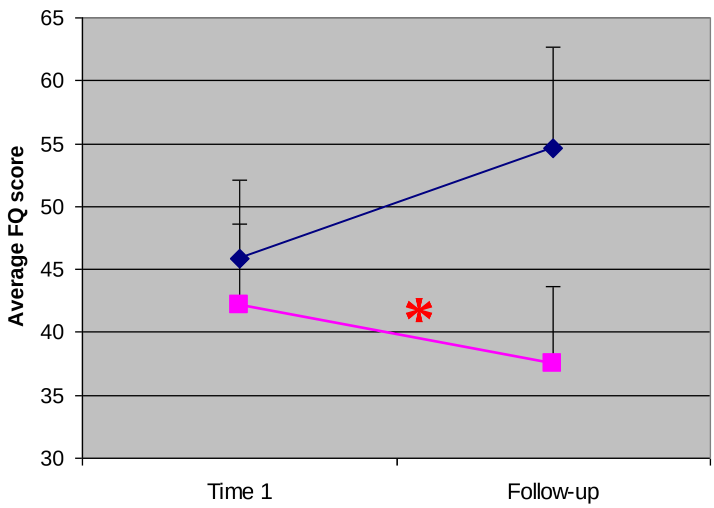
Time 1 measures	Participants from intervention group (N=9)			Participants from AS/HFA control group (N=9)			
	Mean	SD	Range	Mean	SD	Range	t (16)[1]
Age	29.21	9.68	17.6-44.1	33.57	12.10	18.8-52	.84
Verbal IQ	112.89	8.42	102-125	113.0	10.92	95-129	.02
Performance IQ	113.56	10.20	101-129	122.33	13.86	99-140	1.53
AQ	39.56	6.54	28-47	40.89	5.28	31-49	.48
FQ	45.89	18.64	10-67	42.11	19.42	15-80	.42

[1] p>.1 for all t-tests

Table 6.3: Time 1 background averages, standard deviations and ranges for the participants who filled in the follow-up questionnaire – Experiment 1.

6.3.3 Results - FQ

To check for long term distant generalisation effects of the use of *Mind Reading* during the original intervention period, the two groups were compared on the difference between their Time 1 and follow-up FQ scores. However, since participants in both groups received the software after the Time 2 assessment and had the chance to use it, estimated usage time was taken into account. Hence, after ensuring that the measures' distributions do not significantly differ from normal in the two groups (using Kolmogorov-Smirnoff tests), a repeated measures MANCOVA was conducted, with FQ Time 1 and follow-up scores as the dependent variables, group as the independent variable, and estimated usage time of *Mind Reading* since the Time 2 assessment as a covariate. The analysis yielded a significant time by group interaction ($F_{wilks}[1,15]=10.57$, $p<.006$). Analysis of simple main effects revealed that FQ scores of participants who belonged to the AS/HFA control group decreased significantly from Time 1 (M=42.11, SD=19.42) to follow-up (M=37.56, SD=18.23; t[8]=4.03, p<.005), but FQ scores of participants who belonged to the intervention group showed a non-significant increase from Time 1 (M=45.89, SD=18.64) to follow-up (M=54.67, SD=24.12; t[8]=1.94, p=.088). The interaction between group and Time is illustrated in Figure 6.2.



Legend: ◆ Software Users ■ AS/HFA Controls * p<.005

Figure 6.2: Average FQ scores (with error bars) for participants from the intervention and AS/HFA control groups at Time 1 and Follow-up

An interaction between Time and estimated number of *Mind Reading* usage hours (after the Time 2 assessment) was also found ($F_{wilks}[1,15]=10.34$, $p<.01$). A positive correlation was found between the number of usage hours and the difference between follow-up and Time 1 FQ scores ($r=.58$, $p<.02$), suggesting that beyond group, the more time participants spent using *Mind Reading*, the more their FQ scores increased (or the less decreased) from Time 1 to follow-up.

To check whether FQ scores change could be predicted by background, a multiple regression analysis was conducted on follow-up minus Time 1 FQ scores. Predictors included age, verbal and performance IQ, sex, AQ scores, group, and time used *Mind Reading* since the Time 2 assessment. Stepwise method was used, due to the great number of predictors and the small number of participants. Only two predictors entered the regression equation: group ($\beta=.44$, $p<.05$), and *Mind Reading* usage time since Time 2 ($\beta=.58$, $p<.01$). Together these two predictors explained adj. $R^2=56.0\%$ of the variance in the dependent variable ($F[2,13]=10.54$, $p<.005$). No other effects were found significant in the regression or in the bivariate correlation analysis held with the above mentioned variables.

6.3.4 Results – follow-up feedback questionnaire

The follow-up feedback questionnaire required participants who used *Mind Reading* since the Time 2 assessment to estimate how helpful this was for their functioning in the socio-emotional domain. Three subjects were predetermined: Recognition of emotions in everyday life, understanding social situations, and confidence in social situations. Participants were also asked to add any of their own areas they found *Mind Reading* to be useful in, and to comment about the usefulness and relevance of the software for their everyday life. Of the 18 participants who returned the follow-up questionnaires, 14 used the software after Time 2 assessment (6 from the intervention group and 8 from the AS/HFA control group). Since there were no significant group differences on their ratings, they were analysed beyond group. The ratings for each of the three areas appears in Table 6.4. In order to check whether *Mind Reading* was perceived to be more helpful in any particular area, Friedman's non parametric test for related samples was conducted, using ratings on the three areas, and was found significant (χ^2 (2)=13.82, p<.001). Pairwise post-hoc comparisons, corrected for Bonferroni, included three Wilcoxon Signed Ranks tests, each comparing ratings for two of the areas. The contribution of *Mind Reading* for recognition of emotions in everyday life was rated as higher than the software's contribution for understanding of social situations (Z=2.65, p<.01) or its contribution for the respondents' confidence in social situations (Z=2.72, p<.01). There was a tendency for a higher rating of *Mind Reading*'s contribution to the respondents' understanding of social situations, compared to its contribution to their confidence in social situations. However, this comparison did not reach significance (Z=1.93, p=.053).

How helpful was *Mind Reading* for:	Not helpful at all (1)	Not very helpful (2)	Quite Helpful (3)	Very Helpful (4)	Mean
Emotion recognition in every day life	-	7.1%	64.3%	28.6%	3.21
Understanding of social situations	-	28.6%	71.4%	-	2.71
Confidence in social situations	21.4%	42.9%	28.6%	7.1%	2.21

Note: n=14

Table 6.4: Participants' ratings of *Mind Reading*'s contribution for socio-emotional functioning in real-life – Experiment 1

Correlation analysis for the first two items (*MR* helpful for ER in everyday life, and *MR* for understanding of social situations) used Spearman's Rho, as the distributions of ratings of these two items differed significantly from normal (Kolmogorov-Smirnoff test). Neither of the items correlated significantly with age, verbal or performance IQ, AQ, number of hours used *Mind Reading* since the Time 2 assessment, or score differences on the CAM-A battery. For the third item, '*MR* helpful in getting confidence in social situations', Pearson's correlations were used, as this item's distribution did not vary from normal. A positive correlation was found between ratings on this item and the amount of time spent using *Mind Reading* since Time 2 (r=0.53, p<.05), suggesting the more hours participants used the software, the more they felt it helped them feel confident in social situations. However, since a large number of correlations was conducted, the chance of a type 1 error could have increased and this result should be interpreted with caution. No other correlations were found significant.

In addition to the three areas included in the follow up questionnaire, individual participants suggested *Mind Reading* was quite helpful for their ability to "*communicate on a more emotional level*" and for "*taking an interest in facial expressions*".

Comments respondents made with regards to *Mind Reading*'s relevance and usefulness in their everyday life dealt largely with the three following themes:

❖ *Awareness to the importance of emotions and emotional cues*

Five participants commented on the contribution using *Mind Reading* had to their awareness of the importance of emotional and expressions of emotions in social functioning, e.g. "*it made me aware of these things, which I was previously unaware of*". However, one participant added that this awareness has made him more self-conscious, as he realised he does not express such emotional cues, and was worried about being misunderstood.

❖ *Understanding of emotions and their corresponding facial and vocal expressions*

Five respondents found *Mind Reading* helpful with understanding emotional subtleties, e.g. "*differentiating between genuine interest and mere politeness*", "*understanding that conflicting emotions may have similar features in facial expression*". Some commented they were previously able to generally recognise *happy* and *sad*, and that *Mind Reading* was helpful by teaching more than the basic emotions, and by "*helping distinguish emotions by small differences*". Two respondents referred to the software as a 'reference catalogue of emotions' they approached when wondering about the meaning or the associated expression of an emotion.

❖ *Relevance to real life functioning*

Six of the respondents commented that the software helped them recognise emotions in everyday life, e.g. "*what to look for in expressions*", and that it "*increased the range of facial expressions that I can recognise*". Two participants found the software helped them improve their relationships with their partners, as they were more able to decipher changes in their emotional expressions. Two other participants found the software helped them improve their motivation and ability to communicate with others, e.g. "*I now try to hold conversations*". However, five of the respondents argued it was difficult to transfer the knowledge they had acquired using *Mind Reading* to social situations in everyday life. This was related to the difference between the quiet, controlled, training situation which provided immediate feedback on performance, and the intense, anxiety-inducing real-life social situation, during which no feedback is provided on accuracy of ER. Respondents suggested that in order to bridge this gap between the training situation and real-life functioning, information about body language (not covered on *Mind* Reading) and more references to actual social situations might be useful. In addition it was suggested to include *Mind Reading* in social skills groups, in order to enhance generalisation.

6.3.5 Discussion

In this experiment, the use of *Mind Reading* at home was compared to no intervention. Results showed that following 10-20 hours of using the software over a period of 10-15 weeks, users significantly improved in their ability to recognise complex emotions and mental states from both faces and voices, compared to their performance before the intervention, and relative to the control group. This finding is interesting, considering the short usage time and the large number of emotions included in the software, and since participants were not asked to study these particular emotions. The long term follow-up of using the software at home showed that taking part in the intervention group a year before served as a protective factor from deterioration in social relationships, which were reported by the control ASC group. Using *Mind Reading* after the study was over was also associated with improvement on the friendship and relationship measure.

Time 1 results support Hypothesis 1, confirming the ER difficulties from faces, voices, and eyes, and from holistic social situations, reported in the previous Chapters. Since no differences were found between the clinical groups at Time 1 on any of the ER measures or the background variables, any difference at Time 2 can be attributed to the intervention.

Supporting Hypothesis 2, the intervention group improved significantly on close generalisation measures, including faces, voices, and emotions individuals with AS/HFA had particular difficulties with on the CAM-A (see Appendix 1). This improvement was significantly higher than that of the AS/HFA control group, supporting Hypothesis 3. Improvement amongst software users on the ability to recognise mental states such as *intimate, insincere* or *grave* might have a positive effect on their confidence, willingness and functioning in interpersonal situations. This, together with participants' reports of greater attention to faces and emotions, and improved eye contact, suggests that the systematic analysis of emotions using *Mind Reading* allows people with ASC to improve emotion recognition skills and may have a positive effect on their social functioning too.

Against Hypotheses 2 and 3, improvement following the intervention was limited to close generalisation tasks, i.e. to faces and voices taken from *Mind Reading*. No improvement from Time 1 to Time 2, and no difference between the intervention and AS/HFA control group were found on either feature-based or holistic tasks of distant generalisation. Participants' reports of difficulties generalising from the training environment to real-life functioning further support these results. Similar findings of poor generalisation have been found in studies teaching ToM, ER and social skills to individuals with ASC (see Chapter 3). These findings are discussed in more detail in Chapter 9. Software usage time was, however, positively correlated with RMF-A scores, suggesting that the more participants used the software, the higher they scored on the holistic distant generalisation task. It is possible that a longer period of usage would have led to improved generalisation amongst software users.

Interestingly, the control group significantly improved on the CAM-A face task between Time 1 and Time 2, despite having no formal intervention. This could be the result of taking the same task for a second time after a relatively brief interval. In addition, when interviewed at the end of the second assessment, participants in this group reported greater interest in emotions following the assessment at Time 1 (e.g., "*I looked more at faces, though I couldn't tell what to look for*"). Therefore, it is possible that the assessment itself served as a limited short-term intervention, arousing participants' awareness of the importance of faces and emotions. This new awareness was not sufficient to cause an improvement on voices, but did allow for improvement on faces, which might suggest this domain is more easily changed through intervention. However, the improvement in this group's CAM-A face task scores was significantly smaller than that of the software home users' group (in accordance with Hypothesis 3).

Verbal IQ was found to have a significant effect on CAM-A voice task scores. It is important to note that verbal IQ was correlated with performance on the CAM-A voice task beyond Time, i.e. it was not related to the ability to improve or learn from *Mind Reading*. The inclusion of verbal content in the voice task segments of speech (rather than just intonation) might account for this effect, and for the lack of it in the face task. However, no such effect was found for the RMV-R, which also includes verbal content in its items. Experiment 2 was in part intended to help determine how central the role is that verbal IQ plays in the ability to recognise emotions.

Longitudinal and follow-up studies of individuals with ASC report a general improvement in socialising and communication skills from adolescence into adulthood (Howlin, 2000; McGovern & Sigman, 2005; Tantam, 1991). In view of this, the effects found on the follow-up questionnaire, particularly the decline in self-reported ability to enjoy friendships and close relationships amongst participants from the AS/HFA control group, require careful interpretation. Participant samples represented their original groups relatively reliably, as their age, IQ and AQ averages fell within 1 standard deviation from the mean of their groups. However, no data was collected at follow-up about potential changes in participants' socio-economic status, living arrangements, psychopathology, or treatment taken since the Time 2 evaluation. These factors may have affected the follow-up results. In addition, one might wonder whether participants who replied to the follow-up questionnaire were those who were in greater need of social contact (even if just in the form of correspondence), i.e.

whether the decline (or improvement) in FQ scores represents a specific group, with greater social difficulties. It is also important to note that the FQ scores rely on self-report, which may differ from actual social performance. Another possible problem with individuals with ASC filling in the FQ lies in some of the questions being somewhat hypothetical for them, as they are often socially isolated and have no friends (Orsmond, Krauss, & Seltzer, 2004; Tantam, 2003). As one of the participants commented: "*The question assumes one already has friends. I don't have any.., I have to go into my memory to retrieve what I would do if I had them and how I would behave… I have no such recollections*". Finally, the stability of test-retest scores on the FQ has never been checked before, since Baron-Cohen and Wheelwright (2003) reported single administration data on the questionnaire (Baron-Cohen & Wheelwright, 2003). It is possible that a decline in scores would be found amongst controls from the general population as well. Unfortunately, this was not examined in the current study, and would need to be tested in the future.

All of the points made above call for caution in the interpretation of the follow-up results. However, they should apply to all participants with ASC, and therefore fail to explain the significant FQ score differences at follow-up between participants who originated from the intervention group and those from the AS/HFA control group. FQ scores of all the participants in the AS/HFA control group decreased from Time 1 to follow-up, whereas those of participants from the intervention group increased in all but one of the cases. Longitudinal studies describe an increase in awareness to their difference and isolation amongst high functioning young adults with ASC. This awareness, especially in the absence of appropriate support, is often linked to depression, which is the most common co-morbidity in adolescents and adults with ASC (Tantam, 2000, 2003; Tsatsanis, 2003). It is therefore possible that following the assessment, participants in the AS/HFA control group became more aware of their difficulties understanding and interpreting emotions, which affected their perception of their ability to form friendships and intimate interaction. Those who used *Mind Reading* after the Time 2 assessment showed less of a decrease in FQ scores, suggesting the use of the software moderated their level of self-confidence. This is also reflected in the positive correlation between *Mind Reading* usage time after the Time 2 assessment and participants' ratings of the software's contribution to their confidence in social situations.

Amongst participants from the intervention group, using *Mind Reading* regularly between Time 1 and Time 2 may have had a protective effect on their perception and attitudes towards friendship and relationship, as overall their FQ scores did not decrease, and even showed a near-significant tendency to increase from Time 1 to follow-up. Such an increase could have become significant with a larger sample. Though fewer participants carried on using *Mind Reading* after Time 2 in this group (compared to the AS/HFA control group), continuing to use it had a positive effect on their FQ scores too, as was its effect on their feeling of confidence in social situations. If indeed the use of *Mind Reading* is associated with increased awareness of socio-emotional abilities amongst adults with ASC, then this group may have felt the intervention period had equipped them with new skills, which they could use in the social arena. The qualitative feedback showed that this was indeed the case for some of the participants, but that others stressed the difficulty in generalising from the software to real-life functioning. There was no clear pattern of positive or negative comments coming from any of the two groups.

An alternative explanation for the group differences in ER tasks between Time 1 and Time 2, and on FQ score differences between Time 1 and follow-up lies in the design of this experiment. It is possible that the group difference stems from the AS/HFA control group having had no alternative intervention/task to take during the 10-15 week period between assessments. The impact of participants getting some sort of intervention, which usually comes with additional attention from the experimenters, has been shown to bring about positive effects, regardless of the intervention's nature (Brewer, 2002). A more balanced design, should allow both groups to take part in some kind of intervention, in order to limit the effect of this potential confound. Experiment 2, described in the next chapter, used such a design.

Since the use of *Mind Reading* was associated with increased FQ scores (i.e. with increased self-confidence in the area of friendship and relationship), one might wonder why participants did not use it more after Time 2. This could be possibly linked with the difficulties some of the participants in the intervention group had using the software unprompted between Time 1 and Time 2. The executive functioning difficulties individuals with ASC experience (planning and prioritising in particular), described in Chapter 1, may explain why the software was not used more when participants were not prompted to do so for the study. Experiment 2, which examined the effectiveness of using *Mind Reading* with support from a tutor in a group, tested the additional merit of a more structured training environment on participants' use of the software and its effect on their ER and socialising skills.

6.4 Conclusion

Independent use of a computer-based systematic guide to emotions and mental states by high-functioning adults with ASC over a relatively short period of 10-15 weeks brought a significant improvement in their emotion recognition skills. This improvement was significantly higher than that of control participants with ASC who took the assessment tasks twice with no intervention. Whereas this improvement was limited to facial and vocal expressions of emotions covered in the systematic guide and was not found in ER tasks containing material not included in the software, performance on a distant holistic generalisation task improved the more participants used the software. This suggests that a longer intervention period might bring improved generalisation. In addition, using the software was found to have a long-term effect on participants' self-report of friendship and close relationships with others. This effect moderated the deterioration of friendship and relationship status reported by the ASC control group, and showed a tendency to boost participants' friendship and relationships amongst participants who had used the software a year before. The next chapter examines whether these effects are replicated, or are even greater, when the independent use of the software is complemented with tutor and group support, and when this intervention is compared to another, non-systematic intervention.

7 Adult intervention study: Tutor and group supported use of Mind Reading

This experiment tested the effect of using *Mind Reading* with a weekly support of a tutor in group meetings, using activities aimed for enhancing generalisation. This was compared to the effect of taking part in a ten week social skills training course without using the software. Improvement of ER skills at the three levels of generalisation and FQ at follow-up after a year were the target outcome measures. This experiment tested if there was any extra value of using the software compared to simply attending a group training for social skills which had little (if any) systematic method for teaching ER. The experiment's design and hypotheses are detailed in Chapter 4. The instruments used, are the same as those used for Experiment 1, described in Chapter 6.

7.1 Method

7.1.1 Participants

Three groups took part in this experiment: 2 AS/HFA intervention groups (who were independent from those in Experiment 1) and one typical (general population) control group. Participants in the clinical groups had all been diagnosed with AS/HFA in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). They were recruited via two support organisations and two colleges for individuals with autism spectrum conditions, where group meetings were also held. Since they were recruited via organisations that had volunteered to help with the study, participants were not randomly allocated to the groups, but instead were assigned by their recruiting organisation.

Participants filled in the AQ (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), to assess their self reported number of autistic traits. Only 40% of them scored above the cut-off of 32, a result that was unexpectedly low given earlier studies using the AQ (Kurita, Koyama, & Osada, 2005; Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005). To validate these self-report scores, a parent version of the AQ was sent to the parents or tutors of those who scored below cut-off, to check if this reflected under-reporting of symptoms, or if diagnosis was in question. All but 3 parents/tutors returned the questionnaires. Using these reports, 70% of the participants scored above cut-off of 32 (Baron-Cohen et al, 2001) and all participants scored above the more sensitive cut-off of 26 (Kurita, Koyama, & Osada, 2005; Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005), an indicator that diagnosis was reliable and suggesting some under-reporting by this group. Participants were accepted to the study if they had not participated in any related intervention during the last three months and had no plans for engaging in another intervention while the study was on-going. The groups in this experiment were:

<u>1. Software and tutor group</u>: 13 participants (12 males and one female) were asked to use *Mind Reading* alone for two hours a week, over a 10 week period. In addition, these participants attended 10 weekly sessions in small groups of up to six members. A tutor worked with each group, following a protocol which included general guidelines for group structure and activities (e.g. analysis of features in different facial and vocal expressions of emotions, examples of situations from participants' everyday life and the emotions these evoke, and analysis of emotions in pictures from newspapers and scenes from feature films and TV programs). The protocol appears in Appendix 6. Tutors were free to choose the materials they wanted, and were asked to relate lessons to emotion groups in *Mind Reading*, to help associate the software with group activities and with everyday life. Each of the tutors was given a complimentary copy of *Mind Reading* and was asked to become familiar with it before the course. Two of the tutors were support workers and one was a teacher, all experienced in working with adults with ASC. Three such groups were run – two in support centres for individuals with ASC and one in a college for adolescents and young adults with AS/HFA.

As in Experiment 1, participants were included if they completed a minimum of 10 hours using *Mind Reading*. Out of 18 participants who originally started the course, 3 withdrew during the course, and 2 were excluded after failing to complete the 10 hours of minimum use. Two groups were given extra time, so that participants could complete the minimum usage time. As in Experiment 1, no pattern was found for the participants who withdrew or did not complete the program: their age range varied between 19-46; 2 of them had continued studying beyond compulsory education whereas three did not; one was unemployed, 3 were students and one was employed. Their IQ levels and Time 1 assessment task scores were within one standard deviation from their group average. The reasons they gave for not having completed the program included not being able to commit to working with the software for 2 hours a week due to their studies or work. One participant left the group after falling in love with another participant and being rebuked.

2. Social skills course group: 13 participants (10 males and 3 females) took part in 10 sessions of social skills training. 2 courses, with 9 participants in each, were facilitated by a clinical psychologist who specialises in social skill training for individuals with ASC. One course took place in a college for adolescents and young adults with ASC, and the other in a support organisation. In each course a local staff member of the institution, who knew all participants, assisted the facilitator. The facilitators also recruited the participants. The 2 groups followed the same curriculum, which included themes such as conversational rules, emotional expressions and body language, preparing for job interviews, and managing friendship. The facilitators used a variety of techniques, such as stand-up teaching, group discussion, role play and picture analysis. The facilitators were blind to the curriculum of the software and tutor group, and to *Mind Reading*. In addition, the experimenters were blind to the group's curriculum until after the Time 2 assessment. Of the eighteen participants who originally attended the course, two did not complete the course and three others were excluded from the analysis (but not from attending the groups) because their IQ scores fell below 70.

3. Typical (general population) control group: 13 participants (10 males and 3 females) of the typical control group described in Experiment 1 were matched to the two AS/HFA groups in this Experiment.

Intelligence of participants was assessed using the WASI. All participants scored above 70 on both verbal and performance scales. One way ANOVA conducted on groups' AQ scores was significant for both self-report ($F[2,36]=16.98$, $p<.001$) and parent/tutor report ($F[2,36]=69.34$, $p<.001$). Tukey post hoc comparisons revealed that the two clinical groups scored significantly higher on the AQ compared to the typical control group, and did not differ from each other. Participants were also asked to report any comorbid diagnosis. 5 participants in the software and tutor group and 4 participants in the social skills group had co-morbid diagnoses such as depression, learning difficulties or OCD. The 2 AS/HFA groups did not differ on the proportion of participants with comorbid diagnoses ($\chi^2[1]=0.17$, n.s.; see Table 7.1).

The 3 groups were matched on age, verbal and performance IQ, handedness and sex. They spanned an equivalent range of socio-economic classes and educational levels. No significant differences were found between the groups for age ($F[2,36]=0.7$, n.s.), verbal IQ ($F[2,36]=2.36$, n.s.), performance IQ ($F[2,36]=2.70$, n.s.), sex ($\chi^2[2]=1.39$, n.s.), handedness ($\chi^2[2]=2.44$, n.s.), education ($\chi^2[2]=0.22$, n.s.) and employment status ($\chi^2[2]=0.21$, n.s.). In addition, no difference was found between the two AS/HFA groups on the time between the two assessment meetings ($t[17.3]=1.61$, n.s.). Table 7.1 presents the groups' background data.

7.1.2 Procedure

Participants were tested at the local support centres and colleges for individuals with ASC. They were tested in groups of 3 in large, quiet rooms by two experimenters at a time. Computer tasks were delivered on two IBM compatible laptop computers with 15 inch monitors and with headphones for the voice tasks. Background details (Time 1), feedback (Time 2), and the WASI data, were collected individually in a separate room. Two of the experimenters were blind to which group participants belonged.

Participants of the software and tutor group were asked to help in the evaluation of a new intervention program. They were asked to commit to using *Mind Reading* for two hours a week over a period of 10 weeks and to attend all 10 group sessions. *Mind Reading* was briefly introduced to the participants of this group at the first assessment. The group tutors then introduced it in more detail at the first group meeting.

Participants of the social skills group were told the aim of the study was to evaluate how social skills groups teach people to recognise emotions. They were asked to take part in the assessments at the beginning and the end of the course (which was free of charge).

	Software and tutor group (N=13)	Social skills group (N=13)	Typical controls (N=13)	F (2,36)
Age	25.48 (9.31) 17-50	24.41 (6.39) 17-42	25.52 (9.59) 17-51	0.70
Verbal IQ	105.69 (16.08) 76-123	96.54 (15.54) 78-126	109.23 (14.41) 86-128	2.37
Performance IQ	103.92 (19.84) 71-129	95.46 (6.01) 87-105	106.31 (5.99) 92-117	2.72
AQ (self report)	25.08 (7.15) 16-37	29.46 (8.08) 17-45	13.85 (5.71) 8-26	16.98[1]
AQ (with parent/ tutor report[2])	33.69 (4.39) 27-42	34.23 (4.90) 29-45	13.85 (5.71) 8-26	69.34[1]
				χ^2 (2)
% Females	7.7%	23.1%	23.1%	1.39
% Left handed	15.4%	30.8%	7.7%	2.44
% Employed	38.5%	38.5%	46.2%	0.21
% A levels or above[3]	38.5%	30.8%	38.5%	0.22
% with comorbid diagnoses	38.5%	30.8%	NA	0.17[4]

Table 7.1: Means (standard deviations), ranges, and proportions of background variables for the three groups of Experiment 2.

Participants' written consent was obtained. All participants were told they were free to withdraw from the study at any time. In the first assessment background information was obtained, followed by administration of the CAM-A, RME-R and RMV-R tasks and half of the WASI. The testing procedure was similar to that of Experiment 1. Participants filled out the AQ and FQ in advance and brought them to the assessment meeting. Those who needed help filling out the questionnaires, were supported by the experimenters during the first assessment meeting. Meetings took about 3 hours.

During the intervention time between the two assessments, participants' use of the software was monitored by the tutors. The tutors were also in charge of collecting the

[1] $p<.001$. All other test results are not significant ($p>0.05$).
[2] Parent/Tutor filled AQ replaced self report if score was lower than cut-off.
[3] A levels are the first component of non compulsory education in Britain.
[4] df=1 for this test.

log files created by the software to verify usage time. Average usage time of the software in the tutor and software group was 14.9 hours, (SD=3.1, range: 10-23). In the second assessment, participants of both intervention groups took all ER tasks (including RMF-A) and the other half of the WASI. They were also asked for their detailed feedback about the programme. Participants were then debriefed about the aims and design of the study and were rewarded with a complimentary copy of *Mind Reading* (or were allowed to keep the copy they used). This assessment meeting took about 3 hours too.

7.2 Results

One sample Kolmogorov-Smirnov tests were conducted for all task scores in the three groups. Distributions of all scores in both groups did not differ from normal. Hence, parametric analysis was used.

For differences in group performance at Time 1, five one way ANOVAs were conducted on the emotion recognition task scores, using Holm's Sequential Rejective Bonferroni Procedure. Significant differences were found between the groups for the CAM-A face task ($F[2,36]=9.76$, $p<.001$), voice task ($F[2,36]=5.64$, $p<.01$) and the number of emotional concepts recognised ($F[2,36]=7.77$, $p<.005$); as well as for Reading the Mind in the Eyes ($F[2,36]=6.75$, $p<.005$) and Reading the Mind in the Voice-R ($F[2,36]=4.99$, $p<.02$). Pre-planned comparisons with Bonferroni corrections revealed no significant differences between the two clinical groups on any of the task scores, and significantly higher scores by the typical control group on all tasks, compared to the two AS/HFA groups ($p<.005$ for all). These findings support Hypothesis 1. Means and standard deviations of the groups' emotion recognition scores at Time 1 and Time 2 appear in Table 7.2.

Five multivariate analysis of covariance with repeated measures were conducted to test for group differences on the various tasks at Time 1 and Time 2. Verbal IQ was entered as a covariate. Using Holm's Sequential Rejective Bonferroni Procedure, time by group interactions were found for CAM-A voices ($F_{wilks}[1,23]=6.5$, $p<.012$), CAM-A number of concepts recognised ($F_{wilks}[1,23]=6.04$, $p<.016$), and RME-A

(F_{wilks}[1,23]=8.4, p<.01), but not for CAM-A faces (F_{wilks}[1,23]=0.23, n.s), or RMV-R (F_{wilks}[1,23]=0.11, n.s.). Verbal IQ had a significant effect as a covariate on all tasks, beyond time (F[1,23]=17.89 for CAM-A faces, 19.4 for CAM-A voices, 17.05 for CAM-A number of concepts, 11.2 for RME-A, 10.0 for EMV-R, p<.01 for all).

	Software and tutor		Social skills		Typical controls
	Time 1	Time 2	Time 1	Time 2	
CAM-A face task (Max score=50)	32.31 (8.14)	36.23 (8.92)	26.85 (9.75)	29.31 (9.54)	40.77 (5.96)
CAM-A voice task (Max score=50)	33.15 (9.08)	38.92 (7.62)	31.08 (9.12)	31.85 (10.88)	41.08 (5.17)
CAM-A no. of concepts recognised (Max score=20)	10.23 (4.87)	13.46 (5.19)	7.69 (5.79)	8.54 (6.33)	15.08 (3.66)
Reading the Mind in the Eyes-A (Max score=36)	21.62 (6.35)	23.85 (4.18)	21.46 (5.61)	19.23 (6.76)	28.15 (3.55)
Reading the Mind in the Voice-R (Max score=25)	15.08 (2.81)	16.23 (3.47)	13.92 (4.52)	14.69 (4.55)	18.08 (2.75)
Reading the Mind in Films-A (Max score=22)		11.92 (3.71)		10.54 (3.15)	14.85 (2.64)

Note: N=13 for every group.

Table 7.2: Means (and standard deviations) of the emotion recognition measures in the 3 groups of Experiment 2 at Time 1 and Time 2

Simple main effects analyses with Bonferroni correction revealed the software and tutor group improved significantly from Time 1 to Time 2 on CAM-A voice task (t[12]=4.65, p<.01) and CAM-A number of concepts recognised (t[12]=5.2, p<.001), whereas the social skills group did not improve on either CAM-A voices (t[12]=0.56, n.s.) or CAM-A number of concepts (t[12]=1.72, n.s.). Simple main effect analysis of the time by group interaction for Reading the Mind in the Eyes with Bonferroni correction revealed no significant time effects for either the software and tutor group (t[12]=2.01, n.s.) or the social skills group (t[12]=-2.5, n.s.). Since the strong effect of verbal IQ might have overshadowed any main effect or interaction in the CAM-A faces scores, simple main effect analyses were conducted for the two groups despite the lack of a significant interaction. Paired t tests for CAM-A face scores at Time 1 and Time 2 with Bonferroni correction showed a significant improvement in the software and tutor group ($t[12]=4.2$, $p<.005$) but not in the social skills group ($t[12]=2.0$, n.s). Proportions of correct answers of all tasks are shown in Figure 7.1.

The MANOVA conducted for the 20 CAM-A concept difference scores failed to reach significance ($F_{wilks}[20,5]=1.16$, n.s.). This may have been due to the small number of participants in each group and the small range of difference scores for each emotional concept. Significant interaction effects suggesting greater improvement on the software and tutor group, compared to the social skills group were found for two concepts only: *vibrant* ($F[1,24]=4.88$, $p<.05$), and *mortified* ($F[1,24]=10.04$, $p<.01$).

The ANOVA conducted for the Film task scores, testing for holistic distant generalisation at Time 2 only, was significant ($F[2,36]=6.15$, $p<.01$). However, pre-planned contrasts with Bonferroni correction revealed no difference between the software and tutor group ($M=11.92$, $SD=3.71$) and the social skills group ($M=10.54$, $SD=3.15$) for this task ($t[36]=1.10$, n.s.), though the typical control group ($M=14.85$, $SD=2.64$) scored significantly higher than both of the other groups ($t[36]=3.33$, $p<.01$). These results suggest that as seen in experiment 1, the AS/HFA group using *Mind* Reading did not perform better than the AS/HFA group who did not use it, at this level of generalisation.

Non parametric correlation analysis conducted for software usage time with task improvement scores in the software and tutor group revealed significant correlations of software usage time with improvement on the CAM-A voice task ($r_{spearman}=0.60$, $p<.05$), the number of CAM-A concepts correctly recognised ($r_{spearman}=0.53$, $p<.05$) and with film task scores ($r_{spearman}=0.50$, $p<.05$). No other correlations were found significant.

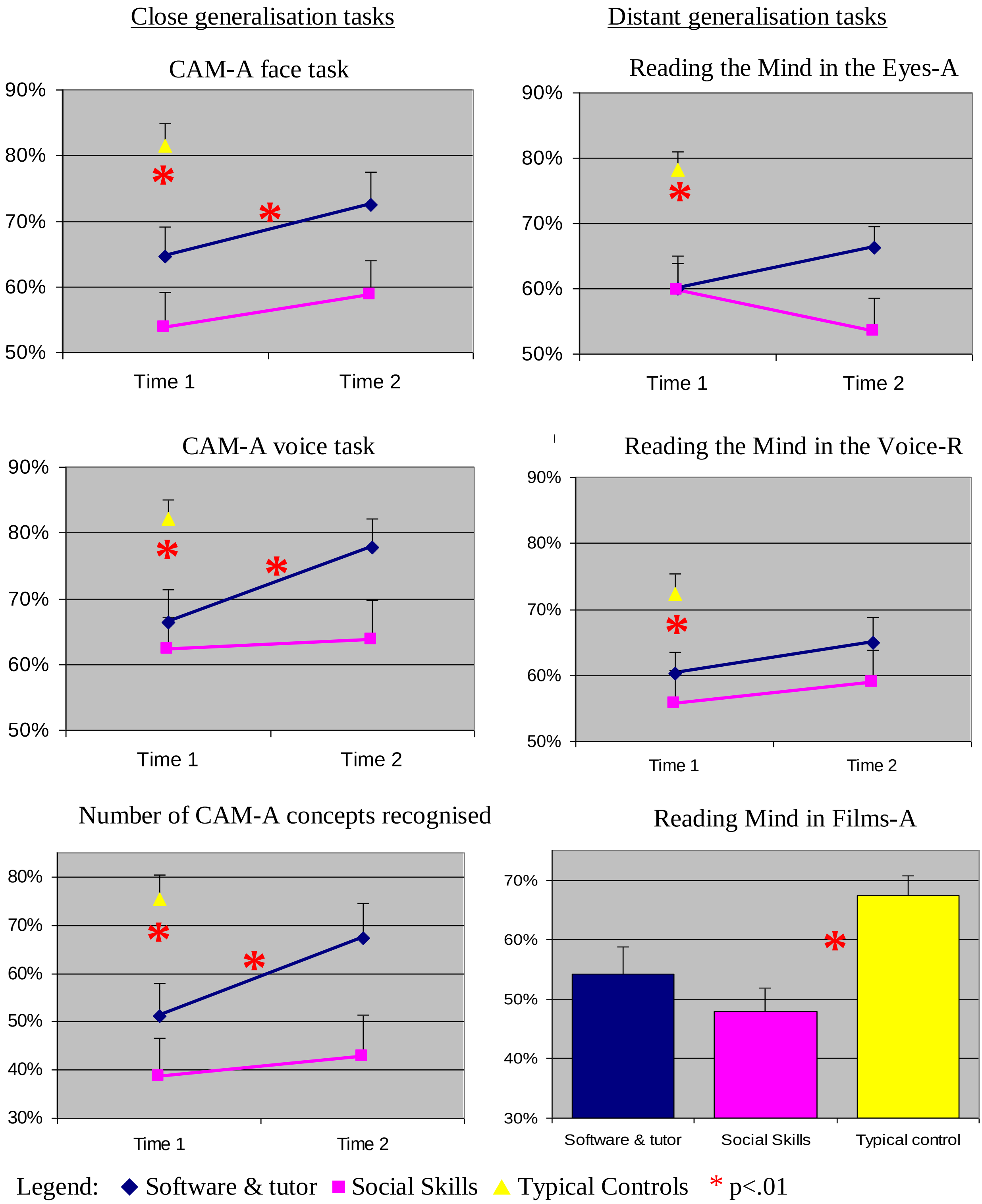


Figure 7.1: Mean proportions of correct responses (with standard error bars) for the 3 groups on the two levels of generalisation – Experiment 2

The feedback at the end of the Time 2 assessment of participants who used *Mind Reading* resembled that of the software users in Experiment 1. They said the software was interesting and enjoyable to use (n=3), comprehensive (n=2) and useful for recognition of facial expressions (n=6) and understanding of emotions (n=3), particularly the more subtle and complex ones. Six participants mentioned the software helped them recognise emotions in everyday life, e.g. "*I know more about emotional clues when I look in someone's face now,*" or "*I use it to better understand my mum*". Four participants said they became more aware of emotions and their expressions and think about them more, e.g. *"I have become more sensitive to emotions*". Two participants also mentioned their emotional vocabulary has grown following the use of the software. Other comments criticised the software for being too repetitive, and becoming boring after a while (n=2), for having expressions that are too extreme and different to real life (n=2), and for different missing aspects in its curriculum, such as body language, ER from context, integration of facial and vocal expressions, and real-life social situations, e.g. "*Characters in the software need to have personalities, describe relationships between them, make them easier to identify with*".

Feedback about the support of the tutor and group acknowledged importance of the integration between software and group for the learning process (n=2). Participants felt the group helped increase their awareness to the importance of emotional cues, e.g. "*I am looking for them more now*". They said the group helped them analyse the emotions on *Mind reading* by looking for similarities and differences between expressions (n=4), e.g. "*The group emphasised elements I wouldn't have thought of or recognised by myself*". Participants also mentioned the additional material taught in the group, which was not covered in *Mind Reading*, such as body language, and integration of emotional cues in the different modalities (n=3). The role of the group in linking the material from *Mind Reading* to real life was also mentioned, e.g. "*Role playing was useful for people to recognise emotions in each other and to practise expressing the emotions*". Finally, the group gave participants a chance to socialise and to support each other, which four of them mentioned was a central reason for them to take part. The groups were criticised for being too short (3 participants argued it should last a year to be effective). Two participants did not feel they would use the material taught in the group in their everyday life, e.g. "since I don't usually think about these things". Finally, one of the participants commented about his experience in the group, saying: "*Most of the time, I was too anxious to learn during the group meetings. Being in the same room with 6 other people for such a long time and having to answer the tutor's questions was too stressful for me*".

Participants in the social skills group found the group very helpful (n=4) and enjoyable (n=2). Many of them (n=9) commented that the best thing about the group was the social get together, the opportunity "*to meet other people like me*", "t*o feel I am not alone in the world*", or "*to learn about the problems other people experience*". Six of them said they wish the group continued for longer. The Participants praised the use of role play in the group (n=4) and said the group "*helped understand people better, why they do things the way they do*". One complaint dealt with the content of some of the meetings, and another with the pace that had to be slowed down for participants with greater difficulties. One participant described his anxiety coming to the group at first, which got better when he got to know the other people better.

7.3 Follow-up

7.3.1 Procedure

About a year (12-14 months) after the Time 2 assessment, participants in the ASC groups were sent the follow-up questionnaire. The questionnaire was posted to participants through the support organisations and colleges for individuals with ASC, where the groups were run. Support workers were asked to go through the questionnaire with the participants, to make sure they understand it well before answering. In the case of participants who had already left college, the questionnaire was forwarded by the college to their home address, and parents were asked to assist them, if needed, when filling the questionnaire out.

7.3.2 Participants

Of the 26 participants with ASC who took part in Time 1 and Time 2 assessments of Experiment 2, eleven completed and returned the follow-up questionnaires. Seven of them originally belonged to the software and tutor group (2 females, 5 males), and four originally belonged to the social skills group (2 females, 2 males). The two groups did not differ on age, verbal IQ, performance IQ, and the AQ and FQ scores from Time 1. Background information for these participants appears in Table 7.3.

Time 1 measures	Participants from software & tutor group (n=7)			Participants from Social Skills group (n=4)			
	Mean	SD	Range	Mean	SD	Range	U[1]
Age	27.29	7.49	17.1-35.5	28.70	10.15	17.9-42.4	14
Verbal IQ	111.86	15.39	89-132	100.0	6.98	91-106	8
Performance IQ	106.0	21.35	71-135	86.0	15.14	64-96	6
AQ (with parent/tutor)	35.86	3.98	30-42	37.25	5.44	33-45	12.5
FQ	77.43	22.82	39-111	66.50	32.13	25-100	11

[1] Mann-Whitney test. p>.1 for all comparisons.

Table 7.3: Time 1 background averages, standard deviations and ranges for the participants who filled in the follow-up questionnaire – Experiment 2

Of the 11 respondents, only four (3 from the software and tutor group and one from the social skills group) continued to use *Mind Reading* after the Time 2 assessment. Discriminant analysis was conducted to find if any of the background variables could predict whether participants continued to use *Mind Reading* after Time 2 or not. Group, sex, age, verbal and performance IQ, and AQ scores (assessed by parents/tutors where required, as described in section 7.1.1) were included as predictors. Stepwise method was used due to the large number of predictors and small sample size. AQ score was the only predictor entered into the analysis ($F_{wilks}[1,9]=8.92$, $p<.02$). Based on this predictor, the significant discriminant function ($\chi^2[1]=5.85$, $p<.02$) successfully classified 90.9% of the participants – all the participants who continued to used *Mind Reading*, and all but one of those who did not. A Mann-Whitney test revealed that the AQ scores of participants who continued to use the software (M=32.50, SD=1.73) were significantly lower than the AQ scores of participants who did not (M=38.57, SD=3.78; U=5, Z=2,57, p<.01).

7.3.3 Results - FQ

To compare the FQ score differences between the two groups (software and tutor, and social skills), the Mann-Whitney non-parametric test was used, as normality could not be assumed with a group of 4 participants (Howell, 1997). No significant difference was found between the groups (U=14, z=0, n.s.). In addition, using the Wilcoxon Signed Ranks non parametric test, the change in FQ scores from Time 1 to Follow-up was not significant among participants from the software and tutor group (Time 1 FQ: M=77.43, SD=22.82; Time 2 FQ: M=76.71, SD=20.50; Wilcoxon Z=.34, n.s.) or among participants from the social skills group (Time 1 FQ: M=66.50, SD=32.13; Time 2 FQ: M=69.50, SD=15.50; Wilcoxon Z=0, n.s.).

Due to the small sample size, regression analysis was not conducted. Instead, correlation analysis of FQ score change for the whole sample of 11 participants was conducted with age, verbal and performance IQ, AQ scores (assessed by parents/tutors where required, as described in section 7.1.1), and improvement scores on CAM-A tasks. The analysis yielded only one significant correlation – a negative correlation between FQ change scores and verbal IQ ($r_{spearman}$=-.76, $p<.01$), suggesting that the higher the participant's verbal IQ, the greater the decrease of FQ scores from Time 1 to Follow-up.

7.3.4 Results – follow-up feedback questionnaire

Since only four participants continued to use *Mind Reading* after the Time 2 assessment, no analysis was conducted on their ratings. The distribution of their ratings appears in Table 7.4. From the ratings, it appears that participants found *Mind Reading* very helpful for emotion recognition in everyday life, and quite helpful for understanding social situations and feeling confident in them. However, a larger sample is required to assess whether the differences between those ratings are statistically significant.

How helpful was *Mind Reading* for:	Not helpful at all (1)	Not very helpful (2)	Quite Helpful (3)	Very Helpful (4)	Mean
Emotion recognition in every day life	-	-	25.0%	75.0%	3.75
Understanding of social situations	-	-	75.0%	25.0%	3.25
Confidence in social situations	-	25.0%	50.0%	25.0%	3.0

Note: n=4

Table 7.4: Participants' ratings of *Mind Reading*'s contribution for socio-emotional functioning in real-life – Experiment 2

A few comments were given regarding *Mind Reading*'s relevance and usefulness in the respondents' everyday life. Three of the respondents mentioned *Mind Reading* was helpful in recognition of emotion and non verbal communication, e.g., "*Mind Reading was very helpful in me realising how subtle a lot of emotions are in real life*" or "*I can recognise non-verbal communication more clearly, for example when someone is bored with what I have been saying*". One respondent wrote the generalisation to real life was difficult, and suggested more social-situations need to be included in the software to enhance that. Two of the respondents argued the software did not help them cope with their social isolation.

7.4 Discussion

This experiment compared the use of *Mind Reading* at home with the weekly support of a tutor in group sessions to participating in a social skills course without the use of the software. Results supported Hypothesis 1, as the two ASC groups performed significantly lower than the typical control group on all ER tasks at Time 1. As in Experiment 1, Hypotheses 2 and 3 were only partially supported: the group using *Mind Reading* improved significantly more than the control group on two measures of close generalisation: recognition of voices and number of emotional concepts recognised, both taken from *Mind Reading.* The close generalisation face task also showed that the software and tutor group improved following training, whereas no improvement was found for the social skills group. Participants commented that the combination of software and tutor led group was helpful, as the two complement each other, allowing to practice material taught on the software through various group activities. Unlike Experiment 1, this Experiment provided the ASC controls with an alternative intervention, which included attention from professional facilitators, social stimulation and a curriculum that was relevant for social functioning. Therefore, any difference in the groups' improvement is likely to be related to the content of the intervention used by the software and tutor group, i.e. systematic analysis of emotional expressions in faces and voices in the software itself.

Despite the group activities, aimed to increase generalisation, software users failed to improve more than controls on feature-based distant generalisation of faces and voices, and did not perform better than controls on the holistic distant generalisation level – the RMF-A. The lack of effect on both feature-based and holistic distant generalisation measures matches the findings of Experiment 1 and could be attributed to the generalisation difficulties which are characteristic of ASC. These are discussed in detail in Chapter 9.

Software usage time in the software and tutor group was correlated with greater improvement on CAM-A voice task and CAM-A's number of emotions correctly recognised, as well as with higher scores on the RMF-A, the holistic generalisation task, though caution should be used when interpreting these results, due to the small group size. However, when the two *Mind Reading* user groups from Experiments 1 and 2 were combined (n=32), the correlation of software usage time with scores on the CAM-A voice task ($r=.55$, $p<.01$) and with film task scores ($r=.41$, $p<.05$) remained significant. This suggests that longer use of the software may lead to improved generalisation. Therefore, despite the developmental delay and long lasting difficulties in emotion and mental state recognition by individuals with ASC (Baron-Cohen, 1995; Frith & Frith, 2003; Frith & Hill, 2004), learning emotion and mental state recognition is possible for them even into adulthood when the intervention harnesses their systemising strengths. It may be the case that even greater improvement would be achieved if intervention was started at a developmentally earlier time-point. The next chapter examines this with 8-11 year old children.

The most striking effect in this experiment was that of verbal IQ. Unlike Experiment 1, its effect was significant on all tasks. Yet, despite this, the software and tutor group improved significantly more than the social skills group on the measures described above. The difference between this Experiment and the home-use Experiment in the potency of verbal IQ could be related to the IQ difference between the groups in the two Experiments. While the groups in Experiment 1 had above average intelligence, the two groups in Experiment 2 had average (and many lower than average) IQ. This could result from the way the groups were recruited: as the groups in Experiment 2 were recruited through organisation for individuals with ASC, it is possible that some of the people who turn to those organisations are lower functioning and in greater need of support. As a result of these IQ differences, the groups of this Experiment might have found the assessments more challenging and the words in the tasks more difficult. Although definition handouts were offered, it is possible that those who have higher verbal IQ coped better with the tasks, as they were less distracted and stressed by the need to use the handouts. However, it is important to note that verbal IQ was not associated with improvement on any of the tasks, suggesting that *Mind Reading* could potentially be helpful for lower functioning users too. A non-verbal assessment task could have been useful in the current experiment, to improve the validity of assessment in lower functioning groups. This should be applied in future studies.

Verbal IQ was also the only variable associated with change on the follow-up measure – the Friendship and Relationship Questionnaire (FQ). Interestingly, this correlation was negative, i.e. the higher participants' verbal IQ, the more their FQ scores decreased from Time 1 to follow-up. A possible explanation of this phenomenon may be similar to the low AQ scores found in the self-reports of some of the participants in this experiment. It is possible that lacking awareness of their difficulties led to an over-estimation of FQ (and AQ) scores of participants with lower IQ. However, this was not tested and requires further investigation. Drawing clear conclusions on this question is difficult, as this finding relies on a very small number of respondents (n=11).

The difficulties in getting participants to complete the required number of working hours on the software during the course, as well as the relatively high dropout rate of participants in the software using groups (21% in the Experiment 1 and 28% in the

Experiment 2), raise the question of motivation, which might have varied between the two Experiments: Whereas participants in Experiment 1 contacted the research team to volunteer for the study, participants in the current Experiment were recruited and trained through their support organisation or college, hence their initial motivation might have been lower. Furthermore, as some of the participants pointed out, joining the study gave them an opportunity for desirable social contact with the experimenters, and (in Experiment 2) with the tutors and the group. Whereas this reason may have been secondary to participants' wish to acquire socio-emotional skills, for some it may have been the main reason to join the study, as some of the verbal feedback shows. It is therefore possible that participants felt the social gain was not strong enough to justify spending so much time using the software on their own.

The reasons individuals gave for not completing the required usage time, or for withdrawing were usually related to their inability to find the time to use the software at home for two hours a week. Assuming that this is not an extensive amount of time to ask for, and bearing in mind that these participants had joined the study willingly, while knowing the work requirements, this finding could relate to the difficulties individuals with ASC have prioritising, planning, and adhering to goals. Such difficulties could be explained by the executive dysfunction theory (Ozonoff, 1995a; Russell, 1997), and suggest that even high functioning adults with ASC may struggle with these issues. Future studies will need to evaluate the effect of possible executive dysfunction on the ability of adults with ASC to benefit independently from computer-based interventions. Individuals who have such difficulties may require help planning their work with the software, in conjunction with group activities aimed at boosting both generalisation and motivation.

Another possible explanation for the low usage time in Experiment 2 could be related to the nature of group-based courses: though the tutor led group activities provided participants with a structured learning environment, they might have also taken the responsibility for the participants' learning process away from them. Indeed, in their summary of the groups' work, tutors commented on the difficulties they have experienced getting some of the participants to complete the required software usage time.

One of the differences between Experiments 1 and 2 is in the proportion of participants who continued to use *Mind Reading* after the Time 2 assessment. Whereas more than 75% of the follow-up respondents in Experiment 1 continued to use the software, only 36% of respondents in this Experiment had done so. As the use of *Mind Reading* beyond the Time 2 assessment was associated with improvement on the FQ scores from Time 1 to Follow-up in Experiment 1, the fact that so few of the participants continued using the software in the current experiment could be another reason for the lack of difference on FQ scores from Time 1 to Follow-up. One possible explanation for the low proportion of software users after termination of the intervention period relates to the perception of both software and tutor and social skill groups as time limited activities. Participants might have associated the software with the group activities, which meant that they felt they did not need to continue using the software when the group came to an end.

Results of the discriminant analysis showed that participants with lower AQ scores (i.e. lower level of autistic traits) were more likely to carry on using the software. As these corrected AQ scores relied on parental/tutor report, they were potentially more reliable and could suggest that participants who have a greater number of autistic traits would be unable to work with *Mind Reading* independently, i.e. that they would need a structured environment that provides them with this training. This is not necessarily due to technical difficulties using the software, but rather due to difficulties initiating, planning and prioritising such activities, as suggested above.

A possible limitation of this experiment lies in the different number of hours the two groups received. Though the two groups received a similar number of group intervention hours, the social skills group had no input at home, to match the homework of the Software and tutor group. This difference in the number of intervention hours, as well as having other subjects (except for emotion recognition) discussed in the social skills group, might have accounted for some of the improvement difference.

Lastly, group size limits the power of findings in this Experiment. This is related to practical reasons of keeping the social skills group and tutor and software groups small, to optimise learning conditions, and to minimise drop out rate. Future studies should evaluate the use of *Mind Reading* in group settings with more participants to further validate findings of this Experiment.

7.5 Conclusion

Findings of this Experiment confirm the results of Experiment 1, in lower (though still relatively high) functioning participants with ASC, who were recruited and trained differently. Results showed improved emotion and mental state recognition skills amongst software users on tasks that included material from the software, with generalisation difficulties on material that was not included in training. As in Experiment 1, software usage time was positively correlated with higher scores on the holistic distant generalisation task (the RMF-A), suggesting that a longer period of training may result in improved generalisation. In addition to the replication of the findings of Experiment 1, this Experiment also showed that the ER skills gained from using *Mind Reading* exceed the effect of attending social skills group training, and that using the software has additional merit beyond simply attending a group (which participants in both groups did). The group course format, as well as the characteristics of the participants in this Experiment, did not promote software use beyond the limited training period of 10-15 weeks. This possibly requires further support for users beyond the group sessions, to help further motivate them to explore the software. Results were found despite the relatively small sample size, and the overshadowing effect of verbal IQ, which was found on all tasks. Experiment 3, presented in the next chapter, assessed whether improved generalisation could be found with younger software users.

8 Child intervention study: Home use of *Mind Reading*

This experiment assessed the effectiveness of *Mind Reading* as a home training instrument for primary school aged children with ASC. It tested if it was possible to replicate the results of Experiments 1 and 2 with a younger sample, with the additional aim of assessing whether this group finds generalisation easier than the adult groups. This is a plausible possibility because interventions with ASC have been found to be more effective the earlier they are utilised, perhaps because brain plasticity is greater earlier in life (Dawson & Zanolli, 2003; Howlin, 1998; Howlin & Rutter, 1987).

Although early autism interventions start long before primary school, interventions into social skills in high functioning children with ASC often commence around the first years of primary school, since these children are often undiagnosed before that age (Howlin, 1998; Jordan & Jones, 1999; Martin, Bibby, Mudford, & Eikeseth, 2003; Rogers, 2000). In addition, in order to test if relatively independent use of the software was possible in a young age group, children were required to be text and computer literate. The performance of children with ASC who used the software over a period of 10-15 weeks was compared to that of children with ASC who did not get any systematic intervention over and above their regular school curriculum; and to typically developing controls. The design and hypotheses of this experiment are detailed in Chapter 4.

8.1 Method

8.1.1 Participants

Three groups took part in this experiment: one AS/HFA intervention group, one AS/HFA control group and one typical control group. Participants in the clinical groups had all been diagnosed with AS/HFA in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). Participants' parents filled in the Childhood Asperger Syndrome Test (CAST; Scott et al, 2002; see description below), to assess the children's number of autistic traits. One participant in the AS/HFA intervention group and two in the AS/HFA control group scored below the suggested cut-off of 15 (max = 31). Since these participants scored below cut-off because of several unanswered items, they were not excluded from the sample. One participant who was originally in the AS/HFA control group scored 10 on the CAST, and this low score was not attributable to unanswered items. Further investigation revealed that his diagnosis was questionable amongst medical staff, so he was excluded from the study.

Participants with AS/HFA were recruited from several sources, including the Autism Research Centre's volunteer database, a local clinic for children with ASC, and an advert in the National Autistic Society magazine *Communication.* They were randomly allocated into the two clinical groups:

<u>(1) Software home-users</u>: 21 participants (20 boys and 1 girl) were asked to use *Mind Reading* at home by themselves for two hours a week over a period of ten weeks, a total of 20 hours. Participants were included in the study if they completed a minimum of ten hours of work with the software. If they did not complete this minimum, participants were given an extension of up to 4 weeks to use the software. Of 24 participants originally recruited to this group, 1 withdrew during the 10 week period due to illness, 2 failed to complete the 10 hours minimum after 15 weeks, and 1 was excluded as his IQ fell bellow 70. The 2 children who failed to complete 10 hours reported they did not find the software interesting and had stopped using it after a short period. Parents were reluctant to pressure their children to use the software.

<u>(2) AS/HFA control group</u>: 21 participants (20 boys and 1 girl) attended the assessment meetings with a 10-15 week period between them, during which they did not use the software, or take part in any new intervention programme related to emotion recognition. Participants who were already involved in social skills training at school continued their course as usual. Of 24 participants originally recruited for this group, two were excluded as their IQ fell below 70, and one was excluded because he found the first assessment too stressful. In addition to these groups, a group of typically developing children was also included:

(3) Typical control group: 22 participants (21 boys and one girl) were recruited for this group from a local mainstream primary school. Parents and school reports were checked to confirm that none of the children in this group had a psychiatric diagnosis or special educational needs, or had family members diagnosed with ASC. The latter criterion was included to avoid the risk of ER difficulties that might be part of the 'broader phenotype' of autism. Of 24 participants originally recruited for this group, two were excluded: one had a sibling with ASC and the other had hearing impairment, which prevented him from performing the auditory tasks. To screen for ASC, participants' parents filled in the CAST. None of the control participants scored above the cut-off point of 15. All participants were given the WASI, and scored above 80 on the verbal scale, and above 75 on the performance scale. This was used to confirm that none of the children had learning difficulties.

One way ANOVA of participants' CAST scores revealed a significant main effect of Group ($F[2,61]=145.01$, $p<.001$). Tukey post-hoc comparisons showed that the two AS/HFA groups scored significantly higher on the CAST compared to the typical control group, but did not differ from each other. Four children in the AS/HFA intervention group and 5 in the AS/HFA control group had other diagnoses such as ADHD or dyspraxia. No difference was found between the AS/HFA groups in the proportion of participants with comorbid conditions ($\chi^2[1]=0.14$, n.s.). Seven children from the AS/HFA intervention group and three from the AS/HFA control group were in special education schooling, either an ASC unit in a mainstream school or a special school for children with ASC ($\chi^2[1]=2.1$, n.s.). Four children in each of the clinical groups were taking part in social skills training at school during the study ($\chi^2[1]=0.0$, n.s.). In addition, no difference was found between the two clinical groups in the time between the two assessment meetings ($t[40]=0.14$, n.s.).

The three groups were matched on age, verbal and performance IQ, handedness and sex. Participants' parents spanned an equivalent range of employment and educational levels. As shown in Table 8.1, no significant differences were found between the groups for any of these background variables.

	Software home users (N=21)	AS/HFA controls (N=21)	Typical controls (N=22)	F (2,61)
Age	9.41 (.98) 8.2-11.7	9.91 (1.08) 8.6-11.8	9.96 (1.15) 8.2-11.7	1.70
Verbal IQ	112.29 (14.39) 85-139	109.14 (15.39) 83-143	112.95 (10.25) 88-129	.49
Performance IQ	108.67 (12.04) 87-141	109.95 (17.29) 79-140	111.68 (13.56) 91-133	.24
CAST	20.40 (3.56) 14-28	18.95 (4.59) 11-28	4.28 (1.87) 1-8	145.01[1]
				χ^2 (2)
% Left handed	10.0%	23.8%	9.5%	2.21
% both parents are employed	66.7%	66.7%	73.7%	0.30
% both parents above A levels	28.6%	23.8%	33.3%	0.43
				χ^2 (1)
% with comorbid diagnoses	19.0%	23.8%	N/A	0.14
% in mainstream education	66.7%	85.7%	N/A	2.10
% social skills training in parallel	19.0%	19.0%	N/A	0.00

[1]p<.001. All other test results are not significant (p>.1).

Table 8.1: Means (standard deviations), ranges, and proportions of background variables for the three groups of Experiment 3

8.1.2 Instruments

8.1.2.1 Close generalisation: The Cambridge Mindreading Face-Voice Battery, Child version (CAM-C)

The CAM-C battery uses face and voice clips from *Mind Reading* and is similar to the CAM-A in its structure and guiding principles. The battery tests the recognition of 15 different emotions, the 6 basic emotions and 9 complex emotions (e.g., *jealous, disappointed, embarrassed*). Due to the inconclusive findings relating to basic ER in

ASC (described in Chapter 2), this structure allows a comparison of basic and complex ER in faces and voices in children with ASC. The recognition of each of the 15 emotional concepts is tested through 6 items, 3 face and 3 voice items. Items were presented on a computer screen, using DMDX experimental software. A handout of definitions of all the adjectives used in the battery was available for the participants at the beginning and through the assessment. There was no time limit for answering. The CAM-C took about 45 minutes to complete, including breaks. The battery provides ER scores for faces and for voices (max=45 for each), as well as for the number of emotions correctly recognised (max=15). These are also available separately for basic and complex emotions. Test-retest correlations, calculated for the ASC control group in the current experiment were r=0.79 for the face scale, r=0.75 for the voice scale, and r=0.77 for the number of concepts recognised ($p<.001$ for all). Creation and validation of the battery is described in detail in Appendix 1.

8.1.2.2 Distant generalisation visual task – Reading the Mind in the Eyes, Child version (RME-C)

The adult RME test was adapted for children by reducing the number of items and simplifying the mental state words available for each item. The RME-C (Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001) includes 28 items. It was administered to the participants by an experimenter, checking the child is familiar with all the mental state words provided, and using a definition handout when necessary. Task items were presented in a random order. One short break was given to the children during the task. The RME-C took about 20 minutes to complete. Test–retest correlation, calculated for the ASC control group in the current experiment, was r=0.64 ($p<.005$).

8.1.2.3 Distant generalisation auditory task – Child Feature-based Auditory Task (C-FAT)

The C-FAT is an auditory task, which assesses complex ER using voice clips that were recorded for *Mind Reading* but were not included in it. Hence, while the actors' voices are similar to those who appear in the software and in the vocal task of the CAM-C, the content of these additional recordings is completely novel for *Mind Reading* users, and was thus used as a distant feature-based generalisation task. All 17 items of the C-FAT include a short segment of speech, followed by 4 emotion and mental state adjectives. Participants are asked to tell how the speaker is feeling by choosing one of the 4 available answers. Foils were selected to match the content of the verbalisations but not the intonation, thus making the task harder to answer. The test items were played on the computer in random order, preceded by an instruction slide and two practice items. Participants were given a definition handout before the beginning of the task. There was no time limit for answering. The C-FAT took about 15 minutes to complete. Test–retest correlation, calculated for the ASC control group in the current experiment was $r=0.71$ ($p<.001$).The creation and validation of this task is described in detail in Appendix 2.

8.1.2.4 Holistic distant generalisation task – Reading the Mind in films, Child version (RMF-C)

The RMF-C comprises 22 short social scenes taken from children feature films. Its structure and guiding principles are similar to those of the RMF-A. The RMF-C took about 30 minutes to complete. Task score is the number of correctly recognised emotions (max=22). Creation and validation of the task are described in detail in Appendix 3.

8.1.2.5 Follow-up measure: Supplementary items to the Vineland Adaptive Behaviour Scales (VABS-S)

The Vineland Adaptive Behaviour Scales (VABS; Sparrow, Balla, & Cicchetti, 1984) assess the behavioural and personal sufficiency of individuals from birth to adulthood via a respondent who is familiar with the individual. Occurrence of the different behaviours included in the VABS is rated as ‘0 - never occurring’, ‘1 - sometimes or partially occurring’ or ‘2 - regularly occurring’. Frith and colleagues (1994) developed a supplement for the VABS (which is referred to here as the VABS-S). They selected items from the communication and socialisation domains of the expanded and survey forms of the VABS, and devised some additional theoretically derived items to form two kinds of socio-emotional functioning scales: the *Active Sociability* scale involves social behaviours which could be learned and do not necessarily require mentalising, e.g. ‘Initiates social contact’ or ‘Takes turns in conversations’. The *Interactive Sociability* scale, involves social behaviours that require mentalising, e.g. ‘Refrains from statements that might embarrass’ or ‘Apologises for hurting other’s feelings’. Each of the scales includes 16 items. Scores on each scale range between 0-32.

The VABS-S provides a concise list of behaviours, related to socio-emotional functioning. It was used with 7-19 years old children with severe autism, who had a verbal mental age ranging between 4-10 years (Frith, Happe, & Siddons, 1994), and with 8-12 year olds with conduct disorder, who had IQs in the normal range (Happe & Frith, 1996). Though it was not used with high functioning children with ASC before, the VABS-S samples different aspects of relevant social functioning, characteristic of children in the age group tested in the child intervention study. In addition, since Frith and her colleagues showed higher VABS-S scores in children who have better ToM abilities, it would be interesting to see whether a change on these behaviours could be affected by an intervention aimed to improve emotion and mental state understanding in children with ASC. Hence, the VABS-S was used as the real-life functioning follow up measure in the child experiment.

Frith et al did not report reliability or validity of the two VABS-S scales. Using the ASC and control groups of children described in Appendix 1, Cronbach’s α was calculated for the two scales. Internal consistency level for *Interactive Sociability* was $\alpha=0.92$, and for *Active Sociability*: $\alpha=0.82$. Discriminant analysis conducted using all items of the *Interactive Sociability* scale successfully classified 100% of the children into their original groups. The analysis for the *Active Sociability* items successfully classified 92.7 % of the children (88.9% of children with ASC and 100% of typically developing controls) into their original groups.

In the current study, the VABS-S was formed as a questionnaire for parents to fill out independently at home. It was posted with the CAST (see description below) to parents of children in both ASC and control groups to fill out in advance and bring with them to the first assessment meeting. The research team was available to explain the meaning of items, if parents needed clarifications, during the first assessment meeting.

In the follow-up assessment, the VABS-S was sent to parents of participants in the ASC groups, a year after their time 2 assessment. As in the adult follow-up measure, it was preceded by a short questionnaire about the use of *Mind Reading* since the second assessment meeting. Parents were asked to estimate how much time their child spent using *Mind Reading* since the second assessment meeting, and to rate how helpful the use of *Mind Reading* was for their child's ability to recognise emotions in every day life, to understand social situations, to use emotion words in his/her speech, and to have confidence in social situations. A Likert scale of 1 (not helpful at all) to 5 (very helpful) was used for these questions. Parents were also asked to describe how relevant and useful their child found *Mind Reading* in everyday life and in which areas. The follow-up questionnaire appears in Appendix 5.

8.1.2.6 The Childhood Asperger Syndrome Test (CAST)

The CAST (Scott, Baron-Cohen, Bolton, & Brayne, 2002) is a parental questionnaire designed specifically to screen school-age populations for behavioural symptoms indicative of autism spectrum conditions. Scores range from 0-31, and the higher the score, the more autism spectrum features the child possesses. In a recent community sample study (Williams et al., 2005), the CAST was validated against the Autism Diagnostic Interview-Revised (Lord, Rutter, & Le Couteur, 1994) and the Autism Diagnostic Observation Schedule-Generic (Lord et al., 2000), and was found to discriminate well between children with ASC and typically developing children, with sensitivity of 100% and specificity of 97%. Its test-retest reliability in a community sample was 0.83 (Williams et al.,2006).

8.1.3 Procedure

Participants from the clinical groups were individually tested at the Autism Research Centre in Cambridge. Six participants were tested individually in their local support centre for individuals with ASC, as they lived too far from Cambridge to be assessed there. Typically developing controls were tested individually in a quiet room in their school. Four trained experimenters individually helped the participants through the assessments. Three of them were blind to which group the participants belonged. Participants in the intervention group were asked to help evaluate *Mind Reading*. They were asked to commit to using the software for 2 hours a week over a period of 10 weeks and to be assessed before and after this training period. Participants of both control groups were asked to take part in an emotion recognition study, helping to validate new ER tasks. For this reason, participants in the AS/HFA control group were asked to come for two assessments, separated by a 10-15 week period. Written consent was obtained from participants' parents. All child participants expressed verbal consent to participate. They were told they were free to withdraw from the study at any time.

Participants with ASC were seen twice: In the first assessment, background information was collected, and parents were asked to hand in the CAST and VABS-S they filled out in advance. Participants were then seated in front of IBM compatible computers with 15 inch monitors to take the CAM-C, and the C-FAT. They were given headphones for the voice tasks. The RME-C was administered manually. All ER tasks were presented in a random order. Two breaks were given during the CAM-C battery, and one during each of the other tasks. In between the tasks, two subtests of the WASI were administered.

After the assessment was completed, participants from the AS/HFA control group were thanked and excused. The participants of the intervention group were introduced to *Mind Reading* in detail. The introduction was similar to that given to the adult participants (see section 6.1.3), but more attention was given to the learning centre and less to the emotion library, since this is more child-friendly. Participants watched a demonstration of a systematic analysis of an emotion (*kind*, which is in a level appropriate for their age), comparing different faces and voices to identify the unique facial/intonation features of this emotion. Children were encouraged to analyse facial and vocal stimuli systematically, using the lessons and quizzes in the learning centre, and the emotions library. They were advised to start using *Mind Reading* at the lower levels, appropriate for their age group.

Children were asked to use the emotions library and learning centre as they wished, but not to use the game zone for more than a third of the time. Participants' parents were taught how to lock the game zone and provide 'time in the game zone' as a reward for answering questions in lessons and quizzes. Parents were asked to support the children's learning by suggesting emotions they might find challenging, or by discussing examples from everyday life, to help them consolidate their knowledge and enhance generalisation. Participants were asked to bring the log file, which documented the duration of their use of the software, to the second assessment meeting, for usage time verification. The whole assessment meeting took about 3 hours. The research team provided technical support for installation and use of *Mind Reading* during the time between the two assessments. In addition, parents of the intervention group were approached by telephone at least once, to check their children were still committed to the study and to working with the software.

In the second assessment meeting children with AS/HFA were given the same ER tasks they had taken in the first assessment, in addition to the RMF-C and the other two WASI subtests. Task administration order was randomised. The log files of participants in the intervention group were checked to verify they had used the software for the required amount of time. Average usage time was 18.0 hours (SD=4.7, range: 11.8-28.2). Participants and their parents were then asked for their feedback about the program, their experience with it and comments about its usefulness. All participants were then debriefed about the aims and design of the study and were rewarded with a complimentary copy of *Mind Reading* (or were allowed to keep the copy they used). This assessment meeting took about 3 hours.

Typically developing controls were seen for one assessment. However, since they had to take all the ER tasks, as well as the whole WASI, the assessment for each child was broken down into two parts, administered in two consecutive days in their local school. Laptop computers with 15 inch monitors were used for this assessment. Headphones were used for the voice tasks. Participants brought the CAST and VABS-S with them to the assessment, filled out in advance by their parents. At the end of the assessment, each typically developing participant was given a £10 book token.

8.2 Results

One sample Kolmogorov-Smirnov tests were conducted for all task scores in the three groups. Distributions of all scores did not differ from normal (with the exception of CAM-C number of basic emotion concepts recognised at Time 1 in the typical group, and at Time 2 in the AS/HFA intervention group). Hence, parametric analysis was used.

First, performance of the three groups on the emotion recognition tasks at Time 1 was explored. Further to findings of differences between basic and complex ER on the CAM-C (see Appendix 1), basic and complex emotions were analysed separately. Seven one way ANOVAs and one Kruskal Wallis non-parametric test (for number of basic emotions recognised, which was not normally distributed) were conducted, testing group differences on the emotion recognition tasks at Time 1. Using Holm's Sequential Rejective Bonferroni Procedure (Holm, 1979), significant differences were found between the groups on CAM-C complex emotion faces ($F[2,61]=9.13$, $p<.001$), CAM-C complex emotion voices ($F[2,61]=6.70$, $p<.005$), and number of CAM-C complex emotions correctly recognised ($F[2,61]=5.89$, $p<.005$), but not on CAM-C basic emotion faces ($F[2,61]=2.41$, n.s.), CAM-C basic emotion voices ($F[2,61]=0.11$, n.s.), or number of basic emotion concepts correctly recognised (Kruskal Wallis $\chi^2[2]=2.47$, n.s.). Significant group differences were also found on Time 1 scores for the RME-C ($F[2,61]=11.03$, $p<.001$) and the C-FAT ($F[2,61]=8.68$, $p<.001$). Pre-planned comparisons with Bonferroni corrections revealed no significant differences between the two clinical groups on any of the task scores. In addition, significantly higher scores were found for the typical control group, comparing to the two AS/HFA groups on CAM-C complex emotion faces, voices and number of concepts recognised, and on RME-C and C-FAT ($p<.001$ for all of these comparisons). These findings support Hypothesis 1 on all measures but the CAM-C basic emotions. Table 8.2 shows the means and standard deviations of the groups' ER scores at Time 1 and Time 2.

	Software home users		AS/HFA controls		Typical Controls
	Time 1	Time 2	Time 1	Time 2	
CAM-C scores					
Basic emotion faces (Max=18)	13.05 (2.29)	16.24 (1.09)	11.86 (3.57)	12.29 (2.69)	13.82 (2.84)
Basic emotion voices (Max=18)	12.52 (2.56)	14.52 (1.47)	12.19 (2.75)	12.33 (2.20)	12.18 (2.72)
No. of basic concepts recognised (Max=6)	4.48 (1.33)	5.67 (0.48)	4.10 (1.38)	4.10 (1.58)	4.64 (1.40)
Complex emotion faces (Max=27)	15.05 (4.31)	21.10 (2.95)	14.57 (4.48)	15.48 (5.98)	19.50 (3.76)
Complex emotion voices (Max=27)	16.48 (3.71)	20.14 (2.22)	16.86 (4.08)	17.33 (4.07)	20.41 (3.90)
No. of complex concepts recognised (Max=9)	4.81 (1.97)	7.10 (1.45)	4.71 (2.31)	5.05 (2.33)	6.68 (2.10)
Reading the Mind in the Eyes-C (Max=28)	16.62 (2.78)	18.62 (3.11)	16.71 (3.95)	16.52 (3.23)	20.86 (3.36)
Child Feature-based Auditory Task (Max=17)	9.95 (2.31)	11.95 (2.16)	10.24 (2.76)	10.05 (2.58)	12.77 (2.27)
Reading the Mind in Films-C (Max=22)		15.48 (2.54)		14.52 (3.61)	17.73 (2.80)

Table 8.2: Means (and standard deviations) of the emotion recognition measures in the 3 groups of Experiment 3 at Time 1 and Time 2

Next, seven MANCOVAs with repeated measures were conducted, to examine the differences between the intervention and AS/HFA control group on the various tasks at time 1 and time 2. Due to the large number of tests, age, verbal IQ and performance IQ were included in the analyses as covariates only if they had a significant main effect beyond time or an interaction with time (i.e. with performance change from time 1 to time 2). The results of the 7 MANCOVAs are presented in Table 8.3.

	CAM-C – Close generalisation					*Distant generalisation*	
	Basic emotions		Complex emotions				
	Faces	Voices	Faces	Voices	Concepts	RME-C	C-FAT
df:	(1,38)	(1,38)	(1,38)	(1,38)	(1,38)	(1,38)	(1,38)
Main effects:							
Time	2.62	0.03	0.96	0.28	0.06	0.02	0.94
Group	24.74^{****}	5.79^{*}	18.47^{****}	3.96	13.52^{****}	3.40	8.46^{**}
Age	12.06^{***}	7.66^{*}	29.43^{****}	12.38^{***}	31.00^{***}	6.13^{*}	39.75^{****}
VIQ	5.25^{*}		10.52^{***}	11.52^{***}	17.57^{****}		37.97^{****}
PIQ						2.19	
Interactions:							
Group X Time	9.31^{***}	5.42^{*}	18.97^{****}	8.10^{**}	9.96^{***}	4.25^{*}	6.75^{*}
Age X Time	1.95	0.01	2.40	0.01	0.03	2.29	2.74
VIQ X Time	0.01		0.08	2.23	0.17		0.89
PIQ X Time						5.67^{*}	

$^{*}p<.05$ $^{**}p<.01$ $^{***}p<.005$ $^{****}p<.001$

Table 8.3: F scores of main effects and interactions of the repeated measures MANCOVAs of Experiment 3

As the table shows, no significant main effects of Time (i.e. differences between time 1 and time 2 measures) were found. Significant main effects of Group were found for all measures, except for CAM-C complex emotion voices, and RME-C. However, these differences can be attributed to the different levels of change from Time 1 to Time 2 in the two groups, analysed in the Group X Time interactions below. Age had a significant main effect as a covariate on all tasks, suggesting older participants scored higher (over and above the effects of Group and Time) on the different tasks. This confirms the effect of age on ER scores, reported for all children's tasks in Appendices 1-3, and separately for RME-C (Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001). Similarly, verbal IQ had a main effect on most tasks, suggesting the higher one's verbal IQ, the higher one's score on tasks, over and above the effects of Time and Group. Verbal IQ had no effect on CAM-C basic emotion voices (As found on the validation study of the CAM-C, see Appendix 1). In addition, verbal IQ had no effect on RME-C, which matches similar findings for the RME-A (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Performance IQ had no main effect in any of the tasks.

Using Holm's Sequential Rejective Bonferroni Procedure, significant Group x Time interactions were found for all ER scores, suggesting the AS/HFA intervention and control groups varied on their change in performance from time 1 to time 2. Verbal IQ and Age had no interaction with Time on any of the measures, i.e. they did not affect improvement on task performance. However, performance IQ had a significant interaction with Time on RME-C scores. A positive correlation was found between performance IQ and RME-C difference scores ($r=.30$, $p<.05$), suggesting that the higher one's performance IQ, the greater their improvement was on the RME-C. This effect was additional to the Group x Time interaction effect.

To learn more about the Group x Time interactions, simple main effect analyses, using paired samples t-tests (with Bonferroni corrections) were conducted for the different task scores. The results of this analysis are shown in Table 8.4.

Group (df=20 for each)	*CAM-C – Close generalisation*					*Distant generalisation*	
	Basic emotions		Complex emotions				
	Faces	Voices	Faces	Voices	Concepts	RME-C	C-FAT
Intervention	6.04^{***}	3.77^{**}	6.92^{***}	4.35^{***}	5.04^{***}	2.87^{*}	3.77^{**}
AS/HFA control	0.73	0.25	1.02	0.84	0.98	0.28	0.43

$^{*}p<.01$ $^{**}p<.005$ $^{***}p<.001$

Table 8.4: Simple main effect analysis: paired samples (Time 1-Time 2) t-test scores for the intervention and the AS/HFA control groups

As Tables 8.4 and 8.2 show, the intervention group significantly improved from time 1 to time 2 on all task scores, in both close and distant feature-based levels of generalisation, whereas the AS/HFA group did not improve significantly on any of the tasks. These findings support Hypotheses 2 and 3.

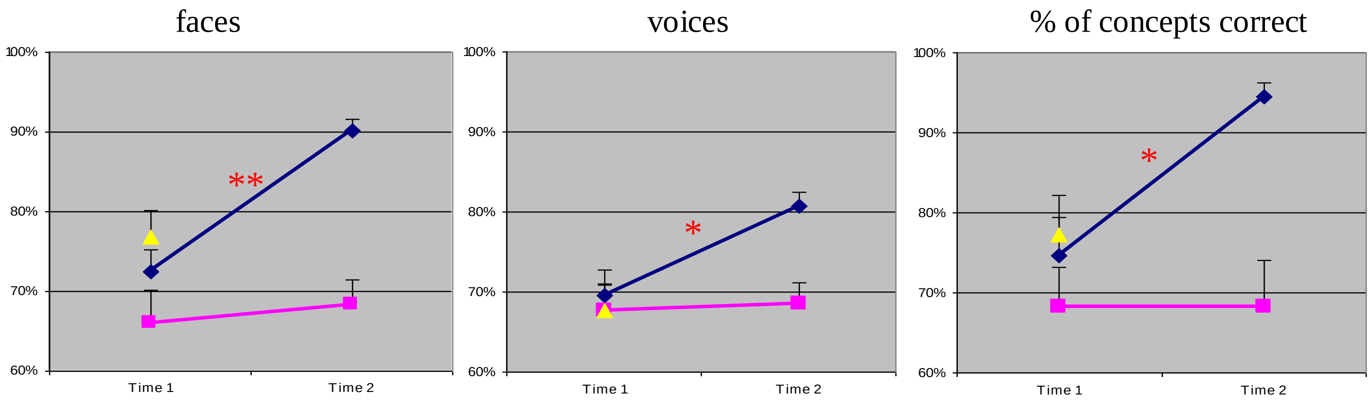
To test differences in the number of basic emotional concepts recognised before and after (which was not normally distributed, and could not be tested using parametric analysis), Time 2 minus Time 1 score differences were computed, and a Mann-Whitney test was conducted for Group differences on this score. The significant result ($U=131.5$, $Z=2.30$, $p<.03$) revealed a Group difference on score changes between

Time 1 and Time 2. Simple main effects were analysed separately for each group using the Wilcoxon Signed Ranks test for the number of basic concepts correctly recognised at Time 1 vs. Time 2. A significant difference was found for the intervention group ($Z=3.07$, $p<.005$) but not for the AS/HFA control group ($Z=0.23$, n.s.). This matches the findings presented above for all other tasks analysed.

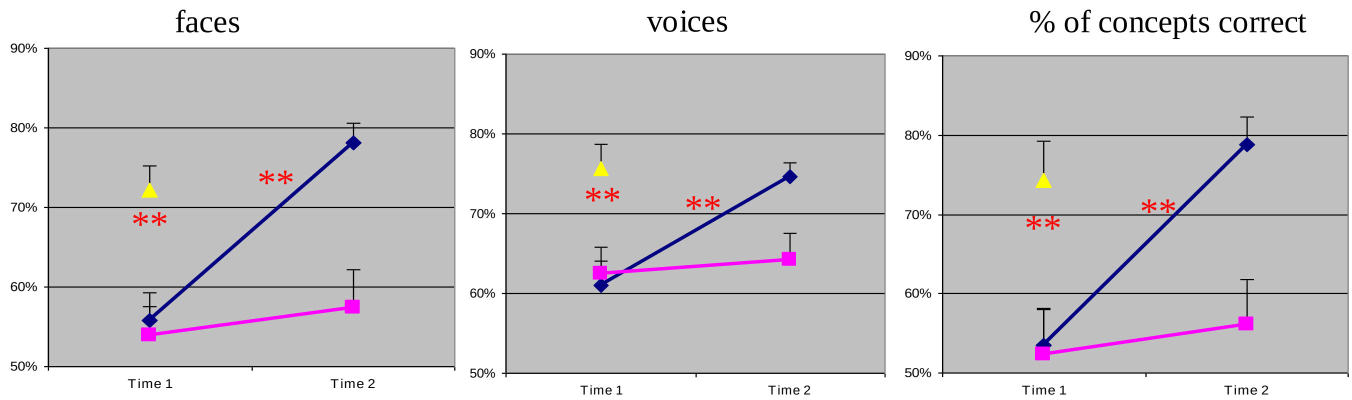
Next, scores of the RMF-C task (which was only taken at Time 2 by the ASC groups) were analysed, to test for group differences on holistic distant generalisation. A one way ANOVA, conducted on the three groups' RMF-C scores, was significant ($F[2,61]=6.43$, $p<.005$). Pre-planned comparisons with Bonferroni correction revealed no difference between the intervention group ($M=15.48$, $SD=2.54$) and the AS/HFA control group ($M=14.52$, $SD=3.61$) for this task ($t[61]=1.02$, n.s.). The typical control group ($M=17.73$, $SD=2.80$) scored significantly higher than the AS/HFA control group ($t[61]=3.48$, $p<.005$), but only marginally higher than the AS/HFA intervention group ($t[61]=2.45$, $p=.052$). These results suggest the intervention group did not perform better than the AS/HFA control group at this level of generalisation.

Figure 8.1 illustrates the group differences on all task scores in terms of proportion of correct answers. Proportions were used in the graphs instead of raw scores, in order to keep the scale uniform for all tasks.

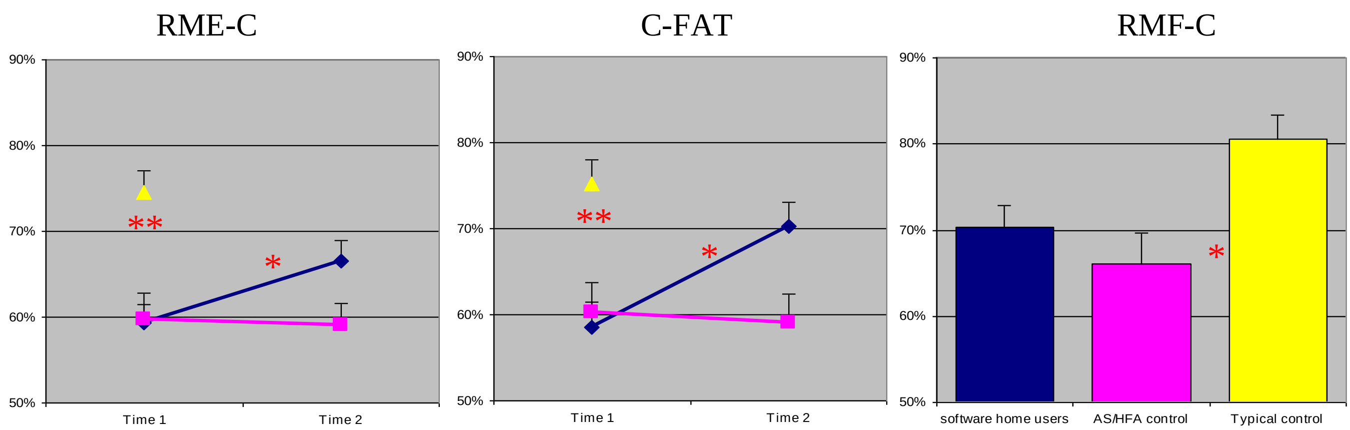
Close generalisation: CAM-C basic emotion score proportions



Close generalisation: CAM-C complex emotion score proportions



Distant generalisation measures



Legend: ◆ Software Users ■ AS/HFA Controls ▲ Typical Controls * p<.01 ** p<.001

Figure 8.1: Mean proportions (with standard error bars) of correct responses to the ER tasks for the 3 groups, Experiment 3

To test for group differences on the 6 basic and 9 complex emotional concepts of the CAM-C, Time 2-Time 1 difference scores were computed. One sample Kolmogorov-Smirnov tests showed that none of the distributions of these scores differed significantly from normal in any of the groups. Two MANOVAs were conducted, one for the basic emotional concepts and the other for the complex ones, with group as the independent variable. The MANOVA yielded a significant Group difference over and above emotional concept for basic ($F_{wilks}[6,35]=2.71$, $p<.05$) and for complex ($F_{wilks}[9,35]=2.64$, $p<.05$) emotions. Significant univariate Group effects were found for the basic emotions *afraid* ($F[1,40]=10.97$, $p<.005$) and *angry* ($F[1,40]=7.93$, $p<.01$) and for the complex emotions *embarrassed* ($F[1,40]=12.47$, $p<.001$), *unfriendly* ($F[1,40]=5.47$, $p<.05$), *disappointed* ($F[1,40]=5.39$, $p<.05$), *amused* ($F[1,40]=5.09$, $p<.05$), *nervous* ($F[1,40]=5.05$, $p<.05$), and *jealous* ($F[1,40]=4.37$, $p<.05$). Recognition of all these concepts improved more in the intervention group than in the AS/HFA control group. However, care should be taken when interpreting these findings, as the analysis was not corrected for multiple comparisons.

Bivariate correlations computed for software usage time in the intervention group with improvement scores of each task and with RMF-C scores yielded no significant results.

Feedback given by *Mind Reading* users focused on the software being entertaining and enjoyable ("*fun*", n=12), though some said it became boring after a while (n=7). Many of the children (n=10) said they worked through all 6 levels of the software, rather than focusing on the levels designed for their age group. Children described different ways they used the software, such as comparing the faces in the software to expressions of characters they like on TV; creating social stories on *Mind Reading*'s scrapbook using facial and vocal clips from the software; coming up with their own examples of emotional situations (e.g. "*I felt excited when I heard that we are going to the Mary Rose in the Summer School I am going to. It is a very exciting adventure as it is all the way in the City of London*"); or adding their own thoughts about emotions in the software's notes tab (e.g. "*Being exploited can get complex. The rules of exploited change if the person who is making you feel exploited offers you a fair reward. I feel this emotion a lot, especially at school*"). Some children said they would continue to use the software in the future (n=6), though others confessed they might need a break from it before trying it again.

Parents of children in the intervention group agreed the software was enjoyable, clear and well designed (n=5). Most parents (n=14) reported they were involved in their child's learning process, discussing emotions and expressions together, and trying to apply them to everyday life. Some of the parents (n=8) said their child's ability to recognise emotions had improved, e.g. "*He recognises expressions from* Mind Reading *in other people*" or "*He is beginning to develop a better understanding of emotional voice tones, especially the more subtle ones*". Others (n=6) described the software's effect in terms of awareness to the existence of emotional cues in faces and voices and their importance in everyday life, e.g. "*We found ourselves looking more at facial expressions, analysing emotions and discussing them in the family*", or "*Mind Reading has given my son the knowledge that there is a point to facial expressions and voice tones and that there are many that he simply hadn't recognised before*". Some parents described more general changes in social behaviour (n=3), e.g. "*His behaviour at school has improved... It was the first time he apologised to the teacher for losing his temper. He told her he was angry*". Interestingly, many parents (n=11) commented about an improvement in their child's emotional vocabulary and use of emotional language, e.g. "*He uses more emotion words when he talks, for example when he tells me about things that happened in school"*, or "*He can express his own emotions better now*". Some parents suggested *Mind Reading* would be more useful if used in school settings (n=3), or if it included more examples of social interaction (n=2). One parent suggested that using the software might affect the children's awareness of their condition and affect their mood: "*Since he started using* Mind Reading*, my son has learned a lot about emotions. On the downside he has picked up that he has to learn these skills whilst his peers already mostly have them. This makes him feel upset as he recognises that this is not a natural thing for him to do and it highlights to him his own differences*".

8.3 Follow-up

8.3.1 Procedure

About a year (12-14 months) after the Time 2 assessment, participants in the ASC groups were sent the follow-up questionnaire shown in Appendix 5. The questionnaire was emailed to participants who had access to email, or posted to those who did not. A second letter was sent a month later to participants who did not respond to the first one.

8.3.2 Participants

Of the 42 participants with ASC who took part in the study at Time 1 and Time 2, twenty six completed and returned the follow-up questionnaires, thirteen from each ASC group. The two groups did not differ on age, verbal IQ, performance IQ, and the CAST and VABS-S scores from Time 1. The percentage of participants (69.2%) from the control AS/HFA group who had used *Mind Reading* since Time 2 did not differ from the percentage of participants from the intervention group who had done so (46.2%; χ^2 (1)=1.42, n.s.). Nor was there any difference between the AS/HFA control group (M=19.06, SD=18.53) in their use of the software after Time 2, compared to the intervention group (M=8.67, SD=9.48; Mann-Whitney U=22, Z=0.60, n.s.). Background information for these participants appears in Table 8.5

Time 1 measures	Participants from intervention group (N=13)			Participants from AS/HFA control group (N=13)			
	Mean	SD	Range	Mean	SD	Range	t (24)[1]
Age	9.52	1.01	8.2-11.7	10.17	1.18	8.6-11.8	1.53
Verbal IQ	109.46	12.84	92-139	107.85	16.05	83-143	0.28
Performance IQ	106.0	11.09	86-129	111.08	18.83	78-140	0.84
CAST	20.08	3.69	14-27	18.08	4.94	11-28	1.17
VABS-S active	19.65	5.33	12-28	20.00	6.01	10-30	0.16
VABS-S interactive	11.23	5.61	4-22	14.15	7.02	3-25	1.17

[1] p>.1 for all t-tests

Table 8.5: Time 1 background averages, standard deviations and ranges for the participants whose parents filled in the follow-up questionnaire – Experiment 3

8.3.3 Results – VABS-S

To check for long term distant generalisation effects of the use of *Mind Reading* during the original intervention period, the two groups were compared on the difference between their Time 1 and follow-up VABS-S *active sociability*, and VABS-S *interactive sociability* scores. After ensuring that the measures' distributions did not significantly differ from normal in the two groups (using Kolmogorov-Smirnov tests), three repeated measures MANOVA analyses were conducted with Time (Time 1 vs. Follow-up) as the repeated measures factor and group as the independent variable. Estimated usage time of *Mind Reading* since the Time 2 assessment was not included as a covariate, as it was not normally distributed, and could not be included in a parametric analysis. Results of the analyses are presented in Table 8.6.

	Intervention group		AS/HFA controls				
VABS-S	Time 1 M (SD)	Follow-up M (SD)	Time 1 M (SD)	Follow-up M (SD)	Time F^1	Group F^1	Time X Group F^1
Active sociability	19.65 (5.33)	19.49 (5.15)	20.0 (6.01)	19.62 (5.47)	0.12	0.01	0.02
Interactive sociability	11.23 (5.61)	14.05 (5.02)	14.15 (7.02)	13.77 (5.60)	2.47	0.37	4.28*

$^*p<.05$ ^{1}df (1,24) for all F scores.

Table 8.6: Means (standard deviations), and main effects and interaction F scores for VABS-S scales repeated measures MANOVAs

As the table shows, no main effects of Group or Time were found, and neither were there Time x Group interactions for VABS-S *active sociability* scores. This suggests neither of the groups had improved on this measure a year after the intervention was held. However, a significant Time x Group interaction was found for the VABS-S *Interactive sociability* scale. Simple main effect analysis, with Bonferroni correction for multiple comparisons, revealed that scale scores of children who came from the intervention group significantly improved from Time 1 to Follow-up (t[12]=2.59,

$p<.025$) whereas no significant change in scores was found amongst participants from the AS/HFA control group ($t[12]=0.35$, n.s).

Correlation analysis of follow-up minus time 1 difference scores on the VABS-S scales with age, verbal and performance IQ, CAST, and time 2 minus time 1 task performance differences yielded no significant results. Similarly, non-parametric correlation analysis conducted for the estimated number of hours *Mind Reading* was used for since Time 2 revealed no significant results. When computed separately for children from the intervention group and children from the AS/HFA group, a negative correlation between age and number of usage hours after Time 2 was found amongst children from the intervention group ($r_{spearman}=-.61$, $p<.05$) but not amongst children from the AS/HFA control group ($r_{spearman}=.25$, n.s.). This finding suggests the older the children who used *Mind Reading* between time 1 and time 2, the less they continued to use it after time 2, whereas in the control group, who did not use the software between time 1 and time 2, there was no such association between age and software usage.

8.3.4 Results – follow-up feedback questionnaire

The follow-up feedback questionnaire required parents of participants who used *Mind Reading* since the Time 2 assessment to estimate how helpful this was for their children's functioning in the socio-emotional domain. In addition to the three areas included in the adult follow-up questionnaire (recognition of emotions in everyday life, understanding social situations, and confidence in social situations) use of emotion words in the child's speech was added, following parents' feedback at time 2 (see above). Parents were also asked to add areas they found *Mind Reading* to be useful in, and to comment about the usefulness and relevance of the software for their child's everyday life. Fifteen participants (6 from the intervention group and 9 from the AS/HFA control group) used the software after the Time 2 assessment. No significant group differences were found on ratings of the feedback items. Therefore, they were analysed beyond group. The ratings for each of the four areas appear in Table 8.7.

How helpful was *Mind Reading* for:	Not helpful at all (1)	(2)	(3)	(4)	Very Helpful (5)	Mean
Emotion recognition in every day life	-	13.3%	20.0%	46.7%	20.0%	3.73
Use more emotion words in speech	6.7%	6.7%	33.3%	40.0%	13.3%	3.47
Understanding of social situations	6.7%	20.0%	40.0%	6.7%	26.7%	3.27
Confidence in social situations	20.0%	33.3%	13.3%	20.0%	13.3%	2.73

Note: n=15

Table 8.7: Participants parents' ratings of *Mind Reading*'s contribution for socio-emotional functioning in real-life – Experiment 3

In order to check whether *Mind Reading* was perceived to be more helpful in any of the four areas, Friedman's non parametric test for related samples was conducted, using ratings on the four areas, and was found significant ($\chi^2[3]=13.62$, $p<.005$). Post-hoc pairwise comparisons included six Wilcoxon Signed Ranks, each comparing ratings for two of the areas. With Bonferroni correction for multiple comparisons, only one of the comparisons reached significance: The contribution of *Mind Reading* for recognition of emotions in everyday life was rated as higher than the software's contribution for confidence in social situations ($Z=2.88$, $p<.005$).

Correlation analysis was conducted between ratings on the four areas, background variables, and Follow-up minus Time 2 difference measures. Only one correlation was found significant: The score difference on VABS-S *active sociability* was positively correlated with parents' ratings of *Mind Reading*'s help in getting the children to use more emotion words in their speech ($r=0.52$, $p<.05$), associating the use of *Mind Reading* with more general social functioning through the acquisition of emotional vocabulary. In addition to the four areas included in the follow up questionnaire, parents also suggested *Mind Reading* was helpful for their child's ability to "*be interested and express ideas about others' emotions*" and to "*help others understand his difficulties*".

In their feedback about the usefulness and relevance of the software for their child's functioning in everyday life, parents focused on the following themes:

- Understanding of emotions and recognising their corresponding facial and vocal expressions: Most of the comments parents made related to the usefulness of the software for their child's ability to understand and to recognise emotions (n=12), (e.g. "*it helped him recognise some of the more subtle emotions and signals*"), and to understand social situations (n=2). This was manifested at home with family members (n=3), as well as in social situations out of the home environment, e.g. "*He started to recognise facial expressions in the school playground and he would mention this to me*". One parent suggested this acquired knowledge had lessened their child's "*confusion and misery*" in social situations. Some of the parents (n=3) said their children still use the software as a reference: "*He has a visual aid which helped him to recognise some of the more subtle emotions and signals he had been missing*".

- Awareness to the importance of emotions and emotional cues: Some parents (n=5) reported the use of *Mind Reading* had increased their child's awareness of the importance of people and their emotions, and increased their interest in these: e.g. "*he is more interested in how people are showing emotions through facial expressions*", or "*He does seem to take more notice of other people's facial expressions and feelings in social situations*".

- Socio-emotional functioning: Parents (n=5) related the use of the software to different improvements in their child's emotional and social functioning: e.g. "*He certainly seems more secure socially than he used to, and more flexible in response*"; "*It increased his self confidence a lot*"; or "Mind Reading *certainly helped him understand a lot more about how to interact with other people*".

In addition to these themes, which were also brought up by adult users of the software (see Chapters 10 and 11), parents added two other themes, which were unique to the child study:

- Emotional vocabulary: Four parents reported an improvement in their child’s emotional vocabulary: e.g. “*He started using more emotion words in his language straight away*”, or “*It has improved his vocabulary relating to emotions*”.

- Family interaction: As the software was used by children at home, with the involvement of parents (and potentially siblings), there were three parental reports of a positive contribution of the software to the communication and interaction around emotions at home: e.g. “*I do think it has improved the framework enabling us as a family to discuss behaviour, feeling and the ways that we express ourselves- both verbally and nonverbally*”.

Comments on the limitations of the software focused on generalisation problems into real-life functioning (n=2): e.g. “*He was good on the software, but this did not cross over into everyday situations*”; and on the difficulties to maintain interest and motivation of children to continue using the software (n=3): e.g. “*It was difficult to get him to use it on a regular basis once his initial interest had gone*”.

8.4 Discussion

Experiment 3 assessed the use of *Mind Reading* by children with ASC at home, and compared their ER performance change to that of children with ASC who did not have any intervention (over and above their regular school curriculum) between the two assessments. Unlike Experiments 1 and 2, the software users showed significant improvement in their performance on close as well as feature-based distant generalisation tasks. In other words, they improved in their ability to recognise emotions from visual and auditory stimuli included in *Mind Reading*, as well as visual and auditory stimuli not included in it. In addition, children who used the software improved in the long run on parental assessment of their *interactive sociability*, an effect which was not found amongst the controls. These findings suggest that overall, child users of the software were able to gain more from it than the adult users in Experiments 1 and 2.

The results of time 1 assessment support Hypothesis 1, as found in the adult experiments. However, the participants in the AS/HFA groups did not perform significantly lower than typically developing controls on the CAM-C basic emotion faces, voices, and number of concepts recognised. These results confirm past findings of successful compensation on basic emotion ER tasks in ASC (see Chapter 2 and Appendix 1).

Hypothesis 2, predicting significant improvement of the AS/HFA intervention group was supported in its entirety, as participants who used the software improved on all ER measures from time 1 to time 2, including CAM-C basic emotion measures. In accordance with Hypothesis 3, the improvement was significantly higher compared to the AS/HFA control group, where no significant change of performance was found on any of the tasks. This is different to the improvement found on CAM-A faces from time 1 to time 2 in the adult AS/HFA control group in Experiment 1. Apparently, the time 1 assessment session did not have such an impact on the children in the control ASC group, and did not increase their awareness to facial expressions or emotions, as it did in the adult AS/HFA control group.

The ER improvement in the intervention group included recognition of basic emotions children and adults with ASC are reported to have difficulties with, such as fear and anger (Dawson, Webb, & McPartland, 2005). More importantly, improvement was found on recognition of complex emotions children with ASC have difficulties recognising, such as jealousy (Bauminger, 2004) or embarrassment (Hillier & Allinson, 2002). The improvement found on the recognition of such developmentally significant emotions, as well as subtle emotions (e.g. *nervous*, *amused*), and complex emotions which require understanding of mental states (e.g. *disappointed*, *unfriendly*) may, if generalised to real life settings, lead to improvement in social understanding and functioning (Attwood, 2000).

Although CAM-C items were harder than questions from *Mind Reading*, included more foils then questions in the software, and were played with no feedback (hence different from the *Mind Reading's* environment), improvement on CAM-C face and voice ER tasks could basically be the result of associative learning, which individuals with ASC find easy, and even excel in (Tsatsanis, 2004; Williams, Goldstein, & Minshew, 2006). Improvement at this level was also expected from the findings of the adult experiments. However, participants in the intervention group also improved significantly on feature-based tasks which stimuli were not included in *Mind Reading*, and which participants had only one encounter with, during the time 1 assessment. Improvement on these tasks beyond that of the control AS/HFA group who took the tasks twice suggests that participants in the intervention group implemented principles acquired through their use of *Mind Reading* to achieve this level of generalisation.

The C-FAT included voice clips which were originally recorded for *Mind Reading*, and were therefore uttered by the same actors who appear in the software. Hence, generalisation from *Mind Reading* to the C-FAT may have been somewhat easier, as the identity and voice of the speakers were known to the software users. However, the content and intonation of the verbalisations were novel, which required participants to rely on the systematic knowledge they acquired from the software to pick the appropriate emotion or mental state. Findings show that this was successful.

The RME-C was probably more challenging than the C-FAT, as it required participants to generalise from colourful, full face *Mind Reading* video clips to black and white, still, eye region only strips in the RME-C. Hence, the intervention group's improvement in this task suggests participants were able to acquire knowledge about the emotional information contained in the eye region from the software and to successfully apply it to this considerably different task. The contribution of performance IQ (measured in the WASI by the block design and matrix reasoning tasks) to improvement on the RME-C suggests that visuo-spatial skills have been utilised in the learning process that led to generalisation from *Mind Reading* stimuli to that of the RME-C.

It is important to note that performance IQ had no effect on any of the ER tasks, but only a significant interaction with time on the RME-C, suggesting it had contributed to the learning process, rather than to ER per se. The systematic presentation of the stimuli in *Mind Reading*, and the analysis recommended by the researchers in the introduction to the software promoted a feature-based learning style. When such a piecemeal strategy for face processing is employed, then pictures of the eye region could relate to videos of the whole face as the single block relates to the holistic geometric shape in the block design (see Chapter 1). Bearing in mind the systemising strengths characteristic of ASC, it is possible that the implementation of such a visuo-spatial strategy on the pictures of eyes, brought the improvement found in the intervention group.

Indeed, improvement on an ER task from strips of the eye region has been previously reported by Bölte and his colleagues (2002), who trained participants with ASC on the FEFA (see Chapter 3). At the brain functioning level, this was related to visuo-spatial processing and visual attention compensatory processes (Bölte et al., 2002; Bölte et al., 2006). However, Bölte et al's eye region stimuli were included in the training program (and therefore could not be considered 'distant generalisation'), whereas software users in the current experiment applied this strategy to stimuli they were not trained on. Furthermore, only basic emotions were included in Bölte et al's training and assessment, whereas the children's improvement on the RME-C in the current experiment included complex emotions and mental states, which may require attribution of mental states in addition to visuo-spatial matching. Hence, the generalisation effect reported in this experiment is unique, as it appears in both visual and auditory modalities, and since it involves complex emotions and mental states, rather than just basic emotions.

As in the adult intervention study, no difference between the AS/HFA intervention and control groups was found on the RMF-C, the holistic generalisation distant generalisation task, suggesting that participants failed to integrate facial, vocal and contextual information, required for successful ER on this task, as in real-life situations. However, unlike the adult study, performance of the typically developing control group on the RMF-C, which was significantly better than that of the AS/HFA control group (at time 2), did not significantly differ from that of the AS/HFA intervention group (at time 2). Whereas this could suggest that the RMF-C is not powerful enough to reveal an existing difference between the groups, this result could also suggest that the AS/HFA intervention group's performance on this task improved to a level of a non-significant difference from the typically developing control group. However, since the RMF-C was only taken once by all groups, there is no way to estimate the improvement the AS/HFA group might have made on this task, and this possible explanation remains hypothetical. Future studies should extend the range of items on the RMF-C, to allow for it to be used before and after the intervention, so that training induced performance changes on a holistic ER task could be tested.

The possibility of a positive effect of software use on participants' mentalising-related socio-emotional functioning, suggested by the improvement on distant feature-based generalisation measures, was supported by the improvement on the VABS-S *interactive sociability* scale in children from the intervention group, but not in children from the AS/HFA control group. This *interactive sociability* scale was designed by Frith and her colleagues (Frith, Happe, & Siddons, 1994) to represent social behaviours that require mentalising (as opposed to the *active sociability* scale, which includes social behaviours that do not require mentalising). The group difference on this scale suggests that using *Mind Reading* contributed to children's mentalising abilities, and to their social functioning beyond ER.

In contrast, no significant group, time or group by time effects were found on the *active sociability* scale of the VABS-S. This matches early findings by Frith et al, who reported no difficulties on these simple and active everyday social behaviours among children with autism (Frith, Happe, & Siddons, 1994). Individual item analysis on the VABS-S was not possible due to the multiple comparisons required and the small group size available. However, a descriptive example for ER related changes in the *interactive sociability* scale is the increase from time 1 to follow-up on the scores of the item '*refrains from statements that might embarrass*', which was found in 31% of the children from the intervention group, compared to 15% of children in the AS/HFA control group. An understanding of the concept of embarrassment, its context and expression are needed in order for the child to improve this aspect of social behaviour, and the use of the software may have contributed to this change.

However, over a year's follow-up, during which other interventions might have taken place, the long-term effect of the training held with *Mind Reading* is by no means certain. In addition, a more direct measure of the children's social functioning (e.g. via observation in real life settings), rather than parental report, may be needed in order to assess the actual long term effect of software usage. This could help validate the effect found on the VABS-S *interactive sociability* scale.

One measure which could have better linked the use of *Mind Reading* with long term improvement found in the follow-up assessment is the estimated number of hours the software was used since the time 2 assessment. In the adult study, this measure was not only correlated with improvement on ER tasks between time 1 and time 2, but also with increased FQ scores amongst adults who continued using the software after time 2. No such effect was found in the child study. The lack of correlation between usage time during the intervention period (i.e. between time 1 and time 2) could be related to less variation (or greater homogeneity) in participants' usage time, which have decreased the covariance between usage time and ER measures, resulting in weaker, non significant correlations (Howell, 1997). A larger sample of participants varying on software usage time may be needed to assess this association. As for the estimated usage time since time 2, this may have been less reliable than the estimation made by adult participants in experiments 1 and 2, simply because it was made by parents, rather than by the users themselves. Since participants were not asked to monitor their usage time past the time 2 assessment, there was no better way to estimate usage time in the current study.

When compared to the effects achieved by adults who have used *Mind Reading* in Experiments 1 and 2, the child study yielded considerably better generalisation effects. This improved level of generalisation following a relatively short training period could be attributed to the greater plasticity, both at the cognitive and the neurological levels, which is characteristic of children (Temple, 1997). A child's brain is less differentiated than that of a mature adult, and may be more capable of adopting new strategies to compensate for existing deficits (Anderson, 2001). In ASC, it is well established that early intervention into communication and social functioning are associated with improved effects and better outcomes, ensuring that children with ASC are able to develop their skills to the full (Howlin & Rutter, 1987; Lovaas, 1987; Rogers, 2000). Early intervention programs help in shaping desirable associations between brain areas, to enhance their connectivity and functionality. Many of them do so by drawing children's attention to social information (e.g. faces, eyes) and rewarding such attention to increase its likelihood in the future (Dawson & Zanolli, 2003). The earliest this could be done, the less the impact of non-adaptive social experience (e.g. focusing on irrelevant details or avoiding attending to people altogether) on the functioning of the developing brain.

At 8-11 years of age, children with ASC who took part in Experiment 3 are likely to have already established some strategies for functioning in the social world. Nevertheless, their ability to generalise knowledge from the software to other visual and auditory emotional stimuli suggest their social perception, and possibly brain systems are still malleable to a greater extent than those of adults. The feedback provided by participants' parents about the effect of *Mind Reading* in raising the awareness of their children to the existence and importance of emotional expressions in faces and voices is somewhat similar to raising of attention to social information in the early intervention studies. However, this reported training-induced increase in attention to socio-emotional stimuli should be validated scientifically, e.g. in a gaze tracking study.

Besides the age of participants, there was another important difference between the child and adult experiments, which could offer an alternative explanation to the difference in results: The adults in Experiment 1 used *Mind Reading* at home alone, and the adults in Experiment 2 used it in conjunction with tutor and group work once a week. However, the children who used *Mind Reading* did this in a nurturing family environment. The support of parents was manifested in several ways: firstly, parents fulfilled an administrative role, by ensuring their children use the software and complete the required amount of usage hours requested. Thanks to that, most of the child participants, unlike the adult participants, were not required to set time for their work with the software, plan it, and execute it independently, thus avoiding potential executive dysfunction difficulties. This may have decreased the attrition rate of child participants (which came down to 12.5%), and allowed children to work with the software for longer.

The more important role of the parents (and siblings) was in allowing the child constant association of the material learned through *Mind Reading* with real life events and experiences in and out of the home environment, thus consolidating the child's knowledge, and supporting generalisation. Tsatsanis (2004) argues that the verbal cues of a caregiver, that transfer across contexts, could be used to foster recall of information and generalisation of skills in ASC, be it in the area of communication or social functioning. Parents of children in the intervention group seemed to be doing that throughout the study. Lastly, parents, who attended the first assessment meeting and watched the presentation of the systematic analysis recommended, might have promoted greater systematising amongst their children, thus increasing the impact and effectiveness of the software above the level of the adults, who had to exercise this strategy with no (or with less) support.

As some of the parents noted, the use of *Mind Reading* by one family member had made the atmosphere and the discourse in the family more attuned to emotions and emotional expression, which the whole family benefited from. The family dynamics that evolve around such a home-based intervention are a matter for a separate study. In addition, as some of the parents suggested, it would be interesting to assess the impact *Mind Reading* could have, if it was used in the school environment, in which peers and teachers could support it with their own real-life practice and challenges. The usability of the software in school settings has been evaluated as a replication of the current study in the USA (LaCava, Golan, Baron-Cohen, & Myles, in press).

8.5 Conclusion

Home-based use of the software by high-functioning children with ASC (who were supported by their families) over a relatively short period of 10-15 weeks brought a significant improvement in their emotion recognition skills not only on complex emotional expressions included in the software, but also on other ER tasks, using stimuli not included in the software. This improvement was significantly higher than that of control participants with ASC who took the assessment tasks twice but with no intervention. Performance on a holistic generalisation task did not differ from the ASC control group or from the typically developing controls, suggesting that an improvement in this level could be visible, had the task been administered twice. Using the software was also found to have a long term positive effect on participants' interactive sociability, compared with participants from the ASC control group, suggesting that in the long run, software usage had also improved mentalising abilities and social skills in users. The improved generalisation found in this experiment is unique in comparison to results of the adult experiments, as well as in relation to previous computer-based intervention studies into socio-emotional skills in ASC. The underlying mechanisms for this change, the studies limitations and future directions, are discussed in the next chapter.

9 General discussion

This closing chapter summarises the findings of this thesis, and discusses joint themes from the three experiments. Two main questions will be discussed:

(1) *Why was improvement limited*? Various models for the generalisation difficulties characteristic of ASC, and their ability to explain the experiments' findings are reviewed.

(2) *What underlies the improvement that has occurred*? Different perspectives on the cognitive processes underlying the behavioural change observed amongst software users are discussed.

Finally, limitations of the work presented in this thesis, and themes for future work are suggested.

9.1 Summary of findings

This study evaluated a new intervention program for systematically teaching emotion recognition in faces and voices to adults and children with ASC. It relied on the systemising strengths individuals with ASC show (Baron-Cohen & Belmonte, 2005; Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002), with the aim of harnessing them to help those individuals compensate for their difficulties recognising emotions and mental states. Amongst high-functioning individuals with ASC, these difficulties were found mostly with more complex emotions and mental states, both in the visual and the auditory channels, and in life-like social scenes. Findings of these difficulties in adults and children with ASC (reviewed in Chapter 2) were confirmed by their performance on the tasks designed for this thesis. As described in Appendices 1, 2, and 3, adults and children with AS/HFA scored significantly lower than matched controls from the general population on complex ER tasks from faces, voices, and multi-modal stimuli.

The effectiveness of ER systematic training using the software was assessed in 3 different experiments, which varied in the age of participants, their level of independence, and the level of support they received in addition to the software.

Experiment 1 evaluated the independent use of the software by able adults with ASC, and compared their progress to that of an ASC control group who were assessed twice but with no intervention. Experiment 2 examined the effects of tutor and group supported home usage of the software by less able (though still with normal range IQ) adults, and compared it to the effects of attending social skills training which did not include a systematic ER component. Experiment 3 tested the effectiveness of the software with a younger user group of high functioning 8-11 year olds with ASC, who used it at home, with the support of their parents. They were compared to a matched group of controls with ASC who were assessed twice with no intervention (except for regular school curriculum). In all 3 experiments, participants were assessed at the beginning and the end of the 10-15 week training period, at 3 levels of generalisation: they were tested on ER from faces and voices from *Mind Reading*, played on a different platform (close generalisation), on photos of eyes and voice recordings not included in *Mind Reading* (feature-based distant generalisation), and on novel holistic social scenes from feature- films (holistic distant generalisation, taken only at time 2). A follow-up after a year assessed more general socio-emotional functioning of participants in real life.

Results of all 3 experiments showed improvement on close generalisation measures amongst software users. They improved on recognition of emotions from facial as well as vocal expressions, including recognition of individual emotional concepts that participants with ASC found challenging before the intervention. This improvement was greater than any change found on these measures between the first and second assessments in the ASC control groups, including when the control group underwent social skills training during the intervention period (Experiment 2). These findings suggest participants were able to learn using the software, and could apply the acquired knowledge to a different and more challenging presentation of stimuli taken from the software.

Applying this systematic knowledge to stimuli not included in the software was more challenging for adult software users in Experiments 1 and 2. These participants did not improve on either visual or auditory feature-based tasks between the two assessments, whether they were supported by a mediator (Experiment 2), or working individually (Experiment 1). These findings confirm prior evidence of limited generalisation in ASC, found in previous studies attempting to teach theory of mind, emotion recognition, and social skills (see Chapter 3 for review). However, children who used the software at home with the support of their families were able to generalise to feature-based distant stimuli not included in the software, suggesting that using this kind of intervention earlier in life, and being supported by mediators who help apply the acquired knowledge to everyday life, extend generalisation beyond the training environment.

In all 3 experiments, performance of software users on the holistic distant generalisation task, measured after the intervention period, was no higher than that of the ASC control group. Though this task was only administered once (hence improvement scores could not be computed), these results suggest that systematic analysis of emotions and mental states from facial and vocal expressions presented separately are not generalised into holistic multi-modal social scenes, which are more similar to real-life situations.

Software usage time was associated with greater improvement on close generalisation vocal measures and with holistic distant generalisation scores, suggesting longer use of the software could bring improved ER performance, and possibly improved generalisation into holistic multi-modal presentations of emotions. These effects were found for the adult experiments only. Verbal IQ was associated with higher scores on ER tasks, especially those which involved verbal dialogue (the voice tasks and the holistic task). However, verbal IQ did not effect the training induced improvement on any of the measures in all 3 experiments.

A follow-up conducted a year after the intervention took place showed that participants from Experiments 1 and 3, who used the software during the intervention period, improved their performance on measures of general socio-emotional performance, compared to participants from the ASC control group. Whereas these effects could be related to other intervening factors which affected participants’ performance during that long period of time, they suggest that systematic ER training has a long term positive effect on participants' socio-emotional functioning.

9.2 Generalisation difficulties in ASC

Central findings of the 3 experiments were the generalisation difficulties participants experienced, either beyond the taught material (in the adult study) or in generalising to holistic multimodal scenarios (in the child study). Generalisation difficulties have been reported to be characteristic of autism spectrum conditions as early as the 1960s (Hermelin & O'Connor, 1970; Rimland, 1965). After four decades, they continue to challenge educators, clinicians, and researchers.

Successful generalisation requires attention to the relevant (and inhibition of the irrelevant) information, creation of rules or categories that are based on this information (or modification of existing ones), and flexible application of these rules to allow one to successfully handle new situations, of a similar nature. For that, generalisation requires a comparison between new input and existing experience, which, in case the two are related, allows one to apply previously acquired strategies in order to deal with the current situation (Plaisted, 2001; Schwartz, Wasserman, & Robbins, 2002). The Weak Central Coherence (WCC) theory of autism would interpret generalisation difficulties in ASC in terms of increased attention/oversensitivity to details, a deficient abstraction ability, and lacking top-down processes which could give the perceived details a bigger perspective and relate it to context (Happe & Frith, 1996; Happe, 1999). The Executive Dysfunction (ED) theory of autism would attribute generalisation difficulties to the limited cognitive flexibility, which results in insistence on sameness and restricted ability to assimilate novel information or to accommodate existing schemata accordingly. Lack of inhibition of irrelevant details may also interrupt the creation of effective schemata (Hill, 2004b; Ozonoff, 1995a).

9.2.1 Generalisation from a systemising perspective

Unlike the WCC and ED theories, the Empathising-Systemising (E-S) theory of autism argues that individuals with ASC have the ability to go beyond the detail level and to create integrative complex rule-based systems (Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002; Baron-Cohen, Wheelwright, Stone, & Rutherford, 1999). According to this theory, the ability to generalise from learned material depends on two factors: The centrality of systemising in the individual's cognitive style, and the degree to which the information is systemisable. The stronger the drive to systemise, the greater is one's need for rules and structure, clarity and predictability and the lower is one's tolerance of variance or ambiguity (Baron-Cohen, 2006). The 'systemisability' of a particular knowledge domain depends on its reliance on a clear system of rules, which could be implemented in order to accurately predict events or outcomes in this domain. A well-defined, rule-based system could be seen as a 'closed' system, as the predictions made on its rules are not affected by any level of chance or error. Alternatively, a system that is more variable, and allows for wrong predictions and errors could be labelled an 'open' system (Klin, Jones, Schultz, & Volkmar, 2003; Lawson, 2003). Strong systemisers who study a closed system would be able to successfully generalise their knowledge and apply its rules to examples that were not present in the learning process (e.g. in mathematics, engineering etc). However, if strong systemisers are faced with variable, loosely structured and less predictable information (an open system), their attempts to structure (systemise) this information may result in a multitude of single information units, which cannot be grouped further until there is reliable evidence that they have no functional differences, since to group them together risks losing key information. For strong systemisers, such a loosely defined set of information units could not be easily generalised out of the examples available in the learning process.

As described in Chapter 1, the socio-emotional domain is an example of such an open system: although it includes some rules (e.g. a smiling person is happy), these are often context-dependent (e.g. a person that has just realised his socks don't match may be smiling in embarrassment) or culture-dependent (e.g. a smiling person could be polite), and leave lots of room for error. Since individuals with ASC are strong systemisers, compared to controls from the general population (see Chapter 1 for review), the likelihood of learning from direct encounters with socio-emotional phenomena is predicted to be low. What *Mind Reading* attempted to do is to insert a level of systematic structure to this open system of emotions, resulting in a more closed system, to improve the chance of learning the ER material included in it. This was done by structuring the emotions thematically, and by providing different examples of the same emotion for comparative systematic analysis. However, this structure was somewhat artificial, as manifested in the separation between facial and vocal expressions, the de-contextualisation of facial expressions, or the allocation of particular emotions to one group rather than to another, even if they may fit in both. Although perhaps less close to real-life emotions, this imposed ‘closure’ of the emotional domain made it easier for strong systemisers to access this domain.

The results of the adult study showed that this systemising of emotions was only partly successful. It enabled software users to improve their ability to recognise complex emotions and mental states from facial and vocal expressions, but limited the generalisation to learned stimuli played on a different platform. However, the children study revealed improved generalisation into feature-based distant generalisation tasks, suggesting a systematic analysis of the emotions included in the software was effective as a mean of improving ER. However, the systemising of emotions on *Mind Reading* was apparently too far from life-like holistic stimuli, as software users failed to improve on these across all experiments.

9.2.2 Alternative explanations for generalisation difficulties in ASC

Other models extend the general explanations provided by WCC and ED theories, focusing on the relation of specific cognitive functions to generalisation in ASC:

9.2.2.1 Rule based vs. prototype-based categorisation

Klinger and Dawson (1995, 2001) argued that generalisation difficulties amongst individuals with ASC are derived from their information processing style, in particular their categorising of information. Individuals with ASC tend to adhere to rule-based categories, while failing to create prototype-based categories of information[1]. Prototype-based categorisation appears in typical development from infancy and continues to develop through childhood (Strauss, 1979; Younger, 1990). Studies have shown that individuals with ASC are capable of forming rule-based categories (e.g. based on shape or colour; Bowler, Matthews, & Gardiner, 1997; Hermelin & O'Connor, 1970). Klinger and Dawson tested low-functioning children with autism and matched controls on two sets of category learning tasks; one that could be solved using a rule-based approach and the other that had no defining rule and required prototype-based categorisation. Children with autism (like typically developing controls) were successful at creating a new rule-based category, but (unlike typically developing controls) failed to learn a new prototype-based category. Since social information does not follow necessary and sufficient rules, and tends to be much more prototype-based in nature (e.g. categorisation of 'a friend' or alternatively 'a stranger'), when individuals with ASC attempt to categorise social information in a rule-based manner, the result is a rigid, inflexible and often socially inappropriate (Klinger & Dawson, 1995, 2001).

Klinger and Dawson's argument could explain some of the generalisation difficulties found in this thesis. The taxonomic structure of *Mind Reading* was aimed at supporting systematic categorisation of emotions. The allocation of emotions into the taxonomy attempted to place emotions in categories in which they most typically belong[2]. However, no rules for the allocation of emotions into categories were given to users. If individuals with ASC fail to create prototype-based categories, they might have accepted the taxonomic structure of the emotions in *Mind Reading* as a rule-

[1] Rule-based, or 'classical' categories, are defined by necessary and sufficient criteria (e.g. all objects with four equal sides that are joined by 90 degree angles are squares). Classical categorisation often fails to explain natural categories (e.g. dogs), which cannot be defined by necessary and sufficient criteria. Instead, these categories are defined by 'typicality' of category members (e.g. a cow is more typical instance of a mammal than a whale). The best example of such a category is the prototype, which is usually an average of all previously experienced category members, rather than an actual member of the category. By storing prototypes in memory, individuals do not need to memorise all instances of a category and can recognise previously unseen members of a category (Rosch & Lloyd, 1978).

[2] Interestingly, there have been debates between scholars who attempted to characterise and taxonomise complex emotions (see Chapter 2) on the rule-based vs. prototype-based nature of these emotions. Prototype theory supporters argue that emotions, like most natural categories, cannot be defined by a set of necessary and sufficient rules and could be better described in terms of proximity to different prototypes (Russell, 1991; Shaver, Schwartz, Kirson, & O'Connor, 1987).

based system, and failed to use it flexibly. Similarly, since categories of emotional facial expressions tend to be prototype-based (Young, 1998; Young et al., 1997), attempting to learn them as rule-based categories might prevent successful generalisation. Whereas recognition of emotion examples taken from *Mind Reading* should not be affected by this deficit (as they could be learned as individual rules within the system) the generalisation of this emotion-related knowledge into other stimuli (either feature-based or holistic) would be hindered. However, this fails to explain the improvement on feature-based distant generalisation found on the children study. Klinger and Dawson's argument is also weakened by a recent study which failed to replicate their findings with high-functioning children with ASC (Molesworth, Bowler, & Hampton, 2005). This finding, in addition to Klinger and Dawson's findings of prototype formation difficulties in a control group of children with Down syndrome, suggest that this deficit may be related to intellectual ability, rather than specifically to autism.

9.2.2.2 Enhanced discrimination and reduced generalisation

Another explanation is provided by Plaisted (2001), who suggested that the attentional and perceptual abnormalities observed in individuals with ASC could be explained by reduced generalisation and enhanced discrimination, i.e. reduced processing of similarities between stimuli and situations, stressing their differences and the uniqueness of each of them. Reduced generalisation was found in studies of category formation, as described above. Support for enhanced discrimination was found in studies of perceptual learning and visual search tasks, on which participants with ASC were superior to controls from the general population in their ability to discriminate between similar stimuli (Plaisted, O'Riordan, & Baron-Cohen, 1998a, 1998b). This perceptual and attentional style is useful on tasks such as the Embedded Figures Test (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983; see Chapter 1), which requires enhanced processing of the unique features of a target, and reduced processing of the features that the target and the background have in common. However, such a style hampers generalisation and transfer of knowledge or skills from one situation to another. Acknowledgement of similarities between past experience and new situations is required in order for the existing knowledge to be transferred and used in new situations. If individuals with ASC process similarities between stimuli and situations poorly, then they will tend to view even similar stimuli as different, which will prevent them from implementing rules or knowledge they have already acquired (Plaisted, 2001). An example for Plaisted's ideas comes from Donna Williams, an adult with autism, who described in her autobiographical book how as a child, a day after she was scolded for writing a graffiti on the church wall and promised she will not do it again, she was caught writing a graffiti on the school wall. To her, this was a completely different situation, hence the rule she learned the day before was not applicable (Williams, 1992).

As with Klinger and Dawson's model, Plaisted's theory can be used to explain the findings of the adult experiments on this thesis: if participants' learning involved enhanced discrimination of the features of individual facial and vocal stimuli presented on *Mind Reading*, then choosing the emotional label with which they were taught should be easier. In fact, such discrimination processing could reduce the chance of making a mistake on the CAM tasks (which include stimuli from *Mind Reading*): no matter how close a foil is to the target emotion, when there is poor processing of similarities between emotions, then the expression could only be matched with the label with which it was learned. This means that learning is only based on associations between the video or audio clips and their verbal labels, and no generalisation to other stimuli (i.e. not from *Mind Reading*) is possible, as there was no search for common features between expressions. Indeed, no training induced improvement was found on distant generalisation tasks. However, Plaisted's theory could not explain the findings of generalisation to distant feature-based visual and auditory ER tasks on the children study. If search for similarities is lacking, children in the intervention group would have not been able to apply any rules to the distant generalisation tasks and improve their performance significantly more than controls who took the task twice.

9.2.2.3 Sensory integration, episodic memory, and cognitive complexity

A possible explanation for the improvement amongst software users on feature-based tasks (close generalisation in adults, close and distant in children) but not on the holistic distant generalisation task is the cross-modal integration this task required. The integration of visual and auditory information in context, which includes input from faces, body language, intonation and language, may be too demanding for individuals with ASC, even if they had learned to recognise emotions from each component separately. At the neurological level, this could be explained by reduced connectivity between sensory and integrative brain areas (Belmonte et al., 2004). Following a comprehensive review of neuro-functional deficits in ASC, Waterhouse and colleagues (1996) suggested that cross-modal sensory information processing for ongoing events and long-term memory records for past events are abnormally fragmented in ASC. The fragmentation is caused by immaturity and over packing density of hippocampal neurons in the autistic brain. This immature system fails to integrate and code incoming sensory information, and to match this with sensory memory records projected from parietal and temporal association cortices. This results in inadequate cued or reflective recall for context and episodes and in poor cross-modal integration of the whole experienced context or episode (Waterhouse, Fein, & Modahl, 1996).

Ben Shalom (2003) carried on this line of thought to argue that people with ASC have specific impairment in episodic memory. Individuals with ASC show impaired performance on tasks that depend on episodic memory, and that cannot be compensated for through the intact perceptual representation or semantic memory systems (the latter mostly in high-functioning ASC). Such tasks include tests of source memory and temporal order memory (Bennetto, Pennington, & Rogers, 1996), and memory for personally experienced events (Millward, Powell, Messer, & Jordan, 2000). At the neurological level, this deficit is arguably related to limited connectivity between the hippocampus and the medial prefrontal cortex (Ben Shalom, 2003).

These two models could explain the lack of generalisation on the distant holistic task, despite improvement on feature-based ER tasks: as long as training is focused on single modalities, learning could be more successful, especially since perceptual processing and semantic memory could be used for compensation, and little association with context is required. Once the task required multimodal integration in context (as real life functioning requires), reliance on perceptual and semantic systems becomes less effective, and retrieval of past contextual experience is required. The episodic memory deficit limits the learning of socio-emotional context through life experience, and restricts their learning to perceptual learning (e.g. the look of facial expressions) and factual knowledge (e.g. smiling people are happy). These deficits prevent flexible processing of complex and subtle socio-emotional information in the autistic brain, thus restricting their learning process and generalisation.

These limitations in processing of complex socio-emotional experience match the findings of Minshew and her colleagues (Minshew, Goldstein, Muenz, & Payton, 1992; Minshew, Goldstein, & Siegel, 1997), who conducted a comprehensive assessment of the neuropsychological profiles of high functioning individuals with ASC on a variety of cognitive functions (including visuo-spatial ability, language, memory, attention, and reasoning). With the exception of intact (and even superior) performance on attention and visuo-spatial tasks, the ASC group showed significant within-domain performance differences between intact performance on tasks that require simple information processing and deficits on tasks that require complex information processing. These results brought the authors to the conclusion that ASC are characterised by a deficit in complex information processing (Minshew & Goldstein, 1998). Whereas these studies included no socio-emotional functioning tasks, difficulties in complex information processing could explain the failure of both adults and children who used the software to show improvement on more complex tasks.

Whereas the models described above may suggest that training individuals with ASC to holistically recognise emotions in real life is futile, because they cannot go beyond the feature-based level, it is important to bear in mind that the training participants received using *Mind Reading* focused on faces and voices separately. This way of teaching might have encouraged an atomised learning style leading to improvement, and made it harder to generalise to holistic material. In addition, the single administration of the holistic task (which limited the assessment to between group differences only) and the relatively short period of training might have prevented significant improvement in this challenging level. The positive correlation between usage time and scores on the holistic task in the adult study, and the lack of difference between software users and typically developing controls on the task in the children study, suggest that this complexity boundary may be crossed with well supported, longer periods of training.

9.3 What caused ER improvement amongst software users?

The E-S theory of autism laid the theoretical foundations of *Mind Reading* as an intervention for individuals with ASC. In addition, the hypotheses of this thesis were based on the E-S model's principles. The premise that individuals with ASC would be able to harness their good stystemising skills to improve their ability to recognise emotions has been partially supported with varying levels of generalisation, as described above. In particular, results of the children study suggest a systemising approach could be effective in teaching features of ER, with good generalisation achieved separately in each sensory channel. Furthermore, the difference between software users and ASC controls on follow-up measures of more general socio-emotional functioning suggest that participants who took this systematic intervention acquired learning principles beyond the immediate association between expressions and emotional labels. In order for an intervention to have a pervasive and long lasting effect, the user needs to learn how to learn, i.e. to learn to focus on materials that are necessary, to be sufficiently motivated to endure the repeated practice that is necessary, and to acquire relevant scaffolding to use when encountering new situations (Gordon, 2000; Vygotskii, 1986). From the studies' results, from participants' feedback, and from the attrition rate, it seems that the systematic training method provided through *Mind Reading* helped participants focus on the relevant stimuli (or some of them, as context and body language were not taught). Motivation was maintained (for children more than adults) via the computerised predictable and friendly environment and via the games and rewards system. It also provided users with some principles that could be used on real life situations, though close human mediation may improve its effects considerably (as seen in the child study).

There are, however, alternative explanations to the causes of improvement amongst software users. One of these is the option of rote or associative learning. Software users might have simply memorised all (or many) face and voice examples included in *Mind Reading*, and their corresponding emotion labels. Rote learning could enable participants to improve their performance on the close generalisation ER tasks. Provided the learner is motivated, associative learning has been reported to be quite successful for individuals with ASC (Howlin & Rutter, 1987; Lord et al., 2005), although generalisation remains a challenge. Therefore, the rote learning option cannot explain the improvement on distant feature-based generalisation measures in the children study, or the follow-up effects found in both studies. None of the participants have reported memorising the emotional expressions this way. However, even if the rote learning explanation is correct, it 'floods' users' cognitive system with emotional expressions, which increases awareness (more on this below), and provides them with a 'library' of expressions they could go through and apply, even if not very flexibly, to expressions they encounter in real life. Temple Grandin described how she has accumulated a bank of visual images in her mind, that helped her learn by rote how to act in certain situations: '*I have a vast data bank, but it has taken me years to build up my library of experiences…I did not know until very recently that most people rely heavily on emotional cues*' (Grandin, 1995, p.137). Using *Mind Reading*, even if merely for rote learning, could help people with ASC realise how important emotional cues are, and provide them with the opportunity to learn lots of different examples of them in faces and voices, thus saving them the need to collect such examples independently over many years.

Another alternative explanation for the improvement software users made is the increase in awareness to the importance of emotions and facial and vocal emotional cues. Increased awareness was reported by many of the adult users and parents of children from the intervention group, and even by adults in the ASC control group, who became more aware following the first assessment meeting. The improvement participants from the latter group made on ER from CAM-A faces demonstrates the power increased awareness had, even when no additional intervention was provided. Amongst software users, as they reported, increased awareness meant that they look more at faces and eyes, try to recognise emotions from faces and voices from the media and in real life situations. The importance of mere awareness (e.g. knowing which parts of the face provide the most salient emotional cues) should not be underestimated: Adolphs and colleagues (2005) reported findings of a facial ER experiment, with SM, a patient with bilateral amygdala damage and controls from the general population. As predicted from past experiments with participants with amygdala damage, SM had poorer ER compared to controls, particularly the recognition of fear. Gaze tracking showed that SM fixates on the eyes far less than controls. However, when specifically asked to pay attention to the eyes (with gaze tracking data confirming increased attention to the eye region), recognition of emotion, fear included, has considerably improved, and did not differ from that of controls (Adolphs et al., 2005). Although the effect did not last without an explicit request to attend to the eyes, this relatively simple intervention, increasing SM's attention to facial features that are pertinent for ER, is quite impressive. The similarities between patients with amygdala damage and individuals with ASC (reviewed in Chapter 2) and the reports of *Mind Reading* users, suggest that increased awareness could definitely account for some of their improvement. Future studies should compare the effects of systematic training with simple increasing of awareness, to discriminate the unique contribution of each aspect.

Finally, an alternative explanation to the improvement children who used the software made could be the support they received from their parents. The combination of exposure to lots of emotional expressions (in a friendly, predictable environment), increased awareness in the child and parents (as some have reported in their feedback), and increased parent-child communication around emotions and emotional expressions, may have yielded greater improvement on distant feature-based ER tasks, and on follow-up socio-emotional functioning amongst children who used the software. The involvement of parents helps motivate the children, keeps them on task, and most importantly – provides them with real-life examples for the taught material, allowing them to improve generalisation. The joint effect of all these may significantly enhance the effect of systematic training.

To summarise, the improvement of participants on ER and other socio-emotional measures in the different experiments could be related to the effect of systematic training, as well as to other, more general effects, to do with the existence of an intervention that deals with emotions, which affects users and their environment.

Separating the different causes for improvement could be useful for academic interest. However, it appears that from a clinical perspective, they have a greater effect when used together. Further suggestions for improvements of the training package appear below.

9.4 Limitations and future directions

Conducting intervention evaluation studies requires constant manoeuvring between structured, standardised protocols of scientific experimentation, and the needs of participants who join the study (in part) to benefit from an intervention that is tailored to their own difficulties (Brewer, 2002). Often, to ensure participants take the intervention in the optimal conditions for them, levels of control over their use of the intervention must be compromised. Whereas the use of *Mind Reading* at home had the advantage of providing participants with better learning conditions, there was no way (but their own reports) to consider the extent to which they followed the instructions for systematic analysis of expressions, or to learn about their individual learning style. Future studies could offer participants to use *Mind Reading* with a tutor (or a teacher at school for children), where individual learning styles and the actual use of systematic learning approach could be assessed.

This study aimed to assess whether good systemising skills of individuals with ASC can be used to improve their ER abilities. However, systemising was not measured, and was rather assumed to be high in participants with ASC, based on findings of previous studies (described in Chapter 1). Since systemising is a relatively new concept, instruments directly measuring it were not available when the experiments described in this study were conducted. An impression of participants' systemising skills in this study could have been achieved through the performance IQ tasks from the WASI, on which individuals with ASC have been shown to have superior performance in the past (Scheuffgen, Happe, Anderson, & Frith, 2000; Shah & Frith, 1993). However, since participants were matched on these parameters, strong systemising of participants with ASC could not be assessed through them. Indirect evidence of the good systemising skills of participants with ASC and their relation to improvement on the ER tasks was found in the performance IQ effect on improvement on RME-C, suggesting the higher one's performance IQ (and indirectly – one's systemising abilities) the more s/he improved on the eyes task. Since an instrument for the assessment of systemising in adults has now been introduced (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003) and one for children is being developed, future studies could assess the effect the level of systemising has on the ability of individuals with ASC to benefit from a systematic guide like *Mind Reading*.

Children and adults from the general population were only assessed once, to obtain baseline measures. The lack of a second assessment for these participants, after 10-15 weeks, prevented a comparison of the progress software users have made to the performance of typically developing participants who took the assessment twice. From the current design no conclusion could be drawn on the post-intervention proximity of performance of participants with ASC to that of typical controls, as it is plausible that typical controls' performance would have improved merely due to taking the tasks twice (as happened with ASC controls in Experiment 1). If a second assessment of typical participants is included in future studies, the training induced progress participants with ASC make towards normal performance could be tested.

The use of computer-based tasks in the evaluation of learning and generalisation in this study has its limitations. Firstly, most of the instruments used were designed especially for this thesis, and were therefore not validated beyond the sample used for it. Although significant differences were found between typical controls and individuals with ASC on the different tasks and good correlations were found between tasks (see Appendices 1-3), validity of these instruments may require further support. Future studies should validate these instruments with a larger sample of the general population and with other clinical groups. Furthermore, although the computer tasks used allow for controlled and structured assessment of emotion recognition skills, testing different modalities separately, they are quite different to real life experience and may not reliably represent real life functioning. Lastly, although follow-up measures showed significant improvement on socio-emotional functioning in software users, compared to ASC controls, the relevance of improvement among software users to real life functioning should be considered with care. Indeed, some participants commented they found recognising emotions and mental states on the computer easier than doing this in real social situations, which requires cross-modal information processing and an immediate reaction in real time. This brings up the question of the appropriate evaluation of emotion recognition abilities, which should, in future studies, be conducted through tasks in the lab, as well as through structured observations in natural settings.

Other limitations of the follow-up assessment include the lack of control on interventions participants have taken since the Time 2 assessment, the lack of ER assessment on follow-up, and the use of questionnaires. It is possible that participants who belonged to the intervention group and experienced its positive effect, got involved in other intervention programmes, which affected the significant improvement these participants showed in the follow-up, compared to the ASC controls. The lack of ER assessment at follow-up, which was not possible due to time constraints, leaves an open question with regards to the sustainability of improvement on ER measures in the long run, and its generalisation to other areas. The use of questionnaires, which were filled in by participants and participants' parents, is helpful in giving an internal view of users' sense of improvement. However, external validation for these improvement reports was lacking. Adults and parents who spent so much time using the software and attending the assessments may be positively biased in their evaluation of the software's effectiveness, to please the experimenters and to avoid a cognitive dissonance (Brewer, 2002; Festinger, 1962). Observations on real-life functioning, or reports by teachers, support workers or co-workers who were blind to the study and its goals may provide more reliable results.

All the tasks used in the child and adult studies required normal range verbal ability. As shown in Experiment 2, which included participants with lower verbal IQ (though still in the normal range), performance on the ER tasks may have been affected by the tasks' verbal requirements. Hence, poor verbal ability may have limited participants' performance on the ER tasks. On the other hand, good verbal ability may have biased the results through compensatory strategies (Ben Shalom, 2003; Grossman, Klin, Carter, & Volkmar, 2000), which bypass the use of emotional cues, despite efforts to block these in the construction of the tasks (see Appendices 1-3). Although verbal ability did not contribute to improvement in performance on any of the experiments, it was only statistically controlled in the studies presented in this thesis. In order to avoid verbal ability confounds altogether, future studies should aim to include ER tasks which limit the use of language to a minimum. Although this is hard to achieve with context dependent complex emotions, there are different kinds of tasks which could provide such a language-free ER assessment. Examples include tasks that use real-life scenarios to provide context, asking participants to choose an answer from several facial expressions that match the situation. A blurring of the verbal content of speech, leaving only the intonation and matching these with facial expressions is another example of such a task. Although different to real life situations (which include the use of language), such tasks could exclude the masking effect of verbal IQ, leaving a purer ER component.

The results of the three experiments call for neuro-imaging studies to examine possible changes in the functioning of brain areas following the systematic use of *Mind Reading*. Such changes may be found in activation or connectivity between areas of the social brain (such as the amygdala, the superior temporal sulcus or the prefrontal cortex), face processing areas (such as the fusiform gyrus) or in visuo-spatial compensatory areas, which might be related to systemising (such as the medial occipital gyrus, or the superior parietal lobule; Bölte et al., 2006). Participants' reports of looking more at faces and engaging in more eye contact following the use of *Mind Reading* should be verified in future studies using gaze tracking. Such studies would throw light on whether the observed cognitive changes reported here are arising from changes in those neural regions that are typically recruited by the non-autistic brain, or if they are due to compensatory strategies by other neural regions.

Other reports by parents of software users in the child study suggest that *Mind Reading* helped increase their children's emotional vocabulary and facilitated its use in their speech. Since these effects were not predicted, they were not assessed in Experiment 3. Future studies could use emotional vocabulary measures (Capps, Yirmiya, & Sigman, 1992; Castelli, Frith, Happe, & Frith, 2002; Tager-Flusberg, 1992; Tager-Flusberg & Sullivan, 1994) to validate the parents' reports in this area.

Finally, it would be interesting to use the design of the studies presented in this thesis to test people with other clinical conditions which may have ER and face processing difficulties, such as ADHD (Singh et al., 1998), conduct disorder (Stevens, Charman,

& Blair, 2001), schizophrenia (Feinberg, Rifkin, Schaffer, & Walker, 1986), and depression (Gur et al., 1992). The assessment of complex ER abilities, using the new ER tasks presented in this thesis, and the assessment of training induced changes following the use of *Mind Reading*, could reveal interesting similarities and differences between these conditions and ASC.

9.5 Future clinical directions

The results of the studies presented in this thesis, while supporting the benefits of computer-based training in terms of users' motivation and structured learning, stress the need for other educational means. These are needed to support generalisation of ER skills into everyday life, but also to address issues that are not taught through *Mind Reading* such as expressing emotions (rather than recognising them in others), using context to recognise emotions, and responding appropriately once an emotion has been recognised.

The use of *Mind Reading* should probably be viewed as a first step in a training programme. Software users should ideally be individually supported by parents or teachers (for children) or by support workers or other tutors (for adults), as this was shown to enhance good generalisation in Experiment 3. After facial and vocal examples have been systematically learned, the next steps would need to deal with the systematic introduction of context and integration of different socio-emotional cues into one (flexible) picture. Each step will need to be explicitly connected to the previous ones, with the main features pointed out, to ease generalisation (Ozonoff, 1995a). This mediation between computer-based systematic training and real life flexibility could be included in the curriculum of social skills groups, as these groups offer a semi-natural and tutor supported setting for socio-emotional interaction. Individuals with ASC who are more socially reserved and may find group activities too overwhelming could mediate between *Mind Reading* and real life functioning through the use of virtual reality. One such intervention, designed for adolescents with ASC, simulates social environments such as the cafeteria or the bus and enables users to learn social scripts and practice social functioning in a predictable environment (Parsons, Mitchell, & Leonard, 2005). The use of such virtual environments to practice ER in a social situation could help generalise ER skills into real life social functioning by gradually and systematically introducing complicating factors, which the user could learn to accommodate. Lastly, integration of principles learned using *Mind Reading* with real life context could be supported by teachers in school settings, by parents at home, and by tutors or co-workers who work with adults with ASC. Such support could be particularly valuable in associating the systemised computer program with the open-system that real life social functioning is.

Appendices

Appendix 1 – Creation of close generalisation measures

This appendix describes the formation and validation of two new ER batteries, one for adults and one for primary school children, which were used as the close generalisation measures in the intervention experiments. These two batteries are also offered as new instruments for assessment of ER in typical development and in other clinical groups. Each battery comprises a face and a voice ER task, using stimuli taken from *Mind Reading*.

As described in Chapter 2, most ER studies focus on the ability to recognise the basic six emotions, usually from Ekman & Friesen's 'Pictures of Facial Affect' test of standard facial expressions (Ekman & Friesen, 1971). However, in order to examine emotion and mental state recognition abilities in high functioning individuals with ASC, there is a need for more fine grained tests, examining more complex and more subtle emotions. Existing instruments, such as the 'Reading the Mind in the Eyes' and the 'Reading the Mind in the Voice' test different complex emotions and mental states, but do not systematically examine the recognition of particular emotions. Rather, these tests include a variety of stimuli, which generate one overall score of ER. The Eyes task uses still black and white pictures of only one part of the face, which are far from the colourful moving whole faces one meets in everyday life. No tasks have studied understanding of complex emotions and mental states via both visual *and* auditory channels, though this has been tested with the basic emotions (Hobson, 1986a, 1986b; Loveland, Tunali Kotoski, Chen, & Brelsford, 1995).

The Cambridge Mindreading (CAM) Face-Voice Batteries for adults (CAM-A) and for children (CAM-C) were created to address these issues. Both batteries were designed to assess a wide emotional repertoire of emotions, including a variety of complex emotions in the CAM-A, and of both basic and complex emotions in the CAM-C. These are examined individually through both visual and auditory modalities, using motion in the visual task.

The Cambridge Mindreading Face-Voice Battery - adult version

Introduction

The CAM-A battery evaluates a selection of 20 emotion and mental state concepts, taken from *Mind Reading*. Since the battery was created to evaluate the software, emotional concepts were selected so that they represent the different emotion groups included in *Mind Reading.* Concepts in the battery represent 18 of the software's 24 emotion groups. This covers a good variety of emotions and mental states, while remaining relatively brief for administration. The battery includes two tasks: ER in the face and ER in the voice. Each of these tasks has fifty questions, in which the participant is either watching 3-5 second silent clips of actors portraying an emotion (face task), or listening to short sentences, spoken in a particular emotional intonation (voice task). After watching the clip/listening to the voice recording, the participant is presented with four adjectives and is asked to "choose the word that best describes how the person is feeling".

In order to make sure that the chosen concepts are taken from the adult emotional repertoire, data from the emotional lexicon developmental survey was used. Each of the 20 emotional concepts selected was known to at least 75% of the fifty 17-18 year olds who took part in the emotional lexicon survey, which is significantly above chance ($p<.05$, binomial test) (Baron-Cohen, Golan, Hill, & Wheelwright, submitted). Some of the mental states included are 'positive' in valence, such as *empathic* and *intimate*, and others are 'negative', such as *guarded* and *insincere*. An attempt was also made to include emotions of varying intensity, i.e. subtle emotions on the one hand (e.g. – *uneasy*, *subdued*) and intense ones on the other (e.g. – *appalled*, *mortified*). The twenty concepts, the emotion groups they represent in *Mind Reading*, and the proportion of 17-18 year olds familiar with them are listed in Table 1.

Concept	Emotion Group	Known to 17-18 year olds (n=50)
Appalled	Surprised	100%
Appealing (asking for)	Wanting	100%
Confronted	Hurt	100%
Distaste	Disgusted	92%
Empathic	Kind	82%
Exonerated	Happy	[1]
Grave	Sad	96%
Guarded	Disbelieving	94%
Insincere	Sneaky	96%
Intimate	Romantic	100%
Lured	Interested	88%
Mortified	Sorry	96%
Nostalgic	Touched	82%
Reassured	Liked	98%
Resentful	Unfriendly	98%
Stern	Unfriendly	98%
Subdued	Sad	88%
Subservient	Unsure	94%
Uneasy	Afraid	98%
Vibrant	Excited	96%

Table 1: The 20 emotional concepts included in the CAM-A

Recognition of each of the 20 emotion concepts was tested by 5 items. The criterion for 'passing' a concept was correct recognition of at least 4 out of 5 items. Achieving 4 or more out of 5 would represent above chance recognition of the concept (Binomial test, $p < 0.05$). Of the 20 concepts, 10 concepts were measured by 3 face and two voice items each. The other 10 concepts were measured by 3 voice items and two face items. This design allowed to keep the battery brief overall.

There are four different scores that can be derived from the CAM-A:

An overall emotion recognition score – defined as the sum of all the correctly answered questions, ranging from 0-100, describing overall emotion and mental state recognition. Any score greater than 35 is above chance at the $p < 0.01$ level (Binomial test).

[1] Developmental data for this concept was missing, as it was included in *Mind reading* but not in the developmental survey.

Facial emotion recognition score: defined as the sum of all items answered correctly from the facial items, ranging from 0-50. Any score greater than 20 is above chance at the $p < 0.01$ level (Binomial test).

Vocal emotion recognition score: defined as the sum of all items answered correctly in the voice items, ranging from 0-50. Any score greater than 20 is above chance at the $p < 0.01$ level (Binomial test).

Concepts correctly recognised: The concepts can be studied in two ways: The sum of concepts correctly recognised, ranging from 0-20 (any score greater than 2 is above chance at the $p < 0.01$ level, according to Binomial test), or the particular concepts correctly answered, analysed individually.

The CAM-A battery was tested with adults with AS/HFA and with matched controls from the general population. The ASC group was predicted to perform worse than the control group on all four measures of the CAM-A. In addition, it was predicted that CAM-A scores would positively correlate with the Revised ‘Reading the Mind in the Eyes’ test (RME) and with the revised version of the ‘Reading the Mind in the Voice’ test (RMV-R), that are described in detail in Appendix 2. Performance on complex ER tasks has been found to correlate negatively with level of autistic symptoms, suggesting that the more autistic traits one possesses, the more difficulties they will experience in recognising emotions and mental states in others (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Dennis, Lockyer, & Lazenby, 2000) Hence, it was predicted that CAM-A scores correlate negatively with scores of the Autism Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), suggesting that the more autistic traits one has, the poorer their ER will be. Lastly, since individuals with ASC used the CAM-A twice in this study (before and after the 10-15 week intervention period) - tests-retest reliability was calculated for the different scales.

Method

Participants

The ASC group comprised twenty-one adults (15 males and 6 females), aged 18-50 (Mean age=30.2, SD=10.5). Participants had all been diagnosed with AS or HFA in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). They were recruited from several sources, including a local clinic for adults with ASC, support organisations, and colleges for individuals with autism spectrum conditions around the UK. All participants were given the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) and scored above 70 on both verbal and performance scales.

Group 2 comprised twenty-one adults recruited from a local employment agency. After screening for autistic spectrum conditions using the AQ (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), four participants were excluded for scoring above cut-off of 32. The remaining seventeen, 12 males and 5 females, matched the clinical group in age, verbal IQ, and performance IQ. They spanned an equivalent range of socio-economic classes and educational levels as that seen in the clinical group. Goodness of fit test for the sex showed no group difference ($\chi^2[1]=.003$, n.s.). As shown in Table 2, t-tests for age, verbal and performance IQ revealed no significant differences between the groups at the $p<0.05$ level. The AS/HFA groups' AQ scores (M=33.62, SD=8.92) were significantly higher than those of the control group (M=13.88, SD=6.27; t[36]=7.7, p<.001).

	AS/HFA group (n=21)			Control group (n=17)			
	Mean	SD	Range	Mean	SD	Range	t (36)[1]
Age	30.23	10.50	17.9-49.9	27.08	10.25	17.6-51.2	0.93
Verbal IQ	114.38	9.02	100-127	118.47	7.42	104-130	1.50
Performance IQ	107.71	15.95	72-134	110.82	8.57	92-125	0.72
Full Scale IQ	112.48	12.26	84-131	116.29	6.71	106-128	1.22

[1]p>0.1 for all t tests

Table 2: Age and IQ scores for AS/HFA and control groups - CAM-A study

The CAM-A: Test development

25 complex emotions and mental states were selected from the 24 emotion groups in *Mind Reading*. A pilot carried out earlier with adult participants with ASC using the basic emotions resulted in ceiling effects. This confirmed that choice of complex emotion concepts was appropriate.

Selection of the 25 concepts followed 3 principles: concepts should be (a) selected from all 24 emotion groups, (b) mainly subtle, and (c) important for everyday social functioning. Although most of the 24 emotion groups were represented in the CAM-A by one emotion concept, the larger emotion groups (*unfriendly* and *sad*) were represented by 2 emotion concepts each. For each concept, six test questions were created, using the quiz module in *Mind Reading* which randomly selected three video films and three audio voice recordings, and matched them with foil words from the same developmental level, ensuring that foils were not from the same emotion group as the target answer. Although choosing foils from other emotion groups possibly makes the tasks easier, it was believed that foils taken from the same categories as targets might be too similar and increase the dependency of performance on verbal ability (i.e. the ability to distinguish one emotion from another within the same group). It was decided that some of the groups might still serve as quite difficult foils to other emotion groups with a similar valence and theme. For example, emotions from the *unfriendly* group were used with targets from the *angry* group, emotions from the *sad* group were used with targets from the *hurt* group, etc.

Two tasks (face recognition and voice recognition) were created and run on an IBM compatible computer, using the experimental software DMDX (Forster & Forster, 2003). Each task started with an instruction slide, followed by two practice items, to ensure understanding of the task. In both tasks the test items were presented in a random order. The face task comprised silent clips of adult actors, both male and female, and of different ethnicities, expressing the emotions in the face (though with shoulders visible). An example of a question, showing one frame from the whole clip is shown in Figure 1. The voice task comprised recordings of short sentences expressing various emotional intonations.

1. oppressive 2. vibrant 3. unenthusiastic 4. uneasy

Figure 1: An example of an item from the CAM-A face task (showing only one frame out of the full clip)

In both tasks four adjectives, numbered from 1 to 4, were presented after playing each stimulus. Participants were asked to press 1, 2, 3 or 4 on a keyboard to select their preferred answer. After choosing an answer the next item was presented. No feedback was given during the task. A handout of definitions of all the adjectives used in the task was available for the participants at the beginning of the assessment. Participants were encouraged to go through it and make sure they were familiar with all the words, and to use the handout in case of any doubt during the task. Hence, response time was unrestricted and not measured.

Item validation was conducted before carrying out any group analysis. The data from the adults in the control group was first analysed as follows: An item was considered valid if at least 50% of these participants selected the target word and no more than 33% selected any one of the foils. Using these criteria, 5 of the concepts, were excluded from the battery. Eight other concepts had one invalid item each, and these items were removed. In order to keep the same number of items for all the concepts and to keep the battery brief to administer, one item was then randomly removed from each of the remaining 12 concepts, so that the final battery comprised 20 concepts with 5 items for each concept. The number of concepts having 3 face and 2 voice items or 3 voice and 2 face items were counterbalanced.

Procedure

Participants were individually tested during the first assessment meeting of the intervention study. Testing took place either at the Autism Research Centre in Cambridge, or at local support centres and colleges for individuals with ASC. Participants were seated in front of IBM compatible computers with 15 inch monitors and were given headphones for the voice tasks. Task order was randomised, and participants were allowed two short breaks during each task. There was no time limit to answer each item. The CAM-A took about forty five minutes to complete. The RME and RMV-R tasks were taken as at the same assessment meeting. Participants filled in the AQ in advance and brought it with them to the assessment meeting.

Results

A calculation of face, voice and overall scores was made by counting the number of correct answers in each of the scales and across the whole battery. In addition, a tally of concepts correctly recognised was made. A minimum of 4 correct answers out of 5 items was considered successful recognition of a concept.

All participants scored above chance ($p<0.01$, Binomial test) on the voice scale and on the overall score. All except for two participants from the ASC group scored above chance on the face scale ($p<0.01$). These two participants scored just below the threshold (20 correct answers out of 50), one of whom scored above chance at the $p < 0.05$ level. They were therefore included in the analysis.

One sample Kolmogorov-Smirnov tests were conducted for all task scores in the two groups. Distributions of all scores in both groups did not differ from normal. Hence, parametric analysis was used.

Univariate analyses of covariance (ANCOVA) were performed on the face scale score, the voice scale score, the CAM-A overall score and the number of concepts correctly recognised. Group (AS/control) and Sex were used as independent variables. Verbal IQ, performance IQ, and age were entered as covariates. A main effect of group was found in all four analyses. The scores of the AS/HFA group were significantly lower than those of the control group on the face scale ($F[1,31]=15.61$, $p<.001$), the voice scale ($F[1,31]=21.26$, $p<.001$), the overall score ($F[1,31]=25.32$, $p<.001$) and the number of concepts passed ($F[1,31]=15.67$, $p<.001$). The means and standard deviations of the 4 scores for the AS/HFA and control groups are shown in Table 3.

In addition, a main effect of sex was found for the face scale ($F[1,31]=5.02$, $p<.05$). Females ($M=40.18$, $SD=7.60$) were significantly better than males ($M=36.19$, $SD=8.78$) at recognising emotions from faces in both groups, regardless of diagnosis. No other effect or covariate contribution was significant.

		AS/HFA (n=21)	Control (n=17)	F (1,31)
a. Face scale	Mean	32.33	43.53	15.61**
(Max=50)	(SD)	(7.96)	(4.03)	
b. Voice scale	Mean	35.71	42.76	21.26**
(Max=50)	(SD)	(6.19)	(3.78)	
c. Overall score	Mean	68.05	86.29	25.32**
(Max=100)	(SD)	(11.69)	(5.99)	
d. Number of concepts passed	Mean	10.76	16.76	15.67**
(Max=20)	(SD)	(4.62)	(2.39)	

** p<.001

Table 3: CAM-A task scores in the AS/HFA & control groups

Next, in order to compare the recognition of individual concepts in the two groups, an 18 by 2 multivariate analysis of variance was performed for the proportions of participants who correctly recognised each concept in the two groups. Due to the large number of dependent variables and since groups were matched on sex, age and IQ, these variables were not included. Two concepts – *uneasy* and *appalled*, which all of the participants of the control group passed, were excluded from the analysis. The analysis yielded a significant overall effect of group ($F_{wilks}[18,19]=2.60$, $p<.05$). Individual concept analyses revealed that the AS/HFA group scored significantly lower than the control group in the recognition of 11 of the 18 concepts. These concepts were: *intimate*, *reassured*, *distaste*, *insincere*, *lured*, *mortified*, *nostalgic*, *resentful*, *subservient*, *grave*, and *exonerated*. The two concepts, which were excluded from the analysis, were then analysed separately with a goodness of fit test. The proportion of participants with AS/HFA who correctly recognised *uneasy* was significantly lower than that of the control group ($\chi^2[1]=8.2$, $p<.01$). There was no difference between the groups in the recognition of *appalled* ($\chi^2[1]=0.83$, N.S.). The proportion of participants of the two groups who passed each of the 20 concepts is shown on Table 4.

Concept	AS/HFA group (n=21)	Control group (n=17)	F (1,36)
appalled	95.2%	100.0%	[2]
appealing (asking for)	52.4%	76.5%	2.37
confronted	71.4%	82.4%	0.60
distaste	57.1%	94.1%	7.60**
empathic	76.2%	76.5%	0.00
exonerated	33.3%	94.1%	22.28**
grave	42.9%	82.4%	6.93*
guarded	42.9%	52.9%	0.37
insincere	28.6%	88.2%	19.90**
intimate	42.9%	94.1%	14.61**
lured	42.9%	82.4%	6.93*
mortified	66.7%	94.1%	4.54*
nostalgic	66.7%	94.1%	4.54*
reassured	42.9%	82.4%	6.93*
resentful	61.9%	94.1%	5.95*
stern	47.6%	52.9%	0.10
subdued	66.7%	70.6%	0.06
subservient	28.6%	70.6%	7.64**
uneasy	61.9%	100.0%	[1]**
vibrant	81.0%	94.1%	1.40

** p<.01 * p<.05

Table 4: Proportion of participants who correctly recognised the CAM-A concepts

Power calculations for the different scales (with $\alpha=0.01$) show that the scales are quite powerful in differentiating the group with AS/HFA from the control group. Power scores were $1\text{-}\beta=0.99$ for the face scale, 0.94 for the voice scale, 0.99 for the tally of correctly recognised concepts and 0.99 for the CAM-A overall score. With a more conservative alpha level ($\alpha=0.001$) the respective power scores were: $1\text{-}\beta=.0.97$

[2] Analysis of these concepts was done using goodness of fit test due to a ceiling effect in the control group.

(face), 0.76 (voice), 0.93 (number of concepts recognised), and 0.99 (overall). Test-retest correlations, calculated for individuals with AS/HFA who participated in the no-intervention control group of experiment 1 and took the tasks before and after the 10-15 week period were r=0.94 for the face scale, r=0.81 for the voice scale, r=0.91 for the overall battery score, and r=0.97 for the number of concepts recognised ($p<.001$ for all).

As predicted, the CAM-A was strongly and positively correlated with the 'Reading the Mind in the Eyes' task (RME-A) and the revised version of the 'Reading the Mind in the Voice' Task (RMV-R) (see Appendix 2 for description). All of the CAM-A's scores correlated positively with these external criteria. Unsurprisingly, the face scale had a stronger correlation with the Eyes task ($r = .74$, $p < 0.001$) compared to the Voice task ($r=.49$, $p < 0.01$). Similarly, the voice scale of the CAM-A correlated more strongly with the Voice task ($r =.62$, $p <.001$) than with the Eyes task ($r=.32$, $p <.05$). The two scales of the CAM-A maintained a positive correlation between themselves ($r =.57$, $p <.001$), as they did with the overall score and the number of correctly recognised concepts.

	CAM-A – overall	Face scale	Voice scale	No. of concepts passed
Facial scale	.92**			
Voice scale	.85**	.57**		
No. of concepts passed	.98**	.90**	.82**	
Reading the mind in the eyes	.63**	.74**	.32*	.65**
Reading the mind in the voice–R	.61**	.49**	.62**	.61**
AQ	-.57**	-.47**	-.56**	-.51**
AGE	-.04	-.12	.08	-.04
Verbal IQ	.24	.26	.14	.27
Performance IQ	.19	.18	.14	.18

** p<.01 * p<.05

Table 5: Correlation of the CAM-A scores with each other, with external criteria and with background parameters

All of the CAM-A's measures were negatively correlated with the AQ score, which means that the more autism spectrum characteristics one possesses, the lower one's

CAM-A scores. All correlations of the CAM-A scores with IQ or age of the participants were found to be non-significant, which suggests that what the CAM-A measures is independent of both verbal and performance IQ, as well as chronological age. Correlations are shown in Table 5.

Discussion

The CAM-A - a new battery for testing recognition of complex emotions and mental states in the face and the voice - expands previous work that found difficulties in this domain among adults of normal intelligence, diagnosed with ASC (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997; Rutherford, Baron-Cohen, & Wheelwright, 2002). Unlike previous studies, the CAM-A battery allowed a test of the recognition of specific emotions and mental states as well as overall performance, and recognition in the two perceptual channels, separately. It also tested recognition of complex emotions and mental states using films of faces rather than still pictures.

Results showed that individuals with AS/HFA, when compared to general population controls, had more difficulties in recognising complex emotions and mental states from both faces and voices. In addition, participants with ASC recognised fewer emotion and mental state concepts than controls. In twelve out of the twenty complex emotions and mental states tested in the CAM-A, a significantly lower number of participants with ASC successfully recognised the concept, compared to age-and IQ-matched controls.

The fact that controls were matched on chronological, verbal and nonverbal mental age, and the lack of correlations between the CAM-A scores and these factors, suggests the independence of complex mental state recognition from verbal and non-verbal ability. The strong negative correlation of the CAM-A scores with the participants' Autism Spectrum Quotient (AQ) score (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001) supports the relevance of emotion and mental state recognition difficulties in high-functioning adults with ASC. Despite their ability to recognise basic emotions, such adults still find it hard to 'mindread' complex emotions and mental states from faces and voices, even when the whole faces is presented in live action. The CAM-A was found to have good test-retest reliability over a period of 10-15 weeks. The correlations of the CAM-A face and voice scales with the RME and the RMV, respectively, provides the task with important measures of external validity. The CAM-A goes beyond these two tasks by using motion in the face scale items and by allowing the opportunity to analyse individual concepts. Its power levels show it is sensitive to group differences across all scales and scores.

The CAM-A tests recognising emotions independent of weak central coherence (Frith, 1989) or executive function (Ozonoff, Pennington, & Rogers, 1991) because there is minimal context or planning, which burden working memory. However, like almost everything, responses do of course require some minimal inhibition (the ability to go through all the answers and choose the best one) and working memory.

A review of the emotions and mental states with which the ASC group had significant difficulties reveals no clear pattern: Of the concepts included in the CAM-A, the groups did not find positive emotions easier to recognise than negative ones. However, emotion valence and subtlety were not systematically studied in the CAM-A, and these could be studied in their own right in future studies.

Despite the small number of female participants in this study, a sex difference was found on the face scale in this study, and this was independent of diagnosis: females recognised emotions in faces significantly better than males. This has been found in typically developed controls using the RME (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001) or the 'Profile for Non verbal Sensitivity' (Baron-Cohen, 2003; Hall, 1984). The lack of studies investigating females with ASC calls for a thorough investigation of their profile. The absence of a sex difference on the voice scale mirrors the lack of such a difference on the RMV task (Rutherford, Baron-Cohen, & Wheelwright, 2002).

Although the dependent variables in this study measured number of correctly recognised emotions, it is nevertheless interesting to examine the errors made by the groups. Due to the wide variety of emotions used as distracters, it was not possible to locate specific error patterns. However, there were some interesting anecdotal examples of errors made only by participants in the AS/HFA group. These errors were mainly about missing subtleties in face items and answering according to the content while ignoring the intonation in voice items. Though in most cases chosen distracters were of the same valence as targets, in some items participants in the AS/HFA group preferred a cross valence distracter upon the target answer (i.e. preferred a negative distracter when target was positive and vice versa). All examples of the errors quoted in the discussion section below were chosen by at least 30% of the participants with AS/HFA.

When looking at the group differences on individual concept recognition, the deficit among the AS/HFA group in recognising *insincerity* is most striking. Less than a third of the AS/HFA group members recognised this mental state successfully, which supports evidence from other studies, showing specific difficulty in understanding deception (Baron-Cohen, 1992; Sodian & Frith, 1992) and in judging trustworthiness and approachability of people from pictures of their faces (Adolphs, Sears, & Piven, 2001). Understanding of hidden intent, masked behind an incongruent facial expression, is a major area of difficulty for individuals with ASC. In her book *Thinking in Pictures*, Temple Grandin described her experience as a high functioning adult with autism: “It is easy for me to understand the concept of deception when it involves playing tricks…but understanding the social cues that indicate an insincere person is much more difficult” (Grandin, 1995, p. 137). Other studies have also reported a specific difficulty in matching incongruent faces and emotion labels among children with AS (Grossman, Klin, Carter, & Volkmar, 2000). It is possible that on the CAM-A items for *insincere*, the participants with AS/HFA were distracted by the content of what was said in the voice items, rather than judging the intonation, which caused them to make errors. Similarly, they might have used more obvious cues in the face (e.g. the fact that the person in the film was smiling) to answer the face items while missing other, more subtle facial cues (e.g. gaze that was incongruent with the smile). An example of such an error is the preference of the distracter ‘spellbound’ when asked to recognise insincerity in one of the face items. The participants might have interpreted the smile and avoiding gaze as a spellbound rather than an insincere emotional state. Since deception is tested in the false belief tasks (which adults with

ASC passed with no difficulty), the CAM-A might provide a more subtle alternative to these tasks.

Another emotional concept the AS/HFA group had particular difficulty recognising was *subservient*. For example, two thirds of the AS/HFA group preferred the label ‘miffed’ for a subservient face item (comparing to 11% in the control group). This could reflect their confusion between dominant and submissive characters in a social situation. Since dominance hierarchies are widely recognised in social primates (De Waal, 1998) it is surprising that people with AS should find this emotion difficult. However, it may be that non-human primates rely on other cues to judge dominance and subservience (e.g. physical size or success in conflict). It is likely that people with ASC would have no difficulty in understanding social hierarchies from non-emotional cues (such as rank). It may therefore be that their deficit arises only when the cues are from emotional expression. Such misunderstanding of hierarchical human systems and social relations might, for example, lead to the use of an inappropriate attitude towards authority.

A similar problem might arise for the misunderstanding of *intimacy* – another clearly interpersonal emotion. The AS/HFA group had difficulties spotting the interpersonal aspect in this mental state. More than 40% of participants with AS/HFA mislabelled ‘*intimate*’ face items as ‘determined’ and ‘carefree’. Similarly, 30% of them mislabelled an intimate voice item as ‘subservient’, possibly relying on its content (“ask me anytime”) while not picking up the intonation. It is easy to imagine how such a cognitive deficit could lead to difficulties in relationships and to difficulties distinguishing genuine closeness and interest on one hand, from politeness or even boredom in an encounter on the other. The AS/HFA group’s difficulties in recognising *resentful*, *grave* or *mortified* also may reflect longstanding difficulties in reading emotions and mental states. The lack of group difference for mental states like *guarded* or *stern* may reflect these items being hard for both groups, since both groups scored relatively low on them. Further examination of these mental states and the groups they represent will be required.

Another mental state which the AS/HFA group had significant difficulties with was *exonerated*. A third of the participants in the AS groups mistook a face item of *exonerated* for ‘remote’, which is a cross valence distracter. Similarly, they confused the positive *exonerated* for a negative label, ‘resigned’ in a voice item, again using the content only (‘now I can get down to work’). This effect could relate to verbal difficulty of this concept (which was not included in the emotional lexicon development survey). However, analysis revealed no significant effect of verbal IQ in the ability to recognise this emotion. This mental state (which concerns relief from blame, like *reassured* which also had a significant group effect), requires the use of a theory of mind for its interpretation, which might have made its recognition especially hard for the AS/HFA group.

One of the emotional concepts that was recognised by all members of the control group but only 62% of the AS group is *uneasy*. This emotion, expressing very subtle degrees of fear, is a good example of the difficulty individuals with ASC might have with picking up subtle emotional cues. Ashwin and colleagues conducted an imaging study, testing individuals with ASC and matched controls on the association of amygdala activation with the detection different intensities of fear. They found that, unlike the controls, the ASC group showed no difference in amygdala activation between intense, moderate and mild degrees of fear (Ashwin, Baron-Cohen, Wheelwright, O'Riordan, & Bullmore, in press). Findings of the current study are consistent with these neuro-imaging findings. Future studies should investigate amygdala activity in individuals with ASC not only when watching but also when listening to fear items of different intensity (e.g. uneasy, afraid, terrified).

Another significant difference in recognition between the AS/HFA and control groups was with the concept *distaste* from the *disgusted* group in the taxonomy. This concept was recognised by nearly all of the controls (94.1%) but only by a little more than half of the participants with AS/HFA (57.1%). More than a third of the participants in the AS/HFA group mislabelled a face item of this emotion as ‘offended’ and two thirds of them mislabelled a voice item as ‘battered’ (‘you’ve done it again’). Surprisingly, not many studies have studied recognition of disgust. Those which have, found difficulties in disgust recognition among severely autistic adolescents with mental retardation (Hobson, Ouston, & Lee, 1988a) and no difficulty in its recognition among high-functioning adults with ASC (Adolphs, Sears, & Piven, 2001). The group difference found in the recognition of faces and voices expressing *distaste* (which is slightly more subtle than *disgust*) suggests that even high-functioning individuals with ASC might have problems recognising this in faces in motion and in voices.

Interestingly, there was no difference between the AS/HFA and control groups in the recognition of empathy (the concept *empathic*), and this was not predicted. The reason for this might have been the easy foils which appeared with this concept. Equally surprising was the lack of difference in recognising mental states such as *appalled*, *vibrant*, *confronted* and *subdued*. Since the item foils were selected by the quiz engine of *Mind Reading*, and since item analysis was conducted retrospectively, these items were not amended. This is possible limitation of the instrument was changed in the children version of the CAM, by involving judges in picking the foils, and by conducting item validation on a separate typically developing group of children in advance.

The CAM-A presents both visual and vocal stimuli, including motion in the face task and requiring word labels for the emotion. As these different kinds of stimuli activate different brain areas, one might wonder which of these underlies the difficulties of the ASC group. Such studies were mainly done with typically developed and with brain damaged participants. In a study assessing a patient with bilateral amygdala lesions, Adolphs and colleagues found that while the patient could not recognise the 6 'basic' emotions from still faces, he could easily recognise them when they were expressed by a live model. They argued that still and motion emotional stimuli activate different areas of the brain: while the fusiform gyrus, as well as the posterior and superior temporal cortex and the amygdala and insula are involved in recognising emotions from still faces, emotion recognition from faces in motion is related to middle temporal and middle superior temporal areas of the brain, together with parietal and frontal sectors (Adolphs, Tranel, & Damasio, 2003).

Kilts and colleagues conducted a PET imaging study assessing intensity of anger and happiness from still and video stimuli in participants from the general population. They found that judgment of anger in dynamic expressions was associated with increased right-lateralised activity in the medial, superior, middle, and inferior frontal cortex and cerebellum, while judgments of happiness were associated with relative activation of the cuneus, temporal cortex, and the middle, medial, and superior frontal cortex. In contrast, the perception of anger or happiness in static facial expressions activated a motor, prefrontal, and parietal cortical network (Kilts, Egan, Gideon, Ely, & Hoffman, 2003).

As with visual stimuli, there are different findings with regards to emotion recognition from the voice. In a study involving over a hundred participants with focal brain damage, Adolphs and colleagues found that lesions in right cortical regions impair recognition of emotion in prosody. Phillips and colleagues found in a functional MRI study that the amygdala, as well as the superior temporal gyrus, were involved with recognition of fear and disgust in both visual and vocal stimuli (Phillips et al., 1998). However, Adolphs and Tarnel found no difficulty in recognising emotions from prosody among amygdala lesioned patients (Adolphs & Tranel, 1999). Evidently, this calls for further inquiry of brain structures involved in emotion recognition from voices.

The above mentioned studies assessed emotion recognition using the 6 ‘basic’ emotions. Further investigation into activation of brain regions in recognition of complex emotions from still, dynamic and vocal stimuli, such as those that appear in the CAM-A is still required. Since this study was purely behavioural, no conclusions could be drawn of relevant brain regions involved in the recognition of the CAM-A’s different emotions. The CAM-A lends itself to studies investigating the neural basis underlying visual dynamic and vocal emotion recognition. In addition, it would be interesting to use the CAM-A with other clinical groups that involve ER or face processing difficulties (or strengths), such as schizophrenia, anti-social personality disorder, or prosopagnosia.

The Cambridge Mindreading Face-Voice Battery - child version

Introduction

The CAM-C is similar to the CAM-A in its structure and guiding principles. Since it was designed for 8-11 year old children, it includes age appropriate emotions. The battery tests the recognition of 15 different emotions, the 6 basic emotions and 9 complex emotions. Due to the inconclusive findings relating to basic ER in ASC (described in Chapter 2), this structure allows a comparison of basic and complex ER in faces and voices in children with ASC. The recognition of each of the 15 emotional concepts is tested through 6 items, 3 face and 3 voice items. The battery provides ER scores for faces and for voices, as well as for the number of emotions correctly recognised. These are also available separately for basic and complex emotions.

The CAM-C was tested with 8-11 years old children with AS/HFA and a typically developing matched control group. It was predicted that the AS/HFA group would have lower scores on the battery tasks compared to controls. As these are high functioning children with ASC, it was predicted their ER difficulties would mainly be found in the complex emotions and mental states, rather than in the basic emotions. As with the CAM-A, it was predicated that the more autistic traits the children possess, the lower their CAM-C scores will be.

Age was hypothesised to positively correlate with CAM-C scores since ER abilities tend to improve with age (Herba & Phillips, 2004). A positive correlation with IQ was also expected since previous studies have reported positive correlations between IQ and emotion and mental state understanding both in typical development and in ASC (Dyck, Ferguson, & Shochet, 2001; Hobson, 1986a). Lastly, correlations with child version of the ‘Reading the Mind in the Eyes’ (RME-C; Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001), were calculated, to validate the CAM-C battery, and test-retest reliability correlations were calculated for children with ASC in the no-intervention group.

Method

Participants

The AS/HFA group comprised 30 children (29 boys and 1 girl), aged 8.2-11.8. Participants had all been diagnosed with AS/HFA in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). They were recruited from a volunteer database and a local clinic for children with ASC. A control group from the general population was matched to the clinical group. It comprised 25 children (24 boys and 1 girl), aged 8.2-12.1. They were recruited from a local primary school. Parents and school reports confirmed that none of the children in this group had a psychiatric diagnosis or special educational needs, and none of them had family members diagnosed with ASC. All participants were given the WASI, and scored above 80 on both verbal and performance scales. This was used to confirm that none had learning difficulties. To screen for autism spectrum conditions, participants' parents filled in the Childhood Asperger Syndrome Test (CAST; Scott, Baron-Cohen, Bolton, & Brayne, 2002). None of the control participants scored above the cut-off point of 15. All but two participants in the AS/HFA group scored above cut-off. These two participants scored below cut-off due to several unanswered items, and were therefore not excluded from the sample. The AS/HFA group CAST scores (M=19.7, SD=4.3) were significantly higher than those of the control group (M=3.4, SD=1.7; $t[53]=18.33$, $p<.001$). The two groups were matched on sex ($\chi^2[1]=0.02$, n.s.), age, and IQ. The groups' background data appears in Table 6.

	AS/HFA group (n=30)			Control group (n=25)			
	Mean	SD	Range	Mean	SD	Range	t (53)[1]
Age	9.7	1.2	8.2-11.8	10.0	1.1	8.2-12.1	0.95
Verbal IQ	112.9	12.9	88-143	114.0	12.3	88-138	0.32
Performance IQ	111.0	15.3	84-141	112.0	13.3	91-134	0.27
Full Scale IQ	113.5	11.8	96-138	114.8	11.9	95-140	0.39

[1]p>.0.1 for all t-tests

Table 6: Age and IQ measures of the AS/HFA and control groups – CAM-C study

The CAM-C: Test development

Fifteen emotional concepts were selected for the CAM-C: Basic ER skills were tested using facial and vocal expressions of the six basic emotions. In addition, nine complex emotions were selected from *Mind Reading.* The selected concepts included emotions that are developmentally significant (e.g. *jealous, embarrassed*), subtle variations of basic emotions (e.g. *amused, nervous*), and emotions and mental states that are important for everyday social functioning (e.g. *unfriendly, loving, disappointed*). Since the evaluation study included 8-11 years old children, most of the emotional concepts were selected if at least 75% of 8-9 year olds in the emotional lexicon developmental survey were familiar with them (according to parental report). This was significantly above chance ($p<.05$, binomial test). However, a few of the selected concepts met this criterion only for 10-11 year olds. This was done in order to prevent ceiling effects and to allow for improvement following the use of *Mind Reading* in the intervention study. The list of the emotions included in the CAM-C appears in Table 7. As the table shows, all but 3 emotional concepts were known to more than 75% of 8-9 year olds in the emotional lexicon study. Two of the complex emotions, *amused* and *undecided* and (interestingly) the basic emotional concept *disgusted* were known to less than 75% of 8-9 year olds in the lexicon study. To overcome this vocabulary obstacle, all emotional concepts were defined and explained to children by the experimenters. Pilot testing revealed that when provided with a definition, or with an alternative wording of the concept (e.g. 'not decided' instead of 'undecided'), even 8 year olds understood the meaning of these concepts.

For each emotional concept, three face items and three voice items were created using 3-5 seconds silent video clips of facial expressions and audio clips of short verbalisations taken from *Mind Reading*. Where possible, presentation by child actors was preferred, to make the tasks more child-friendly. Three foils were set for each item, using *Mind Reading*'s quiz module, assuring they were from the same developmental level as the target emotion (i.e. known to a similar age group in the lexicon study). In order to prevent completely random selection of foils, which could have resulted in items that were too easy (as happened in some cases in the CAM-A), the task questions were reviewed, ensuring that at least one foil in each item was closer to the target item. Foils for vocal items were selected so that they could match the verbal content of the scene, if the intonation was not taken into account. This resulted in potentially harder tasks than those of the CAM-A. The labels and foils were then reviewed by two independent judges to ensure none of them was too similar to the target emotion.

	Concept	Emotion Group	Known to 8-9 year olds		Known to 10-11 year olds	
			Percent	Total N	Percent	Total N
Basic	Happy	Happy	96%	(27)	100%	(34)
	Sad	Sad	100%	(28)	100%	(34)
	Afraid	Afraid	100%	(18)	100%	(20)
	Angry	Angry	94%	(18)	95%	(20)
	Surprised	Surprised	93%	(28)	94%	(34)
	Disgusted	Disgusted	61%	(18)	90%	(20)
Complex	Loving	Fond	96%	(28)	100%	(34)
	Embarrassed	Sorry	94%	(18)	94%	(18)
	Disappointed	Sad	89%	(18)	100%	(20)
	Jealous	Wanting	89%	(27)	88%	(34)
	Nervous	Afraid	79%	(28)	97%	(34)
	Unfriendly	Unfriendly	79%	(28)	94%	(34)
	Bothered	Bothered	78%	(18)	85%	(20)
	Amused	Happy	56%	(18)	90%	(20)
	Undecided	Unsure	46%	(28)	94%	(34)

Table 7: The 15 emotional concepts included in the CAM-C

Two tasks, one for face recognition and one for voice recognition were created using DMDX experimental software (Forster & Forster, 2003). Each task started with an instruction slide, asking participants to choose the answer that best describes how the person in each clip is feeling. The instructions were followed by two practice items, to ensure understanding of the task. In the face task, four adjectives, numbered from 1 to 4, were presented after playing each clip. Items were played in a random order. An example question, showing one frame from one of the clips is shown in Figure 2. In the voice task, the four numbered answers were presented before and while each item was played, to prevent overloading participants' working memory. This prevented randomising item order in the voice task. Instead, two versions of the task were created, with reversed order, to avoid an order effect. A handout with definitions for all the emotion words used in the tasks was prepared.

1. Ashamed 2. Ignoring 3. Jealous 4. Bored

Figure 2: An example of a question from the CAM-C face task (showing only one frame out of the full clip)

The tasks were then piloted with 16 children – two girls and two boys from 4 age groups – 8,9,10 and 11 years of age. Children were randomly selected from a local mainstream school. The tasks were played to them on two laptop computers, using headphones for the voice task. To avoid confounds due to reading difficulties, the experimenter read the instructions and possible answers to the children and made sure they were familiar with all the words, using the definition handout, where necessary. Participants were then asked to press a number from 1 to 4, to choose their answer. After choosing an answer the next item was presented. No feedback was given during the task.

Next, item analysis was carried out. Items were included if the target answer was picked by at least half of the participants, and if no foil was selected by more than a third of the participants ($p<.05$, Binomial test). Items which failed to meet these criteria were matched with new foils and played to a different group of 16 children, until they all met criteria. The final task included 45 items in the face task and 45 in the voice task, representing the 15 emotional concepts.

Procedure

Participants with AS/HFA were tested individually at the Autism Research Centre in Cambridge. Controls were tested individually in a quiet room at a local school. The final version of the tasks was presented to the participants on a laptop with a 15 inch screen, using DMDX experimental software. Headphones were provided for the voice task. The experimenter read the instructions and the questions and answers for all items with the participants, and checked that they were familiar with all the possible answers. If needed, a definition handout was used to familiarise participants with word meanings among the possible answers. Hence, there was no time limit to answer each item. Task order was randomised, as was the version of the voice task used. Participants were allowed two short breaks during each task. Completion of the whole battery took about forty five minutes, including the breaks. The RME task was completed on the same meeting and took about fifteen minutes to complete. Participants' parents filled in the CAST in advance and brought it with them to the assessment.

Results

Facial and vocal scores were calculated by counting the number of correct answers (out of a maximum of 45) in each of the scales. Facial and vocal scores were also calculated separately for basic emotions (max score=18) and for complex emotions (max score=27). The number of emotional concepts correctly recognised was calculated by tallying the number of concepts in which at least 4 out of the concept's 6 items were answered correctly ($p<.05$, Binomial test). All participants scored above chance (i.e. above 16, $p < 0.05$, Binomial test) on the face scale. All but one participant from the AS/HFA group scored above chance on the voice scale. This participant scored just below the threshold (15 correct answers out of 45) and was therefore included in the analysis. There were no ceiling effects. One sample Kolmogorov-Smirnov tests were conducted for all task scores in the two groups. Distributions of all scores in both groups did not differ from normal, with the exception of the number of basic emotional concepts passed by the control group (K-S $Z=1.59$, $p<.05$). Hence, parametric analysis was not used when with this score.

In order to check for group differences on ER according to complexity (basic/complex), two multivariate analyses of variance with repeated measures were conducted on the participants' face and voice task scores. Since the CAM-C has a different number of basic and complex emotion items, proportions of correct answers were used. In both analyses group was entered as an independent variable, complexity as a repeated measure factor, and age, verbal IQ and performance IQ as covariates.

The face task analysis yielded a significant main effect (beyond complexity) for group ($F[1,50]=9.36$, $p<.005$), with significantly worse performance in the AS/HFA group, compared to the control group. A significant main effect was also found for the covariate age ($F[1,50]=12.87$, $p<.005$), suggesting with older participants scoring higher on the face task beyond group. In addition, a significant group by complexity interaction was found ($F_{Wilks}[1,50]=5.69$, $p<.05$). Simple main effect analysis for this interaction revealed no group difference on basic emotion face scores ($t[53]=1.73$, n.s.), and a significant group difference on complex emotion face scores ($t[53]=4.07$, $p<.001$), with the AS/HFA group performing significantly worse than the control group.

In the voice task analysis, a significant main effect (beyond complexity) was found for group ($F[1,50]=5.12$, $p<.05$), again, with the AS/HFA performing worse than the control group. Significant main effects were also found for the covariates age ($F[1,50]=16.00$, $p<.001$) and verbal IQ ($F[1,50]=5.75$, $p<.05$), suggesting older participants, and participants with higher verbal IQ perform better on this task. In addition to these effects, a significant group by complexity interaction was found ($F_{Wilks}[1,50]=8.93$, $p<.005$), with no group difference on basic emotion voice scores ($t[53]=0.52$, n.s.), and significantly lower complex emotions voice score in the AS/HFA group compared to the control group ($t[53]=3.93$, $p<.001$). In addition, a verbal IQ by complexity interaction was found significant ($F_{Wilks}[1,50]=4.93$, $p<.05$). Verbal IQ was positively correlated with complex emotion voices ($r=.42$, $p<.005$), but not with basic emotion voices ($r=.13$, n.s.).

To check whether group differences exist in the proportion of basic and complex emotions recognised, two Mann-Whitney tests were conducted. A significant group difference was found for the proportion of emotions correctly recognised, beyond complexity (Z=3.28, p<.005), and for the proportion of complex emotions correctly recognised (Z=3.67, p<.001), with the AS/HFA group performing significantly lower than the control group. The group effect for the proportion of basic emotions recognised did not reach significance (Z=1.9, n.s.). The average proportion scores of the different tasks in the two groups are presented in Table 8 and in Figure 3.

a. Face scale

		AS/ HFA	Control	t (53)
Basic (% of 18)	M (SD)	.71 (.14)	.78 (.15)	1.73
Complex (% of 27)	M (SD)	.55 (.15)	.71 (.14)	4.07**
Total (% of 45)	M (SD)	.62 (.13)	.74 (.13)	3.47**

b. Voice scale

		AS/ HFA	Control	t (53)
Basic (% of 18)	M (SD)	.66 (.15)	.68 (.16)	0.52
Complex (% of 27)	M (SD)	.61 (.13)	.75 (.13)	3.93**
Total (% of 45)	M (SD)	.63 (.12)	.72 (.12)	2.83**

c. Proportion of concepts correctly recognised

		AS/ HFA	Control	M-W Z
Basic (% of 6)	M (SD)	.72 (.21)	.80 (.23)	1.9
Complex (% of 9)	M SD)	.51 (.19)	.73 (.22)	3.67**
Total (% of 15)	M (SD)	.59 (.17)	.76 (.20)	3.28**

Table 8: Proportion of correct answers on the CAM-C task scores for basic and complex concepts

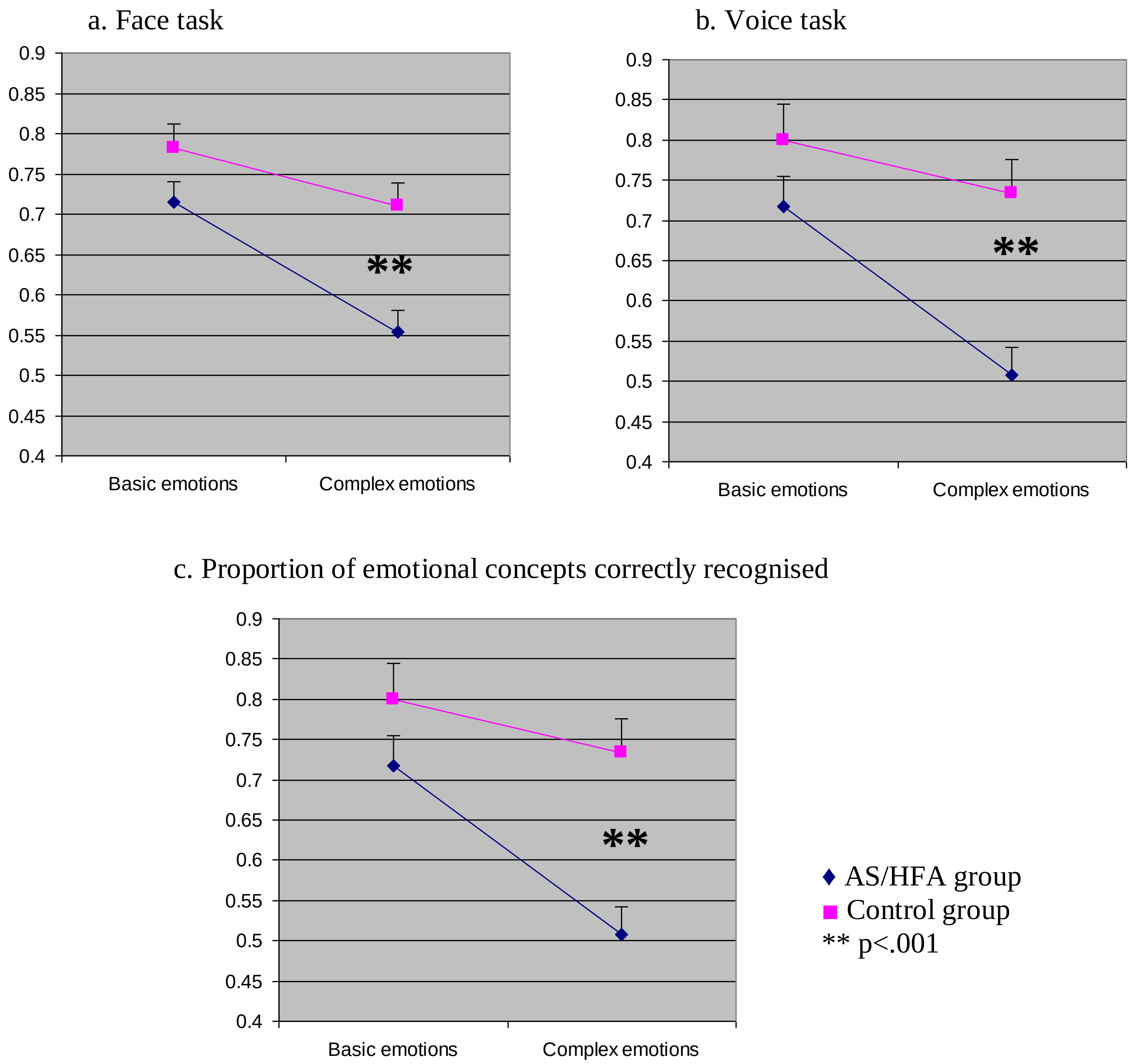


Figure 3: Average proportions (with standard error bars) of correct answers on the CAM-C tasks for basic and complex concepts

From the graph of the voice task, it appears that unlike the decrease in performance of the AS/HFA group from basic to complex emotions (t[29]=2.23, p<.05), there appears to be an increase in the performance of the control group from basic to complex. However, this apparent increase is not statistically significant (t[24]=2.0, n.s.). This lack of difference in performance between basic and complex emotions in the control group, and the relatively small drop in performance from basic to complex voices in the AS/HFA group could be related to the harder foils attributed to the voice task. This will be discussed in more detail later.

In order to compare the recognition of individual emotional concepts in the two groups, a 15 by 2 multivariate analysis of covariance was performed for the proportions of participants who correctly recognised each concept in the two groups. Age and verbal IQ were included as covariates. Since performance IQ did not have a significant effect in any of the above analyses, it was excluded from the MANCOVA. The analysis yielded significant overall effects of group ($F_{wilks}[15,37]=2.58$, $p<.01$) and age ($F_{wilks}[15,37]=2.61$, $p<.01$). Individual concept analyses revealed that the AS/HFA group scored significantly lower than the control group in the recognition of one basic emotion only: *disgusted* ($F[1,51]=4.71$, $p<.05$). However, out of nine complex emotions, the AS/HFA group scored significantly lower than the controls on six concepts: *unfriendly* ($F[1,51]=4.21$, $p<.05$), *disappointed* ($F[1,51]=5.13$, $p<.05$), *jealous* ($F[1,51]=4.95$, $p<.05$), *nervous* ($F[1,51]=4.95$, $p<.05$), *bothered* ($F[1,51]=5.54$, $p<.05$), and *amused* ($F[1,51]=4.95$, $p<.05$). Table 9 shows proportions of participants of the two groups who successfully recognised each of the 15 concepts.

Power calculations for the scales (with $\alpha=0.01$) show that the scales are not very powerful in differentiating the group with AS/HFA from the control group ($1-\beta=0.861$ for the face scale, 0.657 for the voice scale, 0.82 for the number of concepts recognised, and 0.874 overall). However, the power of the complex emotions is higher and distinguishes well between the groups: $1-\beta=0.951$ for complex emotion faces, 0.923 for voices, 0.949 for complex concepts recognised, and 0.986 for all complex emotion items. Test-retest correlations, calculated for children with AS/HFA who participated in the no-intervention control group of experiment 3 and took the tasks before and after the 10-15 week period were $r=0.79$ for the face scale, $r=0.75$ for the voice scale, and $r=0.77$ for the number of concepts recognised. Test-retest correlations for complex emotion items only were $r=0.74$ for complex faces and $r=0.76$ for complex voices ($p<.001$ for all).

Due to the difference in effects between basic and complex emotions, correlation analysis was conducted for simple and complex emotion sub scores, rather than for the full scale scores. The analysis, presented in Table 10, shows the hypothesised negative correlation between CAST scores and ER scores was significant for complex emotions only. Age, however, was positively correlated with face and voice scores, and with number of concepts, on both basic and complex emotions, suggesting the older the participants, the higher their emotions recognition scores. As shown above, verbal IQ was positively correlated only with vocal complex emotion scores and number of complex emotions recognised, suggesting these scores require higher level of language skills. Performance IQ was unrelated to any of the scales.

	Concept	AS/HFA group (n=30)	Control group (n=25)	F (1,51)
Basic	Happy	73%	68%	.20
	Sad	83%	80%	.12
	Afraid	60%	68%	.06
	Angry	77%	84%	.16
	Surprised	77%	92%	2.36
	Disgusted	60%	88%	4.71*
Complex	Loving	73%	72%	.01
	Embarrassed	33%	44%	.30
	Disappointed	53%	84%	5.13*
	Jealous	60%	88%	4.95*
	Nervous	40%	72%	4.95*
	Unfriendly	30%	60%	4.21*
	Bothered	53%	84%	5.54*
	Amused	40%	72%	6.19*
	Undecided	73%	84%	.41

* p<.05

Table 9: Proportions of participants who correctly recognised the CAM-C concepts

		CAST	Age	VIQ	PIQ	RME
Face Scale	Basic	-.28	.32*	.23	.08	.28*
	Complex	-.54**	.53**	.21	.04	.35**
Voice Scale	Basic	-.15	.41**	.13	.00	.17
	Complex	-.48**	.46**	.42**	.00	.40**
Number of concepts recognised	Basic	-.28	.35**	.22	.07	.21
	Complex	-.53**	.57**	.35**	.08	.36**

**p<.01 *p<.05

Table 10: Correlations of the CAM-C scores with background parameters and with an external criterion

Some significant correlations were found between the CAM-C sub scales and the child version of the RME: Being a complex emotion and mental state recognition task from (parts of) faces, RME scores correlated positively with all complex ER scores: faces, voices and number of concepts. Among basic ER scores, the RME correlated significantly only with the face task score.

Discussion

The CAM-C is a new battery for testing recognition of basic and complex emotions and mental states in the face and the voice. Like the CAM-A, it can test the recognition of specific emotions and mental states, in two separate perceptual channels. It also enables one to asses the recognition of basic and complex emotions and mental sates.

Testing of this battery with 8-11 year old children with ASC and matched controls from the general population showed that high functioning children with ASC had more difficulties than controls in recognising complex emotions and mental states from faces and voices, even when controlling for age and verbal IQ. However, they had no such difficulties in recognising basic emotions. In addition, children with ASC recognised fewer mental state concepts then controls. In six out of the nine complex emotions tested in the CAM-C, significantly fewer participants with ASC successfully recognised the concept, compared to controls. Such a difference was found only on one out of the six basic emotions.

These results support previous findings of intact basic ER in ASC, with difficulties in more complex emotions (Baron-Cohen, Spitz, & Cross, 1993; Capps, Yirmiya, & Sigman, 1992; Happe, 1994a; Yirmiya, Sigman, Kasari, & Mundy, 1992). The lack of a group difference in basic ER could be related to the age group tested: Typically developing children recognise most of the basic emotions by 3 years of age. Children with ASC, despite a possible delay in development of ER skills, may have compensated for this deficit by the time they are 8 years old. Similarly, high functioning children with ASC by this age can pass first-order false belief tasks, though younger and lower functioning children with autism struggle with these tasks

(Frith, 2003). Past research has shown children with ASC use their verbal skills and good visual memory to compensate for basic ER deficits (Grossman, Klin, Carter, & Volkmar, 2000), which could explain the lack of a group effect found here.

The positive correlations of all task scores with age, independent of diagnosis, suggest that ER skills continue developing throughout childhood and early adolescence in both typically developing children and children with ASC. As predicted, complex emotion voice task scores were positively correlated with verbal ability. This correlation, reported before, was also found in other complex ER measures developed for this study, which involve speech. These will be described in Appendices 2 and 3. It may be related to the need for integration of content and intonation, which may depend to some extent on verbal ability. Contrary to the prediction, the face task scores and basic emotion voice scores were unrelated to verbal IQ, suggesting the ability to recognise emotions in faces in motion less dependent on verbal ability. This was also found on the CAM-A.

As predicted, CAM-C scores were negatively correlated with the level of autism spectrum symptoms participants possessed. However, this was only found for complex emotions, possibly because of the narrow range of basic emotion scores in both groups. In addition, since the range of CAST scores was quite narrow in both groups, all correlations with level of autism spectrum features were potentially lower than they could be with a wider range of CAST scores. In order to examine the correlations between autism spectrum features and CAM-C scores, a better representation of the entire continuum between typical development and autism would be desirable. Assessing siblings of children with an ASC diagnosis, who have been found to have the broader autism phenotype (Hughes, Plumet, & Leboyer, 1999; Piven, Palmer, Jacobi, Childress, & Arndt, 1997; Yirmiya, Shaked, & Erel, 2001) may allow to better test these correlations.

The positive correlations complex ER scores with the RME task provide the CAM-C with important measures of external validity. These correlations were significant but not very strong (.35-.40), suggesting they may test different aspects of a common skill. The lack of correlation with basic emotion voice task scores may be related to the RME being too different from it, both in modality and in complexity. Test-retest reliability measures for the different task scores suggest good consistency. Power levels of the CAM-C show it is sensitive to group differences across all scales and scores for complex emotions, but less so for basic ones. Testing a younger or lower functioning ASC group with the basic emotion part of the CAM will allow a better test of whether it can distinguish children with ASC from controls. In addition, the CAM-C should be used with other clinical groups, to test its power in distinguishing different clinical conditions from each other and from typical development.

A more detailed look at the individual emotional concepts of the CAM-C reveals that the only basic emotion for which a group difference was found is *disgusted*. This emotion was recognised by 88% all the controls but only 60% of the participants with ASC. A common error in the ASC group was the mislabelling of disgusted faces as *afraid* or *moody*. The mislabelling of disgust as fear could be related to both involving moving the face backwards, away from the scary/disgusting object. This kind of error demonstrates the importance of motion in the face task items, as this would not have been picked from still pictures. Voice items uttered in a disgusted intonation were mistaken for *grumpy* (‘uh, get them out of my sight’) or *hurt* (‘oh, how nasty!’). This suggests that the children with ASC were able to tell the valence of the faces and the voices, but mistook them as expressing other same-valence emotions. In the voice items, it appears that the children used the verbal content of the items, but disregarded the intonation. This reliance on content is characteristic of younger children (Morton & Trehub, 2001), and is often used in ASC to compensate for poor perception of intonation. Interestingly, the adults with ASC in the CAM-A study found it hard to recognise *distaste*, which is related to disgust. Other studies have found inconclusive findings in relation to recognition of disgust (Castelli, 2005; Hobson, Ouston, & Lee, 1988a). Brain imaging studies of patients with insula lesions report specific difficulties with the recognition and experience of disgust, suggesting brain specificity for this emotion (Calder, Keane, Manes, Antoun, & Young, 2000; Calder, Lawrence, & Young, 2001). Children with autism have been found to have decreased regional cerebral blood flow bilaterally in the insula (Ohnishi et al., 2000), which may account for disgust recognition difficulties in ASC.

Typically developing children were found to understand and recognise complex emotions such as *jealous*, *embarrassed* and *disappointed* between the ages of 7-10

(Harris, 1994; Herba & Phillips, 2004). Indeed our findings show that more than 80% of the control group recognised jealousy and disappointment successfully, with the AS/HFA group performing significantly below this level: 60% of the AS/HFA group recognised the concept *jealous*. Common errors included mislabelling facial expressions as *disappointed*, possibly focusing on the mouth area looking unhappy. Relying on the mouth area for ER, while disregarding other areas (such as the eyes) is characteristic of ASC (Baron-Cohen, Wheelwright, & Jolliffe, 1997; Klin, Jones, Schultz, Volkmar, & Cohen, 2002b). Whereas this may sometimes suffice when interpreting basic emotions (e.g. happy or sad), other configural cues are required for recognition of complex emotions like *jealous*. Voice items for this concept were mislabelled as *teasing* ('I can do better than you'), or *bossy* ('I deserve that car more than him'), relying on the content rather than combining linguistic and paralinguistic components of the verbalisations.

A group difference was found for *disappointed* too, with 53% of the AS/HFA group correctly recognising this concept, compared to 84% of the controls. Common errors for the face items included labelling this as *thinking* and *unsure*, possibly due to the gaze being directed downwards, away from the camera. Participants may have failed to integrate this cue with the unhappy mouth expression. *Disappointed* voice items led to common errors such as mislabelling it as *ashamed* ('I should have won') and *hurt* ('I tried so hard').

Interestingly, no group difference was found for the recognition of *embarrassed*. Though a larger proportion of controls (44%) recognised this emotion, compared to the AS/HFA group (33%), this difference was not significant. Common errors for the embarrassed face items in both groups included *sad* and *jealous*. Embarrassed voice items were mislabelled as *disliking* ('I know my shoes are old'), *afraid*, and *asking* ('Do you think anyone saw me?'), and *wishful* ('oh, I wish it hadn't happened'). The lack of difference on this concept and the low scores in both groups could be related to the age range of the sample, the younger participants finding recognition of embarrassment from context-less facial expressions too challenging. It is also common among young typically developing children to rely on content rather than intonation (Morton & Trehub, 2001), as found on the *embarrassed* voice items.

Though the AS/HFA group did not find most basic emotions hard to recognise, they had difficulties with more subtle or complex derivatives of the basic emotions. For example, only 40% of the AS/HFA group successfully recognised the concept *nervous*, a subtle expression of fear. This is similar to the difficulties adults with ASC had with recognition of *uneasy* in the CAM-A study. Common errors the AS/HFA group made when labelling *nervous* were mislabelling a face item as *annoyed* and voice items as *disgusted* ('Don't put that near me'), or *asking* ('How many people are out there?'). These examples show again how intonation is disregarded and verbal content is used to recognise the speaker's emotion/mental state. The lack of differential response in the brain of individuals with ASC to expressions of subtle fear (Ashwin, Baron-Cohen, Wheelwright, O'Riordan, & Bullmore, in press) could explain this difference between successful recognition of *afraid* and the difficulties recognising *nervous*.

Another example of the difficulties with a more complex expression of a basic emotion is the group difference on recognition of *amused*, which was correctly recognised by 40% of the AS/HFA group. Amused faces were mislabelled as *interested*, *keen*, or *cheeky*, possibly because of misinterpretation of smiles. Amusement in the voice was confused with *excitement* ('Imagine that'). This may again reflect the need to process both mouth and eye regions of the face for successful recognition of complex emotions.

Only 30% of the participants with ASC correctly recognised the concept *unfriendly*. Despite being familiar with the verbal label, the AS/HFA group mislabelled unfriendly faces with a variety of labels, including *afraid*, *disgusted*, and *shy*. These errors were probably related to the actors moving their faces away from the camera and looking sideways. Again, this kind of error is unique to video stimuli. The misinterpretation of unfriendliness could have undesirable social consequences. Failure to recognise negative emotions and mental states in others such as *unfriendly*¸ *nervous*, *bothered* or *disappointed*, may lead to a failure to empathise with others in social interaction (Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002), which in turn may lead to social isolation. Interestingly, there was no group difference in the recognition of the positive emotion *loving*, or the 'neutral' mental state *undecided*.

This is in line with past research of specific insensitivity to others' negative emotions among children with autism (Sigman, Kasari, Kwon, & Yirmiya, 1992), and is also common among individuals with amygdala damage (Adolphs, Sears, & Piven, 2001). Theses two similarities between the conditions (specific difficulties with negative emotions, and greater difficulties with complex emotions than with basic ones) support the amygdala theory of autism (Baron-Cohen et al., 2000; Howard et al., 2000).

One possible limitation of the CAM-C is the relatively low proportions of children (in both groups) who recognised the basic emotions. As it was mentioned in the results section, this was especially marked in the voice items of the basic emotions. These lower success rates could be attributed to the creation of harder items for the CAM-C. Unlike the CAM-A, in the CAM-C foils were selectively picked, with the intention of making the task harder by including foils that match the verbal content but not the intonation. This probably made the basic emotion items even harder, as their foils had the same valence and were often more complex then the target answer (e.g. *adventurous* was one of the foils for *happy*). However, whereas this may not reliably represent the balance between basic and complex ER, it could actually help with the intervention study of *Mind Reading*, since it leaves some room for improvement even on the more basic emotions.

The CAM-A and the CAM-C were used for the close generalisation level in the intervention study. The next chapter describes the feature based distant generalisation measures used in the adult and child experiments.

Appendix 2 – Creation of distant vocal generalisation measures

As described in Chapter 4, the second level of generalisation in the intervention study was assessed using stimuli not included in *Mind Reading* and that the participants had not been exposed to before.

Adult auditory task - Reading the mind in the voice-Revised

Introduction

As described in Chapter 2, 'Reading the Mind in the Voice' (Rutherford, Baron-Cohen, & Wheelwright, 2002) is one of the only complex ER tasks using purely vocal stimuli. The original task includes 40 unrelated segments of speech, taken from different BBC dramas. After listening to each segment, participants are asked to judge the speaker's mental state by choosing one of two optional answers. A significant difference on the number of items correctly answered was reported between a group of high functioning adults with ASC and a group of matched controls from the general population. However, pilot testing with this task, with high functioning participants with ASC prior to the intervention study, resulted in some ceiling effects. In addition, Rutherford et al suggested that the task's sensitivity could have been improved by adding foils to each item, and that some of the items did not differentiate between the ASC group and the control groups. By removing these items, the task could have been more sensitive, as well as briefer to administer. These suggested amendments were implemented in the evaluation study, thus constituting a revised version of the 'Reading the Mind in the Voice' task. The revised task (RMV-R) was tested on a group of adults with AS/HFA, and matched typically developed controls.

Based on the results of the original study, and the findings reported for the voice scale of the CAM-A, it was predicted that task scores in the AS/HFA group would be significantly lower than those of the control group. Since emotion recognition difficulties are characteristic of ASC across perceptual modalities (Hobson, 1993), RMV-R scores were predicted to correlate positively with scores on the ‘Reading the Mind in the Eyes’ task (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Based on the findings reported in Appendix 1 with the CAM-A and CAM-C, a negative correlation between RMV-R scores and level of autistic traits, tested using the AQ, was predicted.

Both Rutherford et al (2002) and Baron-Cohen et al (1997, 2001) reported no correlation between RMV or RME scores and IQ. No such correlation was found in the CAM-A study, suggesting that recognition of complex emotions and mental states in adults is independent of intellectual ability. Hence, it was predicted that no such correlation would be found for the RMV-R.

Based on the ‘extreme male brain’ theory of autism, and the reports of female superiority in empathy and emotion recognition tests (Baron-Cohen, 2002, 2003; Baron-Cohen & Wheelwright, 2004; Lawson, Baron-Cohen, & Wheelwright, 2004), one would predict that typical females should perform better than typical males on our task, who in turn should perform better than individuals with ASC. However, no studies have tested sex differences in emotion recognition among individuals with ASC. These were examined on the RMV-R task scores. Finally, the power, reliability and validity of the revised task were calculated and compared with those of the original task.

Method

Participants

The AS/HFA group comprised twenty nine adults (22 males and 7 females), aged 17-50 (M=27.5, SD=8.5). Participants had all been diagnosed in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). They were recruited from several sources, including a local clinic for adults with ASC, support organisations, and colleges for individuals with ASC around the UK. All participants were given the Wechsler Abbreviated Scale of Intelligence (WASI), and scored above 80 on both verbal and performance scales.

The control group comprised twenty-six adults recruited from a local employment agency. After screening for autistic spectrum conditions using the AQ (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), four participants were excluded because they scored above cut-off of 32. The remaining twenty two, 17 males and 5 females, matched the clinical group in age and IQ. They spanned an equivalent range of employment and educational levels as that seen in the clinical group. Goodness of fit test for sex showed no group difference ($\chi^2[1]=.014$, n.s.). As shown in Table 1, t-tests for age, verbal, performance and full scale IQ revealed no significant differences between the groups at the $p<0.05$ level. The AS/HFA groups' AQ scores (M=34.48, SD=9.51) were significantly higher than those of the control group (M=13.59, SD=5.80; $t[47.08]=9.70$, $p<.001$). Table 1 presents the background data of the groups.

	AS/HFA group (n=29)			Control group (n=22)			
	Mean	SD	Range	Mean	SD	Range	t (49)[1]
Age	26.38	8.99	17.9-49.9	24.30	7.71	17.6-51.2	.87
Verbal IQ	112.07	13.71	84-132	114.45	13.49	86-138	.62
Performance IQ	111.86	12.68	88-135	111.50	8.44	92-128	.12
Full Scale IQ	113.45	12.55	92-138	114.45	9.94	97-138	.31

[1] p>.1 for all tests. Equal variances were not assumed for Performance IQ, df=48.29.

Table 1: Age and IQ scores of the AS/HFA and control groups, RMV-R study

Task development

As the original task created by Rutherford et al was recorded on audio cassettes, the first step was to sample the forty items onto a computer and digitally 'clean' tape recording noise. Two items which were completely distorted were excluded at this point. Next, two more foils for each item were chosen. Foils were selected to match the content of the verbalisations but not the intonation, thus making the task harder. For example, the verbalisation "Yeah, well, I know nothing about that", for which the original task two answers were 'Defensive' (correct) and 'Joking' (incorrect), had two foils added: 'Unconcerned' and 'Indignant'.

Age data from the emotional lexicon developmental survey was used to keep word difficulty similar to the existing target and foil of each item. In this way, no foil involved a much easier or harder word than others. Next, items were independently reviewed by two other authors. Five items were removed at this point, either because of disagreement between the authors, or because the correct answers were judged as inappropriate for the utterances. A handout with definitions for all the target and foil words in the items included was prepared for participants' use before and during the task.

The revised items were then played to 15 adults (7 men and 8 women) selected at random from the general population. The items and answers were played in random order on a computer, using DMDX experimental software (Forster & Forster, 2003). An item analysis was then carried out. Items were included if the target answer was picked by at least half (i.e. at least 8) of the 15 participants, and if none of the foils was selected by more than a third (i.e. more than 5) of the participants ($p<.05$, Binomial test). Eight more items were excluded after failing to meet these criteria. Hence, the final task comprised 25 items. Table 2 presents the list of emotions tested in the revised task, and the proportion of 17-18 year olds who were familiar with these in the emotional lexicon survey. Target emotion words in all 25 items were known to at least 75% of 17-18 year olds in the emotional lexicon survey, suggesting that the task is appropriate for adults. Target words for two items were not included in the developmental lexicon survey. The use of a definition handout ensured that recognition of these emotions is not limited by verbal understanding.

Procedure

The final version of the task was presented to participants on a computer screen, using DMDX experimental software. Headphones were given, to improve perception. An instructions slide and two practice items preceded the task. The definition handout was given to participants in advance and was available for participants during the task. Hence, there was no time limit on answering. The items were presented in a random order. For each item, the four possible answers were presented on the screen, followed by the utterance, which was played only once. Participants were then asked to press a number from 1 to 4, to choose the answer which "best describes how the speaker is feeling”. The computer played the next item 500 milliseconds after an answer was selected, to allow participants to prepare for the next item. Task score ranged from 0-25.

Item # in original task	Emotion/ Mental state	Known to 17-18 year olds (n=50)
1	Earnest	90%
2	Friendly	100%
3	Confused	100%
4	Suspicious	100%
5	Worried	100%
8	Apologetic	100%
9	Pleading	100%
10	Perplexed	76%
11	Nervous	100%
12	Irritated	98%
14	Joyous	100%
15	Embarrassed	100%
16	Terrified	100%
17	Enraged	94%
18	Disappointed	100%
20	Sincere	98%
21	Melancholy	[1]
24	Concerned	100%
25	Sincere	98%
27	Derogatory	[1]
28	Stern	98%
30	Defensive	100%
32	Insulted	100%
33	Resigned	94%
37	Hopeful	100%

Table 2: The target emotions covered in the RMV-R

In addition, participants took the RME-A and filled in the AQ (both described in Chapter 6). Participants were tested at the first meeting of the adult intervention study, either at the Autism Research Centre in Cambridge, or at local support centres and

[1] These two emotion words from Rutherford et al’s original task were not included in the emotional lexicon survey.

colleges for individuals with ASC. Participants were seated in front of IBM compatible computers with 15 inch monitors and were given headphones for the RMV-R. Definition handouts for the voice and eyes tasks were given to participants. Administration of the RMV-R and the RME-A took about twenty minutes each. The two tasks were administered in a random order, followed by the WASI. Participants filled in the AQ at home and brought it with them to the assessment.

Results

RMV-R task score was calculated by counting the number of correct answers for each participant. All the participants in the control group and all but four of the participants in the AS/HFA group scored above chance (i.e. above 10, $p<.05$, Binomial test) on the RMV-R. One sample Kolmogorov-Smirnov tests were conducted for task scores in the two groups. Distributions in both groups did not differ from normal. Hence, parametric analysis was used. A univariate analysis of covariance (ANCOVA) was conducted for RMV-R scores, with group (ASC/control) and sex as independent variables. Verbal IQ, performance IQ, and age were entered as covariates.

As hypothesised, a significant main effect of group was found for RMV-R scores ($F[1,44]=39.13$, $p<.001$). The scores of the AS/HFA group ($M=13.76$, $SD=3.39$) were significantly lower than those of the control group ($M=18.77$, $SD=2.41$). No main effect of sex was found ($F[1,44]=1.04$, n.s.), however a sex by group interaction was found significant ($F[1,44]=5.12$, $p<.05$). Post hoc comparisons revealed a sex difference in the AS/HFA group ($t[27]=2.07$, $p<.05$). Task scores of females with AS/HFA ($M=11.57$, $SD=3.74$) were significantly lower than those of males with AS/HFA ($M=14.45$, $SD=3.04$). However, when Bonferroni correction for multiple comparison was used, the sex difference in the AS/HFA group became non-significant ($p>.025$). No sex difference was found in the control group ($t[20]=1.09$, n.s.). The covariate verbal IQ had a marginally significant effect ($F[1,44]=3.85$, $p=.056$). Average task scores of males and females in the AS/HFA & control groups are presented in Table 3.

		Females	Males	Total
AS/HFA	M	11.57[1]	14.45[1]	13.76[2]
	(SD)	(3.74)	(3.04)	(3.39)
	n	7	22	29
Control	M	19.80	18.47	18.77[2]
	(SD)	(2.17)	(2.45)	(2.41)
	n	5	17	22
Total	M	15.00	16.21	
	(SD)	(5.22)	(3.42)	
	n	12	39	

[1] p<.05 [2] p<.001

Table 3: RMV-R average scores of males and females in the AS/HFA & control groups

Bivariate correlations were calculated for the RMV-R with RME-A, AQ, Verbal and Performance IQ. As predicted, RMV-R scores were positively correlated with RME-R scores, (r=0.49, p<.001), and negatively correlated with the AQ (r=-.66, p<.001). No significant correlations were found between RMV-R and verbal IQ (r=.26, n.s.), performance IQ (r=.06, n.s.), or age (r=.03, n.s.).

Power calculations for the RMV-R (2 tailed, α=0.01) revealed a power level of 1-β=0.99. As a comparison, the power level of the original task (two tailed analysis with α=0.01) was 1-β=0.85 for the college control group and 1-β=0.67 for the matched non-college control group (calculated using results reported by Rutherford et al, 2002). Therefore, the revised task is more powerful in distinguishing an AS/HFA group from controls. In general, the percentage of correct answers was lower compared to the original task. This was expected given that two more foils were added to each item. The number of items which 100% of the participants answered correctly decreased from 8 items to 1 in the control group, and from 4 items to 0 in the AS/HFA group, thus minimising ceiling effects. Table 4 presents the percentage of correct responses for each item in the ASC and matched control group for the current study and for Rutherford et al's original task.

Emotion	Revised Task		Original Task	
	AS/HFA	Control	AS/HFA	Control
Earnest	45%	73%	81%	75%
Friendly	41%	68%	74%	100%
Confused	86%	91%	84%	100%
Suspicious	66%	86%	79%	100%
Worried	69%	86%	79%	90%
Apologetic	83%	91%	74%	85%
Pleading	48%	55%	89%	95%
Perplexed	69%	91%	84%	85%
Nervous	31%	68%	68%	90%
Irritated	66%	86%	100%	100%
Joyous	38%	55%	74%	100%
Embarrassed	31%	45%	89%	100%
Terrified	66%	86%	74%	85%
Enraged	55%	45%	84%	95%
Disappointed	55%	100%	89%	90%
Sincere	55%	86%	79%	100%
Melancholy	52%	68%	79%	90%
Concerned	52%	64%	100%	95%
Sincere	59%	82%	63%	95%
Derogatory	28%	77%	84%	85%
Stern	45%	59%	84%	100%
Defensive	38%	73%	100%	95%
Insulted	72%	95%	89%	95%
Resigned	52%	68%	100%	80%
Hopeful	76%	77%	95%	95%

Table 4: Percentage of correct responses to the revised task items in the current and the original samples

Test-retest reliability, calculated for individuals with AS/HFA who participated in the no-intervention control group of experiment 1 and who took the tasks before and after the 10-15 week period, was r=0.8 (p<.001). In a discriminant analysis, the significant discriminant function (χ^2[25]=56.42, p<.001) successfully classified 96.1% of the participants (93.1% of participants with AS/HFA and 100% of controls) into their original groups.

Discussion

The RMV-R is an improved version of Rutherford et al's (2002) 'Reading the Mind in the Voice' task. The results of this evaluation show that individuals with ASC scored significantly lower on the revised task, compared to matched controls. The revised task was found to have good reliability and discriminative validity. It is more sensitive in differentiating the AS/HFA group and general population controls, in addition to being shorter and quicker to administer. Scores on it were positively correlated with scores on the RME-A, a visual mental state recognition test, and with scores on the face and voice tasks of the CAM-A, suggesting a common basis for emotion and mental state recognition abilities across perceptual domains. The RMV-R also correlated negatively with the AQ, suggesting the more autistic traits one has, the lower one's scores on the task. The task was independent of verbal or performance ability, or with the age of participants. These findings match previous reports of complex mental state recognition difficulties from vocal stimuli in ASC (Kleinman, Marciano, & Ault, 2001; Rutherford, Baron-Cohen, & Wheelwright, 2002), as well as the findings of ER difficulties on the voice scale of the CAM-A, reported in Appendix 1.

As described in Chapter 2, the auditory aspect of complex emotion and mental state recognition in ASC has not been extensively studied. Studies which have been conducted focus more on the perception of sound and voice in ASC (Alcantara, Weisblatt, Moore, & Bolton, 2004; Gervais et al., 2004; Gomot, Giard, Adrien, Barthelemy, & Bruneau, 2002), and the recognition of basic emotions (Boucher, Lewis, & Collis, 2000; Loveland, Tunali Kotoski, Chen, & Brelsford, 1995; Phillips et al., 1998) which often present participants with non-verbal intonation. Presentation of non-verbal intonation is less feasible in complex emotions and mental states, and hence the listener is provided with some context through the verbal content. Previous studies reported that individuals with ASC often make use of verbal information to compensate for their ER and mentalising difficulties (Grossman, Klin, Carter, & Volkmar, 2000; Klin, Jones, Schultz, Volkmar, & Cohen, 2002b). However, the

RMV-R makes this harder, as foils were designed to match the semantic but not the prosodic level of stimuli. Hence, successful integration of the linguistic and para-linguistic levels is required in order to perform well on the RMV-R. In light of the findings about the role of the right hemisphere in interpreting emotional intonation and the left hemisphere in semantic content (McNeely & Parlow, 2001; Wildgruber et al., 2005), and further to findings of atypical left hemisphere dominance when participants with ASC make emotional judgement from faces (Ashwin, Wheelwright, & Baron-Cohen, 2005), it would be interesting to investigate whether difficulties in complex emotion recognition from voices in ASC are related to less right hemisphere activity and more left hemisphere activity, compared to controls. Increased left hemisphere activity when processing emotional verbalisations could also relate to attempts to rely on semantic cues in order to compensate for poor processing of intonation.

An interesting result found for the RMIV-R is the difference between males and females with AS/HFA. Female superiority on empathising tasks, found in the general population (Baron-Cohen, 2002, 2003) was not significant in the control group, probably due to the small number of females in this group. However, in the AS/HFA group, males were found to perform better on the task than females. The significant difference between males and females with AS/HFA in AQ scores ($t[21.2]=7.65$, $p<.05$) might explain this sex difference: Females in the AS/HFA group reported having more autistic traits ($M=40.29$ $SD=4.96$) than males ($M=34.64$, 9.94), and since these are associated with lower task scores, this may explain the sex difference in the AS/HFA group. Sex differences in ER or ToM have not been studied within the autistic spectrum. Two studies (Tsai & Beisler, 1983; Volkmar, Szatmari, & Sparrow, 1993) reported that among individuals with classic autism, females show a more severe manifestation of the condition than males. However, these studies did not test for emotion recognition differences and these samples were lower functioning than the present samples. While findings on the RMV-R suggest sex differences may exist, they are limited by the small number of females in both groups. There is need for further examination of sex differences in ER abilities within the autistic spectrum.

In conclusion, the RMV-R is a more sensitive and a briefer task for assessing complex ER in speech segments. It correlates well with an established complex emotion and mental state recognition task (RME-A), with the new tasks of the CAM-A, and with autism spectrum traits. This strong positive association with the other ER tasks used in the study, together with the stimuli which are not included in *Mind Reading,* makes the RMV-R a good task to use as a vocal distant feature-based generalisation measure for the adult intervention evaluation experiments.

Child feature-based auditory task (C-FAT)

Introduction

Unlike the adult feature-based auditory task, no existing task of complex emotion and mental state recognition is available for children. Hence, a new task was created for this study. The auditory material used for this task was taken from recordings that were made for *Mind Reading* but were not included in it. During its production, *Mind Reading* went through some changes in the structure of its emotion taxonomy, which originally included more than 412 concepts. Many of these additional emotions became synonyms to others of the 412, leaving their recorded stimuli out of the final version of the software. This material was used to create the child feature-based auditory task (C-FAT). None of the items selected for this task were rejected from *Mind Reading* for being invalid by the panel of judges as not depicting the target mental state. While the actors' voices are similar to those who appear in the vocal task of the CAM-C, the content of these additional recordings is completely novel for *Mind Reading* users, and could thus be used as a generalisation task.

The C-FAT was tested with a group of 8-11 year old children with ASC and matched typically developing controls. The hypotheses were similar to those predicted for the CAM-C, i.e. significantly lower performance on the task in the ASC group compared to the control group, positive correlation for C-FAT scores with age (as emotional vocabulary and ER skills develop with age) and with verbal IQ (due to the verbal nature of the task), positive correlations with CAM-C scores and with the child version of 'Reading the Mind in the Eyes' (RME-C), and a negative correlation with the number of autistic traits reported by parents, as measured by the CAST. The task's power, reliability and validity, were calculated.

Method

Participants

The participants tested on the C-FAT were the same 30 children with AS/HFA and 25 typically developed controls who were tested on the CAM-C. Thus, their characteristics are described in Appendix 1.

Task Development

The first step in creating the C-FAT was to go through the voice recordings that were left out of *Mind Reading* and select those that were at the right level for children and had good enough sound quality. Twenty voice recordings were selected in this way. As with the CAM-C, it was checked for most of the emotional labels of the C-FAT items that at least 75% of 8-9 year olds in the emotional lexicon developmental survey knew them (according to parental report). However, two of the selected labels (*troubled* and *fond*) met this criterion only for 10-11 year olds. In addition, 3 other labels (*scolding*, *hysterical*, and *romantic*) were not included in the emotional lexicon survey, but were included in *Mind Reading* in a level appropriate for this age group. In order to prevent ceiling effects and to allow for improvement following the use of *Mind Reading* in the intervention study, it was decided to keep these items. To ensure that children understand the words before they answer the questions, all emotional concepts were defined and explained to children by the experimenters. Pilot testing revealed that when provided with a definition, or with an alternative wording of the concept (e.g. ‘telling off’ instead of ‘scolding’), even 8 year olds understood the meaning of these concepts. The list of the emotions included in the C-FAT appears in Table 5.

Next, each item was matched with three foils, ensuring they were from the same developmental level as the target emotion. Foils were selected so that they could match the verbal content of the scene, if the intonation was not taken into account. The labels and foils were then reviewed by two independent judges to ensure none of them was too similar to the target emotion.

Emotion/ Mental state	Known to 8-9 year olds		Known to 10-11 year olds	
	Percent	Total n	Percent	Total n
Displeased	78%	(18)	85%	(20)
Delighted	89%	(18)	100%	(20)
Cold	78%	(18)	75%	(20)
Cheeky	86%	(28)	94%	(34)
Tired	100%	(18)	100%	(20)
Hysterical[2]				
Hateful	96%	(27)	97%	(34)
Helpful	96%	(27)	100%	(34)
Scolding[1]				
Troubled	57%	(28)	85%	(34)
Disbelieving	83%	(18)	85%	(20)
Concentrating	94%	(18)	100%	(20)
Unsure	82%	(28)	97%	(34)
Jealous	89%	(27)	88%	(34)
Fond	67%	(18)	100%	(18)
Judging	48%	(27)	71%	(34)
Romantic[1]				

Table 5: The target emotions covered in the C-FAT

A computerised task was created using DMDX experimental software (Forster & Forster, 2003). An instruction slide was presented first, asking participants to choose the answer that best describes how the person in each voice recording is feeling. The instructions were followed by two practice items. For each question, the four numbered answers were presented before and while each item was played, to prevent overloading participants' working memory. This prevented randomising item order in the voice task. Instead, two versions of the task were created, with reversed order, to avoid an order effect. A handout with definitions for all the emotion words used in the tasks was prepared.

The C-FAT was then piloted with 16 children - two girls and two boys from 4 age groups – 8,9,10 and 11 years of age. Children were randomly selected from a local mainstream school. The tasks were played to them on two laptop computers, using headphones for the voice task. To avoid possible confounds due to reading difficulties, the experimenter read the instructions and possible answers to the

[2] These emotion words were not included in the emotional lexicon survey, but appear in *Mind Reading* in a level appropriate for this age group.

children and made sure they were familiar with all the words, using the definition handout, where necessary. Participants were then asked to press a number from 1 to 4, to choose their answer. After choosing an answer the next item was presented. No feedback was given during the task.

Next, an item analysis was carried out. Items were included if the target answer was selected by at least half the participants, and if no foil was selected by more than a third of the participants ($p<.05$, Binomial test). Items which failed to meet these criteria were matched with new foils and played to a different group of 16 children. After this second round, 3 items which still did not meet these criteria were removed from the task. The final task included 17 items. Any score greater than 7 is above chance at the $p < 0.05$ level (Binomial test).

Procedure

The procedure for the C-FAT is similar to that of the CAM-C described in Appendix 1. The experimenter read the instructions and the questions and answers for all items with the participants, and checked that they were familiar with all the possible answers. If needed, a definition handout was used to familiarise participants with word meanings among the possible answers. The emotional meaning of some of the words, which may be interpreted differently (e.g. *cold*) was explained before the task was taken. There was no time limit to answer each item. Participants were allowed one short break during the task. The C-FAT took about 15 minutes to complete.

Results

C-FAT score was calculated by summing the number of items correctly answered by participants. All children in the control group and all but 4 in the AS/HFA group scored above chance. There were no ceiling effects. One sample Kolmogorov-Smirnov tests conducted for task scores in the two groups showed that the task score distributions in both groups did not differ from normal. Hence parametric analysis was used.

A univariate analysis of covariance (ANCOVA) was conducted for C-FAT scores, with group (ASC/control) as the independent variable and with verbal IQ, performance IQ, and age as covariates. The analysis yielded a significant main effect of group ($F[1,50]=25.56$, $p<.001$). The scores of the AS/HFA group ($M=10.43$, $SD=2.46$) were significantly lower than those of the control group ($M=12.80$, $SD=2.10$). Significant effects were also found for the covariates age ($F[1,50]=31.56$, $p<.001$) and verbal IQ ($F[1,50]=23.58$, $p<.001$), but not for performance IQ ($F[1,50]=0.9$, n.s.).

Bivariate correlations calculated for the C-FAT revealed that as predicted, participants' task scores correlated positively with their age ($r=.57$, $p<.001$) and verbal IQ ($r=.49$, $p<.001$). C-FAT correlated negatively with the level of autistic traits reported on the CAST ($r=-.56$, $p<.001$), suggesting the more autistic traits the child has, the lower his score on the task. Positive correlations were also found between the C-FAT and the RME-C ($r=.40$, $p<.005$), as well as the voice scale ($r=.64$, $p<.001$) and the face scale ($r=.57$, $p<.001$) of the CAM-C, providing it with external validity.

Emotion/ Mental state	AS/HFA	Control
Displeased	73%	88%
Delighted	33%	52%
Cold	53%	72%
Cheeky	53%	68%
Tired	23%	48%
Hysterical	53%	68%
Hateful	83%	88%
Helpful	70%	84%
Scolding	77%	84%
Troubled	60%	80%
Disbelieving	40%	68%
Concentrating	77%	92%
Unsure	90%	92%
Jealous	33%	60%
Fond	40%	64%
Judging	77%	80%
Romantic	87%	92%

Table 6: Percentage of correct responses to the C-FAT items

Power calculations for the C-FAT (two tailed, with $\alpha=0.01$) revealed a power level of $1-\beta=0.96$. In a discriminant analysis, the significant discriminant function ($\chi^2[17]=28.05$, $p<.05$) successfully classified 81.8% of the participants (83.3% of participants with AS/HFA and 80% of controls) into their original groups. Table 6 presents the percentage of correct responses for each item in the AS/HFA and control groups.

Test-retest reliability, calculated for children with AS/HFA who participated in the no-intervention control group of experiment 3 and who took the tasks before and after the 10-15 week period, was $r=0.71$ ($p<.001$).

Discussion

The Child Feature-based Auditory Task was created to assess distant generalisation in the auditory channel. Despite its brevity, 8-11 year old children with ASC still perform significantly lower on it, compared to typically developing children, matched on age and IQ. This finding replicates the finding reported in Appendix 1, of difficulties recognising complex emotions and mental states in children with ASC. Complex ER from voices in children with ASC has not been studied before, which stresses the importance of a similar finding in the CAM-C complex emotion voice and the current task. However, it is important to remember that the two tasks were composed in a similar manner, used stimuli from a similar source, and were tested on similar participant groups. Hence, replication of this finding in other complex ER studies of children with ASC (and with other clinical conditions) is required. In this sense, the correlation of the C-FAT with RME-C scores validates the task using an external criterion of a different modality. The task showed good power, reasonable test-retest reliability and discriminative ability, though not as robust as that of the CAM-C.

As observed in complex emotions of the CAM-C, a child’s age played a role in his or her ability to recognise the emotions correctly from the segments of speech in the C-FAT. These findings, on tasks that strongly rely on intonation support previous results showing greater reliance on intonation as children develop (Morton & Trehub, 2001).

The association between task scores and verbal IQ may be related to the need to integrate both intonation and semantic content to answer items correctly on the C-FAT. Higher verbal IQ may allow better understanding of the spoken meaning and better reading ability of the available answers, resulting in higher task scores. As seen in the adult versions of the CAM-A and the RMV-R, such correlations decrease with age and verbal IQ and become non-significant at older ages, presumably because participants have achieved a minimum verbal level to pass on the task. Therefore, age and verbal IQ effects are probably products of the relatively wide age range of children tested (8-11) and the use of harder items, designed to allow for improvement following intervention. As the ANCOVA showed, no significant interaction was found between group and age, or group and verbal IQ, suggesting the associations of these factors with C-FAT scores appear both in children with ASC and in controls. The effect of these variables on improvement following the use of *Mind Reading* will be monitored in the intervention experiments.

Since the task included a single item for each emotion, an emotion based analysis such as that conducted for the CAM concepts is not possible. An anecdotal example of an item that showed a marked group difference in the proportion of children who correctly answered it is the item depicting *jealous* (‘was she with you?’), which was only recognised by 30% of participants with ASC, versus 60% of children in the typically developing group. Children with ASC commonly mislabelled this item as *surprised*, probably because they recognised a question was being asked. Another example was the mislabelling of *troubled* (in the item ‘I don’t think there’s much else to say”) as *cold*. Whereas this label could fit the content of the utterance, it did not fit both the content and the intonation. Yet, this label was chosen by 33% of participants in the AS/HFA group, who were presumably relying on the semantic content of the item, and who must have failed to make use of the intonation. It would be interesting to see whether the intonation judgements in ER in ASC could be improved among children with ASC following the use of *Mind Reading*, and whether such an improvement could be generalised to material not included in the software, such as that included in the RMV-R and the C-FAT.

Appendix 3– Creation of distant holistic generalisation measures

This chapter describes the tasks designed to assess the third level of generalisation in the intervention study. This level differs from the material taught through *Mind Reading* both in content and in perceptual complexity. In other words, this third level assesses the ability of individuals with ASC, who were trained with faces and voices separately, to recognise emotions from novel stimuli that combine facial, vocal, and contextual information. This level is closer to real life requirements in its holistic presentation of multimodal emotional expression. It can thus be said to have greater ecological validity.

As described in Chapter 2, the breadth of emotional information provided in the different channels hinders individuals with ASC on ER tasks, either due to difficulties integrating the details provided through the different perceptual channels into one coherent picture (the Weak Central Coherence model), or because multimodal emotional information is more imprecise and harder to systemise (the Empathising-Systemising model). *Mind Reading* attempts to improve ER in ASC by removing context and emotional integration (thus taking WCC into account) and by presenting emotional information in a more precise system (thus taking E-S theory into account). To test whether the ER deficit can be compensated, tasks that assess complex ER from such stimuli in children and in adults were required. The only existing task for adults was 'The Awkward Moments Test' (Heavey, Phillips, Baron-Cohen, & Rutter, 2000), which includes 7 social scenes from which participants were asked to tell how a protagonist was feeling. This relatively small number of scenes limits the validity and power of the task, and risks floor or ceiling effects. In addition, the use of adverts, which is what the 7 social scenes were taken from, are by definition exaggerated, and so may make the task easier to answer. Pilot testing with high functioning adults with ASC for the intervention study resulted in ceiling effects for some of the participants. A more subtle collection of social situations may better represent the complexity of socio-emotional interactions, and provide a more valid measure of the ability of adults with ASC to interpret them.

Another task for adults which was recently described is the 'Movie for The Assessment of Social Cognition' (MASC, Dziobek et al., in press). This task uses a longer film about a social encounter between four adults, and stops to ask how protagonists feel at different points. Whereas the MASC has a wider range of scores (and may therefore have improved power), it is limited by the lack of synchrony between lip movements and speech, as it was recorded in German and dubbed into English. No such tasks are available for children, with the exception of the Feshbach and Powell audiovisual test for empathy (Feshbach, 1982), which focuses mostly on basic emotions, with *proud* as the only complex emotion tested.

Hence, there was a need for new multimodal complex ER tasks for the third level of generalisation in the intervention study. This chapter reports the development of these two tasks. Both tasks make use of short scenes sampled from feature films, after which they were named 'Reading the Mind in Films' (RMF-A for adults and RMF-C for children).

The 'Reading the Mind in Films' task – adult version (RMF-A)

Introduction

The RMF-A includes 22 short scenes depicting different complex emotions and mental states that vary in valence, intensity and complexity. They were chosen for their relevance to everyday social interaction.

The RMF-A was tested on adults with AS/HFA and a matched control group from the general population. Based on the ToM, empathising-systemising and Weak Central Coherence models of autism, and on previous findings of ER difficulties in ASC when multimodal stimuli were used (Dziobek et al., in press; Heavey, Phillips, Baron-Cohen, & Rutter, 2000; Klin, Jones, Schultz, Volkmar, & Cohen, 2002b; Yirmiya, Sigman, Kasari, & Mundy, 1992) it was predicted that participants with ASC would score significantly lower than controls on the RMF-A task. As found in all ER tasks reported in Appendices 1 and 2, it was also predicted that task scores would correlate negatively with number of autistic traits.

In order to rule out that difficulties responding to the task are simply due to working memory problems, the associations between task performance and item length, and between task performance and the number of characters appearing in the item, were assessed. Correlations between task scores and age or IQ were also tested. RMF-A scores were also predicted to correlate positively with CAM-A scores, which assesses complex ER from faces and voices separately. Since the RMF-A was only administered once during the intervention study, no test-retest reliability data was available for this task. However, its power and ability to discriminate between the ASC and control groups were tested.

Method

Participants

The AS/HFA group comprised twenty-two adults (17 males and 5 females), aged 17-52. Participants had all been diagnosed with AS/HFA in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). They were recruited from a local clinic for adults with ASC, and from a research volunteer database. All the participants in this group belonged to the no-intervention ASC control group of experiment 1 in the intervention evaluation study.

The control group comprised twenty-two adults (18 males and 4 females) from the general population, aged 18-51. They were recruited from a local employment agency. All participants were given the WASI and scored above 85 on both verbal and performance scales. Participants were also given the AQ to assess their number of autistic traits. The average AQ score of the AS/HFA group (M=38.45, SD=7.81) was significantly higher than that of the control group (M=13.95, SD=5.40; $t[42]=12.11$, $p<.001$). The two groups were matched on sex ($\chi^2[1]=0.14$, n.s.), age, verbal IQ, performance IQ and full scale IQ. The groups' background data appears in Table 1.

	AS/HFA group (n=22)			Control group (n=22)			
	Mean	SD	Range	Mean	SD	Range	t (42) [1]
Age	29.01	9.82	17.4-52.0	25.38	9.46	17.6-51.2	1.25
Verbal IQ	110.64	10.79	87-129	116.36	14.45	86-138	1.49
Performance IQ	114.77	12.52	97-140	113.14	8.37	92-129	0.51
Full Scale IQ	114.18	10.78	91-130	116.50	10.99	97-138	0.71

[1]p>.1 for all measures.

Table 1: Age and IQ measures of the AS/HFA and control groups, RMF-A study

Task development

The scenes selected for the task were sampled from 3 feature films: *Lost for Words* (Bell, 1999), *The Winter Guest* (Rickman, 1997), and *The Turn of the* Screw (Bolt, 1999), and from one mini series: *Perfect Strangers* (Poliakoff, 2000). These films were chosen for their dramatic quality, for the subtlety of emotional expression, and for being relatively unknown to the wide audience (to minimise the chance that participants might know the plot in advance).

Thirty short scenes (5-30 seconds long, M=14.8, SD=9.2), were sampled from these films. The selected scenes involved emotional interaction between 1-4 characters, and the expression of complex emotions and mental states (e.g. *smug, awkward, concerned*). In each scene, a protagonist was identified and their emotion or mental state at the end of the scene was labelled. For each item, three foils, of a similar developmental level to the target answer on the emotional lexicon survey, were selected. Foils were selected so that they matched some of the emotional information in the scene but not all of it, e.g. matching the content of the language but not the intonation, the facial expression, or the context. The labels and foils were then reviewed by two independent judges. A handout with definitions for all the target and foil words in the items included was prepared for participants' use before and during the task.

The items were then played on a laptop computer, using DMDX experimental software to 15 adults (7 men and 8 women) randomly selected from the general population. . The task was preceded by an instruction slide and an example item. Task items were mixed so that no two adjacent items were from the same film. For each film, the order of scenes presented was reversed, so that scenes from the end of the film were played first. This was done to avoid use of the plot for contextual cues to answer the items, and prevented randomising item order. Instead, two versions of the task were created and used randomly with participants. Every item was preceded by the question: 'At the end of the scene, how is the protagonist feeling?', followed by the four emotion labels. This allowed participants to focus on the protagonist and the possible answers. The end of the scene was chosen as the reference point to avoid confusion if the protagonist expressed different emotions during the scene. The question and four possible answers were presented again when the scene finished playing, for participants to make their choice.

Emotion/ Mental state	Known to 17-18 year olds (n=50)	Emotion/ Mental state (contd)	Known to 17-18 year olds (n=50)
Annoyed	100%	Overcome	100%
Awkward	100%	Pleased	100%
Belittled	[1]	Prickly	[1]
Bitter	100%	Reflective	96%
Concerned	100%	Resentful	98%
Disconcerted	74%	Resigned	94%
Disliking	100%	Smug	98%
Embarrassed	100%	Stern	98%
Enjoying	100%	Troubled	100%
Exasperated	86%	Unassuming	[1]
Incensed	68%	Worried	100%

Table 2: The target emotions covered in the RMF-A

An item analysis was then carried out. Items were included if the target answer was picked by at least half the participants, and if no foil was selected by more than a third of the participants. Eight items were excluded after failing to meet these criteria.

[1] *Belittled* and *unassuming* were not included in the emotional taxonomy developmental survey, but appear in *Mind Reading* in the adult levels. *Prickly* was not included in the survey or in *Mind Reading* but was agreed to be the most appropriate label between the three judges of the task.

Hence, the final task comprised 22 items and task score varied between 0-22. The list of emotions and mental states included in the final version of the task and the proportion of 17-18 year olds from the emotional lexicon developmental survey, who were familiar with these labels, appears in Table 2. As shown in the table, three of the emotion labels were not included in the emotional lexicon survey. In addition, two of the emotion labels (*incensed* and *disconcerted*) did not meet the 75% criterion, which was set for the tasks described in Appendices 1 and 2. Whereas 68% who know the word out of fifty 17-18 year olds is still significantly above chance ($p<.01$, Binomial test), it was important to ensure that participants are familiar with the verbal labels before watching the items. This was done by making the definition handout available throughout the assessment, and by presenting the question (and optional answers) before each scene is played.

At the end of the scene, how is the older woman feeling?

1. sociable 2. admiring 3. <u>overcome</u> 4. liked

Figure 1: One item from the adult version of 'Reading the Mind in Films' (showing only one frame out of the full clip)[2]

An example of one item from the task is shown in Figure 1. This item, labelled *overcome*, depicts a young woman complimenting an older woman on the way she

[2] Screenshot taken from 'The Turn of the Screw' (1999). Courtesy of Granada International.

educated the children. The older woman thanks her several times calmly, then runs towards her with tears in her eyes saying ‘Oh miss, I’m so glad you’re here’.

Procedure

Participants were tested in the second assessment session of the adult intervention study, in which they served as controls (i.e., no intervention). They were tested at the Autism Research Centre in Cambridge. Participants were asked in advance whether they were familiar with any of the films presented. None of the participants was familiar with more than one of the films. Five participants in the AS/HFA group, and 7 in the control group were familiar with one of the films. They were told that the scenes from the films are presented out of context, hence there was no point in relying on their previous knowledge of the plot. The final version of the task was presented to the participants on a computer screen, using DMDX experimental software. Headphones were given, to improve sound quality. An instructions slide and a practice item preceded the task. The definition handout was given to participants and was available throughout the task with no time limit to answer each item. Completion of the task took about 20 minutes, including a short break. Participants’ CAM-A and AQ scores were available from the first assessment meeting in the AS/HFA group. In the control group, participants brought the AQ questionnaires with them to the assessment meeting, and the CAM-A was taken during the same meeting as the RMF-C. Task order was randomised.

Results

RMF-A scores were calculated by counting the number of correct answers for each participant. All the participants in the control group and all but three of the participants in the AS/HFA group scored above chance (i.e. above 9, $p<.05$, Binomial test). The proportion of correct responses to task items did not correlate significantly with the items’ length ($r_{spearman}=.05$, n.s.) or with the number of characters appearing in them ($r_{spearman}=.25$, n.s.).

A review of percentage of correct responses for the 22 items of the task, presented in Table 3, shows that no item was answered correctly by 100% of the participants with AS/HFA, and that only one item was answered correctly by all the participants in the control group. In two items, the target emotion label was picked by less than a third of the typically developing participants. In the first, 55% of the control group preferred the foil *intimate* to the target *bitter*. This foil was designed to match all aspects of the emotional information except intonation, which distinguished the target from it and from the other foils. In the second item, participants preferred any of the foils to the target *prickly*, possibly due to it being an uncommon label for an emotion. However, since problems with these items did not occur in the original validation phase, these items were not excluded from the analysis.

Emotion/ Mental state	AS/HFA	Control
Annoyed	77%	73%
Awkward	64%	86%
Belittled	45%	68%
Bitter	18%	18%
Concerned	55%	77%
Disconcerted	45%	91%
Disliking	50%	82%
Embarrassed	64%	68%
Enjoying	64%	100%
Exasperated	64%	82%
Incensed	68%	86%
Overcome	59%	82%
Pleased	77%	91%
Prickly	36%	14%
Reflective	50%	64%
Resentful	36%	41%
Resigned	59%	73%
Smug	73%	86%
Stern	64%	55%
Troubled	50%	59%
Unassuming	73%	95%
Worried	73%	73%

Table 3: Percentage of correct responses to the RMF-A items

To compare the performance of the two groups on the RMF-A, a univariate analysis of covariance (ANCOVA) was performed on task scores, with group as the independent variable and with verbal IQ, performance IQ, and age as covariates. Sex was not entered as an independent variable due to the small number of females in the control group. The analysis yielded a significant main effect of group ($F[1,39]=10.52$, $p<.005$). As hypothesised, the AS/HFA group ($M=12.64$, $SD=3.22$) performed significantly lower than the control group ($M=15.64$, $SD=2.08$). In addition, a significant effect of verbal IQ ($F[1,39]=10.93$, $p<.005$) was found.

Correlation analysis conducted for RMF-A revealed a significant positive correlation with verbal IQ ($r=.48$, $p<.005$). No correlation was found for RMF-A with age ($r=.09$, n.s.), or with performance IQ ($r=-.08$, n.s.). As predicted, RMF-A scores were negatively correlated with AQ scores ($r=-.52$, $p<.001$), suggesting higher numbers of autistic traits are associated with lower task scores. Correlations of the RMF-A with CAM-A scores were, as predicted, positive for the CAM-A face task ($r=.63$, $p<.001$), CAM-A voice task ($r=.62$, $p<.001$) and number of CAM-A emotional concepts recognised ($r=.61$, $p<.001$), giving the RMF-A external validity.

Power calculations for the RMF-A (two tailed, with $\alpha=0.01$) revealed a power level of $1-\beta=0.948$. In a discriminant analysis, the significant discriminant function ($\chi^2[22]=34.52$, $p<.05$) successfully classified 90.9% of the participants (86.4% of participants with AS/HFA and 95.5% of controls) into their original groups.

Discussion

The adult version of the ‘Reading the Mind in Films’ task (RMF-A) is a new ‘ecological’ task for assessing recognition of complex emotions and mental states, using multimodal stimuli in the form of social scenes from feature films. Results show that high functioning participants with ASC scored significantly lower than matched controls in the general population and that this effect was not simply due to the association of task scores with verbal ability. The RMF-A has a wide score range in both groups, with no ceiling or floor effects. Power calculations and discriminant analysis showed it is sensitive and that more than ninety percent of the participants could be correctly allocated to their groups, based solely on their task performance. RMF-A scores significantly correlated with other complex emotion recognition tasks, confirming its validity. Performance on the task was not correlated with length of its items or the number of characters appearing in the scenes, suggesting there was no working memory confound.

The significant group difference on task scores replicates previous findings of difficulties among high functioning adults with ASC on tasks involving recognition of complex emotions and mental states in multimodal stimuli (Heavey, Phillips, Baron-Cohen, & Rutter, 2000; Klin, Jones, Schultz, Volkmar, & Cohen, 2002a, 2002b). Compared to the instruments used in these studies, the RMF-A includes a larger number of items and covers a wider range of complex emotions and mental states.

The RMF-A requires the integration of multimodal socio-emotional information from faces, eye direction, prosody, verbal content and context. As such, performance on it is likely to associate with everyday functioning, as well as with neural connectivity between areas in the social brain network. Previous studies have reported reduced neural activity of social brain areas in ASC during processing of socio-emotional stimuli (Critchley et al., 2000; Dalton et al., 2005; Howard et al., 2000; Wang, Dapretto, Hariri, Sigman, & Bookheimer, 2004) as well as altered connectivity in social brain areas (Welchew et al., 2005). However, most studies of emotion processing in the brain have focused on one modality, and it is possible that when multimodal integration is required, additional brain areas will be recruited.

For example, Iacoboni and colleagues found in a study of typical adults, that watching ecologic social (compared to non-social) scenes increased activation in the medial parietal (precuneus) and dorsomedial prefrontal cortices. This was believed to reflect processing of social relationship watched in the films (Iacoboni et al., 2004). Another study, using still facial expressions in conjuction with vocal emotional expressions, reported increased activation in the medial temporal gyrus in the bimodal, but not in any of the unimodal, conditions (Pourtois, de Gelder, Bol, & Crommelinck, 2005). No brain imaging studies have used ecological multimodal films with individuals with ASC. One study that attempted to increase the salience of emotional faces using supporting prosodic information found that not only did the prosodic information fail to increase performance of participants with autism, their performance actually decreased due to the presentation of stimuli in both channels (Hall, Szechtman, & Nahmias, 2003). These results suggest that when recognising emotion, high-functioning adults with ASC place less emphasis on the integrated emotion than do participants in the general population. Future studies could use the RMF-A in brain imaging studies in the typical and the autistic brain to further explore the differences in processing of such multimodal socio-emotional stimuli.

Unlike the other complex ER adult tasks, described in Appendices 1 and 2, RMF-A scores correlated with verbal IQ. Though this was not predicted, it could be related to the task's verbal nature, including speech in the scenes and potentially harder words as answers. The significant effect of verbal IQ could also relate to the participants' use of semantic content to recognise the protagonists' mental states. Previous studies report that individuals with ASC use semantic content as a way to compensate for their difficulties in theory of mind tasks (Tager-Flusberg, 2000), labelling basic emotions from facial expressions (Adolphs, Sears, & Piven, 2001; Grossman, Klin, Carter, & Volkmar, 2000), and noticing socio-emotional cues in situations (Kasari, Chamberlain, & Bauminger, 2001; Klin, Jones, Schultz, Volkmar, & Cohen, 2002b).

However, on the RMF-A task, since foils were designed to match the semantic content but not other social cues in the scene, relying on language alone may have resulted in misinterpretation of protagonist's mental and emotional state. The following examples show how failing to integrate semantic content with prosody and facial expression in context leads participants with ASC to the mislabel emotions: In the example presented in Figure 1, 59% of the AS/HFA group members labelled the protagonist's emotional state as *overcome*, compared to 82% of controls. 23% of the participants in the ASC group labelled the protagonist's emotion in this scene as *admiring*, whereas only 9% of the participants in the control group gave this label. Participants who chose this distracter might have relied on the words actually spoken only ('Miss, I'm so glad you're here'), missing the emotional component of the protagonist's communication, which could be picked up from her facial expression, intonation and gestures.

Another example was the mislabelling of a protagonist being *concerned*, which was correctly recognised by 77% of controls and 55% of participants with ASC. The scene shows the protagonist's concerned face, though when asked if she is enjoying herself she answers 'Yes, I am', in a concerned tone of voice. Only 5% of participants with ASC relied on the words spoken by the protagonist by choosing *enjoying* as their answer. However, 36% of the participants in the clinical group labelled this contradiction between the words, facial expression and tone of voice as *mysterious*. Whereas this example shows that participants may have spotted a mismatch between verbal and non-verbal communication in the scene, many of them still had difficulties realising which modality better represents the protagonist's emotional state.

These examples demonstrate the subtle errors adults with ASC make when interpreting others' mental states, which in real life situations could lead to wrong interpretation of the interaction and to inappropriate responses. These differences stress the importance of an ecological assessment of socio-emotional understanding in high functioning adults with ASC, as they may be able to pass more basic emotion recognition tasks.

Studies conducted with typically developing preschoolers have shown that children who have more siblings and therefore presumably get more opportunity to practise their mentalising and ER skills, perform better on theory of mind (false-belief) tasks (Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, & Berridge, 1996; Perner, Ruffman, & Leekam, 1994). In addition, anecdotal reports of high functioning adults with ASC describe how they consciously collect examples from past social experiences and implement them in current interactions (Grandin, 1995). Hence, it will be interesting to examine to what extent this knowledge can be acquired through systematic ER training using *Mind Reading*.

To conclude, the RMF-A task allows researchers to quantify complex ER skills in multimodal social scenes, and distinguishes individuals with ASC from controls in the general population. The task was used in the intervention evaluation study, to monitor ER improvements in the distant holistic level.

The 'Reading the Mind in Films' task – child version (RMF-C)

Introduction

The RMF-C is the equivalent of the RMF-A for the children intervention study. It follows the same principles set for the RMF-A, using scenes from children films and assessing recognition of emotions and mental states appropriate for 8-11 year olds. The task includes 22 short scenes, varying in valence, intensity and complexity, chosen for their relevance to everyday social interaction.

The task was administered to a group of 8-11 year olds with AS/HFA, who participated in the child intervention study as no-intervention controls. Their performance on the RMF-C was compared to that of a group of matched typically developing controls. As found with the RMF-A it was predicted that participants with AS/HFA would score significantly lower than controls on the RMF-C. Since ER abilities on all the tasks described in the previous chapters are negatively correlated with number of autistic traits, RMF-C scores were predicted to correlate negatively with CAST scores. Since ER abilities improve with age (Harris, 1989; Herba & Phillips, 2004), and following the results from the child tasks in Appendices 1 and 2, age was predicted to correlate positively with RMF-C scores. Correlation with verbal IQ was also predicted, due to the verbal nature of the task and further to a similar finding reported above on the RMF-A. Item length and the number of characters appearing in the item were correlated with task score in order to test if difficulties responding to the task were simply due to working memory problems. The RMF-C was also hypothesised to correlate positively with the different scales of the CAM-C, to provide it with external validity. As RMF-C was only taken at time 2 of the intervention study, no test-retest reliability measures were available. However, the power and discriminative ability of the task were calculated.

Method

Participants

The AS/HFA group comprised twenty-three children (22 boys and 1 girl), aged 8.3-11.8. Participants had all been diagnosed with AS/HFA in specialist centres using established criteria (American Psychiatric Association, 1994; World Health Organisation, 1994). They were recruited from a volunteer database and a local clinic for children with ASC. A control group from the general population was matched to the clinical group. It comprised twenty-four children (23 boys and 1 girl), aged 8.2-12.1, recruited from a local primary school. Parents and school reports confirmed that none of the children in this group had a psychiatric diagnosis or special educational needs, and none of them had family members diagnosed with ASC. All participants were given the WASI and scored above 75 on both verbal and performance scales. To assess level of parents filled in the CAST. None of the control participants and all but 1 participant with AS/HFA scored above the cut-off point of 15. This participant had a score of 11 on the CAST, due to several unanswered items. Therefore, he was not excluded from the experiment. CAST scores in the AS/HFA group (M=19.43, SD=4.21) were significantly higher than in the control group (M=3.57, SD=1.60; t (45)=16.24, p<.001). The two groups were matched on sex (χ^2[1]=.001, n.s.), age, and IQ. Their background data appears in Table 4.

	AS/HFA group (n=23)			Control group (n=24)			
	Mean	SD	Range	Mean	SD	Range	t (45) [1]
Age	9.97	1.12	8.3-11.8	10.12	1.19	8.2-12.1	0.43
Verbal IQ	110.17	13.76	88-143	111.83	9.92	88-125	0.48
Performance IQ	107.43	18.11	78-140	110.67	12.27	91-133	0.72
Full Scale IQ	110.30	14.79	83-138	112.88	9.54	95-129	0.71

[1] p>.1 for all t-tests

Table 4: Age and IQ scores of the AS/HFA and control groups, RMF-C study

Task development

Twenty seven short scenes (6-30 seconds long, M=16.5, S.D.=7.3) were sampled from four children feature films: *A little princess* (Cuarón 1995), *The Yearling* (Brown, 1946), *The railway children* (Jeffries 1970), and *Anne of Green Gables – The*

Sequel (Sullivan 1987). All films were rated 'Universal'- appropriate for children from 4 years of age, by the British Board of Film Classification. The selected scenes involved socio-emotional interaction between 1-4 characters, and the expression of complex emotions and mental states (e.g. *relieved*, *guilty*, *lonely*). In each scene, a protagonist was identified and its emotion or mental state at the end of the scene was labelled. Next, three foils of the same developmental level as the target emotion (according to the emotional lexicon survey) were selected for each item. Foils were selected so that they match some of the emotional information pertained in the scene but not all of it. The labels and foils were then reviewed by two independent judges. Two scenes were removed at this point due to disagreement between judges. A handout with definitions for all the target and foil words in the items included was prepared for participants' use before and during the task.

The items were then played to 16 typically developing children - two girls and two boys from 4 age groups – 8,9,10 and 11 years of age. Children were randomly selected from a local mainstream school. Teachers confirmed the selected children had no learning difficulties and parents confirmed that no family members were diagnosed with autism spectrum conditions. The items were played to the selected children on two laptop computers, using DMDX experimental software (Forster & Forster, 2003). Every item was preceded by the question: 'at the end of the scene, how is the protagonist feeling?', followed by the four emotion labels. This allowed participants to focus on the protagonist and the possible answers.

To avoid confounds due to reading difficulties or familiarity with the emotional label, the experimenter read each question and the possible answers to the children and made sure they were familiar with all the words before playing the scene. The question and possible answers appeared again at the end of the scene. Participants were then asked to press a number from 1 to 4, to choose the answer which best describes how the protagonist is feeling. Items were mixed so that no two adjacent items were from the same film. For each film, the order of scenes presented was reversed, so that scenes from the end of the film were played first. This was done to avoid use of the plot for contextual cues to answer the items. Since this prevented randomising item order, two versions of the task were created and used randomly with participants.

Figure 2 shows an example of an item from the task. The scene depicts a young woman rushing to the post office and knocking at the door. An elderly woman opens the door, and we can see her face when the young woman says ‘I’m sorry, I know you’re closed’. Participants are then asked to choose the answer that best describes how the elderly woman feels, with *unfriendly* being the target mental state.

At the end of the scene, how is the woman feeling?

1. worried 2. sorry 3. unfriendly 4. interested

Figure 2: One item from the child version of ‘Reading the Mind in Films’ (showing only one frame out of the full clip)[3]

Next, an item analysis was carried out. Items were included if the target answer was picked by at least half of the participants, and if no foil was selected by more than a third of the participants ($p<.05$, Binomial test). Items that failed to meet these criteria were matched with new foils and played to a different group of 16 children. Three items that did not meet the criteria after this second round of validation were excluded

[3]Screenshot taken from Anne of Green Gables - The Sequel (1987), Courtesy of www.sullivanmovies.com.

from the final task, which comprised 22 items. Task score varied between 0-22. The list of emotions and mental states included in the final version of the RMF-C and the proportion of 8-9 and 11-12 year olds from the emotional lexicon developmental survey, who were familiar with these labels, appears in Table 5.

Emotion/ Mental state	Known to 8-9 year olds		Known to 10-11 year olds	
	Percent	Total n	Percent	Total n
Afraid	100%	(18)	100%	(20)
Annoyed	94%	(18)	95%	(20)
Bossy[4]				
Bothered	78%	(18)	85%	(20)
Caring	94%	(18)	90%	(20)
Criticised[1]				
Dreamy	72%	(18)	90%	(20)
Excited	94%	(18)	100%	(18)
Friendly	100%	(18)	100%	(18)
Furious	89%	(18)	94%	(18)
Guilty	75%	(28)	91%	(34)
Lonely	89%	(27)	100%	(34)
Loving	96%	(28)	100%	(34)
Lying	96%	(28)	100%	(34)
Mean	89%	(28)	100%	(34)
Relieved	54%	(28)	85%	(34)
Shocked	68%	(28)	97%	(34)
Tempted	54%	(28)	88%	(34)
Uncomfortable	79%	(28)	88%	(34)
Unfriendly	79%	(28)	94%	(34)
Upset	100%	(28)	100%	(34)
Worried	86%	(28)	100%	(34)

Table 5: The target emotions and mental states included in the RMF-C

As Table 5 shows, most of the target emotions and mental states included in the RMF-C were known to least 75% of 8-9 year olds in the emotional lexicon developmental survey (according to parental report). However, four of the selected labels (*dreamy*, *relieved*, *shocked* and *tempted*) met this criterion only for 10-11 year olds. In addition, 2 other labels (*bossy*, and *criticised*) were not included in the emotional lexicon survey, but were included in *Mind Reading* in a level appropriate for this age group. In order to prevent ceiling effects, it was decided to keep these items. To ensure that

[4] These mental states were not included in the emotional lexicon survey

children understand the words before they answer the questions, all emotional concepts were defined and explained to children by the experimenters. Pilot testing revealed that when provided with a definition, even 8 year olds understood the meaning of these concepts.

Procedure

Participants with ASC were tested individually at the second meeting of the child intervention study, in which they all served as controls (i.e., no intervention). Typically developing controls were tested individually in a quiet room at a local school. Participants were asked in advance whether they were familiar with any of the films presented. None of the participants was familiar with more than one of the films. Nine participants in the AS/HFA group, and eight in the control group were familiar with one film. They were told that the scenes from the films would be presented out of context, and that they should therefore not rely on their previous knowledge of the plot.

The final version of the task was presented to the participants on a laptop with a 15 inch screen, using DMDX experimental software. Headphones were given, to improve perception. An instructions slide and a practice item preceded the task. The experimenter read the instructions, and the questions and answers of all items with the participants, and checked they are familiar with all the possible answers. When needed, the definition handout was used to familiarise participants with answers. There was no time limit to answer each item. Completion of the task took about thirty minutes, including a short break in the middle. Participants' CAM-C and CAST scores were available from the first assessment meeting in the ASC group. In the control group, CAST questionnaires were filled out in advance by parents, and the CAM-C was taken during the same meeting as the RMF-C. Task order was randomised.

Results

Task scores were calculated by counting the number of correct answers for each participant. All the participants in the control group and all but one of the participants in the AS/HFA group scored above chance (i.e. above 9, $p<.05$, Binomial test) on the RMF-C task. No participant scored at ceiling. A review of percentage of correct responses for the 22 items of the task (see table 6) shows that no item was answered correctly by 100% of the participants either in the AS/HFA or in the control group. As in the validation stage, targets were picked by more than 50% of the controls on all items. The proportion of correct responses to task items did not correlate significantly with the number of characters appearing in them ($r_{spearman}=-.02$, n.s. for the ASC group, $r_{spearman}=-.17$, n.s. for the control group) or with items' length ($r_{spearman}=-.008$, n.s. for the ASC group, $r_{spearman}=-.21$, n.s. for the control group).

Emotion/ Mental state	AS/HFA	Control
Afraid	78%	96%
Annoyed	78%	92%
Bossy	83%	75%
Bothered	48%	79%
Caring	39%	67%
Criticised	61%	67%
Dreamy	83%	83%
Excited	70%	83%
Friendly	83%	88%
Furious	52%	83%
Guilty	70%	92%
Lonely	61%	79%
Loving	78%	88%
Lying	35%	79%
Mean	78%	92%
Relieved	70%	83%
Shocked	65%	96%
Tempted	70%	75%
Uncomfortable	70%	88%
Unfriendly	48%	83%
Upset	48%	75%
Worried	52%	79%

Table 6: Percentage of correct responses to the RMF-C items

An analysis of covariance (ANCOVA) was performed on the RMF-C task scores, with group as independent variable and with verbal IQ, performance IQ, and age as covariates. A significant main effect was found for group ($F[1,42]=33.74$, $p<.001$) with AS/HFA group scores ($M=14.17$, $SD=3.45$) significantly lower than the control group scores ($M=18.21$, $SD=2.59$). In addition to the group main effect, the ANCOVA also yielded significant effects of verbal IQ ($F[1,42]=9.91$, $p<.005$) and Age ($F[1,42]=22.72$, $p<.001$). Correlation analysis revealed a significant positive correlation between task scores and verbal IQ ($r=.47$, $p<.005$). A positive correlation was also found between task scores and age ($r=.57$, $p<.001$). Task scores were not significantly correlated with Performance IQ ($r=.20$, n.s.).

As predicted, RMF-C task scores were negatively correlated with CAST scores ($r=-.60$, $p<.001$), suggesting that the more autism spectrum traits one possesses, the lower one's ability to recognise complex emotions from ecologically valid stimuli.

The correlations of the RMF-C task scores with CAM-C scores were positive for the face task ($r=.67$, $p<.001$), the voice task ($r=.58$, $p<.001$) and number of CAM-C emotional concepts recognised ($r=.65$, $p<.001$). These suggest there is an association between the ability to recognise complex emotions separately in faces and voices, and the ability to recognise them when facial, vocal and contextual information is integrated.

Power calculations for the task (two tailed, with $\alpha=0.01$) revealed a power level of $1-\beta=0.961$. A discriminant analysis was conducted, and the significant discriminant function ($\chi^2[22]=35.2$, $p<.05$) successfully classified 87.2% of the participants (87% of participants with AS/HFA and 87.5% of control controls) into their original groups.

Discussion

The RMF-C is the first attempt to assess recognition of a wide range of complex emotions and mental states in children with and without ASC, using multimodal social scenes. Results show that high functioning children with ASC scored significantly lower than matched controls from the general population. This effect was not simply due to the association of task scores with verbal ability or with age. The RMF-C has a wide score range in both groups, with no ceiling or floor effects. Power calculations and discriminant analysis showed the task is sensitive and has a good discriminative validity, with more than 87% of the participants correctly allocated to their groups, based solely on their task performance. RMF-C scores significantly correlated with existing complex emotion recognition tasks, confirming the validity of the task. Performance on the task was not correlated with length of its items or the number of characters appearing in the scenes. Hence, these measures did not indicate there was a working memory confound.

The significant group difference on task scores replicates previous findings of difficulties among high functioning children with ASC on complex emotion and mental state recognition tasks from visual, auditory and contextual stimuli (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999; Baron-Cohen, Wheelwright, Spong, Scahill, & Lawson, 2001; Happe, 1994a; Yirmiya, Sigman, Kasari, & Mundy, 1992). It also replicates findings of studies of adults with ASC, which found difficulties in recognition of complex emotions and mental states when using multimodal social scenes (Dziobek et al., in press; Heavey, Phillips, Baron-Cohen, & Rutter, 2000; Klin, Jones, Schultz, Volkmar, & Cohen, 2002a, 2002b); see also the RMF-A described above). As with the adult task, the RMF-C lends itself to neuro-imaging and gaze tracking studies of ER, mentalising and empathy in typical development, in children with ASC, and in other clinical groups.

As predicted, the task scores' were correlated with age, indicating that the ability to recognise complex emotions and mental states improves with age and experience. However, this association was stronger for the control group ($r=0.78$, $p<.001$) than for children with ASC ($r=.54$, $p<.01$). Typically developing children learn to recognise complex emotions and mental states through constant interaction with family members and peers (Denham, 1998; Harris, 1994; Jenkins & Astington, 1996). The reduced levels of social interaction among children with ASC may to some extent account for their slower learning of complex emotion and mental states recognition. It will be important to examine whether this knowledge can be acquired through training with *Mind Reading*.

RMF-C scores were negatively correlated with the level of autism spectrum features participants possessed. This was predicted and replicated findings on all the other tasks described previously. However, since the range of CAST scores was narrow in both groups, and since the groups differed on their CAST scores, the correlation between RMF-C and CAST scores ($r=-.60$) was actually equivalent to a correlation between RMF-C scores and group (r=-.56). In order to examine the correlation between autism spectrum features and task scores, a better representation of the entire continuum between typical development and autism would be desirable. Assessing siblings of children with an ASC diagnosis, who have been found to have a lesser variant of the condition (Hughes, Plumet, & Leboyer, 1999; Piven, Palmer, Jacobi, Childress, & Arndt, 1997; Yirmiya, Shaked, & Erel, 2001) may allow to better test this correlation.

The task scores' correlation with verbal IQ, which was predicted due to the task's verbal nature and to findings on the RMF-A, suggests participants use verbal cues to pick up the protagonist's mental states. As the following examples show, relying on language alone may result in misinterpretation of protagonist's mental and emotional state.

A more detailed observation of the task reveals that on some items there is a greater group difference in the proportion of correct answers. For example, only 48% of the AS/HFA group labelled the emotional state depicted in figure 1 as *unfriendly*, compared to 83% of controls. 22% of the clinical group labelled the protagonist's emotion in this scene as *sorry*, whereas only 8% of the control group gave this label. Since the protagonist in this scene did not speak, participants who chose this distracter may have relied on what the other character said ('I'm sorry, I know you're closed') rather than use the context set up by this utterance and the protagonist's facial expression.

Another example was the mislabelling of a protagonist *lying*. The scene shows a father, mother and a son having dinner together. The son wants to tell his mother about something he saw today, but his father looks at him, then kicks him under the table. The son then says he only saw a big frog, for which the father answers 'A big frog, eh? A big frog!" Participants were asked about the father's mental state at the end of the scene. In the control group, 79% of the participants spotted the son's initial enthusiasm, the father's disapproval and then his false excitement, which combined together gives the impression the father is trying to conceal some information from the mother. Only 35% of participants in the AS/HFA group correctly labelled this scene, whereas 39% of them thought the father was genuinely interested. This item confirms the difficulties children with ASC have with understanding of deception (Baron-Cohen, 1992; Sodian & Frith, 1992). The social relevance of such difficulties is clear. In her book, Claire Sainsbury describes how she used to skip school on 'April Fool's Day', as she couldn't recognise the other children's tricks or understand their hidden intent, leaving her the victim of their pranks (Sainsbury, 2000). Difficulties recognising deception is still challenging for adults with ASC, as demonstrated by the difficulties recognising *insincere* on the CAM-A by adults with ASC.

A third example is the mislabelling of a protagonist expressing *care* towards another. This scene, which involved no speech, depicts a well-dressed girl looking at a girl her age who is dressed in rags and is busy scrubbing the floor. The camera then moves back to the first girl, who continues to look with a *caring* expression. Only 39% of the AS/HFA group thought the girl was expressing care, compared to 67% of the control group. 43% of the AS/HFA group preferred the label *curious* to describe the girl's mental state. Though this label is plausible in this situation, it suggests they have adopted a more factual, less empathic view of the scene. If the first girl *looked* at the second girl, then she wanted to gather information. *Looking* is not associated with expressing an emotion.

The first two examples demonstrate a tendency among participants with ASC to process part of the given information, both perceptually and temporally. Perceptually, they did not integrate the visual, auditory and linguistic information to arrive at an answer, but rather focused only on one channel, the linguistic. Temporally, they used information which was usually proximal to the time the clip was stopped and the question was raised (the word 'sorry' in the first example, and the last utterance in the second), instead of taking into account the entire development of the situation. This suggests that participants with ASC processed the surface information, perhaps because of difficulties accessing the full depth of the situation (Lawson, 2003). Great attention to detail is characteristic of ASC (Frith, 1989; Plaisted, Saksida, Alcantara, & Weisblatt, 2003). Neurologically, this pattern may result from a skewed balance of local (i.e. within neural assemblies) versus long-range connectivity between functional regions (Baron-Cohen & Belmonte, 2005).

The third example includes information in one channel (visual). Participants cannot rely on language to make sense of the situation. This preferred choice of the AS/HFA group in this item is consistent with the hypothesis of an underlying empathiing deficit in ASC (Baron-Cohen, Wheelwright, Lawson, Griffin, & Hill, 2002; Lawson, Baron-Cohen, & Wheelwright, 2004). Though the label picked by the majority of participants with ASC is relevant, it fails to make use of available cues in the scene, namely the emotion the protagonist expressed with her eyes. Difficulties attributing mental and emotional states on the basis of information from the eyes have been found among children and adults with ASC in several studies (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001; Baron-Cohen, Wheelwright, & Jolliffe, 1997; Klin, Jones, Schultz, Volkmar, & Cohen, 2002b). Neurologically, interpretation of emotions from the eye region is associated with amygdala and superior temporal gyrus activation. Individuals with ASC show reduced activation in these brain areas when attributing emotions and mental states stimulated by pictures of the eye region (Baron-Cohen et al., 2000; Baron-Cohen et al., 1999).

As in the RMF-A, the errors participants with ASC made on the RMF-C demonstrate the subtlety of the errors children with ASC often make when interpreting others' mental states. However, such subtle errors could lead to wrong interpretation of an interaction or to inappropriate responses in real life situations. Using the RMF-A and RMF-C in the 3rd generalisation level of the intervention evaluation study is used to assess whether subtle errors of this kind could be overcome through training that places extra attention to the face and eyes and to the intonation in vocal emotional expressions. The evaluation design is quite challenging to the tasks' power, if the task is able to discriminate between an ASC groups that took the intervention and ASC groups that did not.

Appendix 4 – The adult study follow-up questionnaire

<u>Emotion Recognition Project – Follow up Questionnaire</u>

Name:____________________________________ Today's date:________________

Please try to estimate how much time you have spent using *Mind Reading* <u>after</u> the study was finished and for how long (For example: an hour every month for 6 months, 2 hours a week for a month, 3 hours overall, never):

____

____

When was the last time you used *Mind Reading* <u>after</u> the study was finished? (for example: yesterday, last week, 3 months ago, never):

____

____

If you have been using *Mind Reading*, please try to tell how helpful it has been to your ability to do the following (For each item, please tick the appropriate box):

How helpful was *Mind Reading* to your ability to:	Very Helpful	Quite Helpful	Not very helpful	Not helpful at all
Recognise emotions in every day life				
Understand social situations				
Confidence in social situations				
Other (Please specify): ________________________ ____				

In this space, please try to describe how relevant and useful you found *Mind Reading* in your everyday life and in which areas. If you need more space, please enclose another sheet of paper:

____

____

____

____

____

Following is a questionnaire that relates to your personal life. You have completed this questionnaire before the study, and we would thank you for completing it again as a follow up. Please try to answer all of the questions. Thank you.

For each of the following questions, tick the box next to the statement which most applies to you.

1. a I have one or two particular best friends. ☐
   b I have several friends who I would call best friends. ☐
   c I don’t have anybody who I would call a best friend. ☐

2. a The most important thing about a friendship is having somebody to confide in. ☐
   b The most important thing about a friendship is having somebody to have fun with. ☐

3. a If I had to pick, I would rather have a friend who enjoys doing the same things as me than a friend who feels the same way about life as I do. ☐
   b If I had to pick, I would rather have a friend who feels the same way about life as I do, than a friend who enjoys doing the same things as me. ☐

4. a I like to be close to people. ☐
   b I like to keep my distance from people. ☐
5. a When I talk with friends on the phone, it is usually to make arrangements rather than to chat. ☐
   b When I talk with friends on the phone, it is usually to chat rather than to make arrangements. ☐
6. a I tend to think of an activity I want to do and then find somebody to do it with. ☐
   b I tend to arrange to meet somebody and then think of something to do. ☐
7. a I prefer meeting a friend for a specific activity, e.g. going to the cinema, playing golf. ☐
   b I prefer meeting a friend for a chat, e.g. at a pub, at a café. ☐

8. a If I moved to a new area, I would put more effort into staying in touch with old friends than making new friends. ☐
   b If I moved to a new area, I would put more effort into making new friends than staying in touch with old friends. ☐

9. a My friends value me more as someone who is a support to them than as someone to have fun with. ☐
   b My friends value me more as someone to have fun with than as someone who is a support to them. ☐

10. a If a friend had a problem, I would be better at discussing their feelings about the problem than coming up with practical solutions. ☐
    b If a friend had a problem, I would be better at coming up with practical solutions than discussing their feelings about the problem. ☐

11. a If a friend was having personal problems, I would wait for them to contact me as I wouldn't want to interfere. ☐
    b If a friend was having personal problems, I would contact them to discuss the problem. ☐

12. a When I have a personal problem, I feel that it is better to work it out on my own. ☐
    b When I have a personal problem, I feel that it is better to share it with a friend. ☐
    c When I have a personal problem, I feel that it is better to try and forget about it. ☐

13. a If I have to say something critical to a friend, I think it's best to broach the subject gently. ☐
    b If I have to say something critical to a friend, I think it's best to just come right out and say it. ☐

14. If I fell out with a good friend and I thought that I hadn't done anything wrong, I would
    a do whatever it takes to repair the relationship. ☐
    b be willing to make the first move, as long as they reciprocated. ☐
    c be willing to sort out the problem, if they made the first move. ☐
    d not feel able to be their <u>close</u> friend anymore. ☐

15. My ideal working space would be
    - a in an office on my own, without any visitors during the day. ☐
    - b in an office on my own, with an occasional visitor during the day. ☐
    - c in an office with one or two others. ☐
    - d in an open plan office. ☐

For the next set of questions, please tick the box to indicate your answer.

16. How easy do you find discussing your feelings with your friends?

Very easy ☐ Quite easy ☐ Not very easy ☐
Quite difficult ☐ Very difficult ☐

17. How easy would you find it to discuss your feelings with a stranger?

Very easy ☐ Quite easy ☐ Not very easy ☐
Quite difficult ☐ Very difficult ☐

18. In terms of <u>personality</u>, how similar to your friends do you tend to be?

Very similar ☐ Quite similar ☐
Not very similar ☐ Very dissimilar ☐

19. In terms of <u>interests</u>, how similar to your friends do you tend to be?

Very similar ☐ Quite similar ☐
Not very similar ☐ Very dissimilar ☐

20. How important is it to you what your friends think of you?

Of no importance ☐ Of little importance ☐ Fairly important ☐
Very important ☐ Of upmost importance ☐

21. How important is it to you what strangers think of you?

Of no importance ☐ Of little importance ☐ Fairly important ☐
Very important ☐ Of upmost importance ☐

22. How easy do you find it to admit to your friends when you're wrong?

Very easy ☐ Quite easy ☐ Not very easy ☐
Quite difficult ☐ Very difficult ☐

23. How easy to do you find it to tell a friend about your weaknesses and failures?

Very easy ☐ Quite easy ☐ Not very easy ☐
Quite difficult ☐ Very difficult ☐

24. How easy do you find it to tell a friend about your achievements and successes?

Very easy ☐ Quite easy ☐ Not very easy ☐
Quite difficult ☐ Very difficult ☐

25. How interested are you in the everyday details (e.g. their relationships, family, what's currently going on in their lives) of your <u>close</u> friends' lives?

Completely disinterested ☐ Not very interested ☐
Quite interested ☐ Very interested ☐

26. How interested are you in the everyday details (e.g. their relationships, family, what's currently going on in their lives) of your <u>casual</u> friends' lives?

Completely disinterested ☐ Not very interested ☐
Quite interested ☐ Very interested ☐

27. When you are in a group, e.g. at work, school, church, parent group etc., how important is it for you to know the "gossip" e.g. who dislikes who, who's had a relationship with who, secrets.

Of no importance ☐ Of little importance ☐ Fairly important ☐
Very important ☐ Of great importance ☐

28. Do you work harder at your career than at maintaining your relationships with friends?

Yes ☐ No ☐ Equal ☐

29. How often do you make plans to meet with friends?

Once or twice a year ☐
Once every 2 or 3 months ☐
Once a month ☐
Once every couple of weeks ☐
Once or twice a week ☐
3 or 4 times a week ☐
More than any of the above ☐

30. How would you prefer to keep in touch with friends?

(Please put: 1 in the box next to your most preferred method
2 in the box next to your second preference
3 in the box next to your third preference)

Face to face contact ☐
Email/letters ☐
Telephone calls ☐

31. How easy to do you find it to make new friends?

Very easy ☐ Quite easy ☐ Not very easy ☐
Quite difficult ☐ Very difficult ☐

32. What would be the <u>minimum</u> social contact you would need to get through a <u>day</u>?

No contact – I don’t get lonely ☐
Just being near to people, even if I am not talking to them ☐
A casual chat, e.g., with a shop assistant or hairdresser ☐
A chat with a friend ☐
Two or three chats with friends during the day ☐
More than any of the above ☐

33. What would be the minimum social contact you would need to get through a week?

None – I don't get lonely ☐
Being around people, even if I wasn't talking to them ☐
Casual chats, e.g. with a shop assistant or hairdresser ☐
One chat with a friend ☐
Two or three chats during the week with friends ☐
One chat every day with a friend ☐
Two or three chats every day with a friend ☐
More than any of the above ☐

34. When talking with friends, what proportion of your time do you spend talking about the following:

(*Please put: 1 in the box next to the topic that you talk most about,*
2 in the box next to the topic you talk next most about, etc, through to
7 in the box next to the topic you talk least about.
Use each number only once i.e. there should be no ties.)

Politics and current affairs ☐
Hobbies and interests (eg. sport, TV, music, cinema, fashion, holidays, gardening, DIY etc.) ☐
Personal matters (e.g. life choice decisions, arguments, feelings) ☐
Work ☐
Family and friends ☐
The weather ☐
What you've been doing since last time you spoke ☐

35. At social occasions, when you meet someone for the first time, how likely are you to talk about the following.

(Please put: 1 in the box next to the topic that you talk most about,

2 in the box next to the topic you talk next most about, etc, through to

7 in the box next to the topic you talk least about.

Use each number only once i.e. there should be no ties.)

Politics and current affairs ☐

Hobbies and interests (e.g. sport, TV, music, cinema, fashion, holidays, gardening, DIY etc.) ☐

Personal matters (e.g. life choice decisions, arguments, feelings) ☐

Work ☐

Family and friends ☐

The weather ☐

What you've been doing recently ☐

Thank you for completing this questionnaire

Appendix 5 – The child study follow-up questionnaire

Emotion Recognition Project – Follow up Questionnaire

Child's Name:____________________________ Today's date:____________________

Who is filling in this questionnaire: ____

Please try to estimate how much time your child spent using *Mind Reading* since the study was finished and for how long (For example: an hour every month for 6 months, 2 hours a week for a month, 3 hours overall, never):

____
____
____
____
____

When was the last time your child used *Mind Reading* since the study was finished? (for example: yesterday, last week, 3 months ago, never):

____
____
____
____

If your child has been using *Mind Reading*, please try to tell how helpful it has been to his/her ability to do the following (For each item, please tick the appropriate box, on a scale from Very helpful to Not helpful at all):

How helpful was *Mind Reading* to his/her ability to:	Very Helpful 5	4	3	2	Not helpful at all 1
Recognise emotions in every day life					
Understand social situations					
Use more emotion words in his/her speech					
Have confidence in social situations					
Other (Please specify): _________________________ ______________________________					

In this space, please try to describe how relevant and useful you and your child found *Mind Reading* in his/her everyday life and in which areas. If you need more space, please enclose another sheet of paper:

____
____
____
____
____
____
____
____
____

VABS-S

Please read the following questions carefully, and tick the appropriate answer for your child's CURRENT functioning.

	Does your child do the following:	Yes, usually	Sometimes or partially	No, never
AS 1	Shows a desire to please			
IS 2	Chooses appropriate presents			
AS 3	Takes turns in conversations			
IS 4	Responds to hints and indirect cues in conversation			
AS 5	Shares toys when asked			
IS 6	Makes confidences			
AS 7	Recognises happiness and sadness in others			
IS 8	Recognises surprise and embarrassment in others			
AS 9	Initiates social contact			
IS 10	Initiates conversation of interest to others			
AS 11	Initiates fixed small talk			
IS 12	Initiates flexible small talk			
AS 13	Uses appropriate table manners			
IS 14	Supplies important missing information			
AS 15	Delivers a simple message			
IS 16	Expresses ideas in more than one way			
AS 17	Says 'Please' when asking for something			

IS 18	Refrains from statements that might embarrass			
	Does your child do the following:	Yes, usually	Sometimes or partially	No, never
AS 19	Names favourite TV programs and times			
IS 20	Engages in elaborate make-believe activities			
AS 21	Asks permission to play with a toy			
IS 22	Knows behaviour appropriate for different people			
AS 23	Plays board games			
IS 24	Plays hide and seek or cheat appropriately			
AS 25	Follows time limits set by care-giver			
IS 26	Has realistic long range goals and plans			
AS 27	Responds appropriately when introduced			
IS 28	Keeps secrets for as long as appropriate			
AS 29	Apologises for errors			
IS 30	Apologises for hurting other’s feelings			
AS 31	Returns borrowed items			
IS 32	Weighs consequences of actions			

Thank you for completing this questionnaire.

Appendix 6 – Protocol for the software and tutor group – Experiment 2

<u>Objectives</u>

1. Enhancing generalisation of self-learned material from *Mind Reading* to everyday life.
2. Structuring the learning of emotions and putting it into context.
3. Adding complementary information about emotion recognition, that is not included in *Mind Reading*.
4. Supporting the use of *Mind Reading* by the participants.
5. Encouraging interaction among participants regarding emotional understanding.

<u>Design</u>

- The course will last ten weekly sessions of two hours each.
- There will be maximum six participants in a group, unless more than one tutor is present.
- The meetings will be held in a room that allows small group discussions as well as individual work by the participants on personal computers.
- The first meeting will be dedicated to introduction of *Mind Reading*, its installation and use, and to the working format in the group and meeting dates. Participants will be asked to go through the software at home according to level and define the level or concepts that each finds challenging.
- Meetings 2-9 will deal with emotion recognition issues brought up by the participants following their self-use of *Mind Reading* and by the tutor according to the group's level and needs. The meetings will be divided to three main parts:
  i. Opening group session: will deal with the passing week's homework, and discuss a common issue brought up by the tutor according to group's level. Questions and suggestions raised by participants, which are relevant to the whole group will also be addressed (recommended time for this session is 60 minutes).
  ii. Individual work: participants will individually work with the software. The tutor will spend some time with each participant to address questions,

which are relevant to the participant's personal progress and difficulties. In addition to the help given by the tutor, advanced users will be encouraged to support less able users (recommended time – 30 minutes). Note that the extra merit of this part is the one on one tutoring and the peer tutoring of the group members and not self-use of the software (which is done by participants at home).

iii. Closing group session – the tutor will introduce an issue to be learned and practiced during the week for the next meeting (recommended time – 10 min.).

Note: The time spent on each of these activities depends on the tutor's discretion and the participants' requests.

- ☒ The last meeting will be dedicated to summary of the learning in the group, recommendations for future self-use and participants' feedback.
- ☒ Tutor will be asked to write down his/her impression of each meeting and also to record attendance.

<u>Recommendation for activities at home and at group sessions:</u>
All these activities will be discussed with the group, and preferably brought up by the participants, further to their home work. Some of these activities may be given as homework.

- ❖ Find shared features in face/voice of each emotion.
- ❖ Show various life events where the discussed emotions are experienced.
- ❖ Relate to body posture/gestures that accompany discussed emotions.
- ❖ Relate to the way other people view the discussed emotion and react to it.
- ❖ Distinguish between different emotions in the groups according to:
    i. Feature differences
    ii. Meaning and impact of the emotions
    iii. Intensity of the emotion (eg- shocked more intense than surprised)
    iv. Sort of situation the emotion is experienced in.
- ❖ Collect examples of representation of the different emotions as seen by participants on:
    i. Media: TV, Papers, etc

ii. Personal photographs

- ❖ Discuss participants' personal experiences, related to the discussed emotions. Role-playing is also welcome, though dependent on group's co-operation.

Important note: As a general theme, the tutor should stress the importance of integration of all data sources (ie face, voice, body, context) for correct recognition on one hand and on the other hand, of the inability to always tell the true feeling of a person, due to the "Mind Reading" skill being inaccurate on many occasions (discuss issues of deception, courtesy, "face value" etc, that may prevent us from recognising the correct emotion).

References

Adolphs, R. (2001). The neurobiology of social cognition. *Current Opinions in Neurobiology, 11* (2), 231-239.

Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinions in Neurobiology, 12* (2), 169-177.

Adolphs, R., Baron-Cohen, S., & Tranel, D. (2002). Impaired recognition of social emotions following amygdala damage. *Journal of Cognitive Neuroscience, 14* (8), 1264-1274.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433* (7021), 68-72.

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience, 13*, 232-240.

Adolphs, R., & Tranel, D. (1999). Intact recognition of emotional prosody following amygdala damage. *Neuropsychologia, 37*, 1285-1292.

Adolphs, R., & Tranel, D. (2003). Amygdala damage impairs emotion recognition from scenes only when they contain facial expressions. *Neuropsychologia, 41* (10), 1281-1289.

Adolphs, R., Tranel, D., & Damasio, A. R. (2003). Dissociable neural systems for recognizing emotions. *Brain and Cognition, 52*, 61-69.

Alcantara, J. I., Weisblatt, E. J., Moore, B. C., & Bolton, P. F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *Journal of Child Psychology and Psychiatry, 45* (6), 1107-1114.

American-Psychiatric-Association. (1994). *DSM-IV Diagnostic and Statistical Manual of Mental Disorders, 4th Edition*. Washington DC: American Psychiatric Association.

Anderson, V. (2001). *Developmental neuropsychology : a clinical approach*. Hove: Psychology Press.

Ashwin, C., Baron-Cohen, S., Wheelwright, S., O'Riordan, M., & Bullmore, E. T. (in press). Differential activation of the social brain during emotion processing in adults with autism. *Neuropsychologia*.

Ashwin, C., Wheelwright, S., & Baron-Cohen, S. (2005). Laterality biases to chimeric faces in Asperger syndrome: what is 'right' about face-processing? *Journal of Autism and Developmental Disorders, 35* (2), 183-196.

Asperger, H. (1944). Die "autistischen psychopathen" in kind esalter. *Archive fur Psychiatrie nnd Nervenkrankheiten, 117*, 76-136.

Attwood, T. (2000). Strategies for improving the social integration of children with Asperger syndrome. *Autism, 4* (1), 85-100.

Attwood, T. (2003). Understanding and managing circumscribed interests. In M. R. Prior (Ed.), *Learning and behavior problems in Asperger syndrome* (pp. 126-147). New York ; London: Guilford Press.

Bakeman, R., & Adamson, L. (1984). Coordinating Attention to people and objects in mother-infant and peer-infant interactions. *Child Development, 55*, 1278-1289.

Baron-Cohen, S. (1989). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30* (2), 285-297.

Baron-Cohen, S. (1991). Do people with autism understand what causes emotion? *Child Development, 62* (2), 385-395.

Baron-Cohen, S. (1992). Out of sight or out of mind? Another look at deception in autism. *Journal of Child Psychology and Psychiatry, 33*, 1141-1155.

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

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*, 248-254.

Baron-Cohen, S. (2003). *The Essential Difference : men, women and the extreme male brain*. London: Penguin,.

Baron-Cohen, S. (2006). Two new theories of autism: hyper-systemising and assortative mating. *Archives of Disease in Childhood, 91* (1), 2-5.

Baron-Cohen, S., & Belmonte, M. K. (2005). Autism: A Window Onto the Development of the Social and the Analytic Brain. *Annual Review of Neuroscience, 28*, 109-126.

Baron-Cohen, S., Bolton, P., Wheelwright, S., Scahill, V. L., Short, L., Mead, G., et al. (1998). Autism occurs more often in families of physicists, engineers, and mathematicians. *Autism, 2*, 296-301.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., & Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology, 13*, 379-398.

Baron-Cohen, S., Golan, O., Hill, J. J., & Wheelwright, S. (submitted). Developmental changes in the emotion lexicon.

Baron-Cohen, S., Golan, O., Wheelwright, S., & Hill, J. J. (2004). Mind Reading: the interactive guide to emotions. London: Jessica Kingsley Ltd (www.jkp.com).

Baron-Cohen, S., & Goodhart, F. (1994). The "seeing-leads-to-knowing" deficit in autism: The Pratt and Bryant probe. *British Journal of Developmental Psychology, 12* (3), 397-401.

Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The Autism-Spectrum Quotient (AQ)-Adolescent Version. *Journal of Autism and Developmental Disorders, 36* (3), 343-350.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: evidence from very high functioning adults with autism or asperger syndrome. *Journal of Child Psychology and Psychiatry, 38*, 813-822.

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., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and Intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology, 4* (2), 113-125.

Baron-Cohen, S., O'Riordan, M., Stone, V. E., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children with asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders, 29* (5), 407-418.

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The Systemising Quotient (SQ): An investigation of adults with Asperger Syndrome or High Functioning Autism and normal sex differences. *Philosophical Transactions of the Royal Society, Series B, Special issue on "Autism : Mind and Brain", 358*, 361-374.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Wheelwright, S., Ashwin, C., & Williams, S. C. (2000). The amygdala theory of autism. *Neuroscience and Biobehavioral Reviews, 24*, 355-364.

Baron-Cohen, S., Ring, H. A., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. *British Journal of Psychiatry, 165* (5), 640-649.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., et al. (1999). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience, 11*, 1891-1898.

Baron-Cohen, S., Scahill, V. L., Izaguirre, J., Hornsey, H., & Robertson, M. M. (1999). The prevalence of Gilles de la Tourette syndrome in children and adolescents with autism: a large scale study. *Psychologiccal Medicine, 29* (5), 1151-1159.

Baron-Cohen, S., Spitz, A., & Cross, P. (1993). Can children with autism recognize surprise? *Cognition and Emotion, 7*, 507-516.

Baron-Cohen, S., & Wheelwright, S. (1999). 'Obsessions' in children with autism or Asperger syndrome. Content analysis in terms of core domains of cognition. *British Journal of Psychiatry, 175*, 484-490.

Baron-Cohen, S., & Wheelwright, S. (2003). The Friendship Questionnaire: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 33* (5), 509-517.

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., Hill, J. J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42*, 241-251.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger syndrome. *Visual Cognition, 4*, 311-331.

Baron-Cohen, S., Wheelwright, S., Lawson, J., Griffin, R., & Hill, J. J. (2002). The exact mind: Empathizing and systemizing in autism spectrum conditions. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 491-508). Malden, MA, US: Blackwell Publishers.

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The Adult Asperger Assessment (AAA): a diagnostic method. *Journal of Autism and Developmental Disorders, 35* (6), 807-819.

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*, 5-17.

Baron-Cohen, S., Wheelwright, S., Spong, A., Scahill, V. L., & Lawson, J. (2001). Are intuitive physics and intuitive psychology independent? A test with children with Asperger Syndrome. *Journal of Developmental and Learning Disorders, 5*, 47-78.

Baron-Cohen, S., Wheelwright, S., Stone, V. E., & Rutherford, M. (1999). A mathematician, a physicist and a computer scientist with Asperger syndrome:

Performance on folk psychology and folk physics tests. *Neurocase, 5* (6), 475-483.

Baron-Cohen, S., Wheelwright, S., Stott, C. M., Bolton, P., & Goodyer, I. (1997). Is there a link between engineering and autism? *Autism, 1* (1), 101-109.

Barry, T. D., Klinger, L. G., Lee, J. M., Palardy, N., Gilmore, T., & Bodin, S. D. (2003). Examining the effectiveness of an outpatient clinic-based social skills group for high-functioning children with autism. *Journal of Autism and Developmental Disorders, 33* (6), 685-701.

Bauminger, N. (2002). The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: Intervention outcomes. *Journal of Autism and Developmental Disorders, 32* (4), 283-298.

Bauminger, N. (2004). The expression and understanding of jealousy in children with autism. *Development and Psychopathology, 16* (1), 157-177.

Behrmann, M., Avidan, G., Leonard, G. L., Kimchi, R., Luna, B., Humphreys, K., et al. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia, 44* (1), 110-129.

Bell, A. J. W. (Writer) (1999). Lost for words. UK: Yorkshire Television (YTV) PLC.

Belmonte, M. K., Allen, G., Beckel-Mitchener, A., Boulanger, L. M., Carper, R. A., & Webb, S. J. (2004). Autism and abnormal development of brain connectivity. *Journal of Neuroscience, 24* (42), 9228-9231.

Belmonte, M. K., & Yurgelun-Todd, D. A. (2003). Functional anatomy of impaired selective attention and compensatory processing in autism. *Cognitive Brain Research, 17* (3), 651-664.

Ben-Ze'ev, A. (2000). *The subtlety of emotions*. Cambridge, Mass. ; London: MIT Press.

Ben Shalom, D. (2003). Memory in autism: Review and synthesis. *Cortex, 39* (4-5), 1129-1138.

Bennett, M., & Matthews, L. (2000). The Role of Second-Order Belief-Understanding and Social Context in Children's Self-Attribution of Social Emotions. *Social Development, 9* (1), 126-130.

Bennetto, L., Pennington, B. F., & Rogers, S. J. (1996). Intact and impaired memory functions in autism. *Child Development, 67* (4), 1816-1835.

Bernard-Opitz, V., Sriram, N., & Nakhoda-Sapuan, S. (2001). Enhancing social problem solving in children with autism and normal children through computer-assisted instruction. *Journal of Autism and Developmental Disorders, 31* (4), 377-398.

Berthoz, S., Armony, J. L., Blair, R. J., & Dolan, R. J. (2002). An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain, 125* (Pt 8), 1696-1708.

Bettelheim, B. (1967). *The empty fortress : infantile autism and the birth of the self*: Free P.; Collier-Macmillan.

Bishop, J. (2003). The Internet for educating individuals with social impairments. *Journal of Computer Assisted Learning, 19* (4), 546-556.

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & Dolan, R. J. (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain, 122* (5), 883-893.

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience, 2* (8), 561-567.

Bolt, B. (Writer) (1999). The Turn of the Screw. UK: Granada International.

Bölte, S., Feineis-Matthews, S., Leber, S., Dierks, T., Hubl, D., & Poustka, F. (2002). The development and evaluation of a computer-based program to test and to teach the recognition of facial affect. *International Journal of Circumpolar Health, 61 Suppl 2*, 61-68.

Bölte, S., Hubl, D., Feineis-Matthews, S., Prvulovic, D., Dierks, T., & Poustka, F. (2006). Facial Affect Recognition Training in Autism: Can We Animate the Fusiform Gyrus? *Behavioral Neuroscience, 120* (1), 211-216.

Bolton, P. F., & Griffiths, P. D. (1997). Association of tuberous sclerosis of temporal lobes with autism and atypical autism. *Lancet, 349* (9049), 392-395.

Bormann-Kischkel, C., Vilsmeier, M., & Baude, B. (1995). The development of emotional concepts in autism. *Journal of Child Psychology and Psychiatry, 36*, 1243-1259.

Bosseler, A., & Massaro, D. W. (2003). Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with autism. *Journal of Autism and Developmental Disorders, 33* (6), 653-672.

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

Bowler, D. M., Matthews, N. J., & Gardiner, J. M. (1997). Asperger's syndrome and memory: Similarity to autism but not amnesia. *Neuropsychologia, 35* (1), 65-70.

Bowler, D. M., & Strom, E. (1998). Elicitation of first-order "theory of mind" in children with autism. *Autism, 2* (1), 33-44.

Bretherton, I., & Beeghly, M. (1982). Talking about internal states: The acquisition of an explicit theory of mind. *Developmental Psychology, 18* (6), 906-921.

Brewer, K. (2002). *Research methods in clinical psychology, psychiatry, and the mental health professions*. Grays: Orsett Psychological Services.

Brosnan, M. J., Scott, F. J., Fox, S., & Pye, J. (2004). Gestalt processing in autism: failure to process perceptual relationships and the implications for contextual understanding. *Journal of Child Psychology and Psychiatry, 45* (3), 459-469.

Brothers, L., & Ring, B. (1992). A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience, 4* (2), 107-118.

Brown, C. (Writer) (1946). The Yearling. USA: MGM Home Entertainment.

Buitelaar, J. K., van der Wees, M., Swaab-Barneveld, H., & van der Gaag, R. J. (1999). Verbal memory and performance IQ predict theory of mind and emotion recognition ability in children with autistic spectrum disorders and in psychiatric control children. *Journal of Child Psychology and Psychiatry, 40* (6), 869-881.

Calder, A. J. (2003). Disgust discussed. *Annals of Neurology, 53*, 427-428.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience, 3*, 1077-1078.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Neuroscience, 2*, 352-363.

Calder, A. J., Young, A. W., Rowland, D., Perrett, D. I., Hodges, J. R., & Etcoff, N. L. (1996). Facial Emotion Recognition after Bilateral Amygdala Damage: Differentially Severe Impairment of Fear. *Cognitive Neuropsychology, 13* (5), 699-745.

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*, 1169-1182.

Caron, A. J., Caron, R. F., & MacLean, D. J. (1988). Infant discrimination of naturalistic emotional expressions: the role of face and voice. *Child Development, 59* (3), 604-616.

Castelli, F. (2005). Understanding emotions from standardized facial expressions in autism and normal development. *Autism, 9* (4), 428-449.

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* (8), 1839-1849.

Celani, G., Battacchi, M. W., & Arcidiacono, L. (1999). The understanding of the emotional meaning of facial expressions in people with autism. *Journal of Autism and Developmental Disorders, 29* (1), 57-66.

Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *JAMA, 285* (24), 3093-3099.

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: confirmation of high prevalence. *American Journal of Psychiatry, 162* (6), 1133-1141.

Courchesne, E., & Pierce, K. (2005). Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. *Current Opinion in Neurobiology, 15* (2), 225-230.

Critchley, H. D., Daly, E. M., Bullmore, E. T., Williams, S. C., Van Amelsvoort, T., Robertson, D. M., et al. (2000). The functional neuroanatomy of social behaviour: changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain, 123 (Pt 11)*, 2203-2212.

Cuarón , A. (Writer) (1995). A little princess. USA: Warner Home Video.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*.

Damasio, A. R. (1995). *Descartes' error: emotion, reason and the human brain*. London: Picador.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., et al. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience, 9* (1), 28-30.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray.

Davis, M. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44* (1), 113-126.

Dawson, G., Meltzoff, A. N., Osterling, J., & Rinaldi, J. (1998). Neuropsychological correlates of early symptoms of autism. *Child Dev, 69* (5), 1276-1285.

Dawson, G., Munson, J., Estes, A., Osterling, J., McPartland, J., Toth, K., et al. (2002). Neurocognitive function and joint attention ability in young children with autism spectrum disorder versus developmental delay. *Child Development, 73* (2), 345-358.

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: insights from behavioral and electrophysiological studies. *Developmental Neuropsychology, 27* (3), 403-424.

Dawson, G., & Zanolli, K. (2003). Early intervention and brain plasticity in autism. In G. Bock & J. Goode (Eds.), *Autism: Neural bases and treatment possibilities* (Vol. 251, pp. 266-280). Chichester, UK: John Wiley & Sons Ltd.

De Sonneville, L. M., Verschoor, C. A., Njiokiktjien, C., Op het Veld, V., Toorenaar, N., & Vranken, M. (2002). Facial identity and facial emotions: speed, accuracy, and processing strategies in children and adults. *Journal of Clinical and Experimental Neuropsychology, 24* (2), 200-213.

De Waal, F. B. M. (1998). *Chimpanzee politics : power and sex among apes* (Rev. ed.). Baltimore ; London: Johns Hopkins University Press.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3* (2), 71-100.

Denham, S. A. (1998). *Emotional development in young children*. New York ; London: Guilford Press.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, Mass: MIT Press/Bradford Books.

Dennis, M., Lockyer, L., & Lazenby, A. L. (2000). How high-functioning children with autism understand real and deceptive emotion. *Autism, 4* (4), 370-381.

Deruelle, C., Rondan, C., Gepner, B., & Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders, 34* (2), 199-210.

Dorris, L., Espie, C. A., Knott, F., & Salt, J. (2004). Mind-reading difficulties in the siblings of people with Asperger's syndrome: evidence for a genetic influence in the abnormal development of a specific cognitive domain. *Journal of Child Psychology and Psychiatry, 45* (2), 412-418.

Dyck, M. J., Ferguson, K., & Shochet, I. M. (2001). Do autism spectrum disorders differ from each other and from non-spectrum disorders on emotion recognition tests? *European Child & Adolescent Psychiatry, 10*, 105-116.

Dziobek , I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand, M., et al. (2006). Introducing MASC: A Movie for The Assessment of Social Cognition. *Journal of Autism and Developmental Disorders*, 36(5), 623-636.

Ekman, P. (1999). Basic emotions. In T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion* (pp. 45-60). Sussex, UK: John Wiley & Sons Ltd.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*, 124-129.

Ekman, P., Sorensen, E. R., & Friesen, W. V. (1969). Pan-Cultural Elements in Facial Displays of Emotion. *Science, 164*, 86–88.

el Kaliouby, R., & Robinson, P. (2005). The emotional hearing aid: an assistive tool for children with Asperger syndrome. *Universal Access in the Information Society, 4* (2), 121-134.

Fein, D., Lucci, D., Braverman, M., & Waterhouse, L. (1992). Comprehension of affect in context in children with pervasive developmental disorders. *Journal of Child Psychology and Psychiatry, 33* (7), 1157-1167.

Feinberg, T. E., Rifkin, A., Schaffer, C., & Walker, E. (1986). Facial discrimination and emotional recognition in schizophrenia and affective disorders. *Archives of General Psychiatry, 43* (3), 276-279.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*.

Feshbach, N. D. (1982). Studies of empathic behavior in children. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 315-338). New York: Academic Press.

Festinger, L. (1962). Cognitive dissonance. *Sci Am, 207*, 93-102.

Fisher, N., & Happe, 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.

Flavell, J. H. (1999). Cognitive development: Children's knowledge about the mind. *Annual Review of Psychology, 50*, 21-45.

Folstein, S. E., & Rosen-Sheidley, B. (2001). Genetics of autism: complex aetiology for a heterogeneous disorder. *Nature Reviews Genetics, 2* (12), 943-955.

Forster, K. L., & Forster, J. C. (2003). DMDX: a windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116-124.

Frith, U. (1989). *Autism : explaining the enigma*. Oxford: Blackwell.

Frith, U. (2003). *Autism : explaining the enigma* (2nd ed. ed.). Blackwell Pub.: New York.

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 358* (1431), 459-473.

Frith, U., Happe, F. G., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development, 3* (2), 108-124.

Frith, U., & Hill, E. (2004). *Autism: mind and brain*. Oxford: Oxford University Press.

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology, 1*, 329–342.

Gallese, V. (2003). The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology, 36* (4), 171-180.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119 (Pt 2)*, 593-609.

Gepner, B., Deruelle, C., & Grynfeltt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders, 31*, 37-45.

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., et al. (2004). Abnormal cortical voice processing in autism. *Nature Neuroscience, 7* (8), 801-802.

Gillberg, C., & Billstedt, E. (2000). Autism and Asperger syndrome: coexistence with other clinical disorders. *Acta Psychiatrica Scandinavica, 102* (5), 321-330.

Gillberg, C. L. (1992). The Emanuel Miller Memorial Lecture 1991. Autism and autistic-like conditions: subclasses among disorders of empathy. *Journal of Child Psychology and Psychiatry, 33* (5), 813-842.

Gomot, M., Giard, M. H., Adrien, J. L., Barthelemy, C., & Bruneau, N. (2002). Hypersensitivity to acoustic change in children with autism: electrophysiological evidence of left frontal cortex dysfunctioning. *Psychophysiology, 39* (5), 577-584.

Gordon, B. (2000). Commentary: a neural systems perspective for improving behavioral treatments for autism. *Journal of Autism and Developmental Disorders, 30* (5), 503-508.

Grandin, T. (1995). *Thinking in Pictures*. New York: Doubleday.

Griffith, E. M., Pennington, B. F., Wehner, E. A., & Rogers, S. J. (1999). Executive functions in young children with autism. *Child Development, 70* (4), 817-832.

Griffiths, P. (1997). *What emotions really are : the problem of psychological categories*. Chicago ; London: University of Chicago Press.

Griffiths, P. (2003). Basic Emotions, Complex Emotions, Machiavellian Emotions. In A. Hatzimoysis (Ed.), *Philosophy and the Emotions* (pp. 39-68). Cambridge: Cambirdge University Press.

Gross, A. L., & Ballif, B. (1991). Children's understanding of emotion from facial expressions and situations: A review. *Developmental Review, 11* (4), 368-398.

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

Gur, R. C., Erwin, R. J., Gur, R. E., Zwil, A. S., Heimberg, C., & Kraemer, H. C. (1992). Facial emotion discrimination: II. Behavioral findings in depression. *Psychiatry Research, 42* (3), 241-251.

Hadjikhani, N., Joseph, R. M., Snyder, J., & Tager-Flusberg, H. (2005). Anatomical Differences in the Mirror Neuron System and Social Cognition Network in Autism. *Cerebral Cortex*, 10.1093/cercor/bhj1069.

Hadwin, J., Baron-Cohen, S., Howlin, P., & Hill, K. (1996). Can we teach children with autism to understand emotions, belief, or pretence? *Development and Psychopathology, 8* (2), 345-365.

Hall, G. B., Szechtman, H., & Nahmias, C. (2003). Enhanced salience and emotion recognition in Autism: a PET study. *American Journal of Psychiatry, 160* (8), 1439-1441.

Hall, J. A. (1984). *Non-verbal sex differences : communication, accuracy and expressive style*. Baltimore ; London: Johns Hopkins University Press.

Happe, F., & Frith, U. (1996). The neuropsychology of autism. *Brain, 119* (4), 1377-1400.

Happe, F. G. (1994a). 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 Developmental Disorders, 24*, 129-154.

Happe, F. G. (1994b). Annotation: Current psychological theories of autism: The "Theory of Mind" account and rival theories. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 35* (2), 215-229.

Happe, 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.

Happe, 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.

Happe, F. G. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology, 15* (Pt 1), 1-12.

Happe, F. G. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences, 3* (6), 216-222.

Happe, F. G., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., et al. (1996). 'Theory of mind' in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport, 8* (1), 197-201.

Happe, F. G., & Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology, 14*, 385-398.

Harris, P. L. (1989). *Children and emotion : the development of psychological understanding*. Oxford: Blackwell.

Harris, P. L. (1994). The child's understanding of emotion: Developmental change and the family environment. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 35* (1), 3-28.

Harris, P. L., Olthof, T., Meerum-Terwogt, M., & Hardman, C. E. (1987). Children's Knowledge of the Situations That Provoke Emotion. *International Journal of Behavioral Development, 10* (3), 319-344.

Haviland, J. M., & Lelwica, M. (1987). The induced affect response: 10-week-old infants' responses to three emotion expressions. *Developmental Psychology, 23* (1), 97-104.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry, 51* (1), 59-67.

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*, 225-236.

Herba, C., & Phillips, M. (2004). Annotation: Development of facial expression recognition from childhood to adolescence: behavioural and neurological perspectives. *Journal of Child Psychology and Psychiatry, 45* (7), 1185-1198.

Hermelin, B. M. a. O. C. W. o. M. D., & O'Connor, N. W. o. M. D. (1970). *Psychological experiments with autistic children*: Oxford: Pergamon.

Hetzroni, O. E., & Tannous, J. (2004). Effects of a Computer-Based Intervention Program on the Communicative Functions of Children with Autism. *Journal of Autism and Developmental Disorders, 34* (2), 95-113.

Hill, E. L. (2004a). Evaluating the theory of executive dysfunction in autism. *Developmental Review, 24* (2), 189-233.

Hill, E. L. (2004b). Executive dysfunction in autism. *Trends in Cognitive Sciences, 8* (1), 26-32.

Hillier, 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.

Hobson, R. P. (1986a). The autistic child's appraisal of expressions of emotion. *Journal of Child Psychology and Psychiatry, 27*, 321-342.

Hobson, R. P. (1986b). The autistic child's appraisal of expressions of emotion: a further study. *Journal of Child Psychology and Psychiatry, 27*, 671-680.

Hobson, R. P. (1993). *Autism and the development of mind*. Hove: Lawrence Erlbaum Associates.

Hobson, R. P. (1994). Understanding persons: The role of affect. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds* (pp. 204-227). London: Oxford University Press.

Hobson, R. P., Ouston, J., & Lee, A. (1988a). Emotion recognition in autism: co-ordinating faces and voices. *Psychological Medicine, 18*, 911-923.

Hobson, R. P., Ouston, J., & Lee, A. (1988b). What's in a face? The case of autism. *British Journal of Psychology, 79*, 441-453.

Hoffner, C., & Badzinski, D. M. (1989). Children's integration of facial and situational cues to emotion. *Child Development, 60* (2), 411-422.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6*, 65-70.

Howard, M. A., Cowell, P. E., Boucher, J., Broks, P., Mayes, A., Farrant, A., et al. (2000). Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. *Neuroreport, 11*, 2931-2935.

Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA ; London: Duxbury Press.

Howlin, P. (1998). Practitioner review: psychological and educational treatments for autism. *Journal of Child Psychology and Psychiatry, 39* (3), 307-322.

Howlin, P. (2000). Outcome in Adult Life for more Able Individuals with Autism or Asperger Syndrome. *Autism, 4* (1), 63-83.

Howlin, P., Baron-Cohen, S., & Hadwin, J. (1999). *Teaching children with autism to mind-read : a practical guide for teachers and parents*. Chichester: B J Wiley.

Howlin, P., & Rutter, M. (1987). *Treatment of autistic children*. Chichester: Wiley.

Howlin, P., & Yates, P. (1999). The potential effectiveness of social skills groups for adults with autism. *Autism, 3* (3), 299-307.

Hughes, C. (2001). Executive Dysfunction in autism: Its nature and implications for the everyday problems experienced by individuals with autism. In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 255-275). Mahwah, NJ, US: Lawrence Erlbaum.

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* (5), 705-718.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia, 32* (4), 477-492.

Iacoboni, M., Lieberman, M. D., Knowlton, B. J., Molnar-Szakacs, I., Moritz, M., Throop, C. J., et al. (2004). Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *Neuroimage, 21* (3), 1167-1173.

Izard, C. E., & Harris, P. L. (1995). Emotional development and developmental psychopathology. In D. J. Cohen & D. V. Cicchetti (Eds.), *Developmental psychopathology* (Vol. Vol. 1: Theory and methods, pp. 467-503). Oxford, England: John Wiley & Sons.

Jeffries , L. (Writer) (1970). The Railway Children. UK: Warner Home Video.

Jenkins, J. M., & Astington, J. W. (1996). Cognitive factors and family structure associated with theory of mind development in young children. *Developmental Psychology, 32* (1), 70-78.

Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition and Emotion, 3* (2), 81-123.

Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development : the case of face recognition*. Oxford: Blackwell.

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 38* (5), 527-534.

Jolliffe, T., & Baron-Cohen, S. (1999a). The Strange Stories Test: A replication with high-functioning adults with autism or Asperger Syndrome. *Journal of Autism and Developmental Disorders, 29*, 395-406.

Jolliffe, T., & Baron-Cohen, S. (1999b). 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.

Jolliffe, T., & Baron-Cohen, S. (2000). Linguistic processing in high-functioning adults with autism or Asperger's syndrome. Is global coherence impaired? *Psychological Medicine, 30* (5), 1169-1187.

Jolliffe, T., & Baron-Cohen, S. (2001a). A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate fragments of an object? *Cognitive Neuropsychiatry, 6* (3), 193 - 216.

Jolliffe, T., & Baron-Cohen, S. (2001b). A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate objects in context? *Visual Cognition, 8* (1), 67–101.

Jordan, R. (1999). Evaluating practice: Problems and possibilities. *Autism, 3* (4), 411-434.

Jordan, R., & Jones, G. (1999). Review of research into educational interventions for children with autism in the UK. *Autism, 3* (1), 101-110.

Joseph, R. M., & Tager-Flusberg, H. (2004). The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and Psychopathology, 16* (1), 137-155.

Joseph, R. M., & Tanaka, J. (2003). Holistic and part-based face recognition in children with autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 44* (4), 529-542.

Just, M. A., Cherkassky, V. L., Keller, T. A., & Minshew, N. J. (2004). Cortical activation and synchronization during sentence comprehension in high-functioning autism: evidence of underconnectivity. *Brain, 127* (Pt 8), 1811-1821.

Kadesjo, B., Gillberg, C., & Hagberg, B. (1999). Brief report: autism and Asperger syndrome in seven-year-old children: a total population study. *Journal of Autism and Developmental Disorders, 29* (4), 327-331.

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

Kasari, C., Chamberlain, B., & Bauminger, N. (2001). Social emotions and social relationships: Can children with autism compensate? In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 309-323). Mahwah, NJ, US: Lawrence Erlbaum.

Kasari, C., Sigman, M., Mundy, P., & Yirmiya, N. (1990). Affective sharing in the context of joint attention interactions of normal, autistic, and mentally retarded children. *Journal of Autism and Developmental Disorders, 20* (1), 87-100.

Kilts, C. D., Egan, G., Gideon, D. A., Ely, T. D., & Hoffman, J. M. (2003). Dissociable neural pathways are involved in the recognition of emotion in static and dynamic facial expressions. *Neuroimage, 18*, 156-168.

Kleinman, J., Marciano, P. L., & Ault, R. L. (2001). Advanced theory of mind in high-functioning adults with autism. *Journal of Autism and Developmental Disorders, 31*, 29-36.

Klin, A., Jones, W., Schultz, R., & Volkmar, F. (2003). The enactive mind, or from actions to cognition: lessons from autism. *Philosophical Transactions of the Royal Society, Series B, 358*, 345-360.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. J. (2002a). Defining and quantifying the social phenotype in autism. *American Journal of Psychiatry, 159*, 895-908.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. J. (2002b). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59*, 809-816.

Klinger, L. G., & Dawson, G. (1995). A fresh look at categorization abilities in persons with autism. In E. Schopler & M. G. B. (Eds.), *Learning and cognition in autism* (pp. 119-136). New York, NY, US: Plenum Press.

Klinger, L. G., & Dawson, G. (2001). Prototype formation in autism. *Dev Psychopathol, 13* (1), 111-124.

Klinnert, M. D. (1984). The regulation of infant behavior by maternal facial expression. *Infant Behavior and Development, 7* (4), 447-465.

Krasny, L., Williams, B. J., Provencal, S., & Ozonoff, S. (2003). Social skills interventions for the autism spectrum: essential ingredients and a model curriculum. *Child and Adolescent Psychiatric Clinics of North America, 12* (1), 107-122.

Kuchuk, A., Vibbert, M., & Bornstein, M. H. (1986). The perception of smiling and its experiential correlates in three-month-old infants. *Child Development, 57* (4), 1054-1061.

Kurita, H., Koyama, T., & Osada, H. (2005). Autism-Spectrum Quotient-Japanese version and its short forms for screening normally intelligent persons with pervasive developmental disorders. *Psychiatry and Clinical Neurosciences, 59* (4), 490-496.

LaCava, P., Golan, O., Baron-Cohen, S., & Myles, B. S. (in press). Using Assistive Technology to Teach Emotion Recognition to Students with Asperger Syndrome: A Pilot Study, *Remedial and Special Education*.

Lane, R. D., Reiman, E. M., Ahern, G. L., Schwartz, G. E., & Davidson, R. J. (1997). Neuroanatomical correlates of happiness, sadness, and disgust. *American Journal of Psychiatry, 154* (7), 926-933.

Lawson, J. (2003). Depth Accessibility Difficulties: An Alternative Conceptualisation of Autism Spectrum Conditions. *Journal for the Theory of Social Behaviour, 33* (2), 189-202.

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.

LeDoux, J. E. (1998). *The emotional brain: the mysterious underpinnings of emotional life*. London: Pheonix.

Leekam, S. R., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1997). Eye-direction detection: A dissociation between geometric and joint attention skills in autism. *British Journal of Developmental Psychology, 15* (Pt 1), 77-95.

Leslie, A. M. (1987). Pretense and Representation: The Origins of “Theory of Mind”. *Psychological Review, 94* (4), 412-426.

Leslie, A. M., & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology, 6*, 315-324.

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition, 43* (3), 225-251.

Leslie, K. R., Johnson-Frey, S. H., & Grafton, S. T. (2004). Functional imaging of face and hand imitation: towards a motor theory of empathy. *Neuroimage, 21* (2), 601-607.

Lewis, C., Freeman, N. H., Kyriakidou, C., Maridaki-Kassotaki, K., & Berridge, D. M. (1996). Social influences on false belief access: specific sibling influences or general apprenticeship? *Child Development, 67* (6), 2930-2947.

Liss, M., Harel, B., Fein, D., Allen, D., Dunn, M., Feinstein, C., et al. (2001). Predictors and Correlates of Adaptive Functioning in Children with Developmental Disorders. *Journal of Autism and Developmental Disorders, 31* (2), 219-230.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30* (3), 205-223.

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.

Lord, C., Wagner, A., Rogers, S., Szatmari, P., Aman, M., Charman, T., et al. (2005). Challenges in evaluating psychosocial interventions for Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders, 35* (6), 695-708; discussion 709-611.

Lotspeich, L. J., Kwon, H., Schumann, C. M., Fryer, S. L., Goodlin-Jones, B. L., Buonocore, M. H., et al. (2004). Investigation of neuroanatomical differences between autism and Asperger syndrome. *Archives of General Psychiatry, 61* (3), 291-298.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55* (1), 3-9.

Loveland, K. A., Tunali Kotoski, B., Chen, R., & Brelsford, K. A. (1995). Intermodal perception of affect in persons with autism or Down syndrome. *Development and Psychopathology, 7*, 409-418.

Loveland, K. A., Tunali Kotoski, B., Chen, Y. R., Ortegon, J., Pearson, D. A., Brelsford, K. A., et al. (1997). Emotion recognition in autism: Verbal and non-verbal information. *Development and Psychopathology, 9*, 579-593.

Luna, B., Minshew, N. J., Garver, K. E., Lazar, N. A., Thulborn, K. R., Eddy, W. F., et al. (2002). Neocortical system abnormalities in autism: an fMRI study of spatial working memory. *Neurology, 59* (6), 834-840.

Macdonald, H., Rutter, M., Howlin, P., Rios, P., Le Conteur, A., Evered, C., et al. (1989). Recognition and expression of emotional cues by autistic and normal adults. *Journal of Child Psychology and Psychiatry, 30* (6), 865-877.

Mari, M., Castiello, U., Marks, D., Marraffa, C., & Prior, M. (2003). The reach-to-grasp movement in children with autism spectrum disorder. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 358* (1430), 393-403.

Martin, N. T., Bibby, P., Mudford, O. C., & Eikeseth, S. (2003). Toward the Use of a Standardized Assessment for Young Children with Autism: Current Assessment Practices in the UK. *Autism, 7* (3), 321-330.

McCann, J., & Peppe, S. (2003). Prosody in autism spectrum disorders: a critical review. *International Journal of Language and Communication Disorders, 38* (4), 325-350.

McClure, E. B. (2000). A Meta-Analytic Review of Sex Differences in Facial Expression Processing and Their Development in Infants, Children, and Adolescents,. *Psychological Bulletin, 126* (3), 424-453.

McGovern, C. W., & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry, 46* (4), 401-408.

McGregor, E., Whiten, A., & Blackburn, P. (1998a). Teaching theory of mind by highlighting intention and illustrating thoughts: A comparison of their effectiveness with three-year-olds and autistic subjects. *British Journal of Developmental Psychology, 16* (3), 281-300.

McGregor, E., Whiten, A., & Blackburn, P. (1998b). Transfer of the picture-in-the-head analogy to natural contexts to aid false belief understanding in autism. *Autism, 2* (4), 367-387.

McKenzie, K., Matheson, E., McKaskie, K., Hamilton, L., & Murray, G. C. (2000). Impact of group training on emotion recognition in individuals with a learning disability. *British Journal of Learning Disabilities, 28* (4), 143-147.

McNeely, H. E., & Parlow, S. E. (2001). Complementarity of linguistic and prosodic processes in the intact brain. *Brain and Language, 79* (3), 473-481.

Medical-Research-Council. (2001). *Review of Autism Research: Epidemiology and Causes*. London: Medical Research Council.

Meltzoff, A. N. (2002). Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 6-25). Oxford: Blackwell Publishers.

Millward, C., Powell, S., Messer, D., & Jordan, R. (2000). Recall for self and other in autism: children's memory for events experienced by themselves and their peers. *Journal of Autism and Developmental Disorders, 30* (1), 15-28.

Minshew, N. J., & Goldstein, G. (1998). Autism as a disorder of complex information processing. *Mental Retardation and Developmental Disabilities Research Reviews, 4* (2), 129-136.

Minshew, N. J., Goldstein, G., Muenz, L. R., & Payton, J. B. (1992). Neuropsychological functioning in nonmentally retarded autistic individuals. *Journal of Clinical and Experimental Neuropsychology, 14* (5), 749-761.

Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: profile of a complex information processing disorder. *Journal of the International Neuropsychological Society, 3* (4), 303-316.

Molesworth, C. J., Bowler, D. M., & Hampton, J. A. (2005). The prototype effect in recognition memory: intact in autism? *Journal of Child Psychology and Psychiatry, 46* (6), 661-672.

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., et al. (1999). Face Perception During Early Infancy. *Psychological Science, 10* (5), 419-422.

Moore, D., McGrath, P., & Thorpe, J. (2000). Computer-Aided Learning for People with Autism – a Framework for Research and Development. *Innovations in Education and Training International, 37*, 218-228.

Morton, J. B., & Trehub, S. E. (2001). Children's understanding of emotion in speech. *Child Development, 72* (3), 834-843.

Muller, R. A., Pierce, K., Ambrose, J. B., Allen, G., & Courchesne, E. (2001). Atypical patterns of cerebral motor activation in autism: a functional magnetic resonance study. *Biological Psychiatry, 49* (8), 665-676.

Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: a meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience, 3* (3), 207-233.

Narumoto, J., Yamada, H., Iidaka, T., Sadato, N., Fukui, K., Itoh, H., et al. (2000). Brain regions involved in verbal or non-verbal aspects of facial emotion recognition. *Neuroreport, 11*, 2571-2576.

Nelson, C. A., & Dolgin, K. G. (1985). The generalized discrimination of facial expressions by seven-month-old infants. *Child Development, 56* (1), 58-61.

Nieminen-von Wendt, T., Metsahonkala, L., Kulomaki, T., Alto, S., Autti, T., Vanhala, R., et al. (2003). Changes in cerebral blood flow in Asperger syndrome during theory of mind tasks presented by the auditory route. *European Child and Adolescent Psychiatry, 12* (4), 178-189.

O'Riordan, M., & Plaisted, K. (2001). Enhanced discrimination in autism. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 54A* (4), 961-979.

O'Riordan, M. A., Plaisted, K. C., Driver, J., & Baron-Cohen, S. (2001). Superior visual search in autism. *Journal of Experimental Psychology: Human Perception and Performance, 27* (3), 719-730.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion, 1*, 29-50.

Oberman, L. M., Hubbard, E. M., McCleery, J. P., Altschuler, E. L., Ramachandran, V. S., & Pineda, J. A. (2005). EEG evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive Brain Research, 24* (2), 190-198.

Ohnishi, T., Matsuda, H., Hashimoto, T., Kunihiro, T., Nishikawa, M., Uema, T., et al. (2000). Abnormal regional cerebral blood flow in childhood autism. *Brain, 123 (Pt 9)*, 1838-1844.

Orsmond, G. I., Krauss, M. W., & Seltzer, M. M. (2004). Peer relationships and social and recreational activities among adolescents and adults with autism. *Journal of Autism and Developmental Disorders, 34* (3), 245-256.

Ortony, A., Clore, G., & Foss, M. (1987). The referential structure of the affective lexicon. *Cognitive Science, 11*, 341-364.

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology, 14* (2), 239-251.

Ozonoff, S. (1995a). Executive functions in autism. In E. Schopler & G. B. Mesibov (Eds.), *Learning and cognition in autism* (pp. 199-219). New York, NY, US: Plenum Press.

Ozonoff, S. (1995b). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology, 9*, 491-500.

Ozonoff, S., & Jensen, J. (1999). Brief report: specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders, 29* (2), 171-177.

Ozonoff, S., & McEvoy, R. E. (1994). A longitudinal study of executive function and theory of mind development in autism. *Development and Psychopathology, 6* (3), 415-431.

Ozonoff, S., & Miller, J. (1995). Teaching theory of mind: A new approach to social skills training for individuals with autism. *Journal of Autism and Developmental Disorders, 25* (4), 415-433.

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.

Parsons, S., & Mitchell, P. (1999). What children with autism understand about thoughts and thought bubbles. *Autism, 3* (1), 17-38.

Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research, 46*, 430-443.

Parsons, S., Mitchell, P., & Leonard, A. (2005). Do adolescents with autistic spectrum disorders adhere to social conventions in virtual environments? *Autism, 9* (1), 95-117.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35* (2), 205-220.

Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of Autism and Developmental Disorders, 32*, 249-261.

Perner, J., Frith, U., Leslie, A. M., & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: knowledge, belief, and communication. *Child Development, 60* (3), 688-700.

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of Mind Is Contagious: You Catch It from Your Sibs *Child Development, 65* (4), 1228-1238.

Phan, K. L., Wager, T. D., Taylor, S. F., & Liberzon, I. (2004). Functional neuroimaging studies of human emotions. *CNS Spectrums, 9* (4), 258-266.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological Psychiatry, 54* (5), 504-514.

Phillips, M. L., Young, A. W., Scott, S. K., Calder, A. J., Andrew, C., Giampietro, V., et al. (1998). Neural responses to facial and vocal expressions of fear and disgust. *Proceedings of The Royal Society, B: Biological Sciences, 265*, 1809-1817.

Phillips, W., Baron-Cohen, S., & Rutter, M. (1998). Understanding intention in normal development and in autism. *British Journal of Developmental Psychology, 16*, 337–348.

Pierce, K., Glad, K. S., & Schreibman, L. (1997). Social perception in children with autism: an attentional deficit? *Journal of Autism and Developmental Disorders, 27* (3), 265-282.

Pierce, K., Muller, R. A., Ambrose, J., Allen, G., & Courchesne, E. (2001). Face processing occurs outside the fusiform 'face area' in autism: evidence from functional MRI. *Brain, 124* (Pt 10), 2059-2073.

Pinker, S. (1998). *How the mind works*. London: Allen Lane.

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997). Broader autism phenotype: evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154* (2), 185-190.

Plaisted, K. (2000). Aspects of autism that theory of mind cannot explain. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other* *minds : perspectives from developmental cognitive neuroscience* (2nd ed., pp. 222-250). Oxford: Oxford University Press.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998a). Enhanced discrimination of novel, highly similar stimuli by adults with autism during a perceptual learning task. *Journal of Child Psychology and Psychiatry, 39* (5), 765-775.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998b). Enhanced visual search for a conjunctive target in autism: a research note. *Journal of Child Psychology and Psychiatry, 39* (5), 777-783.

Plaisted, K., Saksida, L., Alcantara, J., & Weisblatt, E. (2003). Towards an understanding of the mechanisms of weak central coherence effects: experiments in visual configural learning and auditory perception. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 358* (1430), 375-386.

Plaisted, K. C. (2001). Reduced generalization in autism: An alternative to weak central coherence. In J. A. Burack, T. Charman, N. Yirmiya & P. R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 149-169). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Plutchik, R. (1980). *Emotion : a psychoevolutionary synthesis*. New York ; London: Harper and Row.

Poliakoff, S. (Writer) (2000). Perfect Strangers. UK: BBC Four.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17* (03), 715-734.

Pourtois, G., de Gelder, B., Bol, A., & Crommelinck, M. (2005). Perception of facial expressions and voices and of their combination in the human brain. *Cortex, 41* (1), 49-59.

Pratt, C., & Bryant, P. (1990). Young children understanding that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61* (4), 973-982.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a "theory of mind"? *Behavioral and Brain Sciences, 1* (4), 515-526.

Prinz, J. (2004). Which emotions are basic? In D. Evans & P. Cruse (Eds.), *Emotion, evolution, and rationality* (Vol. 69-88). Oxford: Oxford University Press.

Prior, M., & Hoffmann, W. (1990). Brief report: neuropsychological testing of autistic children through an exploration with frontal lobe tests. *Journal of Autism and Developmental Disorders, 20* (4), 581-590.

Rajendran, G., & Mitchell, P. (2000). Computer mediated interaction in Asperger's syndrome: the Bubble Dialogue program. *Computers & Education, 35* (3), 189-207.

Reilly, J. S., McIntire, M., & Bellugi, U. (1990). Faces: the relationship of language and affect. In V. Volterra & C. J. Erting (Eds.), *From gesture to language in hearing and deaf children* (pp. 335). New York, NY: Springer-Verlag Publishing.

Rickman, A. (Writer) (1997). The Winter Guest. UK: New Line Home Video

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotiondescriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology, 21*, 901-908.

Rimland, B. (1965). *Infantile autism : the syndrome and its implications for a neural theory of behavior*. London: Methuen.

Rinehart, N. J., Bradshaw, J. L., Brereton, A. V., & Tonge, B. J. (2002). A clinical and neurobehavioural review of high-functioning autism and Asperger's disorder. *Australian and New Zealand Journal of Psychiatry, 36* (6), 762-770.

Ring, H. A., Baron-Cohen, S., Wheelwright, S., Williams, S. C. R., Brammer, M., Andrew, C., et al. (1999). Cerebral correlates of preserved cognitive skills in autism: A functional MRI study of Embedded Figures Task performance. *Brain, 122* (7), 1305-1315.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Brain Res Cogn Brain Res, 3* (2), 131-141.

Rogers, S. J. (2000). Interventions that facilitate socialization in children with autism. *Journal of Autism and Developmental Disorders, 30* (5), 399-409.

Ropar, D., & Mitchell, P. (1999). Are individuals with autism and Asperger's syndrome susceptible to visual illusions? *Journal of Child Psychology and Psychiatry, 40* (8), 1283-1293.

Ropar, D., & Mitchell, P. (2001). Susceptibility to illusions and performance on visuospatial tasks in individuals with autism. *Journal of Child Psychology and Psychiatry, 42* (4), 539-549.

Rosch, E., & Lloyd, B. B. (1978). *Cognition and categorization*. Hillsdale: Erlbaum ; New York ; London : Distributed by Wiley.

Rumsey, J. M. (1985). Conceptual problem-solving in highly verbal, nonretarded autistic men. *Journal of Autism and Developmental Disorders, 15* (1), 23-36.

Russell, J. (1997). *Autism as an executive disorder*. Oxford: Oxford University Press.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–1178.

Russell, J. A. (1991). In Defense of a Prototype Approach to Emotion Concepts. *Journal of Personality and Social Psychology, 40* (1), 37-47.

Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review, 110* (1), 145-172.

Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology, 48*, 1290-1298.

Russell, J. A., & Bullock, M. (1986). On the dimensions preschoolers use to interpret facial expressions of emotion. *Developmental Psychology, 22*, 97-102.

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

Rydin-Orwin, T., Drake, J., & Bratt, A. (1999). The effects of training on emotion recognition skills for adults with an intellectual disability. *Journal of Applied Research in Intellectual Disabilities, 12*, 253-262.

Sacks, O. (1995). *An anthropologist on Mars : seven paradoxical tales*. London: Picador.

Sainsbury, C. (2000). *Martian in the Playground : Understanding the Schoolchild with Asperger's Syndrome*. London, UK: Lucky Duck Publishing.

Scheuffgen, K., Happe, F., Anderson, M., & Frith, U. (2000). High "intelligence," low "IQ"? Speed of processing and measured IQ in children with autism. *Development and Psychopathology, 12* (1), 83-90.

Schopler, E., Mesibov, G., & Hearsey, K. (1995). Structured teaching in the TEACCH approach. In E. Schopler & G. B. Mesibov (Eds.), *Learning and cognition in autism* (pp. 243-267). New York, NY, US: Plenum Press.

Schroeder, U., Hennenlotter, A., Erhard, P., Haslinger, B., Stahl, R., Lange, K. W., et al. (2004). Functional neuroanatomy of perceiving surprised faces. *Human Brain Mapping, 23* (4), 181-187.

Schultz, R. T. (2005). Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience, 23* (2-3), 125-141.

Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., et al. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Archives of General Psychiatry, 57* (4), 331-340.

Schwartz, B., Wasserman, E. A., & Robbins, S. J. (2002). *Psychology of learning and behavior* (5th ed. ed.). New York ; London: W.W. Norton.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test): preliminary development of a UK screen for mainstream primary-school-age children. *Autism, 6*, 9-31.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: a research note. *Journal of Child Psychology and Psychiatry, 24* (4), 613-620.

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 34* (8), 1351-1364.

Shamay-Tsoory, S. G., Tomer, R., Yaniv, S., & Aharon-Peretz, J. (2002). Empathy deficits in Asperger syndrome: a cognitive profile. *Neurocase, 8* (3), 245-252.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology, 52* (6), 1061-1086.

Shriberg, L. D., Paul, R., McSweeny, J. L., Klin, A. M., Cohen, D. J., & Volkmar, F. R. (2001). Speech and prosody characteristics of adolescents and adults with high-functioning autism and Asperger syndrome. *Journal of Speech, Language, and Hearing Research, 44* (5), 1097-1115.

Sigman, M. D., Kasari, C., Kwon, J. H., & Yirmiya, N. (1992). Responses to the negative emotions of others by autistic, mentally retarded, and normal children. *Child Development, 63* (4), 796-807.

Silver, M., & Oakes, P. (2001). Evaluation of a new computer intervention to teach people with autism or Asperger syndrome to recognize and predict emotions in others. *Autism, 5*, 299-316.

Singh, S. D., Ellis, C. R., Winton, A. S. W., Singh, N. N., Leung, J. P., & Oswald, D. P. (1998). Recognition of Facial Expressions of Emotion by Children with Attention-Deficit Hyperactivity Disorder. *Behavior Modification, 22* (2), 128-142.

Sodian, B., & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *Journal of Child Psychology and Psychiatry, 33,* 591-605.

Solomon, M., Goodlin-Jones, B. L., & Anders, T. F. (2004). A social adjustment enhancement intervention for high functioning autism, Asperger's syndrome, and pervasive developmental disorder NOS. *Journal of Autism and Developmental Disorders, 34* (6), 649-668.

Sparrow, S., Balla, D., & Cicchetti, D. (1984). *Vineland Adaptive Behavior Scales - Survey Edition*: Circle Pines, MN: American Guidance Service.

Spence, S. H. (2003). Social skills training with children and young people: Theory, evidence and practice. *Child and Adolescent Mental Health, 8* (2), 84-96.

Steffenburg, S., Steffenburg, U., & Gillberg, C. (2003). Autism spectrum disorders in children with active epilepsy and learning disability: comorbidity, pre- and perinatal background, and seizure characteristics. *Developmental Medicine & Child Neurology, 45* (11), 724-730.

Stevens, D., Charman, T., & Blair, R. J. (2001). Recognition of emotion in facial expressions and vocal tones in children with psychopathic tendencies. *Journal of Genetic Psychology, 162* (2), 201-211.

Stifter, C. A., & Fox, N. A. (1987). Preschool children's ability to identify and label emotions. *Journal of Nonverbal Behavior, 11* (1), 43-54.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10* (5), 640-656.

Strain, P. S., & Hoyson, M. (2000). The Need for Longitudinal, Intensive Social Skill Intervention: LEAP Follow-Up Outcomes for Children with Autism. *Topics in Early Childhood Special Education, 20* (2), 116-122.

Strauss, M. S. (1979). Abstraction of prototypical information by adults and 10-month-old infants. *Journal of Experimental Psychology: Human Learnig and Memory, 5* (6), 618-632.

Sullivan , K. (Writer) (1987). Anne Of Green Gables - The Sequel. Canada: Sullivan Entertainment.

Swettenham, J. (1996). Can children with autism be taught to understand false belief using computers? *Journal of Child Psychology and Psychiatry, 37*, 157-165.

Swettenham, J., Baron-Cohen, S., Gomez, J. C., & Walsh, S. (1996). What's Inside Someone's Head? Conceiving of the Mind as a Camera Helps Children with Autism Acquire an Alternative to a Theory of Mind. *Cognitive Neuropsychiatry, 1* (1), 73-88.

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: deficits in the early acquisition of a theory of mind. *Child Development, 63* (1), 161-172.

Tager-Flusberg, H. (2000). Language and Understanding Minds: Connections in Autism. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds : perspectives from developmental cognitive neuroscience* (2nd ed., pp. 124-149). Oxford: Oxford University Press.

Tager-Flusberg, H., & Sullivan, K. (1994). Predicting and explaining behavior: a comparison of autistic, mentally retarded and normal children. *Journal of Child Psychology and Psychiatry, 35* (6), 1059-1075.

Takahashi, H., Yahata, N., Koeda, M., Matsuda, T., Asai, K., & Okubo, Y. (2004). Brain activation associated with evaluative processes of guilt and embarrassment: an fMRI study. *NeuroImage, 23* (3), 967-974.

Tantam, D. (1991). Asperger syndrome in adulthood. In U. Frith (Ed.), *Autism and Asperger syndrome* (pp. 147-183). New York, NY, USA: Cambridge University Press.

Tantam, D. (2000). Psychological disorder in adolescents and adults with Asperger syndrome. *Autism, 4* (1), 47-62.

Tantam, D. (2003). The challenge of adolescents and adults with Asperger syndrome. *Child and Adolescent Psychiatric Clinics of North America, 12* (1), 143-163, vii-viii.

Tantam, D., Monaghan, L., Nicholson, H., & Stirling, J. (1989). Autistic children's ability to interpret faces: a research note. *Journal of Child Psychology and Psychiatry, 30* (4), 623-630.

Temple, C. (1997). *Developmental cognitive neuropsychology*. Hove: Psychology Press.

Teunisse, J. P., & De Gelder, B. (1994). Do autistics have a generalized face processing deficit? *International Journal of Neuroscience, 77* (1-2), 1-10.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28* (5), 675-691.

Tsai, L. Y., & Beisler, J. M. (1983). The development of sex differences in infantile autism. *British Journal of Psychiatry, 142* (4), 373-378.

Tsatsanis, K. D. (2003). Outcome research in Asperger syndrome and autism. *Child and Adolescent Psychiatric Clinics of North America, 12* (1), 47-63, vi.

Tsatsanis, K. D. (2004). Heterogeneity in Learning Style in Asperger Syndrome and High-Functioning Autism. *Topics in Language Disorders, 24* (4), 260-270.

van der Geest, J. N., Kemner, C., Verbaten, M. N., & van Engeland, H. (2002). Gaze behavior of children with pervasive developmental disorder toward human faces: a fixation time study. *Journal of Child Psychology, 43*, 669-678.

Vicari, S., Reilly, J. S., Pasqualetti, P., Vizzotto, A., & Caltagirone, C. (2000). Recognition of facial expressions of emotions in school-age children: the intersection of perceptual and semantic categories. *Acta Paediatrica, 89* (7), 836-845.

Volkmar, F. R., & Klin, A. (2005). Issues in the classification of autism and related conditions. In F. R. Volkmar, A. Klin, R. Paul & D. J. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders* (3rd ed., Vol. 1, pp. 5-41). New Jersey: Wiley.

Volkmar, F. R., Szatmari, P., & Sparrow, S. S. (1993). Sex differences in pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 23* (4), 579-591.

Vygotskii, L. S. (1986). *Thought and language*. Cambridge, Mass/London: MIT Press.

Walker-Andrews, A. S. (1986). Intermodal perception of expressive behaviors: Relation of eye and voice? *Developmental Psychology, 22* (3), 373-377.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin, 121* (3), 437-456.

Walker-Andrews, A. S., & Grolnick, W. (1983). Discrimination of vocal expressions by young infants. *Infant Behavior and Development, 6* (4), 491-498.

Walker, A. S. (1982). Intermodal perception of expressive behaviors by human infants. *Journal of Experimental Child Psychology, 33* (3), 514-535.

Wang, A. T., Dapretto, M., Hariri, A. R., Sigman, M., & Bookheimer, S. Y. (2004). Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 43* (4), 481-490.

Waterhouse, L., Fein, D., & Modahl, C. (1996). Neurofunctional mechanisms in autism. *Psychological Review, 103* (3), 457-489.

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

Welchew, D. E., Ashwin, C., Berkouk, K., Salvador, R., Suckling, J., Baron-Cohen, S., et al. (2005). Functional disconnectivity of the medial temporal lobe in Asperger's syndrome. *Biological Psychiatry, 57* (9), 991-998.

Wellman, H. M. (1992). *The child's theory of mind*. Cambridge, Mass/London: The MIT Press.

Wellman, H. M., Baron-Cohen, S., Caswell, R., Gomez, J. C., Swettenham, J., Toye, E., et al. (2002). Thought-bubbles help children with autism acquire an alternative to a theory of mind. *Autism, 6* (4), 343-363.

Wellman, H. M., & Lagattuta, K. H. (2000). Developing understanding of mind. In S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen (Eds.), *Understanding other minds : perspectives from developmental cognitive neuroscience* (2nd ed., pp. 21-49). Oxford: Oxford University Press.

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Pychology, 39* (1), 114-128.

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., et al. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage*, *24* (4), 1233-1241.

Williams, D. (1992). *Nobody nowhere*. New York: Time Books.

Williams, D. L., Goldstein, G., & Minshew, N. J. (2006). The profile of memory function in children with autism. *Neuropsychology, 20* (1), 21-29.

Williams, J., Allison, C., Scott, F. J., Stott, C. M., Bolton, P., Baron-Cohen, S., et al. (2006). The Childhood Asperger Syndrome Test (CAST): Test-retest reliability. *Autism*, 10(4), 415-427.

Williams, J., Scott, F. J., Stott, C. M., Allison, C., Bolton, P., Baron-Cohen, S., et al. (2005). The CAST (Childhood Asperger Syndrome Test): Test accuracy. *Autism, 9* (1), 45-68.

Williams, J. H., Whiten, A., Suddendorf, T., & Perrett, D. I. (2001). Imitation, mirror neurons and autism. *Neuroscience and Biobehavioral Reviews, 25* (4), 287-295.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13* (1), 103-128.

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine, 11* (1), 115-129.

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: epidemiology and classification. *Journal of Autism and Developmental Disorders, 9* (1), 11-29.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening Adults for Asperger Syndrome Using the AQ: A Preliminary Study of its Diagnostic Validity in Clinical Practice. *Journal of Autism and Developmental Disorders, 35* (3), 331-335.

World-Health-Organisation. (1994). *ICD-10 - International classification of diseases* (10th ed.). Geneva, Switzerland: World Health Organisation.

Yirmiya, N., Kasari, C., Sigman, M., & Mundy, P. (1989). Facial expressions of affect in autistic, mentally retarded and normal children. *Journal of Child Psychology and Psychiatry, 30* (5), 725-735.

Yirmiya, N., Shaked, M., & Erel, O. (2001). Comparison of siblings of individuals with autism and siblings of individuals with other diagnoses: An empirical summary. In E. Schopler, N. Yirmiya, C. Shulman & L. M. Marcus (Eds.), *The research basis for autism intervention* (pp. 59-73). New York, NY, US: Kluwer Academic/Plenum Publishers.

Yirmiya, N., Sigman, M. D., Kasari, C., & Mundy, P. (1992). Empathy and cognition in high-functioning children with autism. *Child Development, 63*, 150-160.

Young, A. W. (1998). *Face and mind*. Oxford: Oxford University Press.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., & Perrett, D. I. (1997). Facial expression megamix: tests of dimensional and category accounts of emotion recognition. *Cognition, 63* (3), 271-313.

Younger, B. (1990). Infant categorization: memory for category-level and specific item information. *Journal of Experimental Child Psychology, 50* (1), 131-155.

Zhang, J., Quan, H., Ng, J., & Stepanavage, M. E. (1997). Some statistical methods for multiple endpoints in clinical trials. *Controlled Clinical Trials, 18* (3), 204-221.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J., & Szatmari, P. (2005). Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience, 23* (2-3), 143-152.