Non-language based Theory of Mind tests in individuals with autism

Fulvia Castelli

Institute of Cognitive Neuroscience
Department of Psychology
University College London

Thesis submitted for the degree of Ph.D. in Psychology,
University of London.
October 2001

Abstract

The main focus of this thesis was to investigate the nature of stimuli that provoke the pervasive tendency of people to explain behaviour in terms of mental states (Theory of Mind). A series of experimental tasks was designed to test the "Theory of Mind deficit" hypothesis with high-functioning individuals with autism by using non-verbal stimuli in both behavioural and neuroimaging (PET) studies.

The first three experiments explored the most familiar visual inputs that trigger the attribution of mental states, namely, emotional faces. Children with autism were as able as controls to recognise basic emotions. The fourth and fifth experiments explored the simplest forms of visual information for judging agents' intentions, namely, their motion pattern. Children with autism were as able as controls to attribute an intended goal to an agent in the presence of its unsuccessful outcome. However, they responded similarly to younger control children in the presence of a sudden change in the agent's motion direction. This result is compatible with a developmental delay in autism in the representation of goal-directed motion. The last two PET studies were based on the perception of silent computer animations. These animations depicted two interacting characters whose movement patterns evoked descriptions either in mentalistic terms or in behavioural terms. The first PET study identified brain activity in healthy volunteers while watching the animations. The second PET study investigated brain activity in a group of adults with autism during the same task. Verbal descriptions of the animations showed a mentalising deficit in the autism group. Neuroimaging findings revealed that the autism group showed reduced activation and reduced functional connectivity in several areas of the previously identified mentalising network.

These findings are evaluated in the context of the metarepresentational model and the Theory of Mind deficit hypothesis of autism, and suggestions for further research are discussed.

Acknowledgements

My immense gratitude goes to Uta Frith and James Blair, my supervisors, who made my Ph.D. not only possible, but also a truly wonderful experience. My sincerest thanks go also to Chris Frith and Francesca Happé who were constantly there whenever I needed help and suggestions. I believe that I was extremely lucky I had the opportunity to learn from each of them and to work with them.

I also wish to thank my colleagues of the Developmental Group and John Morton for their invaluable comments during the preparation of the thesis. In particular, Eamon McCrory, without whom my Ph.D. years would not have been so enjoyable. Special thanks go to Marilú Gorno Tempini, Lisa Cipolotti and Angelo Maravita for their unfailing support, advice and friendship. I am also very much indebted to Brian Butterworth for offering me a job, thesis unfinished, and not complaining about it. I am grateful to Lorna Wing for her generous help in recruiting participants, and to all the people, children and adults, who took part of my studies. Finally, I acknowledge the financial support from the European TMR Marie Curie training grant.

Above all, I wish to thank Giacomo. This thesis is because of him, and for him.

FCG

15th October, 2001

Table of Contents

List of Tables

List of Figures

Chapter 1

Non-language Theory of Mind tests in individuals with autism

1.1 Theory of Mind

One of the finest examples of the predictive usefulness of Theory of Mind attribution comes from the fictional character of Sherlock Holmes, the great mastermind of detection created by Sir Arthur Conan Doyle in the 19th century. In "A scandal in Bohemia", Sherlock Holmes has been instructed by his client to find a photograph that a lady has safely concealed in her house. Sherlock Holmes' cunning plan is to make the lady think that her living room is on fire by tossing in a smoke rocket and raising a fire alarm. What happens afterwards is exactly what Sherlock Holmes had predicted: the lady's instinct is at once to rush to secure the photograph, which is the object she valued most, and by doing so she accidentally revealed to Sherlock Holmes, who was in the same room, the secret hiding place. The case is brilliantly solved by the detective thanks to his correct attribution of intentionality to his opponent. How can this be explained?

People have the pervasive tendency to explain their own and others' actions in terms of beliefs, desires and goals. The attribution of intentional states represents a useful social strategy to make sense of and predict people's behaviour. The terms "theory of mind" (ToM), "mindreading" or "mentalising" have all been coined to describe people's effortless propensity to search for intentions behind actions.

1.1.1 The notion of attribution of intentionality

Intentionality is, according to the philosopher Frank Brentano (1874) the feature that sets the mental states completely apart from the physical states. Mental states hold the property of "pointing" towards certain objects or states of affairs (intendere in Latin means to aim), but these objects or states of affairs need not exist. Intentionality is the common feature of propositional attitudes, those cognitive states that are described in everyday language with the use of a "that" clause (typically beliefs and desires). Sherlock Holmes attributed to the lady both the desire that the valuable photograph be secure at all costs, and the sequence of beliefs including the belief that if there is smoke and someone cries "fire" there is indeed a fire, the belief that fire burns paper and the belief that photographs are of paper and so forth. Mental states not only have propositional content, but also causal roles: the lady's set of desires and beliefs were indeed the cause of disclosing her secret.

Emotional states are also mental states, and in our everyday descriptions of people and ourselves the use of different emotion terms are as frequent and useful as belief/desire terms. Nevertheless, the general category of "emotion" includes very different psychological states such as sudden reaction responses to people or objects and introspective experiences. The difference between emotional states and propositional states is that their intentional content is different: emotions are mental states that point directly towards a state of affairs, whereas beliefs and desires are attitudes towards propositional content. Since the debate about the nature of emotion is yet to be resolved (cf. Griffiths, 1997), for the purpose of the present work it suffices to underline that emotions are mental states, which can be either caused by autonomic responses to states of affairs or by specific propositional attitudes. In other words, if someone asks me to touch a snake, I would be afraid, and my heart would beat faster even if I was determined to touch it. On the other hand, the sight of a beautiful flower will trigger in me immediate pleasure, and I would like to touch it. However, if I believe that the flower is poisonous, I would be afraid of touching it. Another example of how beliefs may affect emotional responses comes from a real-life situation shown on television during a "candid camera" program. A man with a snake around his neck asked a lady to close her eyes, and then he put around her neck a leather belt. The lady, still with her eyes closed, was terrified because she believed there was a snake around her neck.

An important aspect of propositional attitudes - beliefs and desires, but also intentions, thoughts, doubts, remembrances - is that they may represent actual or possible states of affairs. The states of affairs upon which people's propositional attitudes are directed may actually attain in the real world, but equally they might not: beliefs are often false, desires can be frustrated, and hopes may be dashed. Hence, propositional attitudes are states that "do not exist here and now": they are states that are about or directed upon a nonexistent state of affairs or objects, that can be true or false, and even paradoxical. More specifically, propositional attitudes have three common semantic properties (Quine, 1961):

i) their reference is opaque, that is, they may be distinct or even contrasting beliefs and yet refer to the same state of affairs (e.g. "John believes that drinking water is safe", and "he believes that drinking H_20 is poisonous"),

ii) they do not entail truth (e.g. "John believes that it is raining" says nothing about whether or not it is really raining),

iii) they do not entail existence (e.g. "John believes that the king of France is bald" says nothing about the existence of a French monarchy).

According to Leslie (1987), there is a remarkable similarity between the logical properties of mental states and the ability to pretend or understanding pretence that emerges during the second year of age. Pretence is, for Leslie, good evidence that at 2 years of age infants possess not only the ability to form primary representations, but also metarepresentations. The former ability allows them to understand the world in an accurate, faithful and literal way - transparent representations - whereas the latter ability allows them to understand states that are not necessarily existing and true - opaque representations - e.g., the child is not confused but delighted when his/her mother is pretending to hold a telephone when she picks up a banana saying "Let's call Daddy!" In this example, mother is represented as an agent in relation to a description of reality, as opposed to an aspect of reality in relation to another aspect of reality. Thus, as an agent, mother adopts an attitude (of pretending) to the truth of a description ("it is a telephone") in regard to a particular aspect of reality (the banana). Whether or not Leslie's account of the cognitive process involved in pretend play, and understanding pretence is accepted, it is clear that our everyday use of intentional terms in order to explain and predict behaviour of ourselves and others is based on the representation of thinking agents and representation of the content of their thoughts.

1.1.2 The metarepresentational model

In the last two decades increasing interest in the evolution, development, and breakdown of social cognition has focused on the ability to attribute mental states to self and others. A model of a dedicated, domain-specific and possibly modular cognitive mechanism has been proposed and developed with the aim of describing the cognitive structure underlying the development of the ability to mentalise (Leslie 1987; Fodor, 1992; Leslie and Thaiss, 1992; Baron-Cohen 1995; Roth and Leslie, 1998). According to this model – henceforth referred to simply as "metarepresentational model" – a specialised cognitive mechanism has the competence of providing agent-centered interpretation of behaviour by constructing metarepresentations. This mechanism forms the specific basis for the ability to acquire Theory of Mind, and can be damaged or impaired independently of other processing systems.

Baron-Cohen (1994, 1995) proposed the existence of different sub-components that have evolved to enable the ability to represent mental states. The four specialized systems "roughly reflect four properties of the world: volition, perception, shared attention and epistemic states" (Baron-Cohen, 1995, p. 31). The first two mechanisms, the Intentionality Detector (ID), which interprets motion stimuli in terms of goals and approach/avoidance desires, and the Eyes-Direction Detector (EDD), which interprets visual stimuli in terms of what an agent sees, deploy "dyadic" representations. Dyadic representations specify the intentional relation between an agent and a state of affairs (e.g. "she wants to get away from this"), or an agent looking at an object (e.g. " she is looking at the toy") or at oneself (e.g. "Mummy is looking at me"). The other two mechanisms, the Shared Attention Mechanism (SAM) and the Theory of Mind Mechanism (ToMM), were postulated to account for the infant's development from a simple view of the world, where people act in relation to an aspect of reality, to a more complex view of the world, where people act in relation to their own's and others' beliefs, desires and knowledge. SAM, in particular, is a system dedicated to form "triadic" representations, whose function is to specify the intentional relation among an agent, oneself and a third object (e.g., "I see - Mummy sees Daddy -"). Finally, ToMM, the full-blown metarepresentational system, deploys representations of agents' propositional attitudes, namely, the attitudes towards the truth of the content of their thoughts, beliefs and desires. According to Baron-Cohen (1994), the four components of the metarepresentational system are domain-specific, with obligatory triggering inputs and can be selectively damaged.

It must be noted, however, that the nature of the metarepresentational systems, regardless of its cognitive mechanisms, is still a matter of debate. Contrasting views to the modular domain-specific model favour a more general explanation of the development of Theory of Mind, e.g. simulation or general theory formation (e.g. Goldman, 1993; Bartsch and Wellman 1995; Gopnik and Meltzoff, 1997; Gopnik, Capps and Meltzoff, 2000). Evidence supporting a dedicated cognitive mechanism for Theory of Mind includes the rapid acquisition, largely independent of other abilities and cross-cultural, and neuropsychological cases of a selective impairment in mentalising.

The metarepresentational model has the heuristic power of bringing both normal and abnormal patterns of development within a single neuropsychological explanatory framework, allowing to make a clear distinction between competence and performance in clinical populations. Indeed, one of the first predictions of this cognitive model was that high-functioning children with autism lack the ability to represent mental states (Baron-Cohen, Leslie and Frith, 1985). This landmark study on the core cognitive deficit of autism adapted a Theory of Mind paradigm developed earlier by Wimmer and Perner (1983) with preschool children. However, neither the domain of autism or developmental psychology constitutes the origin of Theory of Mind research. In the following section I will present a brief historical perspective of the paradigms adopted in this research domain with the purpose of providing an illustrative view rather than a comprehensive survey of the Theory of Mind literature.

1.1.3 ToM, Sarah and Maxi

Research into Theory of Mind, or the ability to attribute mental states to others, began with primates. The very term Theory of Mind was coined by Premack and Woodruff (1978) who, in a pioneering paper, asked whether chimpanzees have the ability to infer mental states from others' behaviour. The first non-language based ToM test was used with Sarah, an adult female chimpanzee. She was shown a series of videotaped events in which human actors were struggling to obtain inaccessible objects, and then asked to indicate how she thought the human actor would solve his problem.

One of the scenes, for example, presented the actor in a cage similar to the chimpanzee's, struggling to get bananas that were hanging from the high ceiling. The animal then was presented with a pair of still photographs, one of which depicted the actor engaged in a behaviour that constituted a practical solution to the impediment, e.g. stepping onto a box.

Premack and Woodruff argued that Sarah's consistent choice of the correct photograph can be understood by assuming that the animal possesses Theory of Mind: she solves such problems by being able to attribute to the actor both an intention or purpose and knowledge. According to the authors, when the chimp watches the videotapes showing the human actor jumping to reach the bananas attached to the ceiling, she assumes that the actor wants the bananas and he is struggling to reach them. She also assumes that the actor knows how to attain the bananas so that when she is required to indicate which photograph depicts the solution, she chooses correctly the picture of the actor stepping on the box. However, these results can also be explained with no reference to mental state representations.

In a commentary to Premack and Woodruff's study, the philosopher Dennett (1978) argued that the evidence of the primate's correct choice may equally well be supported by "supposing that [it is] derived from either habits (of thought) or belief about other features of the world (e.g. experienced regularities in the behaviors of others)" (p. 569). Dennett pointed out three crucial requirements for an experiment to establish whether subjects attribute mental states to others or not: a) the actor's anticipated behaviour is a novel action, in the sense of not being habitual or often repeated in the given circumstances, b) the subject's response consists of an action as much as possible from his natural repertoire, rather than a highly trained artificial response, c) the perceived dependence of the subject's response on the other person's action should also be natural and obvious.

The first requirement is fulfilled when the belief attributed to another person is false. Only in that case, will the individual be expected to act inappropriately to the novel circumstance, hence, in all likelihood, not the way the person has typically acted in the past. Correctly reporting someone's belief when it differs from your own cannot be done by simply referring to one's own knowledge of the facts. The other two requirements are fulfilled by communicative acts, such as asking a question or requesting something from the other individual.

Following Dennett's suggestions, Wimmer and Perner (1983) accepted the challenge to develop a test to investigate children's ability to attribute mental states to others. The classic "false-belief task" paradigm investigates the ability to understand that a belief is opaque, namely, that there can be different beliefs about a single event. Hence, the test is based on the assumption that understanding someone else's false belief requires the representation of a mental state that differs from what is known to be true. If a child is able to represent another's mental states, then he or she will be able to predict the ignorant person's behaviour, which is no longer appropriate in the changed circumstances.

The original test was constructed by Wimmer and Perner around a story character, Maxi, who hides a chocolate in a box before he goes out to play. While he is out, mother comes in and moves Maxi's chocolate from the box to the cupboard. Then Maxi comes back in from playing. The experimenter asks: "Where will Maxi look for his chocolate?" The test allows for distinguishing whether the child predicts Maxi's behaviour on the basis of his own knowledge (Maxi will look where the chocolate really is) or whether he understands Maxi's false belief (Maxi will look where the chocolate was left). The child answers without the need of verbalising his response, since a mere pointing gesture suffices. Results indicated that while most 4-year-olds correctly predicted Maxi's behaviour on the basis of his false belief, most 3-year-olds answered on the basis of the new situation, without taking into account that Maxi's belief remained the same regardless reality changed.

In another task the child himself is the holder of a belief which is revealed to be false. In the "Smarties" task (Perner, Frith, Leslie and Leekam, 1989) the child is asked to guess what a Smarties tube – well-known sweets' package - contains. After the child correctly answers "sweets" or "Smarties", the experimenter opens the package and shows the real content, a pencil. The lid is then replaced and the child is asked: "When you first saw this box, before we took off the lid, what did you think was inside?" In this task 3-year-olds do not seem to recall their false belief but answer instead "a pencil". The same result occurs when the child is asked what another child would say when shown the same tube of Smarties: 3-year-olds expect a true belief while 4-year-olds recognise the false belief (Perner et al., 1989; Wimmer and Hartl, 1991).

1.1.4 Normal development of Theory of Mind

It is now well established that, before the age of four, normally developing children do not fully realise the implications of having a false belief. A review of data collected over a period of five years in various different studies has shown that the predicted probability of normally developing children passing false-beliefs tests steeply increases with chronological age (Happé, 1995). The study showed that at 3.6 years the predicted probability of passing false-belief tests was 23%, at 4 it doubles to 55% and at 4-6 the probability of success was 80%. More recently, a meta-analysis of 178 studies adopting the false-belief paradigm showed a consistent developmental pattern between 3 and 5 years of age, even across various countries and various task manipulations (Wellman, Cross and Watson, 2001).

The issue relative to the cognitive developmental "shift" between the age of three and four in understanding false belief has motivated extensive research that bears on issues such as nativism and modularity, which are not the focus of the present work. These topics are discussed in a collection of papers edited by Baron-Cohen, Tager-Flusberg and Cohen (2000). More importantly, the task itself is not a "pure" measure for mentalising, since both language and age affect performance. In addition, Bloom and German (2000) by reviewing almost twenty years of research based on the false-belief paradigm suggested that one of the most important weaknesses of these types of tasks is that they are inherently difficult to perform for young children even if they had the competence. What are the difficulties of false-belief tests?

As Leslie pointed out (1994; Leslie and Polizzi 1998) beliefs are supposed to be true because that is what they are for: "Just as the usefulness of a currency depends upon a default attribution of genuineness, despite occasional forgery, so the usefulness of the concept BELIEF depends upon a default attribution of veracity, despite occasional falseness" (Leslie and Polizzi, 1998, p. 247). Hence, reasoning about a belief with a false content requires an additional process other than the process of belief attribution, namely, the inhibition of a default attribution response based on beliefs with a true content. This inhibition process is associated with a general cognitive component, the "selection processor" (SP), which develops independently of the metarepresentation system (Leslie and Thaiss, 1992; Leslie and Polizzi, 1998; Roth and Leslie, 1998). The selection processor allows the child to select the correct belief content in a false-belief task and to resist the default assumption that belief contents are true. Thus, both components, the metarepresentational mechanism and the selection processor, need to be well developed and cooperating in order to solve false-belief problems.

The metarepresentational model is "admittedly still sketchy" (Roth and Leslie, 1998, p. 3) and does not include other possible components that might be employed in the development of the ability to represent mental states. The underlying specific mechanism is thought to endow the infant with a basic competence necessary for learning about mental states, given the support of other cognitive processes. Consequently, the ability to pass a false-belief task represents only one of the developmental stages of the mentalising ability. In fact, while pretend play is an early sign that metarepresentation must be available to the child between 12 and 24 months, even earlier ostensive communication behaviours such as referential looking - joint attention - and protodeclarative pointing emerging between 6 and 12 months

(Butterworth, 1991) may also signal the emergence of the ability to represent mental states (Leslie and Happé, 1989).

Imitation skills are also thought to require the representation of mental states of others: Meltzoff (1995) tested 18-month-old infants with a "re-enactment" paradigm. Infants watched an adult who performed a series of intended but failed actions based on object manipulation such as pulling a dumbbell toy apart. A control group of infants observed the same actions but successfully completed. The actions were then re-enacted by the infants who were given the same objects. All the infants performed successfully all the intended actions, regardless of whether they observed the failed or the completed trials. The infants were also exposed to an intended but failed action performed by a mechanical device (two mechanical pincers grasping the ends of the dumbbell and then slipping off). The results indicated that infants are more likely to complete an action denoting intention to humans than to non-humans. It must be noted, however, that infants may have learned through observation to apply outward pressure with their hands to the ends of the dumbbell (or other similar movements relative to the other actions), whereas in the case of the mechanical device, they have learned not to touch the dumbbell. This interpretation is also reinforced by the fact that during the mechanical device trial the adult is standing by the device refraining from touching it. A review of the literature on imitation, which is rich and of multidisciplinary origins, is beyond the scope of the present chapter. Some causes and consequences of imitation are discussed in a paper by Heyes (2001). Heyes concludes that although it is plausible that imitation contributes to the development of theory of mind, (e.g. Whiten and Brown, 1999; Meltzoff and Moore, 1999) "there is not currently a well-supported theory specifying the nature of the contribution" (p. 260).

A longitudinal study (Charman, Baron-Cohen, Swettenham et al., 2000) investigated a sample of 13 infants on a series of measures of joint attention, imitation and spontaneous play at 20 months and of Theory of Mind tests at 44 months of age. The joint attention measures included the gaze switch between a toy with salient features (e.g. a moving car, a flashing robot, a noisy pig), and looking up towards the experimenter's eyes after he prevented the child from playing with a toy (e.g. covering the child's hand with his own, or offering a toy and withdrawing it when the child was reaching out for it). The imitation task consisted in imitating actions with unfamiliar objects previously modelled by the experimenter. The presence of spontaneous functional and pretend play was measured while the child was presented with various toys (e.g. toy tea set, animals, and box). Unfortunately the infants in the sample were unable to understand the false-belief tests, so that easier tasks were adopted. The first one was a visual-perspective-taking task with cards held between the child and the experimenter. Each card depicted a different object on each side, and the child was asked to say what the experimenter could see. The second ToM test was a "seeing-leads-to-knowing" task with two dolls holding a box, but only one opening the lid. The child was asked to say which of the two dolls knew the content of the box. The last mentalising test consisted in attributing the emotion of sadness or happiness to a blank face of a cartoon character, after the child was told whether the character attained what he wanted or not (e.g. "Steven wanted a banana, and he has got a banana. Does he feel happy or sad?"). Interestingly, joint attention ability, but not imitation or play, measured before the second year of age was the best long term predictor of Theory of Mind during the fourth year. It is possible that both imitation and play measures may index also children's ability to understand the specific object features they have to manipulate. The study allowed the researchers to address the interesting question of how specific representational abilities are affected by development before the critical age of 4 years when the child is able to pass false-belief tests. Further studies are needed to address a different question, namely, about the nature of the mechanism by which joint attention behaviour at 20 months is related to later Theory of Mind development.

In conclusion, I have outlined the intrinsic weakness of the false-belief paradigm concerning the fact that it does not allow for exploring developmental stages in the acquisition of Theory of Mind ability. In fact, reasoning about false beliefs requires the support of cognitive components developing independently of the metarepresentational system. In addition, although it is plausible that putative behaviours, in particular, joint attention, contribute to the development of Theory of Mind, the nature of their contribution is still unclear. Nevertheless, the false-belief paradigm has motivated some of the most exciting research in autism under the assumption that failure at attributing a false belief to oneself or others, at a mental age above four, reflects some serious deficit in understanding of mental states of themselves and others – a deficit in Theory of Mind.

1.2 Autism: definition, diagnosis and biological basis

The term autism was coined by Bleuler (1911) from the Greek "autos" which means "self" to describe the egocentric thinking of individual with schizophrenia. Leo Kanner in Baltimore, (1943) and Hans Asperger in Vienna, (1944) independently used this term to name the developmental disorder affecting the children they described as lacking social responsiveness, with poor eye contact, showing stereotyped movement and marked resistance to change. Although Kanner and Asperger described the same type of disturbed children, the eleven cases that Kanner described showed severe language impairment, whereas the four cases that Asperger reported were able to use language, but with "originality" (Asperger, 1943 tranlated in Frith, 1991, p.71).

In 1981, Lorna Wing introduced the label Asperger syndrome to describe those individuals with autism who have a higher degree of social functioning than those with a typical diagnosis of autism. It is unclear whether the superior functioning pertains more to language than to global cognitive skills. The Asperger diagnosis is used in clinical cases of individuals who somehow give the impression to be "too normal" for a full diagnosis of autism: they are relatively high-functioning on IQ tests with social deficits and superficially good language skills (Gillberg and Coleman, 2000). Although the validity of a differential diagnosis of autism and Asperger syndrome is still a matter of debate (cf. Frith, 1991; Klin, Volkmar, and Sparrow, 2000), there is a wide consensus that they lie on a social deficit continuum with the latter in the higher range of abilities.

It was again Lorna Wing (1988) who introduced the concept of a continuum, or spectrum, to describe the heterogeneity of the manifestations of autism, from the most profoundly physically and mentally retarded individual, who has a social impairment coexisting with a host of other problems, to the most able, highly intelligent person with social difficulties. Interestingly, one of the strongest arguments supporting the notion of an autism continuum is the first reported clinical case (Wing, 1981) where the same individual was typically autistic in his early years but made progress and as teenager showed the characteristics of Asperger syndrome (the case of the "late talker", reported also in Frith, 1991).

1.2.1 Explaining autism at a behavioural, biological and cognitive level

Autism is a biologically caused developmental disorder, which is diagnosed on a behavioural basis, and defined at a cognitive level. Morton and Frith (1995) developed a causal modelling approach that allows one to look at different levels of disorders within a unified model, and can be used to examine relationships across biological, cognitive, behavioural, and environmental factors. The causal modelling approach for explaining a developmental disorder follows important rules: do not ignore the biological origin, even if the precise causes are very rarely known, and then build-up causal chains to account for all behavioural signs; specific deficits have to be demonstrated over and above general deficits, and correlation does not imply causation. A brief summary of the description of autism at a behavioural and biological level is given below.

What are the behavioural signs of autism? The necessary and sufficient features for the diagnosis of autistic disorder currently agreed upon by most authorities (DSM-IV, 1994; ICD-10, 1993) are the following:

i) qualitative impairment in reciprocal social interaction,

ii) qualitative impairment in verbal and nonverbal communications and in imaginative activity,

iii) markedly restricted repertoire of activities and interests.

The triad of social, communication and imagination impairments is also referred as the Wing's triade (Wing and Gould, 1979). This triad of impairments, though of primary importance, is not the only variable involved in the clinical picture. Abnormal responses to sensory stimuli, gross and fine motor coordination problems, attention and memory problems are also associated with the core features. Different combinations of skills and disabilities may be found at any level of general intelligence.

What is the prevalence of autism? The most conservative estimate of the prevalence of this developmental disorder, as defined according to the above criteria, is about 1 in 1,000 children (Gillberg and Coleman, 2000). This figure refers to the autistic disorder/childhood autism only, and does not include Asperger Syndrome or cases of "atypical autism" in which one or more of the symptoms are presented atypically. A few studies have suggested a prevalence of at least 4-5 in 1,000 children (Gillberg and Coleman, 2000). A recent review on epidemiological surveys of autism published in the English language between 1966 and 1998 (Fombonne, 1999) revealed that prevalence rates increased with publication year, reflecting changes in case definition and improved recognition. Based on 11 surveys conducted since 1989, the median rate was 7.2 in 10,000 subjects, which is slightly lower than the rate suggested above, by Gillberg and Coleman (2000).

What is the biological origin of autism? Reviews of the biology of autism conclude that evidence for an organic cause is overwhelming (Gillberg and Coleman, 1985,1992, 2000; Schopler and Mesibov, 1987). Many different neuro-disciplines point towards different specific abnormalities. It suffices to mention that evidence for a biological basis of autism includes the high incidence of epilepsy, the association with learning disabilities, and the increase in the incidence of autism with progressively lower IQ. The strong heritability of autism is beyond doubt: twin studies have shown that concordance rates of identical twins with autism exceed the concordance rate of fraternal twins (Folstein and Rutter, 1977; Ritvo, Freeman, Mason-Brothers et al., 1985; Steffenburg, Gillberg, Helgren et al., 1989). A twin study by Bailey, Le Couteur, Gottesman et al. (1995) reported a 60 per cent concordance for monozygotic twins and no concordance for dizygotic twins, giving rise to an estimated heritability of 91 to 93 per cent. Furthermore, studies investigating the extended phenotype of autism have shown autistic-like behaviour in non-autistic relatives (Le Couteur, Bailey, Goode et al., 1996; Bailey, Palferman, Heavey et al., 1998; Baron-Cohen and Hammer, 1997; Happé, Briskman and Frith, 2001). However, no precise findings are as yet available to put together a coherent picture of the origin of autism. What is important to stress, is that the symptoms of autism are probably caused by many different etiologies. In most cases, autism appears to be due to a "cascade" of neurodevelopmental abnormalities that leads to a common pathway resulting in impaired behaviours (Gillberg and Coleman, 2000).

Thus, the behavioural clinical picture of autism differs between individuals and changes markedly over the course of development, and no single biological origin has been pinpointed, but rather multiple causes. The question to be answered is then: "If there is no single origin nor any single kind of damage that can be used as [a criterion to set apart autism from other disorders], what, then, justifies the application of a single label?" (Morton and Frith, 1995, p. 363). The triad of social, communication and imagination impairments provides only an accurate description of the behavioural picture rather than an explanation of autism. The answer to the question can be found if the gap between the heterogeneity at both the biological and behavioural levels is bridged at the cognitive level. The explanation of autism at the cognitive level concerns a single cognitive deficit, namely, a Theory of Mind or mentalising deficit. The hypothesis is that in autism, the cognitive system responsible for computing mental state representations is selectively impaired. In the next section I will summarize the specific tasks adopted to test the lack of Theory of Mind ability in individuals with autism.

1.2.2 Theory of Mind deficit hypothesis

In 1985, Baron-Cohen, Leslie and Frith tested the hypothesis that children with autism, with a mental age above 4 years, fail to take into account other people's beliefs. They used a variation of the Wimmer and Perner paradigm with a group of normally developing children (mean age 4.6 years), a group of Down syndrome children (mean age 11 years) and children with autism (mean age 12 years). Children were presented with a scene with two dolls, "Sally" and "Ann": Sally has a basket and Ann has a box. Sally puts her marble in the basket and leaves the scene. While she is out, naughty Ann takes the marble out of the basket and puts it in her box. Question: where will Sally look for her marble? The results showed that 80% of the children with autism failed, pointing to the location where the marble actually was, whereas the majority of both control groups passed (85% of the normal developing children and 86% of the Down syndrome children). This finding has been replicated in a number of studies, using the verb "think" in the question rather than "look" and real people rather than dolls (Leslie and Frith, 1988). Furthermore, Perner et al. (1989) replicated the findings from the "Smarties" task with a control group of children with specific language impairment, so that the possibility that children with autism failed to understand false-belief tasks on the basis of a general language deficit was ruled out.

An important variation on the tasks described above was made again by Baron-Cohen, Leslie and Frith (1986) in order to rule out the possibility that children with autism were failing false-belief tasks on the basis of a weak reasoning ability with story sequences. The same children from the "Sally and Ann" study were shown frames of a cartoon strip depicting different kind of events, and asked to put the pieces together in order to make a story. In both the "mechanical" sequence (e.g. "the balloon goes up in the sky, gets stuck in the branches of a tree and explodes") or the "behavioural" sequence (e.g. "a girl enters a shop, buys some sweets and leaves") there was no need to understand mental states in order to make a coherent story. By contrast, the "mentalistic" story was constructed around the mental states of a protagonist following the model of the Sally-Ann test (e.g. "a boy puts a chocolate in a box and goes out to play. Mother eats his chocolate while he is away. When he comes back, he is surprised to find the box empty"). The results showed that children with autism performed poorly only in ordering the pictures of the mentalistic story, but not on the other two stories, suggesting that the ability to understand cause-effect sequences on the basis of people' mental states is very different from the ability to understand causality on the basis of people's behaviour or on the basis of objects' physical properties.

Children with autism fail a whole array of other false-belief tasks, with different controls and methodologies. Alternative reasons for task failure, such as motivation, language or memory deficits have been ruled out through a "fine-cuts" approach, namely, using pairs of extremely similar tasks that differ only in the need for mentalising. One example is the ability to understand and engage in sabotage, based on the use of physical means to prevent someone from doing something, as opposed to deception, based on the manipulation of mental states. As predicted from the fine-cut methodology, children with autism found it more difficult to engage in deception than in sabotage (Sodian and Frith, 1992).

While it was important to define in autism the deficits resulting from a lack of Theory of Mind (Baron-Cohen 1987, 1988, 1989 a,b, 1991, 1992; Happé, 1993, 1994; Leslie and Frith, 1988; Frith, Happé and Siddons, 1994; Mundy, Sigman and Kasari, 1993; Tager-Flusberg, 1993) it was also important to investigate the nature of the developmental shift between 3- and 4-years in normally developing children. The issue of Theory of Mind deficit in autism became interconnected with the issue of the failure of 3-year-old children. Paradigms first introduced to explore mentalising in normal children were also used with children with autism. In fact, it is crucial to understand the reason why the two groups fail the same task: could it be that young children are similar to children with autism at the cognitive level?

A particular paradigm based on the understanding of a false state of the world rather than a false belief is of interest for the purpose of providing evidence to support the metarepresentational model and in particular, the Theory of Mind deficit account of autism. Indeed, it was nicely demonstrated that children with autism do not find difficulties with any "out-of-date" representations, but only with those based on mental states. One paradigm (Leekam and Perner, 1991; Leslie and Thaiss, 1992) was based on the "false photograph" task devised by Zaitchik (1990), in which a Polaroid photo is taken of a scene, then placed face-down in front of the child while the scene is changed (e.g. an object is moved from the scene, or is swapped with another object). Children are then asked: "In the photo, where is the object?"

The test is of particular importance because it not only provides a test for the specificity of the mentalising deficit in children with autism, but also defines a critical distinction of the causes determining a failure. Indeed, Zaitchik (1990) showed that normally developing 3-year-old children find this task as difficult as standard false-belief tasks, whereas 5-year-olds find no difficulties in both tasks. Leslie and Thaiss (1992) demonstrated that all the children with autism with a verbal mental age ranging from 4.4 to 14.5 years passed the false photograph tests and the majority of them (77%) failed the false-belief tests. By contrast, normally developing children of 4.5 years passed all tests. This finding has been replicated with analogous tasks, e.g. using line drawings and maps, thus dismissing concerns that the children's inexperience with cameras might effect their response (Charman and Baron-Cohen, 1992, 1995; Leslie and Thaiss, 1992).

Another relevant study has shown that younger normally developing children not only fail "out-of-date" mental state tests (e.g. false-belief tasks) but also "out-of-date" non-mental states tests (e.g. "screen task"), with no metarepresentational content, that older children with autism are able to pass (Roth and Leslie, 1998). In the screen task children are first presented with a scene, then the objects are moved behind an opaque screen, and a new set of objects is placed in *front* of the child who watches a change of object location. The child is then asked to indicate the location of the object *behind* the screen. Both children with autism older than 4 years and normal 3-year-old children fail to represent mental states (false-belief test), but only the 3-year-olds fail to represent an out-of-date state of affairs (screen test), due to difficulties with executive function demands of the test.

Taken together these results support the metarepresentation model and the Theory of Mind deficit account of autism. According to the model, beliefs must be understood as representations of propositional attitudes. Children aged 3 years do cannot demonstrate understanding of propositional attitudes in false-belief tests for reasons of immature executive functions; children with autism do not pass false-belief tests because they have problems in understanding propositional attitudes. In other words, the model makes a distinction between competence and performance in mentalising ability. The failure of 3-year-olds with either mentalising tasks (false-belief), misrepresentation tasks (out-of-date photographs, drawings, maps) or executive control tasks (scenes hidden by a screen) all reflect a performance limitation associated with executive functions not yet fully developed. By contrast, failure on false-belief tasks and success on all other tasks by children with autism reflect a selective impairment in representing the mental states of others. This is a competence limitation due to an impaired mentalising system, rather than to executive functions processing problems. In this regard, 3-year-old normally developing children and older children with autism are very different from each other, even though both fail false-belief tasks. Strong corroborative evidence comes from the subtle but effective manipulation of task instructions. Several investigators have modified the false-belief task so as to make it simpler – for instance, by making the question simpler, more specific and more pragmatically natural (e.g. Siegal and Beattie, 1991; Surian and Leslie, 1999). Surian and Leslie showed that 3.5-year-old normally developing children who failed the standard question "Where does Billy think the ball is" passed the modified question "Where is the *first* place Billy will look for his ball"? However, children with autism who performed at the level of the normal 3.5-year-old children on the "think" question, were not facilitated by the "look first" question. According to the authors, the format of this question enables the prepotent response, which maintains that Billy has a true belief (as a default assumption), to be inhibited.

1.2.3 Individuals with autism who pass false-belief tests

It is quite paradoxical that one of the challenges to both the metarepresentational model and the ToM deficit account of autism emerges directly from the evidence supporting the deficit account. In fact, the study by Baron-Cohen et al. (1985) showed that 80% of children with autism did not understand a false belief, but the remaining 20% were able to understand it. A study by Happé (1995) reported that in most studies, a proportion of children with autism that varies between 15% and 60% passed first order false-belief tasks, requiring the inference of one person's mental state. These children were in general more verbal and older than those who failed. Furthermore, ceiling effects have been reported with higher-functioning individuals with autism or Asperger Syndrome (e.g. Dahlgren and Trillingsgaard, 1996).

Baron-Cohen (1989) suggested that people with autism are grossly delayed in their acquisition of Theory of Mind, so that they are able to pass simple but not more advanced tests, such as the second order false-belief task requiring the inference of a person's false-belief about what another person believes. The task was first devised by Perner and Wimmer (1985) who demonstrated that normal developing children between 5 and 7 years understood a story involving two characters, Mary and John, and an ice-cream van. Mary and John find out independently that the van moved to a new location. Subjects are asked to indicate where Mary thinks John will go to buy an ice cream. The correct answer requires the representation of Mary's false beliefs about John's belief. Baron-Cohen administered this test to older individuals with autism who did pass the Sally–Ann test, and found that they were unable to represent Mary's false belief.

Nevertheless, a proportion of individuals with autism also pass second order ToM tests (Ozonoff, Pennington and Rogers., 1991; Bowler, 1992). These individuals are usually adults with autism or Asperger Syndrome who are high-functioning in terms of IQ and language, but are still socially impaired. Those who can pass second order ToM tests however, may have difficulties in even more advanced test involving complex mental states such as bluff and double bluff (Happé, 1994), decoding complex mental states from the expression in the eye region of the face (Baron-Cohen, Joliffe, Mortimore et al., 1997; Baron-Cohen, Wheelwright, Joliffe et al., 1997), or detecting a faux-pas (Baron-Cohen, O'Riordan, Stone et al., 1999). Furthermore, it is important to stress that autism has a developmental course, so that the appropriate way to disprove the developmental delay hypothesis for autism would be to find a case of a child with autism who passes all available Theory of Mind tests at the appropriate age/mental age. Since no such case has yet been documented, the hypothesis of impairment in mentalising, involving at least a delay, is still tenable for all cases of autism.

1.3 Testing individuals with autism with non-language based ToM paradigms

In the first sections of this chapter I have briefly summarized how the Theory of Mind deficit account of autism and the metarepresentational model, from which it stems, have provided cognitive psychologists with a specific and powerful theory that can explain the social and communicative impairment in autism.

However, an interesting question regarding the ToM hypothesis has emerged from studies indicating that a proportion of individuals with autism are successful in passing Theory of Mind tests. It appears that these high-functioning people, usually with superficially good language skills, acquire a Theory of Mind with time and experience. Nevertheless, they show persistent social difficulties just as other individuals with autism. This suggests that the tasks are not sufficiently sensitive to capture the persistent problem, and that autistic individuals are able to bypass the core impairment by adopting compensatory strategies. Verbal ability appears to be a key factor for successful performance on language-based ToM tasks, without necessarily promoting real-life social adaptation. Several studies have shown that ToM task performance is correlated with level of verbal skills (e.g. Bowler, 1992; Fombonne, Siddons, Achard and Frith, 1994; Happé, 1995). The simple, standardized verbal presentation of the ToM tasks is likely to facilitate the understanding of the experimental situation. In fact, the problem to be solved in the ToM tasks is explicitly defined by the question posed, e.g.: "Where will Sally look for her marble?" Moreover, the verbal response, which is either correct or wrong, may not reflect the degree of the subject's mentalising ability. Thus, the challenge for creating experimental mentalising tasks is to bypass learned strategies and tap real-life impairments by using non-verbal stimuli.

According to the metarepresentational model, mentalising is a "property of our cognitive apparatus that comes into action when triggered by particular stimuli, and it makes sense of other people's and our own behaviour fully automatically" (Morton and Frith, 1995, p. 363). It is of particular interest, therefore to investigate the nature of stimuli that provoke the pervasive tendency of people to explain behaviour in terms of mental states. We are constantly bombarded with sensory inputs that alert our system to the presence of an agent, inputs that make us adopt the so-called "intentional stance" (Dennett, 1981) towards that agent in order to understand and predict its behaviour. Thus, if Theory of Mind failure persists even in able individuals, then it should be possible to investigate obligatory, non-verbal, triggering inputs for mentalising. It is this issue that is the main focus of this thesis.

The first three experiments explored the most familiar visual inputs that trigger the attribution of mental states, namely, facial expression of emotions. The fourth and fifth experiments focus on the perception of moving abstract shapes, exploring the ability to attribute to an agent an intended goal in a specific context. The last two experiments investigate the neural basis of the ability to attribute complex mental states, beliefs, desires and feelings to two interacting agents. In all these studies, the question is asked whether the paradigm adopted reveals differences between autism and control groups. A brief introduction to the three paradigms used in the present experiments is given below.

1.3.1 Understanding emotions from facial expression

One characteristic of emotions is that they are clearly visible on the face of a person. They provide therefore salient stimuli to investigate the ability to attribute mental states to others in autism. There has long been a view that autism stems from "an innate disturbance of affective contact" (Kanner, 1943). Hobson, in particular, proposed that impairments in the expression and comprehension of affect are primary in autism, and do not arise from a cognitive deficit (Hobson, 1986, 1989). The fine-cut method adopted to define subtle differences, distinguishable only by the need of mentalising, within the same behavioural domain, predicts that children with autism have difficulties in recognizing only the emotions that are triggered by beliefs, and not the emotions that are triggered by states of affairs.

With the tremendous growth of the neuroscience approach to developmental disorders, new hypotheses are formulated relative to specific brain abnormalities and the symptomatology of autism. The wide range of emotion stimuli adopted in the first study allows for testing several predictions relative to specific emotion recognition impairments in autism. Chapter 2 is devoted to these studies.

1.3.2 Understanding intention from an agent's goal-directed motion

What is the minimal requirement for attributing mental states to an agent? Surprisingly, the answer is not that the agent needs to be human or human-like (e.g. puppets, cartoon characters). The social environment in which we live provides us with abundant visual stimuli that are in motion. Again, we understand mental states by reading micro-movements of the facial muscles or the changes in body posture, or the motion of the whole body in relation to another body or object. However, one of the simplest forms of visual information for judging agents' intentions comes from observing their motion trajectory alone. The early studies of Michotte (1946/1963) and Heider and Simmel (1944) provided evidence that people are able to automatically translate from the domain of pure physical movement into the domain of intentions and desires. These pioneers of research on the perception of causality and animacy used simple small moving 2D geometric shapes, e.g. triangles, squares, circles. The importance of this type of research is that it revealed that the visual system could recover causality and intentionality from the kinematics of minimal stimuli. This work has inspired two animation paradigms, the first aimed at investigating the perception of goal-directed intention in children with autism (Chapter 3), the second exploring the neurocorrelates of the attribution of intentionality, or complex mental states, in healthy controls (Chapter 4) and in adults with autism (Chapter 5).

Some clarifications are needed for the terms of "intention" and "intentionality". In order to simplify the matter, which pertains to the philosophy of mind and of language, I will refer to the clear definition of semantic properties of propositional attitudes provided by Quine (1961) and see whether they apply to intention. As already mentioned at the beginning of the chapter, "intentionality" refers to all mental states, so that mental states are *intentional*. These include propositional attitudes (beliefs and desires), emotional states, and intentions. Their difference can be expressed in terms of their intentional contents. As stated above, beliefs and desires "point towards" a proposition whereas emotional states point towards a state of affairs. Intentions or goals and aims point towards a *future* state of affairs.

The most important distinction is between desire and intention (and indeed, this is where confusion often arises). The simplest form of intention implies an action, whereas desire implies always an attitude towards a proposition, e.g. "Mary *wishes* that she goes to Mars", is different from "Mary *intends* to go to Mars". Indeed, if Mary intends to go to Mars she has to set off and organise the interplanetary journey, whereas if she wishes to go to Mars, she can simply daydream about it. Obviously in our everyday language such a distinction is not easy to apply: the terms "wanting" and "trying" are used frequently and are more ambiguous. If Mary says: "I want to go to Mars", it is not clear whether she is just uttering a metaphor, or whether she will really try to go to Mars. However, from the very pragmatic view of creating tests that tap the ability to understand different types of intentional states it suffices to clarify that the subject is required to make inferences regarding different intentional contents. Whether different contents of intentional states are reflected in a different cognitive organization, is a matter for investigation.

In conclusion, the test for the perception of intention is based on the perception of the movement of an agent that points towards a future state of affairs (to be next to a stationary goal) attempting to change the physical circumstances (moving from one location to another). In this sense, the action of the agent is perceived simply as goal-directed and does not require representation of propositional attitudes (beliefs and desires). The challenge for creating the new stimuli consisted in defining the context in which the agent was moving and the agent's movement properties. All these aspects will be presented in detail in the introduction to the experiment.

1.3.3 Understanding mental states from an agent's complex patterns of motion

Imagine the following scenario: two children playing at a distance from their parents. The adults cannot hear what they are saying to each other, neither can see their facial expressions. Nevertheless, they are able to distinguish whether the children are playing or fighting, or whether one is trying to convince the other or is upsetting the other. Indeed, by perceiving the motion cues, parents are able to understand the children's intentions.

In Chapters 4 and 5 I will present two neuroimaging studies on the attribution of intentionality based on the perception of the motion pattern of two interacting agents. The underlying assumption is that, as outlined above, the perception of motion provides a salient triggering input for mentalising.

Following the tradition of early work on attribution of intentional behaviour in simple moving stimuli by Heider and Simmel (1944), Blythe, Todd and Miller's (1999) study investigated people's accuracy in attributing intentionality with an original computer-based method. Interestingly, Blythe et al. created the stimuli with the assumption that animate motion tends to fall into a few rather stereotyped categories that can be derived from basic evolutionary and ecological principles. They first defined six types of motions that were expected to be distinct and clear because of their functional importance: pursuit/evasion and fighting (survival domain), courtship (reproductive domain), leading/following, guarding/invading and play (survival and reproductive domains). The motion trajectories of two agents' interacting was then created on-line by two subjects operating on separate computers (e.g. one subject generated the movement of one coloured shape as it was pursuing the other shape. At the same time, the other subject generated the movement of the shape as it was evading the other). A third subject categorised the motions as they were generated into one of six events. The different motion patterns created via this simulation were then plotted against time: faster motions appeared as more horizontal lines, slow motions as more vertical lines. These plots were then shown to a new group of adults, who categorised them into one of the six original intentional actions with great accuracy (77%).

The paradigm adopted in this thesis was developed by Uta Frith and Francesca Happé with the aim of testing on-line mentalising ability in children with autism using stimuli that would selectively evoke mental state attributions by their kinematic properties alone. In a study by Abell, Happé and Frith, (2000) children were presented with different type of computer-based animations depicting two interacting triangles. The task was designed according to the fine-cut approach to create similar motion patterns that differ only in the mentalising demands. In the Theory of Mind animations the interacting movement of the two characters evokes descriptions of the agents as if they had in mind the other's mental state. In the Goal-directed animations the interaction between the two triangles evokes a description in behavioural terms, without referring to specific mental state processing. In the Random animations the absence of interaction between the triangles simply evokes descriptions in terms of non-deliberate actions. As in the adults' study by Blythe, Todd and Miller (1999) results indicated that different types of motions selectively evoked in children the attribution of mental states and goal-directed actions as opposed to non-deliberate actions. Furthermore, high-functioning children with autism used mentalistic descriptions less often than normally developing 8-year-olds, but as often as children with general intellectual impairment. However, children with autism frequently used mental states that were inappropriate. This finding provided a promising start for creating a sensitive ToM paradigm for able adults with autism.

The advantage of these new animations compared to the classic Heider and Simmel (1944) movie is that they are based on a clear-cut distinction between stimuli eliciting ToM based descriptions and non-ToM description. This type of test is particularly appropriate for neuroimaging techniques, based on the principle of contrasting brain activity associated with minimally different tasks, which contrast in terms of their cognitive demands.

1.3.4 Preview of the findings

The thesis examined the possibility of using non-language based ToM paradigms in both behavioural and neuroimaging (PET) studies.

The first set of three experiments explored the ability to understand the emotional states of others through their facial expression based on standardized photographs (Ekman and Friesen, 1976) in children with autism in comparison with normally developing children. The findings indicated that the perception human faces displaying emotional expressions may not represent a sensitive test to capture mentalising deficits in autism.

The second study investigated the ability to attribute goal-directed behaviour to agents by using animated sequences of a moving agent. The results showed that children with autism and normally developing children performed equally well in understanding an agent's intention on the basis of goal-directed motion. The only difference was revealed in an ambiguous condition of the experiment, when children with autism and younger normally developing children tended to attribute the agent's intended goal on the basis of the agent's proximity to a target, whereas older normal children and adults attributed the agent's intended goal on the basis of the persistent direction of the agent's movement towards a target.

The two neuroimaging studies investigated the neural basis of on-line attribution of mental states in healthy individuals and in individuals with autism. The stimuli consisted of silent animations depicting two interacting geometrical shapes. The movement patterns selectively evoked either mental state attribution or simple action description. The results indicated that the regions activated form a network for processing visual-kinetic information about intention in action, and that individuals with autism showed reduced activation in several regions of the mentalising network, but showed normal activation in extrastriate cortex. Unlike the behavioural data of both the emotion judgement and the goal attribution task, the verbal descriptions of the animations suggested that this paradigm is sensitive to Theory of Mind difficulties even in very able individuals with autism.

Chapter 2

Understanding emotions from facial expressions

In this chapter, I will present three experiments on the ability of children with autism to perceive facial expressions of emotion. Although the emotion category includes both simple and complex mental states, the present paradigm investigates only the so-called "basic" emotions. In the first part of the chapter I will underline the relevance of the basic emotions' approach for current research on emotion processing by introducing the psychoevolutionary framework originating Darwin's works and continued in Ekman's research. I will then specify some perceptual aspects of emotional expression processing, with particular attention to the fine-grained paradigm (Calder, Young, Rowland et al., 1996; Young, Rowland and Calder, 1997) that has been adopted in the present study. The hypothesis of a general affective deficit in autism has been investigated by using various types of stimuli in different emotion-recognition tests. A different selective emotion recognition deficit has been suggested on the basis of a neurocognitive model of autism that implies abnormal amygdala functioning. The ToM deficit account of autism made an important breakthrough in this area by suggesting a selective impairment in recognising belief-based emotions as opposed to reality-based emotions. A number of experiments have tested different hypotheses with somewhat inconsistent results. The present investigations attempt to replicate and extend these findings.

2.1 Recognising basic emotions

The paradigm adopted for the present experiments is based on the recognition of six "basic" emotions (anger, fear, disgust, happiness, sadness and surprise) characterized as rapid, fail-safe responses to stimuli correlated with basic survival needs. This classification has its roots in Darwin's evolutionary approach to emotional phenomena and in the recent work of Paul Ekman and his collaborators who have confirmed and extended many of Darwin's results. The strongest evidence for distinguishing one emotion from another comes from research on facial expressions showing high agreement across literate and preliterate cultures in judging what these expressions signal (Ekman, 1989). Since the basic emotion approach has provided evidence for the utility of facial expression as defining characteristics of specific emotions, it seems particularly appropriate to use a recognition test of the six basic emotions in order to investigate specific impairment in emotional processing of children of autism.

2.1.1 Psychoevolutionary research on facial expression recognition

Research on emotional expressions has its roots in Darwin's "The Expression of the Emotions of Man and Animals" (1872, 1998) where he argued that human expressions of emotions are evolutionarily ancient, reflex-like responses with adaptive functions, and they are homologous with expressions in related species. The starting point for Darwin's theory was the production of a detailed morphology of facial expressions of various emotions displayed by humans and animals and their possible functions. He then proceeded by collecting information on the recognition of emotion by members of other cultures using photographs of people displaying different expressions. Whereas Darwin himself obtained reliable data in England, the cross cultural information was collected through questionnaires dispatched to missionaries and traders, yielding not necessarily trustworthy results. Nevertheless, Darwin thought he had the evidence that the same emotional expressions were found throughout a wide range of populations, and this allowed him to infer that emotional expressions were "innate or "instinctive traits" (Darwin, 1872, 1998, p. 22). Despite the lack of rigorous modern techniques, Darwin's approach established a solid foundation for contemporary research on emotion. In particular, Paul Ekman investigated the hypothesis of the universality of facial expressions within the evolutionary framework that characterises basic emotions as physiological responses with rapid onset, short duration, unbidden occurrence, automatic appraisal, and coherence among responses. The set of changes that constitute the emotional responses allows humans to begin to deal quickly with fundamental life-tasks in ways that have been adaptive in the evolutionary past (Ekman, 1992, p.195).

Over the past 30 years numerous studies of literate and in preliterate cultures have shown that anger, disgust, fear, happiness, sadness and surprise are universally recognised (Ekman, 1989; Ekman, Friesen and Ellsworth, 1972; Ekman and Oster, 1979; Izard, 1971, 1977). The emotion of contempt has also been studied, but there is no general consensus on the collected data to be considered as an emotion underlying universal facial expressions. Interestingly, the expression of surprise was found to be recognised only in one out of two isolated cultures studied by Ekman and Friesen (1971): the South Fore people of Papua New Guinea, and the Dani people of West New Guinea.

In the study of the South Fore people, Ekman and Friesen utilised a judgement task designed for working with children. Subjects were given three photographs, each showing a face, and told a short story (for example: "A man has learned that his child has just died"; "Your friend has just come and you are happy") which was designed to involve only one emotion. This method has the advantage of avoiding the necessity of translating emotion terms. The pictures to be matched to the stories depicted facial expressions displaying happiness, sadness, anger, fear, surprise and disgust. In addition, participants were asked to show how their own face would look if they were the person in the emotion story, and their expressions were videotaped to be shown to U.S. college students. Results of face recognition confirmed the extraordinary agreement among literate cultures with only one exception: the expressions of surprise and fear were not perceived as distinct, although both were distinguishable from angry, sad, happy and disgusted expressions. Interestingly, the South Fore people's spontaneous display of fear was often judged by the western cultured group as a display of surprise. Ekman's explanation is that the confusion between these expressions reflects the fact that fearful events are usually also surprising. More specifically, surprise has been defined by Davidson (1992) as an "approach emotion", being associated with a call for further processing. It would, therefore, arise in contexts where the organism requires additional information prior to a final decision about an appropriate action. In this sense, the surprised expression is more likely to be followed by another emotion, as soon as the organism processes the extra information required to solve the discrepancy between what was at first believed and the current reality.

The procedure adopted with the South Fore people was repeated with the Dani people, who showed to recognise all emotions with the only exception of disgust and anger which were not differentiated. However, this outcome was predicted because Dani avoid expressions of anger, and often mask it with disgust (Ekman, 1972).

It must be noted that one of the weaknesses of this recognition test lies in the experimental task. It is possible that people from illiterate cultures found difficulties in comprehending the nature of the story, or most importantly, that the stories were specific to Western and did not appropriately convey the association of the expected emotion. In addition, the forced-choice across the range of basic six emotions might have limited the investigation on other possible frequent associations between events and emotional expressions in a particular illiterate culture. The inclusion of control stimuli depicting neutral faces or emotions that do not conform to the definition of basic emotions might have controlled for cultural specific associations of events (or specific emotion label used in the story) and expression.

In sum, the evidence for the universality of facial expression, while robust, is limited to just six emotions defined as involuntary stereotypical responses involving physiological changes. However, for the particular occasions that will most rapidly call forth an emotion, attitudes about emotions, display rules and behavioural consequences are all factors dependent on social and cultural variation. According to Ekman, the complexity of individual and social experience creates different and malleable aspects of emotional expressions, but they are perfectly compatible with the universality hypothesis. Ekman's view is essentially twofold, making a distinction between universal facial expressions, and learned, cultural specific expressions (Ekman, 1998). When he refers to universal expressions across cultures he refers to involuntary muscular movements displayed in relation to an inner emotion. Nevertheless, facial expressions can be controlled deliberately, as when people exhibit them according to social rules for displaying or concealing certain feelings. The signal in the deliberate, social use of facial expressions is not the same as expressions occurring involuntarily, or in the deliberate attempt to conceal them (Ekman and Friesen, 1969). Evidence of how display rules greatly affect the expressions of emotions is represented by Ekman's cross-cultural study involving subjects from Japan and U.S.A. watching alone or with a presence of a stranger a video-clip designed to elicit negative emotions. Results indicated that both groups displayed negative emotions while alone, but only the Japanese subjects displayed positive expressions when a stranger entered the room: they were masking their negative emotional responses by a polite smiling expression (reported in Ekman, 1972).

Finally, the implications of the psychoevolutionary theory of emotional phenomena relative to the concept of innateness need to be taken into consideration. Indeed, the experiments reviewed here on facial expression of emotions might be thought to show not only that they might be given evolutionary explanations, but also that they are innate (see, e.g., Ekman, 1998). However, the notion of innateness is an ambiguous concept in psychological theory. The fact that expressions have an evolutionary history has no implications about the nature of the process by which it develops, except that the outcome of the very process is consistent enough to allow stability. In fact, it is possible for the expressions of emotions to be at the same time pan-cultural and stable over many generations and yet being sensitive to environment inputs (Griffiths, 1997).

The lack of commitment on how emotions are built in ontogeny does not reduce the explanatory power of the psychoevolutionary theory of emotions. Indeed, the basic emotion approach has guided much of the recent research on emotion specific neuro-physiology providing, for example, the basis for developing distinct paradigms on both the perception and production of emotions.

2.1.2 A paradigm for emotion recognition

The paradigm adopted in the present study stems from step-by-step refinements of stimuli used to investigate the categorical perception of facial expression of emotions. In a pioneering study, Etcoff and Magee (1992) used computer-based stimuli rather than standardised pictures for the first time. The images were based on line drawings of stereotypical expressions of the six basic emotions, transformed ("morphed") in order to capture the physical transformation which occurs naturally in a human face that changes expression from one emotion to another. The study demonstrated that the six basic expressions of emotions, with the exception only of surprise, are perceived categorically. This means that the visual system underlying the ability to recognize expressions of emotion uses continuously transforming facial cues in combination to create perceived categories. For example, if I see a person who is angry at first and then becomes fearful, I would perceive a sudden shift between the two expressions without a region of uncertainty or mixture. However, in the case of a person changing from surprised to fearful, I would not find sharp boundaries between the two expressions. This result, combined with the lack of unequivocal evidence from preliterate cultures, suggests that the expression of surprise, associated potentially with both negative or positive emotions as such as fear or happiness, is somehow perceived differently from the other basic emotions

(Ekman, 1984; Lazarus, 1991). Nevertheless, surprise fits the psychoevolutionary model of emotion criterion of being a short-term involuntary response with autonomic arousal, and as such should be included in investigations of basic emotions.

The study presented in this chapter adopted a paradigm that was originally created to replicate and extend the Etcoff and Magee's study (Young, Rowland, Calder et al., 1997), and that was used a study investigating emotion processing in brain lesioned patients (Calder, Young, Rowland, et al., 1996a).

The paradigm is based on computer-generated stimuli, obtained by transforming the standardized pictures of the Ekman and Friesen series (1976) containing the images of the six basic emotions that have been found to be accurately recognised, and sometimes confused, in most cultures of the world. Young et al. (1997) privileged the "negative" cross-cultural results of emotion recognition, and created a "confusion matrix" with Ekman and Friesen's data in order to quantify the degree of uncertainty across emotions. Results indicated that happiness and surprise are the least confused emotions (0.8%), whereas surprise and fear the most confused (5.8%) along with disgust and anger (6.4%). Relatively less confusing are fear with sadness (2.4%) and disgust with sadness (2.7%).

On the basis of the confusion matrix, Young et al. (1997) created an "hexagon" of facial expressions: each of the six emotions was placed next to the one it was most likely to be confused with, creating a sequence, starting with happiness and terminating with anger. The two extremes were then joined to form the perimeter of the hexagon. All possible combinations of pairs of emotions (e.g., fear-surprise, disgust-anger, and happiness-surprise) were then obtained by connecting the emotions on the perimeter. Each emotion pair constituted a continuum for testing categorical perception (a more detailed description of these stimuli is given in the study's material section). The results obtained from Young et al. (1997) study confirmed Etcoff and Magee's (1992) findings, providing strong evidence that facial expression of emotions are perceived categorically. Interestingly, the expression of surprise was sometimes identified as fear, and disgust as anger, but when this happened, it was usually because one subject consistently did this.

2.1.3 Perceptual basis of facial emotion recognition

Perceptual processing of facial expressions has also been investigated in relation to the different ability to process them holistically (perception of full face) or analytically (perception of distinct facial features) (Kestenbaum, 1992). Ekman (1979) has hypothesized that isolated features alone may be sufficient for identifying whether an emotion is positive or negative (global emotion categories), but not for identifying discrete emotions. For example, certain eyebrow movements may indicate negative expression in general, but cannot discriminate sadness and fear that are similar in that respect. In addition, some expressions share similar features and at the same time differ in small portions of the facial display: the upper eyelids are raised in both surprised and fearful expressions, whereas the lower eyelids are tense in fear and relaxed in surprise. In addition, a single feature might change within the same expression, as in the case of the mouth in angry faces: it could be either open with teeth barred or closed with tight lips in anger. In the case of disgust, where mouth and nose are contracted simultaneously, it is necessary to consider individual features jointly in order to identify it as distinct from anger (Ekman and Friesen, 1975).

Kestenbaum (1992) investigated categorical (negative-positive emotions) and discrete recognition of emotions in different age groups by showing expressions of happiness, surprise, fear and anger to 5 and 7 year-olds, and adults. Whereas fear and anger are clearly negative emotions and happiness positive, surprise is more ambiguous. Since the study explored also the differences in the ability to processing the stimuli analytically and holistically, the emotional stimuli were displayed as full face, mouth alone, eyes alone, or mouth and eyes combined.

In general, Kestenbaum's (1992) study supported Ekman's (1979) suggestion that distinct emotional expressions may not be easily identified from a single feature (because some expressions share features) but that positive-negative category could be identified. All four emotions had a dominant aspect that facilitates recognition of individual emotion: the mouth for happiness, the eyes for anger, fear and surprise. In contrast, only happiness and anger had a dominant feature for the negative-positive categorization (mouth and eyes respectively). Interestingly, some differences in perceptual processing of anger and fear, which were generally more difficult to recognize than happiness, appeared across groups. Younger children (5-year-olds) discriminated these expressions on the basis of single features (eyes or mouth alone), whereas older children and adults improved their performance using multiple features (eyes and mouth combined). Surprise was the most difficult emotion to recognize for all age groups, in particular for the 5-year-olds who were just as likely to include expressions of happiness as expressions of surprise when the target term was surprise. However, surprise appeared to be featurally based, being recognized as often from the eyes as from the full face or combination eyes-mouth. Interestingly, the expression of surprise was more often identified by all groups as a negative emotion ("feeling bad" versus "feeling good") much like the visually similar expression of fear. This indicates that the ambiguity of surprise is twofold: it is perceptually confused with the positive emotion of happiness, and at the same time it is associated with the semantic category of "feeling bad". The authors suggest that the ability to recognize an expression and the understanding of a particular emotion concept may not develop concurrently, even though the same vocabulary term may be used for each.

2.2 Autism and emotions

Since the original clinical description of children with autism first described by Kanner (1943) included a profound lack of affective contact with other people, psychologists have been evaluating the primacy of the social and affective impairments in autism. A variety of studies investigating emotional processing were inspired by clinical observations of children with autism attesting to an ability to recognise emotions in others and to use them in a communicative fashion (Wing, 1981). The empirical research on affective impairment of individuals with autism is wide and heterogeneous so that it is not surprising that the findings are extremely mixed. I will only present a selection of studies investigating a *general* affective impairment focusing on the ability to recognise facial expressions of emotions.

Along with the growth of the neuroscience approach to developmental disorders, some models of autism have suggested a direct relationship between specific brain abnormalities and social impairment. In particular, one distinct neurocognitive model of autism has been suggested based on the hypothesis of abnormal brain functioning specific to the amygdala (Baron-Cohen, Ring, Wheelwright et al., 1999; Baron-Cohen, Ring, Bullmore et al., 2000; Howard, Cowell, Boucher et al., 2000). Unlike both the general affective and the neurocognitive deficit hypotheses, the Theory of Mind deficit account of autism allows for investigating a *selective* emotion processing deficit by contrasting abilities that do and do not necessitate mentalising ability.

2.2.1 A general affective deficit hypothesis

Hobson's extensive work suggests that the primary deficit in autism is based on a lack of basic perceptual-affective abilities and propensities that are required for an individual to engage in interpersonal relations (Hobson, 1993). The hypothesis of an affective deficit has been investigated by using various types of stimuli in different emotion-recognition tests. In a cross-modal task children with autism were asked to match sequences depicting people (videotapes) with a series of drawings and photographs depicting gestures and facial expressions of happiness, sadness, anger and fear. In addition, children were asked to match the visual stimuli with vocal expressions of emotions (Hobson 1986a,b; Hobson, Ouston, Lee, 1988a). Children with autism showed more difficulties than control groups in all matching tasks. Similar findings have been reported by Tantam et al. (1989) in either naming or discrimination tasks based on the standardised pictures of the six basic emotions from the Ekman and Friesen series (1976). Another study (Weeks and Hobson,

1987) investigated the tendency to process different types of descriptive information relative to people. Children were asked to sort out a set of pictures showing people (males and females of different age) wearing either a woolly or a floppy hat and displaying either happiness or sadness. Children with autism spontaneously sorted the pictures according to the type of hat, whereas the control group sorted the stimuli according to the facial expressions. It is unclear whether children with autism privileged the hats on the basis of an affective deficit or on the basis of the perceptual salience of the stimuli. Unfortunately, these findings are not sufficient to support the primary affective deficit hypothesis in autism, since all the tasks described above tap more perceptual matching processes than emotion recognition processing. However, they offer the opportunity to explore general behaviour patterns induced by salient social cues (facial and vocal expressions, gestures). In addition, Ozonoff, Pennington and Rogers (1990) showed, by reviewing and criticising the relevant literature on affective deficits in autism, that the choice of control groups may influence the likelihood of obtaining group differences. They concluded that children with autism do not show a primary deficit in emotion perception if compared with controls of the same language level.

Sigman and her colleagues adopted a different experimental approach to investigate young autistic children's response to affective states by testing them in semi-naturalistic settings. In a study of responses to the negative emotions of others, Sigman, Kasari, Kwon and Yirmina (1992) observed young children with autism and normal developing children matched on mental age during three different affect contexts: when an adult pretended to hurt herself with a toy hammer (distress), when she pretended to be afraid of a toy robot which appeared suddenly in front of them (fear) and when she lay down pretending to feel ill (discomfort). Children's behaviour during the events was coded along four dimensions: a) attention to emotional display, b) behaviour towards the adult, and c) towards a toy, d) facial affects. All children with autism failed to look attentively to the adults showing negative emotions. In addition, their behaviour (playing more with their toy, withdrawing from the situation without seeking parental comfort) indicated they were overall less concerned when the adults displayed distress than the controls, but equally concerned when the experimenter pretended to feel unwell or afraid. All three emotional events elicited neutral facial expressions in all children, indicating that the test situations appeared to be interesting to them but not upsetting.

Another study from the same group investigated the response of children with autism to positive affect situations (Kasari, Sigman, Baumgartner and Stipek, 1993). They found that children with autism responded less than controls when their mother praised them for successful completion of a puzzle, indicating that strong positive affect attracts no more attention in autistic children than strong negative affect.

Two follow-up studies of the original sample of children with autism who took part in the above studies (Sigman et al., 1992; Kasari et al., 1993a) and Kasari et al. (1993b) re-assessed the children at school-age to investigate short and long-term stability of responses to other person's affect (Dissanayake, Sigman and Kasari, 1996). In the first follow-up study, 17 months after children's initial visit to the laboratory, children were exposed to the distress of an adult who pretended to hurt her finger by hitting it with a toy hammer. Results indicated that individual differences in early age to the distress of others remained the same after more than a year. In the second follow-up study, 5 years after the first visit, children were again tested with two different experiments, the "tea party" and the "telephone conversation". In the first experiment, children were engaged in playing at a tea party when an adult, who entered the room, pretended to hurt her knee and began to cry in pain. In the second experiment, children were exposed to two staged telephone conversations while they were sitting at the table playing with the tea-party toys. During one of these conversations, an adult conducted a pleasant, neutral chat, whereas in the other, he/she simulated an angry discussion over a faulty water connection. The "telephone conversation" test is of particular interest because it allows comparing children's ability to discriminate adult's affective behaviour on the basis of their vocal and facial expressions. Interestingly, results indicated that children with autism (between 6.6 and 12.4 years of age) looked longer and showed more empathy (measured with a score 0-6 ranging from not interest to intense affective involment and/or comforting behaviour) to the adult displaying anger than to the adult chatting normally over the phone. Unlike the previous studies, this evidence shows that autistic children and adolescents with autism respond differentially within affective context.

Overall, results from both follow-up studies demonstrated that autistic children's responses to another person's affect at pre-school age predicted their responses to similar displays over five years later. Furthermore, their degree of concern for the distressed adult was correlated with the level of cognitive functioning of the children.

Another study by Corona, Dissanayake, Arbelle, Wellington, and Sigman (1998) investigated a group of pre-school children with autism with the "distress paradigm" adopted by Sigman et al. (1992) by adding a control condition. Thus, the experimenter pretended to hurt herself and showed either distressed or neutral expressions. Results were in line with Dissanayake et al.'s (1996): children with autism were able to distinguish between negative and neutral affect but looked at the adult less often and for a shorter duration of time than control children.

In all the above studies conducted by Sigman and colleagues (Sigman et al., 1992; Kasari et al., 1993a,b; Dissanayake et al., 1996; Corona et al., 1998) a possible reason why children with autism fail to respond to others' emotions, in comparison to control children, may stem from a lack of understanding of these events, rather than an insensitivity to the display of emotions. More specifically, an impairment in mentalising could be an alternative explanation, since these tasks require that children understand that other people may have different mental states from their own. In addition, it is plausible that the level of intellectual functioning of these children may mediate these difficulties.

A very similar study investigated the capacity to respond to the distress of an adult and to a non-social orienting stimulus in two groups of pre-school children with high-functioning and low-functioning autism, and three control groups of children with mental retardation or developmental language disorder, and normally developing children (Bacon, Fein and Morris et al., 1998). At the beginning of the experiment the child was encouraged to play with a toy while the experimenter was chatting with a familiar adult (parent or teacher). During this time a loud animal-like honking sound was emitted from a speaker in the room. After this, the familiar adult left and the experimenter pretended, while playing with the child, to hurt her/his knee or hand. The orienting stimulus condition is of particular interest because the ability of refering to an adult when experiencing an ambiguous event implies an ability to monitor other people's knowledge and to understand that it might be different from one's own knowledge. Interestingly, results showed that only the low-functioning group displayed significantly less response than control groups to the adult's facial and vocal expressions, and did not look directly at him or her. The high-functioning children with autism were as likely as the controls to respond to the adult's simulated distress. However, both high- and low-functioning children with autism looked at the adult significantly less than the normal children when they heard the loud noise. This result indicates that children with autism, regardless of their intellectual ability, did not refer to the adult in order to gain additional information in the presence of the potentially distressing signal. In sum, the Bacon et al.'s study (1998) shows that the level of intellectual functioning mediates affective processing, and furthermore, that mentalising ability is independent from the ability to understand distress in others.

All the above studies based on semi-naturalistic settings show somehow some evidence of relative lack of responsiveness to others' emotions in children with autism. Children with autism did not respond to an adult displaying negative emotions (Sigman et al., 1992; Bacon et al., 1996) and positive emotions (Kasari et al., 1993a) but they were able to distinguish displays of anger, or distress, from a neutral expression (Dissanayake et al., 1996; Corona et al., 1998). In addition, they did not show the tendency to refer to an adult in salient situations regardless of affective display (Bacon et al., 1998). While all these studies offer the opportunity to measure social referencing abilities in children with autism, they are not sufficiently controlled to separate affective processing problems from other potential problems. Thus, the evidence of a lack of responsiveness to others' emotions in autism could be consequence of other core problems. For example, as already mentioned above, a lack of mentalising ability might account for the lack of concern for others' affect. In addition, difficulties in executive function abilities, e.g., an inability to switch attention, might also account for the lack of attention to an adult displaying emotions. In all the semi-naturalistic paradigms, pre-school children are expected to pay attention to an adult feigning distress, or anger, or even praising them, while being completely absorbed in playing with a toy. The inability to switch attention from a salient toy to relatively distant vocal and visual cues may to some extent explain the lack of interest to emotional displays (in particular if the child is of low intellectual ability).

The importance of controlling for executive function abilities also applies to computer-based paradigms and to high-functioning children with autism. In this respect a recent study by Grossman, Klin and Volkmar (2000) is of interest. The study investigated the ability to process facial expressions of emotion in a group of high-functioning children and adolescents with autism (7-18-year-olds). The authors held that the type of information available modulates emotion recognition impairment in autism. Specifically, they predicted that visual-verbal information would be more salient to individuals with autism than visual-affective information conveyed by facial expressions. Subjects were presented with standardised pictures of adults (Ekman and Friesen, 1976) displaying emotions (happiness, sadness, anger, fear and surprise). Each picture was matched with an emotion label. The association between visual and verbal stimuli was either correct (e.g., happy expression matched with "happy" label) or incorrect (e.g., happy expression mismatched with "afraid" label) or irrelevant (e.g., happy expression associated with "orange" label). Subjects were required to select the emotion label that "best says how the person is feeling" by touching a specially designed keypad with five response keys labelled with a different emotion name. Results showed that individuals with autism performed at ceiling when the visual and verbal stimuli were matched, indicating that their ability to recognise facial expressions of emotion is intact. However, they performed significantly worse than nonautistic control children at recognising emotions when the facial expressions were associated with an incorrect emotion label, but performed the same when it was associated with a non-emotion word. The authors argue that individuals with autism use explicit verbal strategies to solve emotion related tasks. This compensatory, language-based mediation allows them to pass tests that tap a basic cognitive level, e.g., simple recognition of facial emotions, but it is not sufficient at a more complex cognitive level, e.g., situations involving specific verbal biasing conditions. However, a possible deficit in executive function processing can also explain the impaired performance in the mismatched condition. Indication of executive function difficulties in participants with autism was indeed revealed by the significant correlation between their total number of errors in the mismatched emotion expression-emotion label condition with the total numbers of errors on the Wisconsin Card Sorting Task, a common measure of executive function deficit. A condition presenting mismatched visual and verbal information with non-emotions stimuli would have been therefore important. For example, one might present the subjects with pictures of neutral faces, or familiar animals or objects associated with both matching and mismatching labels. Only if a specific performance decrement on emotional stimuli were shown would it be warranted to talk of a difficulty in emotion processing. In sum, the Grossman et al.'s (2000) study made an interesting attempt to investigate the ability of people with autism to decode the emotional information beyond the basic ability to recognise basic emotions. On the other hand, it also showed how important it is to control for executive function deficit when subjects with autism are required to perform complex tasks.

In conclusion, the hypothesis of a general affective deficit in autism has been tested with different paradigms, but none of them was sufficiently well controlled to rule out other explanations. Therefore the question of whether individuals with autism have emotion processing impairments remains open to further investigations.

Psychophysiological responses to emotional stimuli

Another approach to investigating general deficit in responses to emotion stimuli in autism consists in examining autonomic responses to emotionally charged stimuli. A skin conductance study investigated the responsiveness of children with autism to negative emotions (Blair, 1999). Subjects were presented with pictures depicting distress cues (a crying face of a child) and threatening cues (e.g., an angry face, a pointed gun, a shark). The emotional stimuli were compared to neutral cues (e.g., an open umbrella). Children with autism (mean age 6 years) showed greater skin conductance response to the distress cues than to the neutral stimuli, just as the verbal mental age matched controls. This implies that their autonomic response activity in response to sadness is intact (this result will be considered again in the section discussing the hypothesis of a specific emotion deficit). Unlike the controls, children with autism did not show greater skin conductance to the threatening stimuli than to the neutral stimuli. Although the study did not address specifically the autistic children's response to frightening pictures, Blair suggested that the low skin conductance response to these cues might reflect a more general hyporesponsiveness of this population compatible with frontal lobe pathology.

The study by Corona et al. (1998) described above, investigated the hypothesis that pre-school children with autism show aversion to others' distress by measuring cardiac responses. Children were exposed to an adult displaying both distress and neutral expressions after she accidentally hurt herself. The behavioural data showed reduced attention and interest in the affect of others in children with autism than in children with mental retardation. Likewise, the autonomic response data showed that the heart rate of children with autism did not change across conditions, that is, there was no increase of heart rate as would be expected to an aversive stimulus. This result, dissimilar to Blair's study, indicates that the physiological measures of autistic children's reaction to emotion are abnormally low. However, as the authors noted, it is unclear whether the abnormal autonomic response was relative to the adult's affect display or to the particular social context. In fact, unlike the "telephone conversation" paradigm used by these authors, the setting did not simulate a familiar situation of an adult talking on the phone: adults do not frequently hurt themselves in presence of children. These findings need to be replicated with different types of social situations.

The studies exploring autonomic responses to the display of emotions in children with autism are of particular interest, but few as yet exist and it is not clear whether they indicate a general impairment in affective processing.

2.2.2 *A specific affective deficit hypothesis*

The neuroscience of emotion is beginning to be understood and neuroimaging studies have investigated differential neural responses to emotionally relevant material in adults with autism. The amygdala, the prefrontal lobes as well as the orbito-frontal cortex were suggested by Brothers (1997) to form part of a system that is assumed to underpin social abilities including mentalising.

The orbitofrontal region has been shown to have increased activation in healthy volunteers during a task involving judgement of mental state words in an early study by Baron-Cohen, Ring, Moriarty et al. (1994) using single-photon emission computed tomography (SPECT). Since no other neuroimaging studies have found this region associated with other tasks involving mentalising, I will focus on the neuroimaging evidence supporting the amygdala theory of autism.

Baron-Cohen, Ring, Wheelwright et al. (1999) adopted the "eyes task", an advanced mentalising test, in a fMRI study using complex mental states stimuli with a control task of gender judgement. Six adults with high functioning autism and twelve controls subjects were asked to make judgements about mental states or gender on the basis of stimuli that consisted of an individual's eyes. In contrast to the comparison group, the individuals with autism showed reduced activation of the inferior frontal gyrus (Broadman area 44/45) and no amygdala activation when the participants were asked to infer the mental state underpinning the eyes region. There were no significant group differences when the participants were asked to infer the gender of the stimuli. On the basis of this result, the authors suggested that the amygdala is a key neural region that is abnormal in autism (also Baron-Cohen, Ring, Bullmore et al., 2000).

The amygdala hypothesis has in fact acquired further support from structural imaging and some support from behavioural evidence. A structural MRI study by Abell et al. (1999) using the method of a voxel-based volumetric analysis, indicated an enlargement of the amygdaloid area. A more recent structural Magnetic Resonance Imaging (MRI) study by Howard, Cowell, Boucher et al. (2000) confirmed this finding showing bilaterally enlarged amygdala volume in ten individuals with high functioning autism (HFA) compared to controls. In addition, the same subjects were presented with the six basic emotion stimuli (standardised stimuli by Ekman and Friesen, 1976) and asked to choose the appropriate label out of six possibilities. Results indicate that, as a group, individuals with autism made more errors in recognising facial expressions of fear than all other five basic emotions. They were also impaired in an eye-gaze direction detection task. The behavioural finding is in line with the cognitive profile of patients with amygdala lesions (Adolphs et al., 1999; Fine and Blair, 2000). Moreover, interestingly, a single case study reported profound difficulty in mentalising ability in a patient with longstanding or congenital lesion to the left amygdala (Fine, Lumsden and Blair, 2000). However, more recently, Adolphs, Sears and Piven (2001) have investigated eight subjects with HFA on the ability to process emotional and social information from faces, and compared their performance to that of amygdala patients. The study used tasks identical to those previously used in studies with bilateral amygdala lesion patients, permitting direct comparison of the two groups. On the simple emotion recognition task, where expressions were rated with respect to the intensity of each of the six basic emotions, the autism group performed better on all emotions than the amygdala patients. Only two subjects with autism showed a pattern similar to the amygdala patients: one subject with impaired performance on fear, disgust and surprise, and the other on disgust only. This result is at variance with the results by Howard et al.'s study on the same task, but with different subjects. More interestingly, all subjects with autism in the Adolphs et al.'s (2001) study showed severe impairment, like the amygdala patients, in a task involving the judgement of trustworthiness and approachability of a person by watching their faces only. It must be noted that the complexity of this task may rely on the fact that it implies mentalising ability. It would appear to require to understanding that others have mental states, desires and intentions that are different from one's own. This impaired social judgement disappeared, as in subjects with amygdala damage, when the information was presented in a more explicit lexical format (personality adjectives and stories). This study indicates therefore that early perceptual processing of affective signals may be intact in individuals with autism, but the normal retrieval of social knowledge triggered by facial cues is impaired. It seems therefore that the amygdala hypothesis of autism does not predict a selective impairment in emotion recognition, but more general difficulties with complex social judgement, e.g., judgement of facial expressions' trustworthiness, judgement of subtle emotions from eye gaze. This conclusion is compatible with the ToM hypothesis.

2.2.3 The Theory of Mind deficit hypothesis

Baron-Cohen, Spitz and Cross (1993) adopted a new methodological approach that assumes that an affective deficit might be secondary to a Theory of Mind deficit - the ability to understand and predict behavior of others on the basis of their mental states such as beliefs and intentions. The paradigm used to test young children with autism and verbal mental age matched controls is based on a specific aspect of emotions: their cause. The clear-cut distinction was between "simple" and "cognitive" emotions. We experience a simple emotion when something happening in the real word causes our emotional reaction and a cognitive emotion when beliefs and desires interact directly with reality. Emotions such as happiness, sadness, anger and fear are typically interpreted as simple emotions because a situational cause would be sufficient to explain the occurrence of the emotional expression. On the other hand, embarrassment and pride are typically explained as cognitive emotions because they are socially and culturally derived and it is necessary to maintain a series of specific beliefs in order to either express or understand them. Surprise is a less complex cognitive expression, for it is caused when we discover that the word is different than expected: our belief does not coincide with reality.

Baron-Cohen et al. (1993) tested children with autism's ability to understand cognitive emotions by comparing the expressions of surprise with happiness and sadness. Results indicated that they were impaired in understanding surprise (belief-based emotion) but do not show difficulty in recognizing happy and sad faces (situation-based emotions). However, the stimuli were not matched for features complexity, since happiness and sadness have been shown to be expressions that are easy to recognize whereas surprise is one of the most difficult (Ekman, Friesen, Ellsworth 1972).

In a later study, Baron-Cohen, Wheelwright and Jolliffe (1997) adopted a wide range of stimuli from a new perspective. The ability of individuals with autism to read others' mental states from their facial expressions has been investigated with a paradigm based on the distinction between the information conveyed by the whole face and the eyes region alone. Since people with autism show difficulties in monitoring gaze direction (e.g., Phillips, Baron-Cohen, Rutter, 1992; Baron-Cohen, Baldwin and Crowson, 1997), the study predicted a specific impairment in autism in decoding the "language of the eyes" as opposed to the whole facial expression. Subjects were presented with a picture depicting either the whole face or the eyes alone of an actress displaying all six basic emotions (happy, sad, surprised, angry, disgusted, distressed, afraid) and complex mental states (admiring, arrogant, bored, flirting, guilt, interested, quizzical, scheming, thoughtful). The task consisted in judging the mental state or basic emotion expressions by selecting one of the two mental states terms presented below each stimulus. Results indicated that adults with autism differed from controls in reading basic emotions from eyes alone, but not from the whole face. Furthermore, the autism group's impairment was more marked when presented with both the eyes and the whole face displaying complex mental states.

It is important to underline that this study with adult subjects, unlike the previous study with children (Baron-Cohen et al., 1993) investigating the recognition of surprise versus happiness and sadness, has compared all six basic emotions versus complex mental states. Therefore, stimuli included expressions that were more difficult to recognise (fear, anger and disgust) than happiness and sadness. Unlike children with autism, the adults with autism did not show any difficulties in recognising surprise. However, subjects were asked to make a forced-choice between "surprise", which is a more difficult emotion, and "happiness", which is an easier emotion, and never between the most confusable emotions of "surprise" and "fear".

A recent study (Buitelaar, Van der Vees, Swabb-Barneveld, Van der Gaag, 1999) adopted the distinction between simple and complex emotion with children with autism, children with pervasive developmental disorder-not otherwise specified (PDD-NOS), and children with psychiatric disorders. The study comprised a matching task and an emotion recognition task in a social context (using pictures displaying emotional situations) using eight facial expressions of emotions (standardised images by Ekman and Friesen, 1976). The stimuli were divided into simple basic emotions (happiness, sadness, anger and fear) and more difficult to recognise emotions (surprise, disgust, shame and contempt). Results indicated no significant group differences on both recognition tasks, with either the simple or the more complex emotions.

It seems that the Theory of Mind deficit hypothesis applied to emotion recognition tasks has adopted different paradigms with different sets of stimuli, and different age groups, yielding some inconsistent data relative to the recognition of basic emotions in individuals with autism. This area is in need of clarification by further experiments.

2.3 Three experiments on emotion recognition in children with autism

The aim of the present study is to investigate specific emotion recognition processes in children with autism by using fine-grained visual stimuli depicting facial expressions of all six basic emotions (anger, disgust, fear, happiness, sadness, surprise) derived from a standard set of pictures of facial effect (Ekman and Friesen, 1976; Calder et al., 1996a; Young et al., 1997). As discussed earlier, some behavioural studies on autism have pinpointed different selective impairments in recognising facial expression of emotions. Children with autism have been shown to be impaired in matching facial expressions displaying surprise as opposed to sadness and happiness (Baron-Cohen et al., 1993). However, adults with high functioning autism showed general impairment in processing simple emotions and complex mental states from the eye regions only, and no impairment in recognizing simple emotions from the whole face (Baron-Cohen et al., 1997). Buitelaar et al. (1999) showed no impairment in children with autism in recognising both simple and complex emotions. On the other hand, Howard et al. (2000) found evidence for selective impairment in recognizing fearful expression associated at the anatomical level to a significant amygdala volume enlargement. However, Adolphs et al. (2001) have shown that individuals with autism performed better than patients with amygdala damage with all simple emotions, including fear.

The present study adopted the fine-cut approach of the Theory of Mind hypothesis, predicting that children with autism have difficulties in recognising only the emotions that are triggered by propositional attitudes and no difficulties with emotions that are triggered by states of affairs. The stimuli consist in the facial expressions displaying the so-called "basic emotions" defined as involuntary stereotypical responses involving physiological changes. The paradigm is based on fine-grained computer-generated stimuli, obtained by transforming the standardized pictures of the Ekman and Friesen series (1976) containing the images of the six basic emotions that have been found to be accurately recognised, and sometimes confused, in most cultures of the world (Young et al., 1997; Calder et al., 1996a).

Thus, it is predicted that children with autism fail to recognise the belief-based emotion of surprise as opposed to the reality-based emotions of anger, disgust, fear, happiness and sadness. Unlike the Baron-Cohen et al. (1993) paradigm that was based on the recognition of the expressions of surprise, happiness and sadness, the present study adopted a broader range of stimuli, including expressions that are more difficult to recognise (fear, anger and disgust) than happiness and sadness.

In addition, the wide range of stimuli allows for monitoring children's performance in relation to the amygdala hypothesis of autism. This hypothesis suggests a correlation between amygdala abnormality and socioaffective impairments. Since the studies reported above (Howard et al., 2000; Adolphs et al., 2001) were contradictory in their findings of face recognition, the additional purpose of the present study is to observe the children's performance in relation to the expression of fear.

The study comprises three different experiments. The first investigates the ability to discriminate facial expressions of all six basic emotions (anger, disgust, fear, happiness, sadness and surprise). It is based on the very simple task of sorting and matching pictures of emotional expressions of a male adult (test stimuli) with the expressions of a female adult (target). The challenge for the child is determined by the difference in intensity level of the emotions displayed in the stimuli expressions. The task requires therefore a more abstract ability to extract the salient invariance of the six expressions across different degrees rather than matching fixed stereotypical features. The higher intensity of the facial expression is expected to facilitate the matching task across all emotions, whereas the expressions combining equal intensity of two emotions are expected to elicit responses regularly distributed between the two possible targets. However, possible preferential biases towards some expressions when equally combined with others might occur, e.g., surprise combined with happiness seen more as happiness, sadness combined with disgust seen more as sadness, anger combined with disgust seen more as disgust, fear combined with surprise seen more as fear. The evidence that normal adults tend to recognise more one of the two equally combined expressions has been reported in a study which has adopted the same stimuli to investigate emotion processing in brain lesioned patients (Calder, Young, Rowland et al., 1996a). It is therefore of interest to investigate whether children with autism show a preferential bias towards only one target emotion within each combination pair.

The aim of the second and third experiments was to investigate children's semantic ability to discriminate emotions from a wide range of individuals' facial expressions. Indeed, the ability to appropriately attribute an emotional state, that is, to "read" an emotion from a facial expression, is distinct from the ability to correctly name an emotion. Children are presented with the target and are asked to say how the person in the picture is feeling. Both the absence of a target and the heterogeneity of the test stimuli control for the possibility that children's performance relies on perceptual matching strategies without a clear understanding of the meaning of the expression. The third study has been added in order to combine the "fine-grained" stimuli of experiment 1 – facial expressions of emotions with different levels - with the naming task of experiment 2. Children are presented with the same stimuli of the matching task and are asked to give a name to the emotion displayed by the person in the picture.

In all three experiments it is expected, on the basis of the "confusion matrix" created by Young et al. (1997) by assembling all the data of Ekman and Friesen (1976), that happiness is the easiest emotion to recognise, that happiness and surprise are almost never confused, whereas the expressions of surprise and fear are the most confused emotions, together with anger and disgust. Interestingly, Young et al.'s (1997) study on categorical perception of facial expressions indicated that the expression of surprise was sometimes identified as fear, and disgust as anger, but when this happened, it was usually because one subject consistently did this. It is therefore of interest to investigate whether the error pattern relative to surprise and fear of children with autism differ from that of normally developing children.

2.3.1 Experiment 1: Discriminating facial expressions of emotions with different intensity levels

Design: The experiment involves a 2 (Group) x 6 (Emotion type) x 3 (Intensity type) design. The Emotion type independent variable consists of interpolated ("morphed") facial expressions derived from prototypes of 6 emotions: anger, disgust, fear, happiness, sadness and surprise. The Intensity type independent variable consists in three intensity levels for each emotion: 90%, 70% and 50%. The two groups consist of children with autism and normally developing children. The experimental task consists in matching each single facial expression with one of the six displayed emotion targets (100% intensity level). The dependent variable is the number of correct matches of emotion stimuli with the emotion target. The facial expressions with 50% intensity level represent the combination of two different emotions, hence two different emotions targets are equally correct for these stimuli.

It is predicted that children with autism find it more difficult to recognise the expression of surprise than normally developing children. Higher intensity of facial expression is expected to facilitate the matching task across all emotions. The expressions combining two emotions at 50% level are expected to elicit responses regularly distributed between the two possible targets, with possible preferential biases in the case of surprise combined with happiness seen more as happiness, sadness combined with disgust seen more as sadness, anger combined with disgust seen more as disgust, fear combined with surprise seen more as fear (Calder et al., 1996a).

Subjects: A group of 20 children with autism resident in a special school, and 20 normally developing children attending mainstream schools were tested. Table 2.1 shows subjects' chronological age (CA) and verbal mental age (VMA). Children with autism were assessed using the VIQ score of the WISC (Wechsler Intelligence Scale for children, third edition UK, 1992), whereas normally developing children were assessed by the BPVS II test (British Picture Vocabulary Scale, 1997) chosen for its brevity. The test was not used as a matching criterion, but rather to check that the control subjects were at a developmentally normal language level. For matching purposes, the group of children with autism was divided into subgroups according to their VMA (6-7 years, 7.1-9 years, 9.1-12 years, 12.1-14 years) and matched with the same number of controls of the same CA in each subgroup, assuming that the chronological age of the nonautistic children was roughly equivalent to their verbal mental age. However, the results were analysed relative to the performance of the whole group.

Table 2.1: Subjects' verbal ability score, chronological age (CA), and verbal mental age (VMA). The age mean and standard deviation are expressed in calendar years and months.

Group	Score *Mean (sd)*	CA (yrs) *Mean (sd)*	VMA (yrs) *Mean (sd)*
AUTISM (N=20)	WISC= **75.2** (16.9)	**12.3** (2.3)	**9.2** (2.6)
CONTROL (N=20)	BPVS= **98** (18.3)	**9.2** (2.4)	**9.11** (2.7)

Materials: The stimuli consisted in laminated cards of cm 6.5 x 9 representing photographic quality images of "morphed" facial expressions developed by Calder et al., (1996a) and derived from prototypes of the six basic emotions in the Ekman and Friesen (1976) series. The facial expressions were ordered by placing each of them adjacent to the one it was most likely to be confused with. The sequence started with happiness, followed by surprise, fear, sadness, disgust and anger. The ends of the sequence - and anger - were then joined to create a hexagon as shown in Figure 2.1.

Figure 2.1: The hexagon stimuli. From top clockwise: happiness, surprise, fear, sadness, disgust and anger.

The expression of one single actor, "JJ" (Ekman and Friesen, 1976) was chosen for morphing the images of the emotional expressions because of their consistent quality across different images. Morphed images were created for the six continua that lie around the perimeter of this hexagon: happiness-surprise, surprise-fear, fear-sadness, sadness-disgust, disgust-anger, anger-happiness. The percentage recognition rates for each prototype image as its corresponding emotion in the Ekman and Friesen (1976) norms were: happiness 100%, surprise 97%, fear 97%, sadness 93%, disgust 88%, and anger 76%.

Computer-manipulated photographic-quality images were created along continua comprising five morphed expression of JJ. The continua were prepared by blending between two prototype expressions (e.g., surprise and fear) in the following proportions: 90:10, 70:30, 50:50, 30:70, 10:90. These proportions provide images of expressions with three different levels of intensity for all emotions (e.g., happiness at 90% intensity morphed with anger at 10%; sadness at 70% morphed with fear at 30%; surprise at 50% morphed with happiness at 50%).

The procedure for the creation of each continuum involved three stages: a) delineation, b) shape interpolation, c) producing a continuous-tone image. Details of the procedure are given below, taking as an example the continuum of surprise-fear.

Delineation: The photograph of JJ's face posing surprise was marked by 186 points positioned manually around the dominant features (e.g., mouth, eyes, eyebrow, nose). Each facial feature was represented by a set number of points (e.g., the mouth was represented by 22 points). These points were then joined to produce a delineated representation comprising 50 points. The same was done to the expression of fear. Hence, across the 2 prototype expression of surprise and fear there was conformity with respect to the anatomical positioning of the 186 points on each face, but not always their exact spatial positions.

Shape interpolation: A continuum of face shapes was generated between the two delineated prototype face shapes (JJ surprised and JJ afraid). This was achieved by taking the delineation data for the two prototype images and calculating the vector difference for each landmark. For example, consider the point describing the tip of the nose: this has a location on the JJ surprised face (x_1, y_1) and a location in the afraid face (x_2, y_2). Equations describing the vector from (x_1, y_1) to (x_2, y_2) were used to obtain positions for the point at the tip of the nose which moved 10%, 30%, 50%, 70%, and 90% along a straight line from the location of that point in JJ's surprised face to the location of that point in JJ's fearful face. This process was repeated for each of the 186 features points, to generate the 5 face shapes which would interpolate at 10%, 30%, 50%, 70%, and 90% distances between the two prototype facial expressions.

Producing a continuous-tone image: The final stage allowed for creating the photographic quality (continuous-tone) of each morphed image. This was achieved by taking the original images of JJ surprised and afraid, and adjusting them to the new shape as they were printed on a rubber sheet. In this way all points representing the same feature were aligned across images. The two faces, now with the same intermediary face shape, were then blended with the appropriate proportion. For example, in the morph comprising 90% of surprise: 10% fear, the pixel intensities in each tessellation were arrived at deforming the afraid face at 10% towards the surprise prototype, and in the surprised face towards the afraid face, and then blending the gray levels in these two contributory images in the ratio 9 parts from the happy prototype to 1 part from the surprised prototype.

Subjects were presented with a total of six different continua: anger-happiness, anger-disgust, sadness-disgust, surprise-fear, surprise-happiness. The target cards represent expressions at 100% intensity level (original prototypes) of six different models -adult females- (Ekman and Friesen, 1976). There are four stimuli cards in each emotion at level 90% and 70%, and two cards each pair of emotion at 50% of intensity. For each subject there are two sets of 30 cards each. The target cards were pasted on plastic boxes sized cm.10 x 15 x 3. Examples of the stimuli are shown in Appendix 2A.

Procedure: Each child was tested individually in a separate room of the school. In the training phase the experimenter showed one target picture for each expressions to the child, and asked him/her to name each expression (e.g.: "This a picture of a woman: how does she feel?") and to provide an example of the displayed emotion (e.g.:"Tell me about a time when you were surprised"). If the child showed uncertainty, giving no example, the experimenter provided a standard example of a time in which she felt the emotion (e.g.: "I was surprised when I opened my birthday present", "I was sad when my dog died", "I was angry when somebody hit me on purpose", "I was disgusted when I drank milk gone bad", "I was afraid when I was left alone in a dark room", "I am happy when I am on holiday"). If the child provided a different but correct label for the emotion target (e.g.: "smiling" instead of "happy", "crying" or "upset" for "sad", "sick" instead of disgusted) the experimenter accepted them as substitutes. Each target emotion was then fixed on an empty box in front the child, making sure that she/he has a full view of all targets. The display of the six emotions was randomly arranged across subjects.

After the initial instruction phase, children went through a training session with 12 cards representing expressions of the six emotions at 90% intensity level. The experimenter showed one at a time the expressions of happiness or sadness asking the child: "How does she feel?", and then: "Put this card in the box with the happy/sad face". The practice was followed by two trial sessions with the same task (for a total of 60 cards). In each session the child was given 30 cards of facial expressions randomly arranged to be matched against the targets. The cards were given one at the time, and the experimenter kept asking "How does she feel?" and "Where does it go?" until the routine was established. At the end of the first session, the display of the targets on the boxes was randomly rearranged in order to control for biases due to a preference of a particular position (e.g., central positions versus lateral). After the two sessions, the cards in each box were counted and coded.

Results: The analysis of the score with emotions at 90% and 70% intensity level was carried out separately from the score with 50% emotions (ambiguous stimuli). An additional analysis was carried out on the overall performance regardless of intensity level, allowing for exploring consistent confusions across emotions. The error patterns of each group have also been analysed to identify consistent mismatches between the expressions of surprise and fear. Note that the stimulus cards that that children had to sort out are defined henceforth as "Emotion-Tests" to be matched against the cards defined as "Emotion-Targets" placed on top of the boxes (each Emotion-Test at 90% and 70% intensity level has one correct Emotion-Target to be matched with, whereas the Emotion-Test at 50% has two correct Emotion-Targets).

i) Analysis of correct performance with emotion stimuli at 90% and 70% intensity levels:

A parametric analysis was performed on the correct matching of each Emotion-Test with its Emotion-Target, with emotions split into higher (90%) and lower (70%) levels of intensity allowing for investigating groups' sensitivity to different degree of emotional display. Table 2.2 shows the correct score relative to each emotion. A repeated measures ANOVA 2 (Group) x 6 (Emotion) x 2 (Intensity) revealed a significant main effect of Emotion ($F_{(5,38)}=10.8$, $p<.001$). Post-hoc analyses revealed that, as predicted, the correct score for Happiness was the highest (planned comparison, $F=14.5$, $p=.0002$) and the scores for Fear and Surprise the lowest (planned comparison, $F=21$, $p=.0001$). The interaction of Group x Intensity ($F_{(5,38)}=9.7$, $p<.01$) was significant, and a post hoc analysis revealed that the autism group performed better in discriminating the stimuli at 90% intensity level than at 70% (paired t-test, $t=3.1$, $p<.01$). The interaction Group x Emotion x Intensity ($F_{(5,38)}=2.5$, $p<.05$) was also significant, and a post-hoc analysis revealed that children with autism found the expression of Sadness at 90% intentity level easier to recognise than Sadness at 70% intentity level (paired t-test, $t=2.3$, $p<.05$). The results failed to support the prediction of a specific impairment in children with autism in recognising the emotion of Surprise from facial expressions. However, children with autism were shown to be more aided by increased intensity of facial expressions than normally developing children.

Table 2.2: Groups correct performance in discriminating emotions at 90% and 70% intensity levels. Mean and standard deviation (max score = 4).

Emotion	Anger		Disgust		Fear		Happiness		Sadness		Surprise		TOTAL	
Intensity level	90%	70%	90%	70%	90%	70%	90%	70%	90%	70%	90%	70%	90%	70%
Autism														
mean	**3.2**	**3.1**	**3.3**	**3.2**	**3.0**	**2.6**	**3.7**	**3.8**	**3.3**	**2.9**	**2.8**	**2.4**	**3.2**	**3.0**
sd	*1.2*	*1.1*	*1.1*	*1.2*	*1.1*	*1.3*	*0.9*	*0.5*	*1.3*	*1.5*	*1.3*	*1.7*	*0.8*	*0.8*
Control														
mean	**2.9**	**3.1**	**3.5**	**3.3**	**2.5**	**2.9**	**4.0**	**3.8**	**3.5**	**3.5**	**2.4**	**2.7**	**3.1**	**3.2**
sd	*1.2*	*0.9*	*0.9*	*1.1*	*1.2*	*1.3*	*0.2*	*0.4*	*0.9*	*0.8*	*1.4*	*1.2*	*1.7*	*0.6*

ii) Analysis of correct performance with ambiguous stimuli (emotion at 50% intensity level):

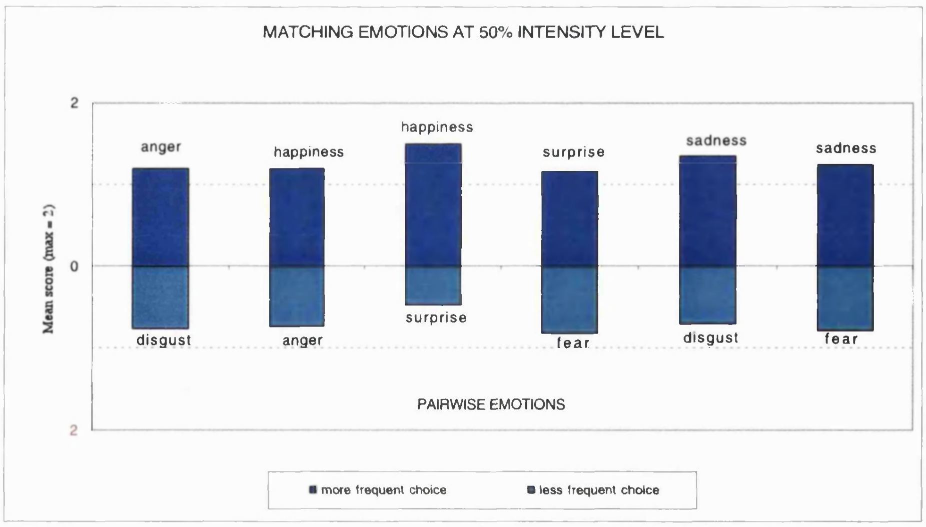
Table 2.3 shows the score for the groups' correct matches of the Emotion-Test at 50% intensity level. Since each observation is dependent on each other, this score did not meet the parametric analysis requirement of non-sphericity. Two nonparametric group-comparison analyses (Mann-Whitney test) were carried out separately on the groups' performance to investigate the possibility of preferential matching towards only one Emotion-Target in children with autism. The results indicated no group differences in matching facial expressions that blend together two different emotions at equal intensity.

Table 2.3: Groups correct performance in discriminating ambiguous expressions (emotions at 50% intensity level). Mean and standard deviation (max score = 2).

Combination of two emotions Intensity level 50%	Anger Disgust	Anger Happiness	Surprise Happiness	Surprise Fear	Sadness Disgust	Sadness Fear
Autism						
mean	**1.8**	**1.6**	**1.7**	**1.6**	**2**	**1.7**
sd	*0.4*	*0.5*	*0.5*	*0.7*	*0.2*	*0.7*
Control						
mean	**1.6**	**1.6**	**1.9**	**1.9**	**2**	**1.7**
sd	*0.7*	*0.6*	*0.4*	*0.3*	*0.2*	*0.6*

An additional analysis was carried out on children's overall performance with the ambiguous stimuli with the aim of investigating the possibility of consistent preferential biases in matching emotions at 50% intensity level (Figure 2.2 shows groups correct performance). A nonparametric analysis of variance (Friedman test) on six pairs of correct matches between Emotion-Target and Emotion-Test revealed significant differences across all responses (chi-square=27.7, p=.004). Nonparametric paired-comparisons (Wilcoxon test) on correct matches of each Emotion-Test with the two possible Emotion-Targets revealed that only two facial expressions elicited a significant preferential bias. The facial expression combining Sadness and Disgust was matched more often with Sadness (z=2.4, p=.02) and the facial expression combining Happiness and Surprise was matched more often with Happiness (z=3.2, p=.001).

Figure 2.2: Correct performance with emotion at 50% intensity level: Subject's score (max = 2) for each pair of facial expressions combining two equally intense emotions (50% intensity level). Children matched significantly more often the expression combining Happiness and Surprise with the Emotion-Target of Happiness, whereas the facial expression combining Sadness and Disgust was matched more often with Sadness. No significant group differences were found.



iii) Analysis of groups' error pattern in matching emotions regardless of intensity level:

Since possible confusions between emotions are also likely to occur with morphed expressions matching emotions at 90%:10% and 70%:30%, a further analysis was carried out on performance relative to the incorrect matching between Emotion-Test and Emotion-Target regardless of intensity level. Table 2.4 shows the matrix of all types of matching for all targets. Note that the types of errors are defined henceforth as "*name* of Emotion-Test in *name* of Emotion Target" (e.g.: "Fear in Surprise" means that the card representing the expression of Fear has been incorrectly matched with Surprise).

Table 2.4: Matrix showing all possible matches, correct and incorrect, between emotion stimuli (regardless of intensity levels) and their targets. Mean and standard deviation (in parentheses) of total score (max = 8). Correct scores shown with the symbol ✓. Incorrect scores shown with ✗. No significant group differences were found.

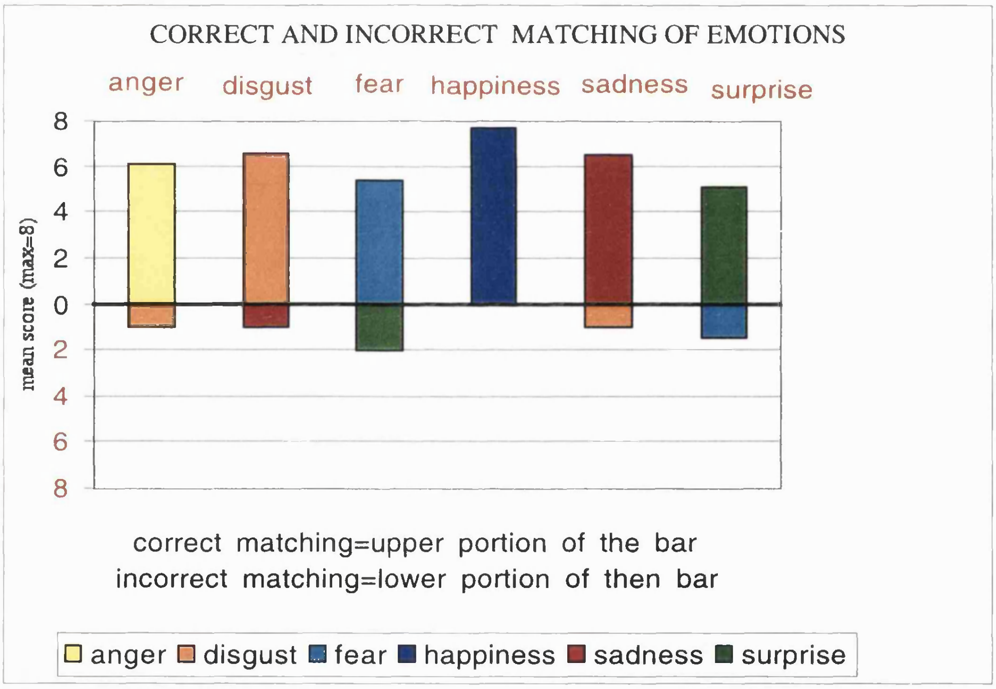
Emotion Test	**Emotion Target**					
	Anger	**Disgust**	**Fear**	**Happiness**	**Sadness**	**Surprise**
Anger	6.2 (2) ✓	0.7 (1.4)	0.4 (0.9)	0.1 (0.4)	0.2 (0.5)	0.5 (1)
Disgust	0.5 (1.2) ✗	6.6 (2) ✓	0.0	0.03 (0.2)	0.8 (1.2) ✗	0.03 (0.2)
Fear	0.3 (0.6)	0.2 (.8)	5.5 (2.2) ✓	0.1 (.5)	0.4 (1.2)	1.5 (1.6) ✗
Happiness	0.1 (0.5)	0.0	0.1 (0.4)	7.7 (1) ✓	0.1 (0.3)	0.0
Sadness	0.0	1.2 (2) ✗	0.2 (0.4)	0.0	6.5 (2.2) ✓	0.1 (0.4)
Surprise	0.2 (0.6)	0.1 (0.5)	2.2 (2.4) ✗	0.3 (0.6)	0.1 (1)	5.1 (2.6) ✓

Note:
The most frequent error with the stimuli representing Fear was to match it against Surprise:
Fear in Surprise > Fear in Sadness, z=2.7, p=.006
Fear in Surprise > Fear in Happiness, z=3.9, p=.0001
Fear in Surprise > Fear in Disgust, z=3.6, p=.0003
Fear in Surprise > Fear in Anger, z=3.5, p=.0005
The most frequent error with the stimuli representing Surprise was to match it Fear:
Surprise in Fear > Surprise in Anger, z=3.9, p<.0001
Surprise in Fear > Surprise in Disgust, z=4.0, p<.0001
Surprise in Fear > Surprise in Happiness, z=4.2, p<.0001
Surprise in Fear > Surprise in Sadness, z=3.7, p=.0002
A direct comparison of the two scores indicating confusion between Surprise and Fear indicated significant more mistakes of Surprise matched against the Target representing Fear (z=2, p=.05).

In addition to the type of error involving the most commonly mistaken expression, the analysis revealed other more common confusions involving Disgust, Sadness and Anger. In particular: The most frequent error with the stimuli representing Disgust was to match it either against Sadness and Anger:
Disgust in Anger = Disgust in Sadness, z=1.3, p= not significant
Disgust in Sadness > Disgust in Surprise, z=3.2, p=.0015
Disgust in Sadness > Disgust in Happiness, z=3.2, p=.001
Disgust in Sadness > Disgust in Fear, z=3.3, p=.001
The most frequent error with the stimuli representing Sadness was to match it against Disgust:
Sadness in Disgust > Sadness in Anger, z=3.5, p=.0004
Sadness in Disgust > Sadness in Fear, z=3, p=.003
Sadness in Disgust > Sadness in Happiness, z=3.5, p=.0004
Sadness in Disgust > Sadness in Surprise, z=3.2, p=.001

Nonparametric group-comparison tests (Mann-Whitney) were carried out separately on each type of error made with each Emotion-Test. The results indicated no differential error pattern in the two groups. A series of nonparametric paired comparisons (Wilcoxon tests) were carried out in order to investigate the type of errors occurring most frequently in both groups. Figure 2.3 shows the correct responses in each Emotion target along with the most frequent type of errors.

Figure 2.3: Errors pattern with each Emotion test regardless of intensity level in both groups. Each bar represents the content of each box (Target emotion) where children dropped the stimuli cards (test emotion). The most frequent errors in both groups occurred when the Fear was matched against Surprise followed by the opposite mistakes of matching Surprise with Fear and matching Disgust with either Sadness or Anger and matching Sadness with Disgust. No group differences in the correct or incorrect matching were found.



Conclusions: Children of both groups found no particular difficulty in performing the task of discriminating emotions by matching morphed facial expressions with pictures of actresses displaying six basic emotions. The group of children with autism performed slightly better at sorting the expressions displaying a

90% intensity level, in particular the expression of Sadness than the same emotions at a 70% intensity level. As predicted, all children found Happiness very easy to recognise amongst all other basic emotions and found both the expressions of Surprise and of Fear the most difficult to discriminate. The most frequent error was matching the expression of Surprise with the target of Fear, followed by matching Fear with Surprise, Disgust with both Anger and Sadness and Sadness with Disgust. The ambiguous stimuli representing expressions combining two emotions with 50% of intensity level were equally distributed in all children with only two exceptions. The facial expression blending together Sadness and Disgust was recognised more as a sad expression and the expression displaying both Surprise and Happiness was recognised more as a happy expression. There was no group difference in recognising or confusing any of the six basic emotions. These results failed to support the Theory of Mind hypothesis, which predicted a selective deficit in children with autism in recognising Surprise.

2.3.2 Experiment 2: Naming facial expression of emotions with natural intensity

Design: The experiment involves a 2 (Group) x 7 (Expression type) design. The Expression type independent variable consists of six facial expressions of emotions: anger, disgust, fear, happiness, sadness, surprise, and one neutral expression. The groups are the same as in study 1. The experimental task consists in naming each single facial expression with one of the six emotions, or the neutral expression. The dependent variable is the number of correct responses. The aim of this second experiment is to investigate children with autism semantic ability to discriminate emotions from a wide range of individuals' facial expressions. It is predicted that children with autism have more difficulties in naming the expression of surprise than normally developing children. A certain degree of confusion is expected between fear and surprise.

Subjects: The same groups of participants were tested as in Experiment 1.

Material: The stimuli consisted in laminated cards of cm 9 x 11.5 representing photographic quality images of facial expressions derived from Benson and Perrett series (1991). The pictures represented the facial expressions of 10 different models (4 adult males and 6 adult females) each displaying the emotions with natural intensity. The cards are arranged randomly into two decks of 35 cards, each composed with pictures of 7 expressions displayed by 5 models (3 females and 2 males). The two decks of cards were given to subjects one after the other in a counterbalanced order across subjects. All stimuli are shown in Appendix 2B.

Procedure: Experiment 2 was carried out always after experiment 1. The child was told that she/he was going to work again with pictures of different emotions, but this time she/he had to say how the people in the picture were feeling. She/he was also informed that there was an additional expression that showed no particular emotion. The experimenter showed first a card representing a neutral expression and then a full set of cards depicting an adult female expressing the six basic emotions (100% intensity level) and a neutral expression and asked the usual question: "This a picture of a woman, how is she feeling?". The name/s given to each expression that was correct but different across children (e.g.: "sick" instead of "disgusted", "annoyed" or "cross" instead of "angry") were recorded for each subject. Definitions pertaining to non-basic emotional states were classified as unclear (e.g.: "grumpy", "disappointed", "bored" "confused"). The pre-naming session was followed by the session during which the child was presented with one card after the other and asked to say how the person in the picture was feeling until routine was established. Each answer was recorded and coded as "correct", "incorrect" or "unclear".

Results: An initial analysis of the correct naming score was carried out, followed by an additional analysis of the erroneous labelling performance, with the aim of exploring consistent errors across emotions. Table 2.5 shows the correct performance for the two groups. As in Exp.1, the stimulus cards that subjects have to name are defined as "Expression Test", whereas the labels they provide are defined as "Expression Label".

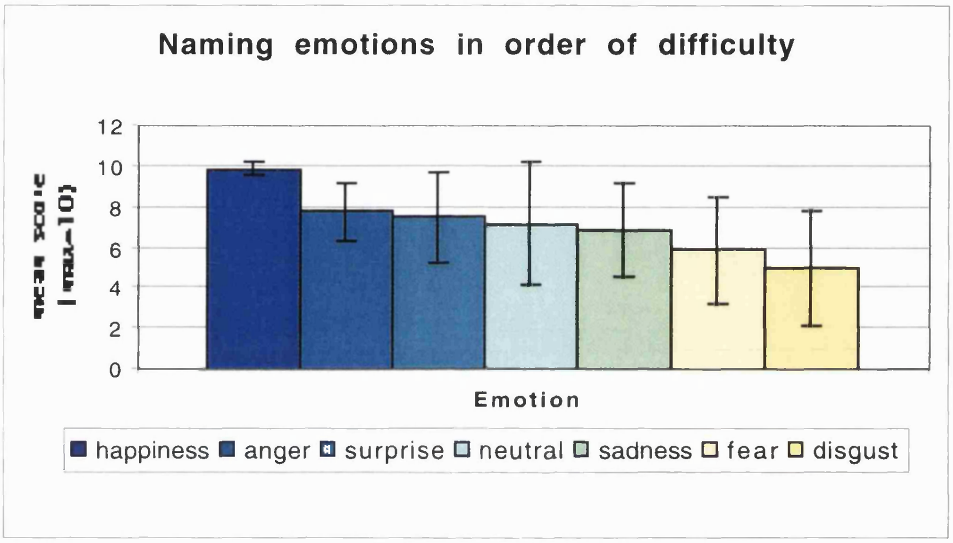
Table 2.5: Groups' correct score for naming neutral and emotional expressions (max = 10). Mean and standard deviation.

	Target Label							
	Neutral	**Anger**	**Disgust**	**Fear**	**Happiness**	**Sadness**	**Surprise**	**Total**
Autism								
mean	**7.8**	**8.0**	**5.7**	**5.9**	**9.9**	**7.7**	**7.6**	**7.5**
sd	*2.4*	*1.5*	*3.3*	*2.4*	*0.4*	*1.9*	*2.6*	*2.6*
Control								
mean	**6.7**	**7.7**	**4.4**	**6.0**	**9.9**	**6.2**	**7.4**	**6.9**
sd	*3.6*	*1.3*	*2.1*	*2.8*	*0.3*	*2.4*	*2.0*	*2.8*

i) Analysis of correct labeling of emotions:

A repeated measures ANOVA 2 (Group) x 7 (Expression) showed a significant main effect of correct expression ($F_{(6,38)}$= 21.7, p=.0001). No group main effect or interaction was revealed. Post-hoc analysis showed that in line with the prediction for the discriminating task, all children found the expression of Happiness the easiest emotion to name (planned comparison, F=72, p=.0001). Figure 2.4 shows the correct score for each emotion in order of difficulty: contrary to the previous experiment showing both Surprise and Fear as the most difficult emotions to recognise, children named Surprise correctly more often than Fear (paired t-test, mean diff.=1.6, t =3.4, p=.001) and named correctly equally often Fear and Disgust (paired t-test, mean diff.=0.9, t=1.9 p=.07, not significant). They also named correctly equally often Surprise and Anger (paired t-test, mean diff.=0.4, t=.95 p=0.3, not significant), Neutral and Sadness (paired t-test, mean diff.=0.3, t=0.5 p=0.6, not significant), Neutral and Anger (paired t-test, mean diff.=0.6, t=2.2 p=0.2, not significant). They found it easier to name Anger than Sadness (paired t-test, mean diff.=0.9, t=2.2 p=.035). Since no difference was found between groups, this experiment did not indicate any particular verbal difficulty of children with autism with labeling the emotional expression of surprise. A further analysis was performed to investigate the possibility of consistent error patterns across emotions.

Figure 2.4: Correct score for each emotion in order of difficulty: Happiness was the easiest emotion to name, followed by Surprise and Anger, Neutral and Sadness. Fear and Disgust were equally the most difficult expression to label.



ii) Analysis of errors pattern with labels of emotions:

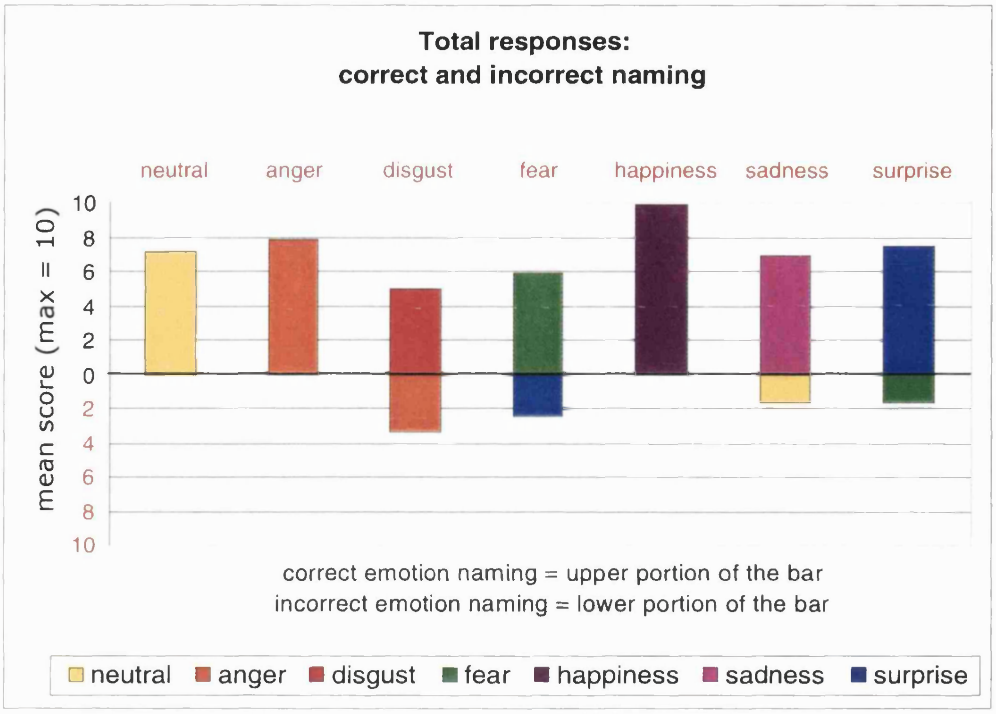
Table 2.6 shows the score for the groups' correct naming along with the erroneous labels that children provided for each expression. Since each score was not independent from each other, the data did not meet the parametric analysis requirement of non-sphericity. Nonparametric group-comparison tests (Mann-Whitney) were carried out separately on each type of error made with each Expression-Test in order to investigate any possible difference in the errors pattern of children with autism and controls. In line with the results of the previous experiment, there were no differences between the two groups. An additional analysis was carried out to explore the type of errors occurring most frequently in both groups. A series of nonparametric paired comparisons (Wilcoxon tests) were carried out in order to investigate the type of errors occurring most frequently in both groups. Figure 2.5 shows the correct responses in each Expression Label along with the significantly most frequent type of errors).

Table 2.6: Matrix showing both the correct and incorrect Expression Label for all expressions. Mean and standard deviation (in parentheses), max score=10. Correct score shown with the symbol ✓; most frequent errors shown with ✖. No significant group differences were found.

Emotion Test	**Target Label**							
	Neutral	**Angry**	**Disgusted**	**Fearful**	**Happy**	**Sad**	**Surprised**	**Unclear**
Neutral	**7.2** (3) ✓	0.6 (0.8)	0.3 (1.1)	0.2 (0.5)	0.8 (1.5)	0.4 (0.8)	0.1 (0.4)	0.5 (1.2)
Anger	0.6 (.8)	**7.8** (1.4) ✓	0.3 (0.6)	0.2 (0.5)	0 (0.2)	0.5 (0.6)	0.2 (0.5)	0.4 (0.8)
Disgust	0.4 (.5)	3.4(2) ✖	5 (2.8) ✓	0.1 (0.6)	0.1 (0.4)	0.2 (0.2)	0.1 (0.2)	0.8 (1.6)
Fear	0.1 (0.4)	0.5 (1.3)	0.1 (0.3)	**5.9** (2.6) ✓	0.4 (0.7)	(0.7)	2.5 (1.9) ✖	0.1 (0.3)
Happiness	0.1 (0.3)	0	0	0	**9.9** (.3) ✓	0	0	0.3 (0.2)
Sadness	1.6 (1.8) ✖	0.3 (0.7)	0.3 (0.5)	0.3 (0.6)	0 (0.6)	**6.9** (2.3) ✓	0.1 (0.4)	0.5 (1.2)
Surprise	0.2 (0.5)	0	0.1 (0.3)	(1.7) ✖	0.6 (1.5)	0.1 (0.3)	**7.5** (2.2) ✓	0.1 (0.2)

Note:
The most common mistake occurring with the expression of Fear was to name it "surprised":
Fear as "surprised" > Fear as "neutral", z=5, p<.0001
Fear as "surprised"> Fear as "angry", z=3.8, p=.0002
Fear as "surprised" > Fear as "disgusted", z=5, p<.0001
Fear as "surprised"> Fear as "happy", z=4.3, p<.0001
Fear as "surprised"> Fear as "sad", z=4.7, p<.0001
Fear as "surprised" > Fear as "unclear", z=5, p<.0001
The most frequent error with the stimuli representing Surprise was to say it was a "fearful" face:
Surprise as "fearful" > Surprise as "neutral", z=4, p<.0001
Surprise as "fearful" > Surprise as "angry", z=4.6, p<.0001
Surprise as "fearful">Surprise as "disgusted", z=4.3, p<.0001
Surprise as "fearful" > Surprise as "happy", z=2.5, p=.01
Surprise as "fearful" > Surprise as "sad", z=4.3, p<.0001
Surprise as "fearful"> Surprise as "unclear", z=4.4, p=.0001
The most frequent error with the stimuli representing Disgust was to say it was an "angry" face:
Disgust as "angry" > Disgust as "neutral", z=5.2, p<.0001
Disgust as "angry"> Disgust as "fearful", z=5.2, p<.0001
Disgust as "angry" > Disgust as "happy" z=5.2, p<.0001
Disgust as "angry" > Disgust as "sad", z=5.2, p<.0001
Disgust as "angry" > Disgust as "surprised", z=5, p<.0001
Disgust as "angry" > Disgust as "unclear", z=5.2, p<.0001
The most frequent error with the stimuli representing Sadness was to say it was a "neutral" face:
Sadness as "neutral" > Sadness as "angry", z=3.5, p=.0005
Sadness as "neutral" > Sadness as "disgusted", z=4, p<.0001
Sadness as "neutral" > Sadness as "fearful", z=3.8, p<.0002
Sadness as "neutral" > Sadness as "happy" z=4.4, p<.0001
Sadness as "neutral" > Sadness as "surprised" z=4.1,p<.0001
Sadness as "neutral" > Sadness as "unclear" z=2.9, p<.004
The most common mistake of all occurred in the cases of naming incorrectly both Disgust and Fear:
Disgust as "angry" = Fear as "surprised", z=1.7, p=.09 ,not significant
Fear as "surprised" > Sadness as "neutral" z=2.4, p=.015
Disgust as "angry" > Sadness as "neutral" z=3.5, p=.004

Figure 2.5: Errors pattern in naming facial expression of emotions in both groups. Each bar represents children's responses for each target emotion. The most frequent mistakes were to say that the expression of Disgust was an angry face and that the expression of Fear was a surprised face. Other errors consisted in saying that a Sad expression was neutral and a surpised expression was fearful. No group differences in naming the emotions were found.



Conclusions: Children with autism performed as well as the controls in this second task based on the semantic ability to discriminate basic emotions. All children found the expression of happiness the easiest to name, and the expressions of fear and disgust equally the most difficult to label. The score between these two extremes is distributed with small differences. In particular, the highest score of Happiness was followed by the score for Surprise and Anger, then by the score of Neutral and Sadness. The difference between the scores of Anger and Neutral was too small to be significant, but the score for Sadness was lower than for Anger. The most frequent mistakes made by all children was to say "angry instead of "disgusted and by the mistake of naming a fearful face as "surprised". Less frequent mistakes consisted in labeling a sad face as "neutral" and a surprised face as "afraid". Contrary to the previous discriminating task, all subjects performed better in naming the expression of Surprise than the expressions of Fear. Since it is possible that these results be due to a lack of test sensitivity, in the next experiment the difficulty of the stimuli was increased.

2.3.3 Experiment 3: Naming facial expressions of emotions with different intensity levels

The aim of this third part of the study was to compare the ability of children with autism and controls to name emotions using more difficult stimuli than in the previous experiment. It was predicted that the expression of surprise is the most difficult emotion to name for children with autism.

Design: The experiment involves a 2(Group) x 6(Emotion type) x 2(Intensity type) design. The Emotion type independent variable consists of the same six basic emotions of the discriminating task of experiment 1. The Intensity type consists in only two emotion levels of intensity (90% and 70%). The experimental task consists in naming each single facial expression with one of the six emotions. The dependent variable is the number of correct labels given to each facial expression of emotion.

Subjects: Participants of this study were the same as in experiment 1 and 2, except that the number of children with autism was reduced to 19.

Material: The stimulus cards were the same as that of the discriminating task (Exp.1), with the exclusion of the facial expressions that combined emotions at 50% intensity. The cards were assembled in 2 blocks of 24 cards (2 stimuli each level for 6 emotions). The presentation order of the two blocks was counterbalanced across subjects.

Procedure: children performed this experiment immediately after the naming task of experiment 2 and therefore there was no need to repeat the pre-labeling procedure. Children were invited to take a pause between the two experiments and then were simply asked to do the same as in the previous naming task. Each answer was recorded and coded as "correct", "incorrect" or "unclear".

Results: As in the two previous experiments, the data were analysed relative to the groups' correct and incorrect performance. Definitions of "Emotion Test" and "Emotion Label" were maintained.

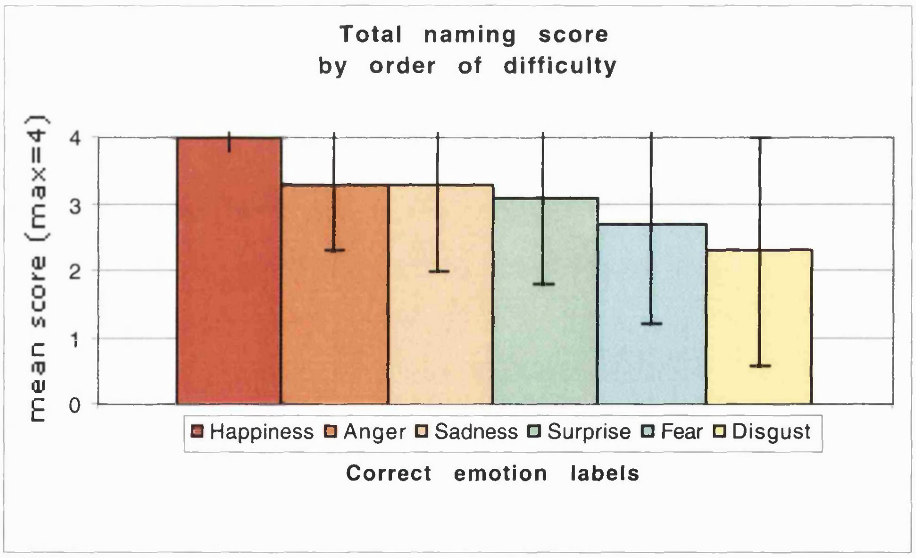
i) Analysis of correct naming emotions with different levels of intensity:

Table 2.7 shows the correct naming for each emotion and level of intensity for both groups. A repeated measure ANOVA 2 (Group) x 6 (Emotion) x 2 (Intensity) revealed a significant main effect of Correct Emotion ($F_{(5, 38)}= 9.5$, $p<.0001$) and of Intensity ($F_{(1, 38)}= 8.6$, $p< .006$) with the score for the emotions at 90% higher than emotion at 70%. Figure 2.6 shows the correct Label Emotion for each Test Emotion in order of difficulty, regardless of intensity level. As predicted, all children performed at ceiling in naming the expression of Happiness (Happiness compared to all other emotions, planned comparison, F=25.5, p< .0001). The scores for Surprise and Fear did not differ significantly (paired t-test, mean diff.=0.8, t=1.6, p=.1, not significant) as well as Fear and Digust (paired t-test, mean diff.=0.7, t=1.1, p=.3, not significant). The equal scores for Anger and Sadness were both higher than Fear (paired t-test, mean diff.=1.2, t=2.7, p=.01) and Disgust (paired t-test, mean diff.=1.9 t=3.6, p=.0009). Finally, the mean difference between the scores of Surprise and Disgust made the latter the most difficult expression to name (paired t-test, mean diff.=1.5, t=3.6, p=.0009). Again, as in the previous task, children with autism did not performed differently from the controls, indicating that the expression of surprise, and furthermore no emotion in particular, constitutes a verbal obstacle at any level of intensity for both groups. To further investigate children with autism's performance, a new analysis was carried out on the erroneous responses of both groups.

Table 2.7: Groups' correct score for naming emotions with different levels of intensity. Mean and standard deviation (max = 4).

Emotion	Anger		Disgust		Fear		Happiness		Sadness		Surprise		TOTAL	
Intensity level	90%	70%	90%	70%	90%	70%	90%	70%	90%	70%	90%	70%	90%	70%
AUTISM														
mean	**3.5**	**3.3**	**2.3**	**2.1**	**2.8**	**2.7**	**4**	**4**	**3.2**	**3**	**3.2**	**3**	**3.1**	**3.0**
sd	*0.9*	*0.9*	*1.8*	*1.8*	*1.6*	*1.5*	*0*	*.2*	*1.6*	*1.5*	*1.2*	*1.1*	*1.4*	*1.4*
CONTROL														
mean	**3.4**	**3.2**	**2.6**	**2.5**	**2.8**	**2.5**	**4**	**4**	**3.5**	**3.6**	**3.3**	**2.9**	**3.3**	**3.1**
sd	*0.9*	*1*	*1.8*	*1.6*	*1.5*	*1.5*	*0*	*0.2*	*1.1*	*1*	*1.5*	*1.4*	*1.3*	*1.3*

Figure 2.6: Correct score for each emotion in order of difficulty regardless of intensity level.



ii) Analysis of errors pattern in naming emotions with different levels of intensity:

Nonparametric group-comparison tests (Mann-Whitney) were carried out separately on each type of labeling errors with each Emotion-Test. The results indicated no differential errors pattern in the two groups. Additional series of nonparametric analyses (paired comparison, Wilcoxon tests) were carried out in order to investigate the type of errors occurring most frequently in both groups. Table 2.8 shows the matrix of Test-Emotion by Label-Emotion.

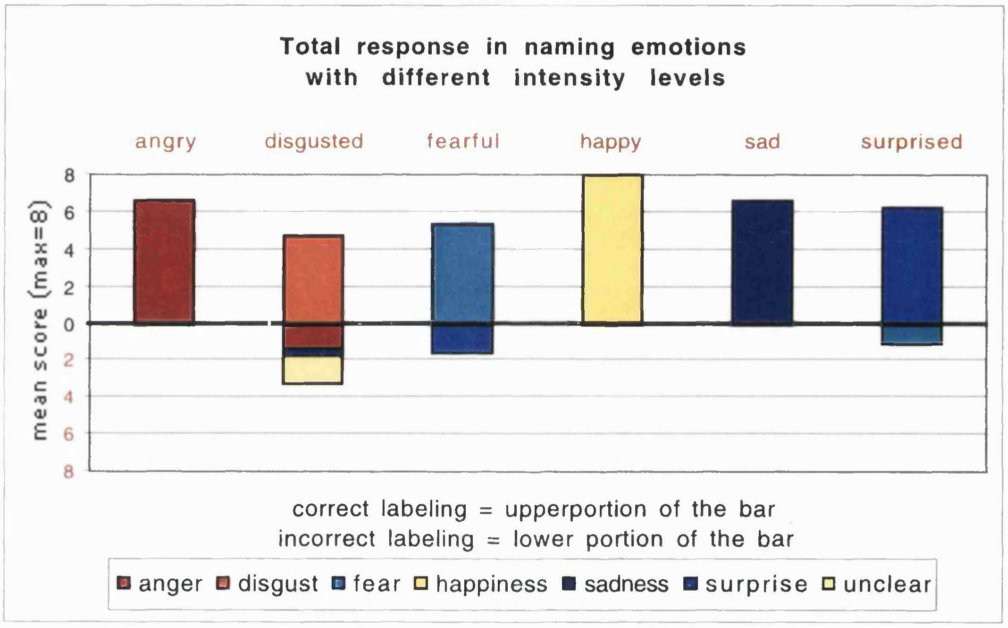
Table 2.8: Matrix showing both the incorrect and correct emotion labels (max score=8) regardless the emotions intensity levels. Correct score shown with symbol ✓, most frequent mistakes shown with ✖. Mean and (standard deviation).

Emotion Test	**Emotion Label**						
	Angry	**Disgusted**	**Fearful**	**Happy**	**Sad**	**Surprised**	**Unclear**
Anger	**6.6** (1.7) ✓	0.1 (0.3)	0.3 (0.6)	0 (.2)	0.1 (0.4)	0.4 (1.4)	0.3 (0.9)
Disgust	1.3(2.6) ✖	**4.7** (3.4) ✓	0	0 (0.1)	0.4 (0.8) ✖	0	1.5 (2.8) ✖
Fear	0.2 (0.5)	0.1 (0.7)	**5.4** (2.8) ✓	0.2 (0.8)	0.4 (0.9)	1.6 (2.3) ✖	0.2 (0.6)
Happiness	0	0	0	**8** (0.2) ✓	0	0	0.1 (0.2)
Sadness	0.1 (0.5)	0.5 (1.5)	0.1 (0.4)	0	**6.6** (2.6) ✓	0 (0.2)	0.7 (2)
Surprise	0.7 (0.3)	0.2 (1.1)	1.1(2) ✖	0.2 (0.5)	0	**6.2** (2.4) ✓	.02 (0.8)

Note:
The most common mistake occurring with the expression of Fear was to name it "surprised":
Fear as "surprised"> Fear as "angry", z=3.3, p=.0009
Fear as "surprised" > Fear as "disgusted", z=3.4, p<.0007
Fear as "surprised"> Fear as "happy", z=3.3, p<.0009
Fear as "surprised"> Fear as "sad", z=2.8, p<.005
Fear as "surprised" > Fear as "unclear", z=3.6, p<.0004
The most frequent error with the stimuli representing Surprise was to say it was a "fearful" face:
Surprise as "fearful" > Surprise as "angry", z=3.2, p=.001
Surprise as "fearful" > Surprise as "disgusted", z=3.1, p=.001
Surprise as "fearful" > Surprise as "happy", z=2.2, p<.03
Surprise as "fearful" > Surprise as "sad", z=3.7, p=.0002
Surprise as "fearful"> Surprise as "unclear", z=2.7, p=.007
Three errors were equally frequent with the stimuli Representing Disgust: to name it with unclear labels, to say it was a "sad" or a"surpised" face:
Disgust as "angry"= Disgust as "unclear", z=.06, p= not sig.
Disgust as "sad" = Disgust as "angry" z=1.6, p= not sig.
Disgust as "sad" = Disgust as "unclear", z=1.5, p= not sig

The analyses revealed that Disgust was the most commonly mistaken expression, named either unclearly, or "sad" or "angry". As in the previous tasks, children also made significantly more errors of saying that Fear was a surprised face and vice versa, Surprise to a fearful face. Figure 2.7 shows the most frequent type of errors in labeling each expression of emotion.

Figure 2.7: Errors pattern in naming facial expression of emotions regardless of intensity in both groups. Each bar represents children's responses foremotion label target. One of most frequent mistakes consisted in saying that the disgusted face was angry or sad or an unclear label was given. Onother error was to say that a fearful face was surprised and viceversa. No group differences were found.



Conclusions: Children of both groups performed equally well in naming facial expressions showing emotions at 90% or 70% level of intensity. This result indicated that children with autism do not have either a selective or general impairment for labeling any particular emotion. Furthermore, their performance was consistent with the non-verbal discrimination task with the same type of morphed stimuli. Again, the score was slightly improved when the emotions were displayed at a higher intensity level for both groups. Children with autism made the same type of errors than the controls. Surprise was among the most commonly mistaken emotions along with Fear and

Disgust. The error pattern consisted of mutual confusions between Surprise and Fear, and with the expression of Disgust that was described with inappropriate labels, or as "sad" or "angry".

2.3.4 Discussion

The present study sought to investigate both perceptual and semantic abilities in children with autism to recognise basic emotional states of others through their facial expressions. The present results failed to support the prediction of a selective impairment in children with autism. They did not show impairment in recognising the belief-based emotion of surprise as opposed to the reality-based emotions of happiness, sadness, anger, fear and disgust. Moreover, it failed to support a general impairment hypothesis predicting affective impairment in response to emotion stimuli.

The study revealed that children with autism were as able as controls to recognise all six basic emotions from facial expressions. This was shown not only when they were required to match pictures of emotional expressions with different intensity levels, but also when they were asked to provide a label for expressions with a normal intensity. In addition, consistent with cross-cultural data (Ekman, 1982; Young et al., 1997) the results showed that the emotion of happiness was very easy to recognise, whereas the expressions of surprise and fear were difficult to distinguish along with disgust and anger. In particular, all children found both surprise and fear the two most difficult expressions to discriminate at a perceptual level (Exp.1). However, this difficulty did not occur at a semantic level, since all children found surprise the second easiest expression to name after happiness (Exp.2). Furthermore, the difficulty relative to discriminating surprise at a perceptual level with 90% and 70% intensity level did not occur when children were asked to name the same stimuli (Exp.3). Indeed, surprise was among the emotions that were easiest to name after happiness.

There are several points of discussion concerning the confusion between surprise and fear. The close relationship between the two expressions was already noted by Darwin who considered these emotions as part of the same continuum, pointing out that fear is often preceded by and sometimes mixed with surprise. Ekman and Friesen (1971) cross-cultural studies indicate the same type of confusion in both literate and illiterate cultures. Data from the performance with the ambiguous expressions combining surprise with fear showed no particular preferences towards one expression or the other, indicating that the two expressions display ambiguous/unambiguous features to the same extent. Confusion cannot be attributed to the ambiguity of the 50% stimuli in general, since when children had to decide to match the 50% expression of surprise/happiness, they chose more often the expression of happiness, as expected on the basis of the categorical perception study by Etcoff and Magee (1992) and on results from Calder et al.'s (1996) study. Overall, the close resemblance between surprise and fear has confused the children with autism no more than the controls.

Interestingly, the analysis of the most frequent errors made by all children indicated that the confusion fear/surprise is not symmetrical: at a perceptual level (Exp.1) surprise was more often matched with fear than fear matched with surprise, whereas at a semantic level (Exp. 2) the error bias was in the opposite direction, that is, subjects confused more often fear with surprise than surprise with fear. Clearly, the possibility of unidirectional bias needs further investigation.

The negative findings of the present study are in contrast with at least two different previous studies indicating specific emotion recognition impairment in autism. In fact, Baron-Cohen et al.'s study (1993) showed a specific impairment in autism in recognising the expression of surprise, and Howard et al. (2000) showed a specific impairment in recognising fear? I will first discuss how these different results can be reconciled with the Baron-Cohen et al.'s (1993) data.

Some differences between the Baron-Cohen et al.'s study and the present have to be taken into account. First, there was a difference in the children's verbal mental age. In Baron-Cohen et al.'s study children with autism had a considerably lower verbal mental age (CA mean: 12.6 yrs, sd.: 3.5 yrs, VMA: 5.3 yrs, sd.: 1.0 yrs) than those in the present study (CA mean: 12.3 yrs, sd.: 2.3 yrs, VMA: 9.2 yrs, sd.: 2.6 yrs). It could be that younger children are more familiar with the facial expressions of happiness and sadness rather than surprise. Indeed, Kestenbaum (1992) reported that 5-years-old children found surprise the most difficult emotion to recognise, and it was perceptually confused with happiness, but at the same time was associated with the semantic category of "feeling bad".

A second difference between Baron-Cohen et al.'s study and the present study was in the type of stimuli presented to the subjects. In the former, the stimuli were both drawings and pictures of individuals showing facial expressions. In the current study, the participants were presented with stimuli either developed by, or derived from, Ekman and Friesen's (1976) standardized series of facial affect. It is possible that the images of surprise presented in the former study may have been less easily discriminated than the images of happiness and sadness.

A third difference between the two studies was in the range of expressions investigated. Baron-Cohen et al.'s study only investigated the ability of children with autism to recognise happiness, sadness and surprise. In contrast, the present study investigated the ability relative to all six basic emotions. This allowed for controlling the effect of featural complexity of surprise. Indeed, the expression of surprise involves coding the eyes and the mouth together (although the crucial difference lies in the lower eyelid), whereas happy and sad expressions can be coded simply by reference to the mouth alone. It might be that in earlier study children with autism had more difficulty identifying the expression of surprise compared to happiness and sadness simply because of the complexity of its features. In the present study, the presence of all six emotions, including the expression of fear, which has been found highly confusable with surprise, might have contributed to eliminate an effect of featural complexity.

Children with autism might have performed worse than the controls in the Baron-Cohen et al.'s study because they attended less to the eye region. In fact, the type of mistakes they made more often was matching surprised faces with happy faces suggesting that they were discriminating on the basis of similar mouth features. The hypothesis of an autism specific impairment in recognizing the "language of the eyes" has been formulated by Baron-Cohen et al. (1997). This study found that adults with autism had difficulties in processing the information carried by the eye region of facial expression of complex mental states rather than for basic emotions. More research is needed to determine possible developmental factors in the acquisition of the "language of the eyes" from simple emotion expressions to complex mental state expressions.

One purpose of the present study was to monitor the performance of children with autism relative to the hypothesis that amygdala abnormality could contribute to some of the neuropsychological impairment in social and emotional processing of individuals with autism (Baron-Cohen et al., 1999; 2000; Howard et al., 2000). These negative findings in both matching and naming tasks are inconsistent with Howard et al.'s study showing concomitant evidence of amygdala abnormality and selective fear recognition impairment (in a forced-choice naming task) in a group of high-functioning individuals with autism. How can the present findings be explained in relation to the Howard et al. (2000) study?

First of all, it is important to underline that so far no other study has pinpointed a specific impairment in basic emotion recognition, and in particular fear, in individuals with autism. Three studies, Adolphs et al.'s (2001), Grossman et al. (2000) and Buitelaar et al. (1999) used the same type of standardized stimuli (Ekman and Friesen, 1976), which have been used in the present study and in Howard's et al.'s study. All three studies reported negative findings with different age groups of individuals with autism - adults, adolescents and children – respectively.

In addition, the same set of standardized stimuli used in the present study have been shown to be sensitive enough to reveal selective emotion impairments on different clinical populations, other than individuals with autism. Calder et al. (1996) found that two patients with amygdala damage failed to identify the expression of fear, and a collaborative study (Adolphs, Tranel, Hamann et al., 1999) between several laboratories, which investigated nine individuals with bilateral amygdala damage, reported that the most common impairment was in recognition of fear. Blair and Coles (2000), and Stevens, Charman and Blair (2001) found that psychopathic individuals show a selective impairment in processing fearful and sad facial expressions. Thus, it seems that the negative findings on fear recognition impairment on autism of the present study compared to the Howard et al. (2000) are not attributable to the reliability of the type of test that has been adopted.

At this point, it might be appropriate to ask why only one single study has pinpointed a selective impairments in fear recognition while several studies did not. It is plausible that Howard et al.'s study showed impairment in a particular subgroup of high-functioning autism. In fact, the Howard et al.'s study is consistent with the finding that subjects with damage to the amygdala also show abnormal emotional processing of the expression of fear (Adolphs et al., 1994; Calder et al., 1996). However, the collaborative study by Adolphs et al. (1999) reported above, indicated that the individual patients' deficit in recognition of fear and anger ranged from extremely impaired to almost normal. Similarly, only four subjects with autism out of ten in the Howard et al.'s study seem to be extremely impaired in fear recognition, whereas the performance of the remaining subjects overlapped with some, and at least one, of the control subjects. It must be noted that some subjects of the autistic group are also extremely impaired in recognition of sadness, disgust and anger, a pattern of impairment similar to that seen in subjects with bilateral amygdala damage (Adolphs et al., 1999). Unfortunately, the individual subjects' data are only shown in a graphic form. In conclusion, a replication of these findings, along with a description of individual subjects' performance is clearly needed to gain more information on the link between amygdala functioning and fear recognition impairment in autism. In particular, it would be of interest to investigate the incidence of a subgroup of high-functioning individuals with autism with specific fear recognition deficit, similar to Adolphs et al.'s (1999) study with amygdala patients.

A more general point of discussion relative to the amygdala hypothesis of autism concerns the evidence of impairment on tasks based on recognition of complex mental/emotional states as opposed to basic emotions. In a recent study Adolphs, Sears and Piven (2001) compared subjects with amygdala lesions to subjects with autism in the ability to process emotional and social information from faces. The autism group found no difficulties in basic emotion recognition, but performed poorly in a social judgement task (judging trustworthiness and approachability of people from their faces). The same pattern was found in the group of patients with bilateral amygdala damage. The authors suggest the negative performance may result from an impaired ability "to link perception of faces to the retrieval of social knowledge, and that this impairment may result in part from dysfunction involving the amygdala" (p. 239). Two other studies reported no basic emotion recognition deficit, but difficulties on more demanding tasks. Grossman et al. (2000) found that high-functioning subjects with autism were impaired on emotion recognition by watching people's face associated with irrelevant labels. Baron-Cohen et al. (1997) found that high-functioning adults with autism had more difficulties if asked to "read" people's basic emotions by watching their eyes alone. They were also impaired in reading complex emotions/higher-order mental states from both the whole face and the eye region. Thus, how could it be explained the convergence of findings that individuals with autism, like patients with amygdala damage, pass basic emotions recognition test but fail to recognise more complex stimuli involving the perception of faces or part of faces?

One plausible explanation is that subjects are able to recognise basic emotions by adopting compensation strategies that allow retrieval of knowledge about the threat, danger and distress signals using anatomical routes other than via the amygdala. These strategies would not support tasks based on high-cognitive processing demands. Interestingly, the social judgment of faces task adopted by Adolphs et al.'s (2001) study, which was failed by both the autism and the amygdala groups, relies on the ability to understand the mental states of others, and to predict their behaviour on the basis of their appearance. It seems, therefore, that subjects failed the task on the basis of a deficit on mentalising. The same explanation applies to the Baron-Cohen et al.'s (1997) eye tasks, but not to the Grossman et al.'s identification task (as already discussed, executive function processing was not controlled). As far as I am aware, there has been no systematic investigation of Theory of Mind performance in large sample of individuals with amygdala lesions. Fine et al. (2001) reported an interesting single case of a patient with left amygdala damage who showed poor mentalising ability, but normal executive control functioning. However, the role of the amygdala in the normal development of ToM ability, and its possible role within the distributed neural system involved in mentalising (Fletcher et al., 1996, Gallagher et al., 2000, see also chapter 4 and 5 of the present thesis) remains unclear and needs further investigation.

In conclusion, a final point of discussion concerns the theoretical clear-cut distinction between belief-based emotions (cognitive) and reality-based emotions (simple), upon which the present study based its prediction. It seems that although emotional expressions are clearly displayed on people faces, the distinction remains elusive. Indeed, Baron-Cohen et al. (1993) acknowledged that the distinction is quite subtle and it is primarily based on what is *typically* interpreted as simple and cognitive emotions. The expression of surprise can be primarily characterized as a "basic emotion" in the sense of a rapid reaction response to certain events. It could be suggested that in order to understand someone else's surprised reaction, it is necessary *first* to represent the event that has caused it, and *second* to represent the individual's expectation that it has been violated. If the expression of surprise is considered an "approach emotion" (Davidson, 1992) which is associated with a call for further information, then it would be of interest to explore the ability of individuals with autism to associate the expression of surprise with an appropriate follow-up behaviour. It would be expected that surprise, contrary to the other basic emotions, is more likely to be followed by another expression (e.g., fear or happiness) before taking an appropriate action.

Although the Theory of Mind hypothesis was not supported by the present findings on *basic* emotion recognition, other studies have found impairments on tasks that tap Theory of Mind processing. Thus, as mentioned above, it is plausible that individuals with autism are able to bypass the impairment in recognizing basic emotions by compensatory strategies. In fact, it seems that learning to associate stereotypical facial expressions to few emotions is not a particularly difficult task. The everyday exposure to such stimuli allows a constant source of information for learning the association between facial expressions and feelings or needs. In addition, it is common that adults provide reinforcement cues to children in order to decode emotional signals, e.g., a parent pretending to be crying because he/she wants the child to understand he/she did something upsetting. On the other hand, it is possible that individuals with autism have no impairment in recognizing emotion displays that have evolved with adaptive functions, but have difficulties in linking the perceptual level of emotion recognition with the higher level of understanding the social meaning of different expressions. In this sense, it is important that research on basic emotion recognition should be oriented towards investigating very young children, and possibly with measures that bypass compensatory strategies. Since imposing time constraints on the task is problematic with individuals with autism who generally have attention deficits, it would be necessary to manipulate the type of visual display, for example, creating kinetic expressions. Considering that movements of facial muscles are crucial in displaying emotions, it is possible that recognition tasks based on kinetic images of emotions tap possible general or selective impairments. The study by Adolphs et al. (2001) and Grossman et al. (2000) discussed above have made an attempt to investigate the ability of people with autism to decode the emotional information beyond the basic ability to recognize basic emotions. Although this type of investigation into emotional stimuli processing appears to be very interesting, it is important to control for both executive function deficit and mentalising deficit, which are theoretically distinct to the process to be investigated.

Chapter 3

Understanding intention from goal-directed motion

In this chapter, I will present a study on the ability of children with autism, normally developing children and adults to attribute to an agent the goal of reaching a stationary target. The paradigm is based on a computer-animated sequence that has been specifically created for the present project.

I will first describe the theoretical background to the study by summarizing the tripartite theory of agency by Leslie (1994). According to this theory, the cognitive system dedicated to understanding agency is constituted of three components, each specialized to attend to a type of property that distinguishes an agent from a non-agent. The first component deals with physical events, the second with goal-directed actions, and the last with propositional attitudes. I will then describe the empirical background, focusing on infant studies that have explored the development of psychological reasoning using visual paradigms displaying moving abstract shapes (Schlottmann and Surian, 1999; Rochat, Morgan and Carpenter, 1997; Gergely, Nadasdy, Csibra and Biro, 1995; Csibra, Gergely, Biro, et al., 1999). Finally, I will present a paradigm (Montgomery and Montgomery, 1999), adopted in a study with preschool children, which inspired the present experiment. The study was targetted at the perception of an agent's goal-directed motion that is actually failing to achieve his intended goal. A test that investigates the ability to attribute to an agent the intention to reach a target by ignoring the salient fact of an unsuccessful outcome seemed appropriate to use with young normally developing children and children with autism.

The first question was whether children with autism have particular difficulties in attributing goal-directed intention. The second question was whether developmental changes would occur in the perception of an agent's motion when it reaches a target: is the perception of motion more salient than the perception of the outcome? For example, younger children might attribute goals on the basis of the agent's proximity to the target (outcome-based representation) whereas adults attribute the intended goal on the basis of the perceived motion of the agent towards the target (motion-based representation, regardless of outcome). The study comprises two aims and will be presented in two sections. The first aim was to investigate the performance of children with autism, normally developing children, and adults in the ability to identify an agent's intended goal in the presence of its unsuccessful outcome. The second aim was to explore the developmental changes in the ability of identifying an agent's intended goal in the presence of an outcome which was on the opposite side of where the goal-directed motion originally "pointed-towards".

3.1 Theoretical background

According to Leslie (1994), the ability to detect agency is the result of domain-specific learning. The architecture of the human brain is the result of evolutionary adaptation, more specifically, the development of distinct information processing systems which reflects different properties of the world. Understanding agency involves three distinct hierarchically arranged processing levels of a dedicated system. Below I will describe the three sub-systems of the cognitive mechanism.

3.1.1 Leslie's tripartite theory of agency

The first level of the architecture of the system reflects the understanding of agency in a mechanical sense. Important functions of this component are distinguishing agents from other physical bodies and describing their mechanical interactions. The other two components reflect the understanding of agency in an intentional sense. In particular, one component is concerned with agents and the goal-directed actions they produce, and the other is concerned with agents' mental states and their behaviour. Whereas the first level can be thought of as part of a "theory of body mechanism", the other two levels are part of a "theory of mind". In this respect, there are three different categories of events that are processed by the three distinct levels of the mind/brain system that deals with agency. The first level of the system ("theory of body mechanism") represents only mechanical relations that are displayed locally and contiguously in space and time (states that are here and now). However, it cannot represent the event when an agent acts in pursuit of a goal (a state of affairs that an agent tries to bring about, e.g., that is in the future), nor when an agent acts in relation to propositional attitudes such as believing and pretending (e.g., states that are in the mind). Leslie uses the term "fictional" circumstances to describe non-contingent events causing the agent's behaviour. Thus, the second level of the system that operates with agents and actions requires a representation of relations between agents and events that are at distant times and places. By contrast, the third level, the ability to understand agents and their mental states requires the representations of relations that are beyond spatiotemporal circumstances, namely, representation of mental states, or propositítonal attitudes. Although these two latter systems are components of a theory of mind mechanism, they are distinct in the sense that the notion of actional agency (as opposed to mechanical agency), includes only a "weak form of the fictional causes problem", whereas the notion of agency and propositional attitudes includes the "full-blown fictional causes problem" (Leslie, 1994, p. 139).

This brief summary of Leslie's theory of agency served the purpose of delineating a theoretical background to the creation of the paradigm of the present study (and in the following chapter 4). It is worth noting that in Leslie's view the tripartite theory of agency does not entail a sequential development: each subsystem can develop in parallel, although they can begin their development in sequence according to the maturation of their subserving neural system and the availability and quality of inputs. An important aspect of this theoretical approach is that each of the three components constitutes a "learning device" with a specific way of organizing the inputs it receives. Thus, an interesting question concerns the nature of their obligatory triggering inputs. In this respect, Leslie hints at an important distinction between the notion of agency and animacy: although most objects that are agents are also animate, "the notion of animateness is external to agency and proprietary to the biological domain" (Leslie, 1994, p.121). Moreover, "separating agency from animateness [...] allows us to apply without obstacle our commonsense understanding of Agency *a*) to inanimates and *b*) without having to know whether or not something is animate" (Leslie, 1994, p.145). This notion is important since in the literature some confusion is caused by the lack of a precise terminology regarding the distinct movements that an object may display. In particular, it is not clear whether self-propelled movement is a property of animacy or of agency, or whether the movement that characterizes living things, namely, biological motion (e.g. expansion and contraction of body surface, as an effect of breathing) is part of the domain of agency or animacy. Research on infants' perception of animated agents would benefit from clarification of these issues.

3.1.2 Triggering inputs to the representation of agency

The first level of the agency system concerned with understanding the physical world deals with representations of mechanical events and three-dimensional objects (the only physical bodies that possess mechanical properties). The two principal inputs from vision are the recognition of three-dimensional objects and the recognition of motion patterns. Indeed, motion is the principal source of information about mechanical events. However, motion analysis takes place independently of the processing of objects' properties.

The perception of physical causality based on motion cues alone was investigated by the French psychologist Michotte (1946/1963), whose work has strongly influenced contemporary research with infants on physical and psychological causation. Michotte discovered that adults had an immediate impression of cause and effect when viewing simple animated sequences displaying the patterns of motion of geometrical shapes. Chief among these sequences is the "launching" event: Two small squares are resting on a line, separated by several inches. The first square A moves in a straight line until it reaches the second square B, at which point A stops moving and B starts moving along the same trajectory. By viewing this sequence, one has the impression of two distinct items and a single motion that is transferred between them, resulting in the perception that A causes the motion of B. The importance of this phenomenon is that the visual system recovers the causal structure of the world in a way that is typically associated with higher levels of cognitive processing, while the processing remains at the level of kinematics. The bulk of Michotte's demonstration and the extension of his work consisted in discovering the spatiotemporal parameters that mediate causal representations, such as the objects' relative speed, speed–mass interactions, overall path lengths, and spatial and temporal gaps. Small manipulations made to the displays can make causal representation to disappear. In addition, these phenomena seem to be culturally universal (Morris and Peg, 1994). The relevant experimental studies have been comprehensively reviewed by Scholl and Tremoulet (2000).

The second level of the agency system, concerning the property of agents to pursue a goal or state of affairs distant in time and place, seems to be triggered by stimuli indicating a change in physical circumstances. Leslie gives some examples of what might be interpreted as a sign of an action-based representation in infants: the ability of 6-month-old infants to follow eye gaze (Butterworth, 1991), the ability to attend to the uses an agent makes of objects (Abravanel and Gingold, 1985), and the ability to appreciate the role structure of some simple goal-directed action, such as "give and take" behaviours (Bruner, 1976). More suggestions relative to infants' sensitivity to goal-directed motion come from an influential paper by Premack (1990). According to his theory of intentionality, "the perception of intention, like that of causality, is hard-wired perception, based not on repeated experience but on appropriate stimulation" (p. 2). The appropriate triggering input for attribution of intentionality is the self-propelled movement. However, the crucial parameter is not self-propelled motion per se, but clear changes in movement. More specifically, Premack and Premack (1994) suggest that the perceptual triggering inputs for intentional attribution to moving agents are the following: *a*) motion directed towards the same single item, *b*) repeated motion (failing and trying again), and *c*) variable motion patterns. In sum, the perception of repeated attempts to overcome failure is a strong cue for attributing an intention with a goal to a moving agent. It is important to stress that the perception of the agent's outcome (e.g. goal-attained or goal-missed) in relation to the repeated goal-directed motion towards a target remains to be explored. Whereas Premack's account (1990; Premack and Premack, 1994) does not explore the possibility of a learning effect in goal-attribution, Leslie acknowledges that the information relative to the agent's outcome can be useful for representing later actions of agents that are directed to the same kind of goal, even when the intended outcome is not achieved.

It must be noted that Baron-Cohen (1994, 1995) described a mechanism of the multi-components metarepresentational system, the Intentionality Detector (ID), which is similar to the second level of agency postulated by Leslie (1994) and to the infant's perceptual intentionality theory postulated by Premack (1990). More specifically, the ID is triggered by a wide range of stimuli that imply direction (e.g. a touch, a push, a jump, a shout, an arrow) or that manifest self-propulsion. The function of detecting eye gaze, or any eye-like stimuli is specific to another mechanism, the Eye-Direction Detector (EDD), which is a specialized part of the human visual system (whereas the ID mechanism is sensitive to visual, auditory and tactile inputs).

Finally, the third level of the agency system deals with propositional attitudes: according to Leslie the clearest early sign of the employment of this sub-system is the ability to understand pretense and implicit meaning during ostensive communication behaviours (Leslie, 1987). This level of the agency system is not further discussed in this chapter since it is the aim of the studies described later in this thesis (chapters 4 and 5).

3.2 Developmental studies on the perception of moving shapes

Work on the developmental origins of the ability to process psychological events as opposed to physical events have focused on different types of agency. An agent's behaviour can be caused by the presence and the behaviour of another agent or by its own intentions, desires and beliefs. For example, Rochat and colleagues (1997), and Schlottmann and Surian (1999) have investigated agency with visual stimuli depicting goal-directed movements of two agents. Gergely, Csibra and colleagues (1995, 1999) examined the ability of infants to distinguish a rationally appropriate goal-directed movement of an agent from one that is not appropriate. Finally, Montgomery and Montgomery (1999) investigated the ability of young children to distinguish an agent's intended outcome from an accidental outcome.

3.2.1 The role of motion cues in infant's perception of moving objects

Rochat et al. (1997) showed movie sequences to 3-month-olds and adults: one sequence displayed an interaction between two geometrical figures (one triangle chasing the other) and another displayed the same figures moving about independently, in an uncorrelated fashion. By analyzing the infant's looking times they found that at a very young age infants are able to distinguish movements that specify chase from those that do not. Since the two movie sequences differed greatly in the type of motion patterns displayed by the two agents, it is unclear whether children looked more interested to the chasing sequence than the random sequence because they appreciated the display of a more continuous motion pattern or because they appreciated the display of "psychological causality".

In a study based on a reversal paradigm and an infant-controlled visual habituation-dishabituation technique Schlottmann and Surian (1999) investigated infants' ability to understand causation-at-a-distance. The habituation-dishabituation technique involves recording the infant's looking time to different displays. It is assumed that when the infant is shown the same display repeatedly, his/her spontaneous visual attention will decline. By contrast, the infant will look longer at a novel display. This effect has been observed as early as the first days of life (Slate, Morison and Rose, 1984). In order to measure their looking times, infants are first habituated with a display, and then are tested with two other distinct events. The event that elicits longer looking times is interpreted as unexpected, whereas the event that elicits shorter looking times is interpreted as expected from the representation that the infant has formed.

In Schlottmann and Surian's study two groups of 9-month-old infants were shown two different types of sequences displaying two non-rigidly moving squares. The experimental group saw a "reaction event" in which a green square moved before the red one stopped, so that the red is perceived as causing the motion of the green. The stimuli were modelled after Michotte's "caterpillar" movement in which two squares moved successively and without contact. The control group was shown a "pause event" obtained by inserting a pause between the movements in the caterpillar sequence, so that the two squares appear to move independently of each other. Infants were habituated to one of the two sequences. After habituation, infants were presented with the same sequence played in reverse. Both reaction event and pause event sequences involve identical spatiotemporal changes that are maintained in the reversed sequence. However, only the reversal of the reaction event involves a change in the causal role of the two objects (e.g., one square "chasing" the other, which in the reversed causal role becomes the victim "retaliating" by hitting back). Both the spatiotemporal relation between each event and the causal roles of the reversed versions were tested with adults subjects. Results showed that infants who were habituated to the reaction event dishabituated more to the reversed sequence than the group habituated to the pause event. The changes in looking times suggest that 9-month-old infants can represent causation-at-a-distance in which one agent appears to affect another without physical contact, and are surprised when their expectation is not fulfilled. According to the authors, the perception of causality-at-a-distance could provide a "blueprint" for early understanding of psychological causality.

In another habituation paradigm Gergely et. al (1995) showed that 12-month-olds develop specific visual expectations about the type of approach of a computer-animated circle towards a stationary goal. The experiment was designed in order to meet the requirements of agency categorization and intentional attribution, under the assumption that "the infant's theory of agency contains, as one of its foundation components, an assumption of rationality of action" (p. 172). In the authors' view, a domain-specific system is identified by the successful application of a principle of reasoning specific to the domain, so that the domain of naïve psychology is based on the principle of "pure reason" (cf. Csibra et al., 1999) as opposed to a system that is specified in terms of features or kinetic-behavioural properties as in the agency system described by Leslie, (1994) and by Premack (1990).

The study on perception of "rational action" with one-year olds was replicated successfully in a more recent study by Csibra et al. (1999) with infants 9-month-old, but not with 6-month-olds. Two groups of 9-month-olds were habituated to a visual event in which they observed a small circle repeatedly approach and contact a large circle with a "jumping" trajectory. The experimental group saw the circle jumping over a wall situated in the middle of the screen separating the two circles. In the control group the wall was situated near the edge of the screen, appearing as if "behind" the small circle. The sequence of the experimental group evoked the interpretation of an instrumental action to achieve a goal-state (reaching the large circle) by overcoming an obstacle (jumping over the wall separating the two circles). Infants of both groups were tested with two different sequences displaying a novel situation, in which the obstacle is no longer present. In one sequence the small circle approaches the large circle with a direct straight trajectory, a novel action but consistent with the principle of least effort and hence rational action. In the other sequence, the small circle approaches the large one exhibiting the same jumping trajectory, which can no longer be justified in the new situation as a rational action to achieve the goal. Since the experimental group looked less at the "straight line approach" sequence and more at the "jumping approach", the authors concluded that 9-month-old infants were able to predict the most likely future goal-directed action of a rational agent.

Another experiment was designed to investigate whether the early ability to interpret rational goal-directed behaviour can be linked to the perception of specific movement cues. The design of these tests was similar to the one described above, with a baseline event, a novel rational event and a novel non-rational event. The circle's behaviour was stripped of animacy cues (it did not exhibit surface contractions), and agency cues (it did not change direction and followed an inert pathway as if subject to gravitation). In particular, the sequence displayed a large circle beyond a wall (similar to the previous test) with the small circle "flying" over the obstacle (its motion trajectory starts above the ground - in the air -) and landing next to the large circle. In the test events, the obstacle was no longer present: in the rational event the circle "flew" as in the habituation event, whereas on the non-rational event, the small circle approached the large circle through the shortest straight-line route. The control group saw the same test events, but was habituated to a different event, which could not be construed as a rational goal-directed action. The small circle flew over a rectangular bar "hanging in the air" so that the bar did not form an "obstacle", as it was possible to pass under it. Results replicated the findings of the "jumping" approach experiment: infants of the experimental group discriminated between the rational and non-rational approach, looking less at the rational approach. Furthermore, since the stimuli of this test showed no movement cues of agency such as self-propulsion, the authors interpreted the results as indicating that the perception of kinetic agency cues is not an obligatory precondition for the attribution of goal-directed behaviour as suggested by the cue-based model of agency described above (Leslie, 1994; Premack, 1990).

However, it must be noted that the circle in the "flying" test may still be perceived as a "flying agent" as opposed to a "flying object". It is unclear whether a flying circle that suddenly enters the visual field and after landing stops immediately, provides the necessary visual information for identifying it as a moving object with an external source of energy (non-agent), as opposed to a moving object with an internal source of energy (agent). In fact, despite the authors' claim that the object's motion looked more like a tennis-ball thrown by an invisible person than a self-propelled object, the visual parameters that distinguish a self-propelled motion from an externally caused motion have yet to be explored in detail.

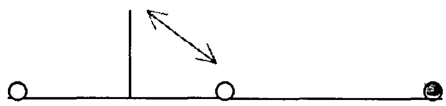
Two general concerns about the paradigm used in the Gergely and Cibra's studies need to be mentioned. Unfortunately, unlike the study by Schlottmann and Surian (1999) no adult subjects were tested to explore the elusive notion of rationality applied to the agent's approach to the goal. For example, given a new situation, it could be less effortful and hence more "rational" to make a conservative choice by following a familiar path rather than following a novel path. In addition, the choice of dependent variable (rational and irrational approach to a target) raises another concern. Whereas it allows for testing the ability to distinguish different means of approaching a target, it does not allow for investigating the ability to attribute a goal *per se*, which would be useful as a baseline condition. Nevertheless, these studies show that movement cues are utilized in the representation of goal-directed action. This ability appears in children as young as 9 months of age.

3.2.2 The role of motion and outcome cues in children's perception of moving objects

Another type of paradigm has been used with preschool children to investigate the attribution of intentionality within a cue-based conceptual model of agency (Montgomery and Montgomery, 1999). The paradigm was designed to conform to the requirements suggested by Premack and Premack (1994) described above. Thus, the critical cues for triggering goal-attribution are the persistence of motion (e.g., when a movement is repeated in the attempt to overcome an obstacle), and consequently, the variability of the motion (the attempts ideally should improve the distance between agent and target). The test is of particular interest mainly because it gives the opportunity to analyze children's understanding of an agent's intention on the basis of a clear distinction between intentional and unintentional outcome, and secondly, because it has the potential of providing evidence for some specific motion features as triggering inputs for intentionality.

Similar to the "rational goal-directed action" paradigm described above, 3- and 5-year old children were presented with computer-animated events displaying a circle, a wall, and a target beyond the wall (Figure 3.1).

Figure 3.1: Stimuli used in the Montgomery and Montgomery's (1999) study.



The different conditions consisted in *a*) the presence of another potential target (non-target) resting on the same side of the circle, but further away from the wall, and *b*) the repeated attempts by the circle to reach the target beyond the wall. The manipulated variable was the outcome of the persistent goal-directed movement. The circle jumps a total of three times: twice it bounces off the wall because it jumps too low, and the third time one of the following three outcomes occurs: 1) the circle overcomes the wall and lands next to the target (goal attained condition), 2) the circle fails to overcome the wall and lands close to the circle standing opposite the target (non-goal condition), 3) the circle fails and lands at a point equidistant to the target and the non-target (neutral condition). Children were asked to indicate the intended goal of the persistent circle by pointing either to the target or the non-target.

The results provided support for Premack and Premack's (1994) suggestion that perceived persistence relative to an agent's movement towards a target can elicit accurate judgements of the goal intended by the agent. Preschool children's success was not only limited to the occasion in which outcome and persistence matched (goal-attained condition), but it also extended to situations in which the outcome of the action was inconclusive (neutral condition). However, the non-goal outcome condition prompted some 3-year olds to attribute the goal to the agent on the basis of the outcome, and not on the basis of the movement cues. Indeed, after bouncing off the wall, the circle landed progressively closer to the non-target allowing for interpreting its movement as aimed at reaching the non-target. Therefore this condition was modified in a second experiment in order to reduce ambiguity: the agent initiates its goal-directed movement close to the non-target, and when it fails, it lands exactly where it started. By watching this new sequence, children in both groups inferred correctly that the agent wanted to reach the target beyond the wall, despite the fact that it landed close to the other target.

In the second experiment an additional condition was presented to examine whether children judged that the direction of the first movement was sufficient for determining the circle's goal. In this condition the wall is not present and the agent jumps towards both potential targets an equal distance and an equal number of times. Children appeared to respond randomly when inferring the agent's intended goal by watching only a single movement towards the target. However, it must be noted that the sequence was very dissimilar to the other conditions: the absence of the obstacle and the agent's motion pattern split between the two targets might have been generally confusing for the subjects.

In summary, Montgomery and Montgomery's study investigated whether preschoolers' attribution of intention was more influenced by patterns of motion characterizing persistence or by the outcome of the motion. The authors interpreted the evidence as an indication that children as young as 3 tend to rely more heavily on the persistent movement cue than to the spatial cue of the outcome to infer an actor's intended goal. However, it must be noted that the paradigm includes only the manipulation of the outcome variable (intentional, unintentional and inconclusive) whereas the motion pattern of the agent is the same in all three conditions (jumping three times towards the direction of the target). Thus, the results indicated only that children were able to attribute an intended goal to an agent in the presence of a persistent movement towards a goal when ignoring the agent's final unsuccessful outcome. Thus, there is no doubt that kinetic patterns are highly important in inferring an agent's goal. It is no clear, however, whether motion is just a sufficient cue, or also a necessary one. The interesting question of whether children can switch their attention either to the motion pattern of the agent or to the final outcome depending on the context, has yet to be investigated.

3.3 Exploring the ability to represent intended goal in children with autism

Behavioural studies on infants and young children have provided evidence for an early emergence of the ability to process psychological causation as opposed to physical causation. Infants are able to process causation-at-a-distance (Schlottmann and Surian, 1999) and to discern goal-directed motion from random movement (Rochat, Morgan and Carpenter, 1997) or from rationally inappropriate movement given the reality context (Gergely et al., 1995, Csibra et al., 1999). Children between the ages of 3 and 5 are able to distinguish an intended outcome from an accidental outcome involving two agents (Montgomery and Montgomery, 1999). However, the ability to suspend causal principles and expectations – i.e. naïve physics - and the development of awareness of special properties of the domain of agency have rarely been investigated in children with autism, although several studies on folk physics in adults with autistic disorder have been carried out (e.g. Baron-Cohen, Wheelwright, Stott et al., 1997; Baron-Cohen, Wheelwright, Stone et al., 1999). A host of studies have investigated the difficulties of individuals with autism to represent mental states at different cognitive developmental stages: from early ostensive communication behaviours such as orienting towards others' focus of attention, joint attention, to the subsequent ability to understand the concept of false belief or deception. This field has been discussed critically in two volumes edited by Baron-Cohen, Tager-Flusberg, and Cohen (1993, 2000). If, as suggested by Leslie (1994) there are two distinct levels for representing agency beyond its mechanical properties, namely, one for its actional properties, and the other for its attitudinal properties, then it is of particular interest to investigate the ability in people with autism to represent intentionality with respect to these sub-systems. Distinguishing different cognitive levels, each dealing with representation of agents' intentional behaviour has theoretical plausibility (Leslie, 1994; Baron-Cohen, 1994, 1995). At a "lower" level, there might be a system concerned with the pursuit of a future state of affairs (goal-directed action) and, at a "higher" level, there might be a system concerned with states that are "about something", namely, propositional states (e.g. "believing that", "hoping that"). The difference between the two levels is that the former is concerned with the representation of an agent as "ACTING to bring about [a state of affairs]" (Leslie, 1994, p. 139). The latter represents an agent as "actively holding an attitude to the truth of a proposition" (p. 139). Hence, it is sufficient for the child to be able to represent possible or future states of affairs in order to understand an agent's goal-directed behaviour without necessarily representing a mental state. As shown in many studies following Baron-Cohen, Leslie and Frith (1985), in the case of people with autism the development of the ability to represent mental states at a "higher" level is compromised. However, it is less clear whether the ability to represent intentional states at a "lower" level of the intentional representation system is unimpaired or not.

3.3.1 Self attribution of intention

The distinction between desire (propositional attitude) and intention (pursuit of a future state of affairs) has been investigated by Phillips, Baron-Cohen and Rutter (1998) in a study with children with autism. Since most of the time desire coincides with intention, the authors developed a paradigm aimed at disentangle the two mental states. The task was based on the personal experience of children who were holding an intention and were asked to compare this intention with the outcome over which they had no control. They were asked to play a target-shooting game with a water gun firing at colourful canisters, some of which contained a prize. At the beginning of the game they were asked to say which coloured canister they wanted to hit. Both the outcome and the desirability of the outcome were manipulated by the experimenter so that there were both intentional and accidental outcomes which either produced desire satisfaction (winning a prize) or not (no prize). The discrepant conditions between desire and intentions consisted in the child missing the intended target but winning a prize, and vice versa, when he/she hit the intended target but won nothing. Results indicated that children with autism did understand their own intended goal when they hit the target regardless of whether this satisfied their desire of winning the prize. By contrast, they had difficulties in self attributing their original intended goal when the target was missed, regardless of prize. The authors pointed out that executive function impairment might account for this result: in order to correctly answer the question "what colour did you mean to shoot?" when they missed the intended target, children had to suppress the prepotent but incorrect response that they wanted to hit the canister that they were actually holding in their hand. However, children were also reminded of their intended target by a coloured card that was placed in front of them at the beginning of the game. It remains an open question whether children's response was influenced more by the inability to inhibit a prepotent response or by the inability to understand unintentional outcome. Interestingly, Phillips et al.'s study showed that children with autism are able to understand their own intended goal when it was attained but discordant with their desire. Furthermore, they seemed overall less impaired in this test than in tests of false beliefs, which were also administered during the experiment.

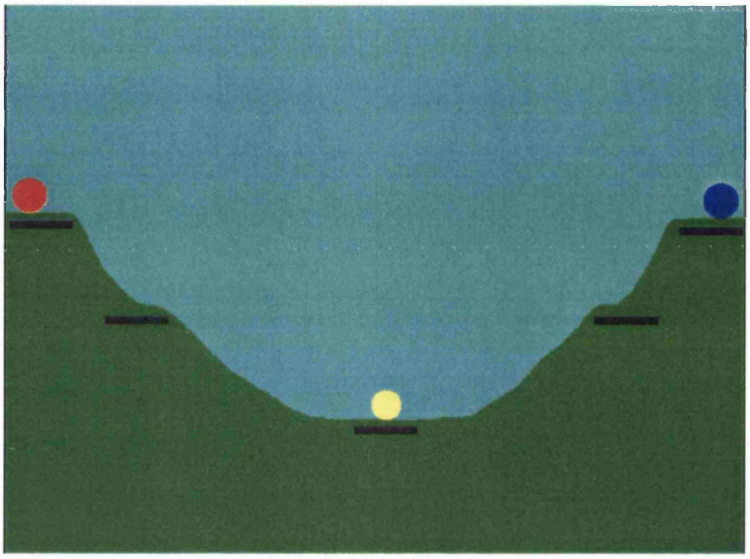
3.3.2 Attribution of an intended goal to a non-human agent

The present study was developed with the aim of testing the ability of children with autism to attribute an intended goal to an agent by using simple kinetic visual stimuli. A new paradigm was created taking inspiration from those used in research reviewed above concerning young children's perception of moving geometrical shapes. The test is consistent with the cue-based model of the cognitive system for representing agency (Leslie, 1994), and in particular, with the suggestion that the perception of persistent motion towards a target triggers the attribution of intent to an agent (Premack, 1990, Premack and Premack, 1994). In line with the Montgomery and Montgomery (1999) study, the paradigm tests the ability to suspend the perception of an agent's failed outcome for ascribing correctly the intended goal. By manipulating both the type of outcome, and the direction of the motion line towards the target, it is possible not only to measure the accuracy of the goal attribution, but also to explore whether subjects rely more on the outcome (end-state) of the goal-directed motion, or on the motion persistence when inferring the agent's intention. With reference to Leslie's agency model, the study investigates the ability in children with autism to "metarepresent" the actional properties of an agent (lower component of the metarepresentation model) as opposed to the ability to metarepresent propositional attitudes of an agent (higher component of the model). In particular, children are tested on the ability to suspend the representation of the agent's contingent situation (the final outcome of its goal-directed motion) and to "metarepresent" the agent's intended goal on the basis of a non-contingent state (the agent's persistent motion towards the target). The details of how the paradigm was developed are described in the next section.

3.3.3 Developing a new paradigm

The paradigm of the present study is based on a computer animated sequence depicting a small circle rolling up and down a valley trying to reach one of two targets resting on top of either sides of the valley (Figure B shows an original still frame from the beginning of the sequence).

Figure 3.2: Still frame of the computer animated sequence of the goal-directed motion test.



The context of a U-shape valley as opposed to a flat baseline with a wall (see paradigms by Gergely et al., 1995; Csibra et al., 1999; Montgomery and Montgomery, 1999, described above) was created in order to provide the observer with a naturalistic sequence. If a critical cue for triggering goal-attribution is the perception of a repeated movement towards a target, it is important that both the target and the obstacle be credible from the point of view of the agent. In the previous paradigms, the target was displayed beyond a wall, and the agent had to overcome that obstacle to reach the target. The problem with these types of stimuli is that they might induce the observer to wonder first of all whether the agent *sees* the target, and secondly, *why* the agent wants to reach the target beyond the wall. In the new paradigm the "agent's credibility problem was solved by creating a sequence depicting an agent trying to overcome the force of gravity in order to reach a visible target on the top of the valley. Another advantage of this display is that the motivation of the agent's goal-directed motion is somehow more obvious in the naturalistic context (e.g. the agent is stuck at the bottom of the valley and wants to get out by reaching the top) or at least, it does not require an explicit explanation of the motivation for the agent to reach the target (e.g., in the Montgomery and Montgomery's study children were informed that: "Charley [the agent circle] wants to go and see his friends. He wants to be right next to his friend. He is going to try and go to his friend"). Finally, the valley context made also explicit that the circle trying to overcome gravity was an agent whose movement was internally caused. The circle displayed also an animacy feature - a surface contraction as if "pulsing"– when resting at the bottom before launching in another attempt to reach the top. Note that this feature was added with the sole purpose of keeping constant the observer's attention on the circle across the whole sequence.

The small circle rolls repeatedly towards one of the two targets resting on each top of the valley, and at the end of its goal-directed motion one of three outcomes occurs: (1) it stops at the "Bottom" of the valley, or (2) it stops "Midway" between the bottom and the top, or (3) it stops at the "Top", having reached the target. The Top outcome displays the agent's intention matched with the outcome: it is a baseline condition to control for the representation of goal-attribution in the presence of a successful outcome. The Bottom outcome displays a clear failure to attain the goal, with the agent ending exactly where it started its motion. The Midway outcome is a more ambiguous condition: it is far from where the agent started its movement and it is close to the top, but not enough. This condition controls for the relevance of the perception of the agent's movement that improves in relation to the target. In fact, the circle's attempts improve in the sense that it finishes closer to the top (note that it reaches repeatedly a point above the midline of the valley).

Unlike all previous studies, this paradigm investigates specifically the effect of the motion pattern on the correct goal-attribution. In fact, the circle rolling up and down the valley was designed for two different types of movement, simple or complex. During the simple movement, the circle starts off from the bottom and rolls upwards in the direction of one target (either to the left or right side of the valley). There was no backwards momentum. The consistent direction towards the target position could provide the observer with a salient spatial cue for a correct attribution of its goal. In fact, the constant direction towards one target can be seen as a "pointing" sign indicating the target to be reached. For comparison with this spatially based cue, weak but still possible, a more complex motion pattern was used in a second condition. During the complex movement pattern, the circle starts its movement by jumping vertically, then rolling backwards and forwards as if gaining momentum to get to the top. Note that the motion properties of persistence towards the top and of improved effort were the same as in the simple movement pattern.

Another difference with other goal-directed motion paradigms is that the present paradigm allows for a wider range of responses, besides target and non-target. The subject can choose the agent's intended goal from among five possibilities: not only the target and non-target, but also the bottom of the valley, and the two symmetrical locations between each top and the bottom (see Methods for details). This widened range of possibilities allows for controlling the very premise of the paradigm, namely, that *a*) the perception of persistent motion towards a goal provides sufficient information to judge correctly the intention of an agent regardless of the outcome of its motion, and *b*) the perception of the outcome that is distant from the visible target induces the error of confusing the outcome with the intended goal. The choices available for the subject to decide the agent's intended goal include visible locations that are neither the potential targets, nor the failed outcome. It is therefore possible to test whether the perception of persistent movement and of the goal-directed outcome (either failed or attained) are sufficient but not necessary cues to attribute intention to a moving agent. In the hypothetical scenario of subjects making the consistent error of judging that the agent wanted to go to the side of the valley opposite to the persistent movement, and towards locations where the agent never stopped at, then the findings would indicate that the very premises of the present paradigm are false, namely, that the perception of persistent motion and outcome do not represent crucial perceptual cues for goal-attribution as suggested by Premack and Premack (1994).

Finally, in part II of the present experiment, another condition was added to test the influence of the final outcome and of the persistent motion in goal-attribution: the agent rolls repeatedly towards one target, and suddenly it changes direction and rolls all the way down and up towards the top on the opposite side, landing next to the stationary target. It is an open question whether subjects perceive this end-state of the movement as an intended outcome or an accidental outcome. If subjects choose as intended goal the target towards which the agent repeatedly attempted to reach, then the influence of the motion persistence is greater than the influence of the final outcome. By contrast, if subjects choose as intended goal the target that the agent eventually reached, then the representation of motion persistence is overridden by the contingent representation of the outcome of the movement.

The issue concerning the wording of the question presented to the subjects needs some clarification. In the Montgomery and Montgomery (1999) study children were asked the following question: "Was Charley trying to go to the blue circle or the green circle?" meaning: "was the agent trying to go to the target or to the non-target?" Unfortunately, this question could be misleading because of the presence of the verb "trying", which implies that sometimes Charley – the circle - did not succeed. Consequently the child is prompted to believe that there is the possibility of a failed outcome. One other concern is that the question was asked when the last frame displaying the final outcome was visible as a still image: the representation of the movement preceding the outcome could therefore have been outbalanced by that of the final outcome.

The question used in the present study was meant to be as neutral as possible by neither providing the subject with helpful verbal cues nor misleading spatial cues. The question: "where did the yellow circle want to go?" appeared on the screen at the end of the sequence and after a still image of the outcome had disappeared. The verb "wanting" was preferable to "trying" because there is no implicit reference to the success or failure of the attempt. Another concern relative to the use of mentalistic terminology was that the paradigm aimed at investigating intentionality at a level that does not require the representation of propositional attitudes. Although the term "wanting" refers to a mental state, it can be processed in different ways, not necessarily involving a propositional attitude. In fact, wanting implies desire and desire could be understood simply as a drive towards something, as a relation between the agent and a state of affairs. In this sense, children with autism should be able to represent a drive towards something, even if they do not represent propositional attitudes. In fact, Baron-Cohen, Campbell, Karmiloff-Smith et al. (1995) have shown that children with autism have more difficulties than controls to understand people's desires by detecting their gaze direction, but are more sensitive to non-human ostensive signals (e.g. a pointing arrow).

Children in the Montgomery and Montgomery study were presented with a pre-test to ensure that they were able to indicate that a person is trying to do something different from what really happens. Thus, prior to the experiment, children watched a video depicting a room with a table, a man walking in with a pile of books, and the man tripping over, dropping all the books on the floor. Following the video, the experimenter asked a few questions to ascertain that the child understood the event as "trying but failed action". Similar to the experimental question, this pre-test allowed children to familiarize themselves with the very condition that is under investigation, which could have biased their responses. In the present study the children were presented, after the experiment, with a test to ascertain whether they were able to understand questions relative to "wanting something" and "someone wanting". Children were shown different pictures depicting one standing woman reaching out for one of two objects placed on top of a shelf too high for her, and two people, only one of which was reaching out for an object on the shelf.

3.4 A study on autism and the attribution of intended-goal

Design: The experiment involves a 3 (Group) x 2 (Motion) x 2 (Motion Direction) x 5 (Outcome Location) design. The three groups consist of children with autism, normally developing children and adults. The Motion type independent variable consists of two levels: Simple Motion pattern (Simple), and Complex Motion pattern (Complex). In the Simple condition the circle rolls up and down the valley with a movement pattern aiming towards one of the two stationary goal circles resting on top of either side of the valley. The circle's trajectory is between the bottom of the valley and the goal circle at the top. In the Complex condition, the circle's movement is more ambiguous: the circle aims repeatedly towards one of the two stationary goal circles, but its trajectory includes vertical jumps and backwards rolling.

The Motion Direction independent variable consists of two levels: Constant and Changing. In the Constant condition, the direction of the motion of the circle towards its intended goal remains constant towards one of the two tops of the valley. The circle repeats three times its attempt to reach the top, and then it reaches a final end state on the same side of the valley where the three attempts occurred. In the Changing condition, the direction of the motion changes suddenly after the three repeated attempts. The circle rolls three times towards one of the tops of the valley, and then it rolls all the way down and up to the top at the opposite side of the valley, ending next to the other stationary goal circle.

The Outcome Location independent variable consists of four possible final outcomes that the circle reaches when it finally stops: 1) at the top of the valley next to the stationary goal circle, 2) midway the side of the slope between the bottom and the top of the valley, at a small distance from the goal, 3) at the bottom of the valley, at a far distance from the goal, 4) and at the top of the valley, next to the stationary circle at the opposite side of the top where the circle was repeatedly moving towards. In addition, there is the outcome location 5) at midway the side of the slope opposite to the agent's goal, where the agent never stops at (it is not an actual outcome). Despite the fact that the circle never stops midway on the side of the slope opposite to its goal, the condition was added to keep the symmetry of the outcome locations on either side of the valley. All the locations are marked with a platform (see Figure 3.1).

The experimental task consists in deciding about the final goal of the moving circle by clicking (with the computer mouse) on one of the five marked locations along the sides of the valley (note that there are four possible outcomes and five possible locations that subjects may choose to click on). The dependent variable is the type of response provided by the children at the end of the sequence when asked: "Where did the yellow circle want to go?" Responses were coded according to the five locations the subject clicks on: "Top-Same side", "Midway", "Bottom", "Midway-Opposite side", and "Top-Opposite side".

The study has two distinct aims: a) exploring the ability in children with autism to attribute an agent's intended goal in the presence of its unsuccessful outcome, b) exploring whether the attribution of intended goal is overall more influenced by the agent's proximity to the target (outcome-based representation) or by the perceived motion of the agent towards the target (motion-based representation, regardless of outcome), and whether the performance of children with autism reflects a developmental delay in their ability to represent intended-goal on the basis of outcome or persistent motion.

The study is therefore divided in two parts. Part I investigates subjects' responses when the direction of the agent's goal-directed motion is constant towards a top of the valley. In this condition (henceforth "Constant Direction") the correct response is always that the circle wanted to reach the top of the valley towards which it kept rolling (Top-Same side). Part II investigates subjects' responses when the direction of the agent's goal-directed motion towards a top of the valley suddenly changes, and the circle ends at the top of the opposite side of the valley (Top-Opposite side). In this condition (henceforth “Changing Direction”) the correct response is either that the circle wanted to go to the top where it was originally rolling towards (Top-Same side), or to the opposite top where it eventually stopped, having changed direction during the last attempt (Top-Opposite side).

In sum, the study is based on a single test depicting a circle that, after rolling up and down a valley trying to reach a target resting on top of either side of the valley, reaches different final outcomes. The analysis of the data is split into two parts. The Constant Direction analysis (Part I) is based on subjects’ responses when the circle reached three actual outcomes (one successful and two unsuccessful). Note that there are five possible responses the subjects can give, and only one correct response. The Changing Direction analysis (Part I) is based on subjects’ responses when the circle reached one actual outcome (Top-Opposite side): there are five possible responses, but only two are correct.

Subjects: A group of 26 children with autism resident in special schools were tested. Their mean chronological age was 12.2 years, with a range between 9.11 and 17.5 years. Their mean verbal age (VMA) was assessed using the BPVS II test (British Picture Vocabulary Scale, 1997) on 14 subjects, and the VIQ score of the WISC (Wechsler Intelligence Scale for children, third edition UK, 1992) on the remaining 12 children. The choice between the IQ tests was entirely determined by the time available for testing each child. The largest group consisted of 50 normally developing children and teenagers attending mainstream schools. Their mean chronological age was 8.11 (sd=3.5) years, with a range between 5.6 and 15.9. More specifically, the group of “under-six” children consisted of 17 subjects with a mean age of 5.2 years (sd=0.4), ranging from 5.6 to 6.2. The group of “young” children consisted in 21 subjects with a mean age of 9.5 years (sd=1.0), ranging from 8 to 11.9. The group of “teenagers” consisted of 12 subjects with a mean age of 14.2 (sd=4.5), ranging from 13.4 to 15.9. The VMA of all normal children was assessed with BPVS II test (4 teenagers were not available for this test). The autism group and the control group of all children were matched on VMA but not on CA or IQ score.

A group of 18 adults, with a chronological age of 31.3 years, ranging from 19 to 40 years. Their VMA was not assessed. All subjects passed the "wanting test" (see Materials) aimed at ascertaining whether they were able to understand questions relative to "wanting something" and "someone wanting" in the absence of verbal stimuli.

Table 3.1: Subjects chronological (CA) and verbal mental age (VMA). The mean and standard deviation (in parentheses) are reported in calendar years and months.

Group	CA (yrs) *mean (sd)*	VMA (yrs) *mean (sd)*
Autism – Children (N=26)	**12.2** (2.0)	**9.7** (3.0)
Control – Children (N=50)	**9.1** (3.6)	**9.7** (3.10)
Control – Adults (N=18)	**31.3** (6.3)	-

Materials: Sixteen goal-directed movie sequences were created with Micromedia Director4™ and presented on a portable computer. All the sequences featured one character, a small yellow circle, rolling up and down a green valley trying to reach one of two circles resting on top of either sides of the valley. The circle's attempts are either successful or unsuccessful: it reaches the top (stopping next to the circle), or rolls down to the bottom of the valley, or midway the side of the valley between the bottom and the top. Each sequence lasts 13 seconds. When the circle reaches its final location, it stays still for 2 seconds, then disappears and a question "Where did the yellow circle want to go?" appears on the top of the screen. The image disappears as soon as the child clicks with the mouse on one of the five possible locations, to be replaced by an interval image – a light blue screen and a message: "Get ready! Click to continue". Every four trials the message was: "Very good! Get ready for level 1. Click to continue". The same message appeared for level 1, 2, 3 and 4.

There are in total 16 movie sequences divided into 4 blocks (Simple and Complex motion pattern for each goal circle: Red and Blue) with 4 trials each (the goal-directed outcome: Top-Same, Midway, Bottom, Top Opposite). Total duration: 13.5 sec frames in motion plus 2 sec still frame, after which the experimental question appears on the screen, and disappears when the subjects clicks on one of the marks.

The material for the "wanting test" consisted in four different photographs (see Appendix 3A). Two of them depicted a standing woman in front of a shelf too high for her, on which there were a teapot and a clock. In one picture she is reaching out for the teapot, in the other for the clock. The experimenter asked: "What does she want? The teapot or the clock?" The other two photographs depicted a woman and a man standing in front of a high shelf with a clock on it. In one picture it is the man reaching out for the object, in the other it is the woman. The experimenter asked: "Who wants the clock? The man or the woman?"

Procedure: Each subject was tested individually in a separate room of the school. In the training phase the experimenter explained to the child that he was going to do a very simple computer game in which he /she had to decide where a yellow circle rolling up and down a valley wanted to go. This simple explanation was followed immediately by showing an example of one sequence (the practice sequence was the Simple motion, Blue goal-circle, bottom outcome). If the subject answered incorrectly, the experimenter showed the practice trial again. No feedback on the performance was given. After the example, the experimenter read aloud the following instructions (displayed on the computer screen): "You are going to see several animations showing a circle rolling up and down a valley. When the circle stops a question will appear on top of the screen. You have to click on the spot where the circle wanted to go. Thank you for your participation". The subject was then motivated in watching with attention all the sequences by saying that the game had four levels of difficulty. Although the speed of the circle was the same in all animations, the experimenter explained that the circle rolls up and down the valley faster and faster so that the "player" had to be very careful in keeping watch of its movement. The subject was also encouraged to stop anytime he/she wanted between trials. The whole experiment lasted an average of 15 minutes.

Each subject was presented with a total of 32 movies (8 with simple motion and 8 with complex motion, each repeated twice) displayed in quasi-random order. There were 4 different sequences with the following order (B=bottom, T=top-same, M=Midway, O=top-opposite): BTMOOMTB, TGOMMOGT, OMTBBTMO, and MOBTTBOM. The order of presentation of the sequences was counterbalanced across subjects. At the end of the experiment the "wanting test" was administered.

3.4.1 Part I: Intended goal-attribution in presence of constant goal-directed movement

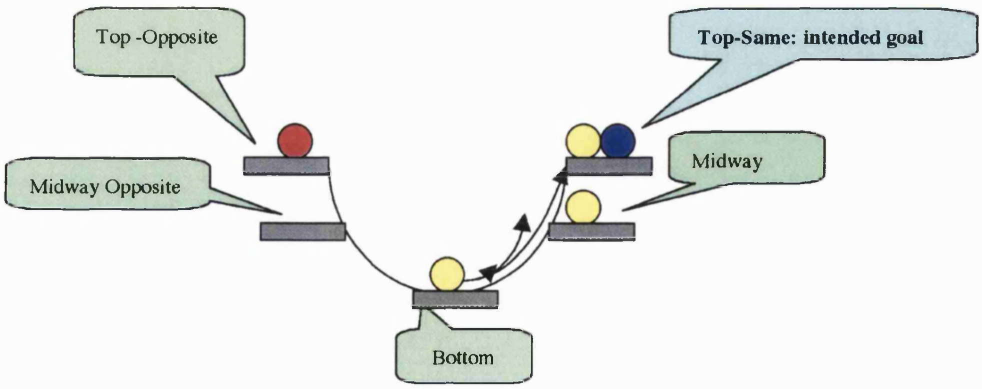
The first part of the study analysed subjects' responses when the goal-directed motion of the agent was constant towards one of the tops of the valley, and its outcome is either successful (the agent reaches the top) or unsuccessful (the agent ends at the bottom of the valley or midway on the slope). Figure 3.3 shows the conditions in a schematic way. The correct response was always that the circle wanted to reach the top of the valley towards which it kept rolling (Top-Same side) regardless of where the circle came to rest. The aim was therefore to explore whether children with autism rely more on the agent's contingent situation (unsuccessful outcome), or on the motion persistence of the agent toward a target when inferring the agent's intention.

After the trial in which the outcome was either Top-Same, Midway or Bottom subjects were asked the following question: "Where did the circle want to go?" They answered by clicking with the cursor of the mouse on one of the five marked locations in the valley, namely, Top-Same, Midway, Bottom, Midway-Opposite, and Top-Opposite.

The Top-Same condition constitutes the least difficult condition, that is, when the agent attained its intended goal. Correct responses to Top-Same condition were expected to be higher than the other four conditions. By contrast, the Midway-Opposite location is the condition that was expected to be chosen less frequently than the others, since the agent's outcome never occurs in this location. The prediction relative to the type of agent's goal-directed motion pattern was that the Complex one should elicit fewer correct responses than the Simple motion pattern.

The minimum requirement for a valid performance was answering correctly to the baseline condition, namely, when the agent attained its goal in the simple motion pattern (the circle ends at Top-Same next to the goal-circle) in at least one out of two trials. Consequently, seven subjects were eliminated from the analysis: one from the autism group, four from the control group of children (2 teenagers, 1 young, 1 under-six), and two from the control group of adults. These subjects were considered either not to have understood the test, or not to comply with the instructions.

Figure 3.3: The agent's goal-directed motion of the "Constant Direction condition in a schematic way. The yellow circle rolls towards one of the top of the valley (Top-Same side) three times, and then it reaches one of three Outcomes: Bottom, Midway, or Top-Same. Subjects can choose to click on each of the five gray platform in response to the question: "Where did the yellow circle want to go?". In this example, the correct answer is to click on the platform underneath the blue circle, where the yellow circle's motion was constantly directed, regardless of where it came to rest.



Results:

i) Analysis of correct performance in the Constant Direction condition:

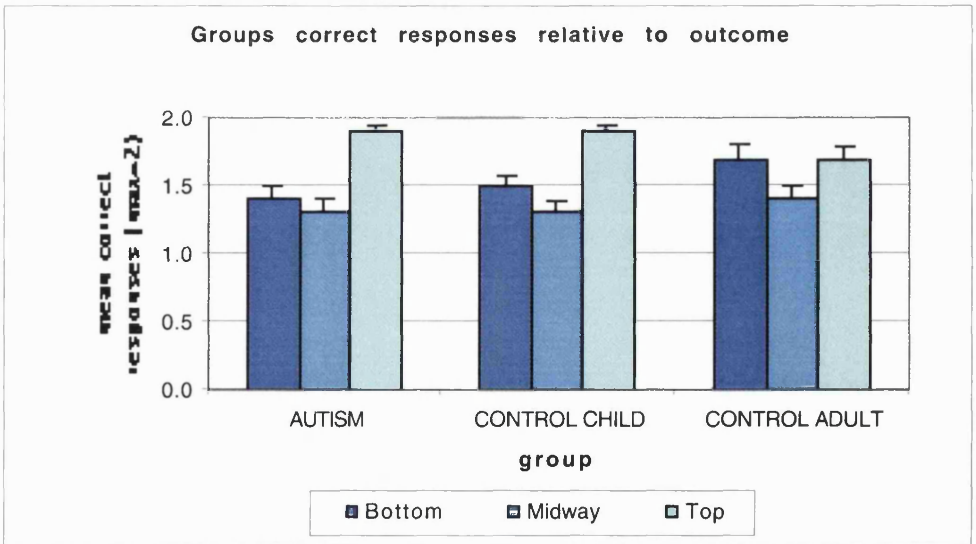
A parametric analysis was performed on the correct responses (Top-Same location) regardless of the final outcome of the agent goal-directed motion. Table 3.2 shows the mean and standard deviation of the correct score relative to each outcome (agent fails its attempt when it stops at Bottom or Midway; agent reaches its goal when it ends at Top-Same), and Figure 3.4 shows the same data in graphic form. A repeated measure ANOVA 3 (Group) x 3 (Outcome) x 2 (Motion) revealed a significant main effect of Outcome ($F_{(2,84)}$=18.2, p<.0001), a significant trend of Motion pattern ($F_{(1,84)}$ = 3.2, p=.06), i.e., Simple motion (mean=1, sd.=.09) and Complex motion (mean=0.96, sd.=.09). The analysis revealed no effect of Group ($F_{(2,84)}$ = 0.1, p=not sig.) . Post-hoc analysis revealed that, as predicted, all subjects gave more correct responses (planned comparison, F=30.9. p=.0001) when the agent's goal-directed motion was successful (Top-Same outcome) rather than when it was unsuccessful (Bottom or Midway outcomes).

No significant interactions were revealed: Group x Outcome ($F_{(4,84)}$=1.2, p= not sig.), Group x Motion ($F_{(4,84)}$=0.4, p= not sig.). These results did not show any impairment in children with autism in the attribution of intention on the basis of perceived goal-directed motion.

Table 3.2: Groups's correct responses split into type of motion pattern and final outcome of the circle's goal-directed motion. The type of motion pattern had no effect on subjects' correct score. Mean (max score=2) and standard deviation (in parentheses).

	Correct responses					
	When the agent fails to reach the target and stops at:				**When the agent reaches the target at:**	
	Bottom		**Midway**		**Top-Same**	
	Motion pattern:		*Motion pattern:*		*Motion pattern:*	
Group	Simple	Complex	Simple	Complex	Simple	Complex
Autism (N= 25)	**1.4** (0.8)	**1.4** (0.7)	**1.4** (0.9)	**1.2** (1.0)	**1.9** (0.3)	**1.9** (0.3)
Control – Children (N= 46)	**1.6** (0.7)	**1.4** (0.8)	**1.4** (0.8)	**1.3** (0.9)	**1.9** (0.3)	**1.8** (0.5)
Control – Adults (N=16)	**1.8** (0.6)	**1.6** (0.6)	**1.5** (0.6)	**1.4** (0.9)	**1.7** (0.4)	**1.8** (0.6)

Figure 3.4: Correct responses of each group relative to the final outcome of the goal-directed motion (regardless of the motion pattern condition). All groups performed equally well in both the successful and unsuccessful outcomes. Mean (max score=2) and standard error.



ii) Analysis of error pattern in the Constant Direction condition:

Subjects made errors (22.5% of the total responses) when, instead of selecting the correct location (Top-Same), they selected one of the following locations: Midway, Bottom, Midway-Opposite, and Top-Opposite. The data were analysed with nonparametric tests since the choice between the four possible errors is reciprocally exclusive, and did not meet the parametric analysis requirement of non-sphericity.

The following hypotheses were investigated:

a) Since the choice of the subjects consisted of five possible locations (Bottom, Midway, Top-Same, Top-opposite, Midway-Opposite) it was predicted that subjects made more frequent errors clicking on one of the two unsuccessful outcome locations (Bottom, Midway) than on the two outcomes that never occurred (Top-Opposite, Midway-Opposite).

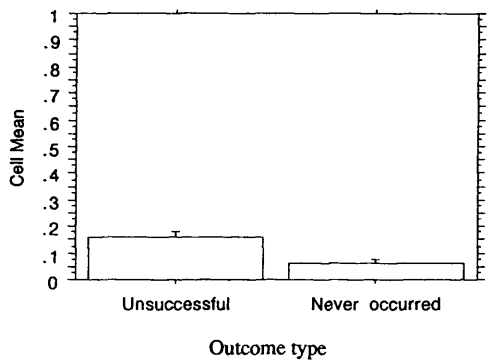
b) Since the paradigm tested the ability to override the contingent spatial representation of the agent's outcome in order to correctly represent the agent's intention, it was predicted that the most common error was to select the locations where the circle actually ended (select Bottom when the circle ended at Bottom, and on Midway when the circle ended at Midway).

c) Since the motion pattern of the agent towards the goal was manipulated in order to control for the visual cue a trajectory constantly "pointing towards" the goal, it was predicted that the Complex motion pattern (the circle rolls forwards, jumps and rolls backwards) would induce more errors than the Simple motion pattern (the circle rolls forwards).

The following results were obtained for each hypothesis:

a) The effect of the outcome type (Bottom, Midway, Midway-Opposite, Top-Opposite) on the subjects' errors was analysed with a nonparametric two-way analysis of variance (Friedman test). Results revealed a difference in the number of errors made by clicking on the four different outcomes (chi-square=11.5 $p<.01$). Nonparametric paired comparisons (Wilcoxon test) revealed that subjects clicked with equal frequency on the Bottom and the Midway locations (z=.04, p=not sig.), and on the Midway-opposite and Top-opposite (z=.4, p= not sig.). By combining the errors of clicking on the two (unsuccessful) outcomes, and the errors relative to the two (never-occurred) outcomes at opposite side of the agent's goal-directed motion (Figure 3.5), it was revealed, as predicted, that the more frequent errors were made with both Bottom and Midway outcomes rather than with Midway-Opposite and Top-Opposite outcomes (z=3, p<.003).

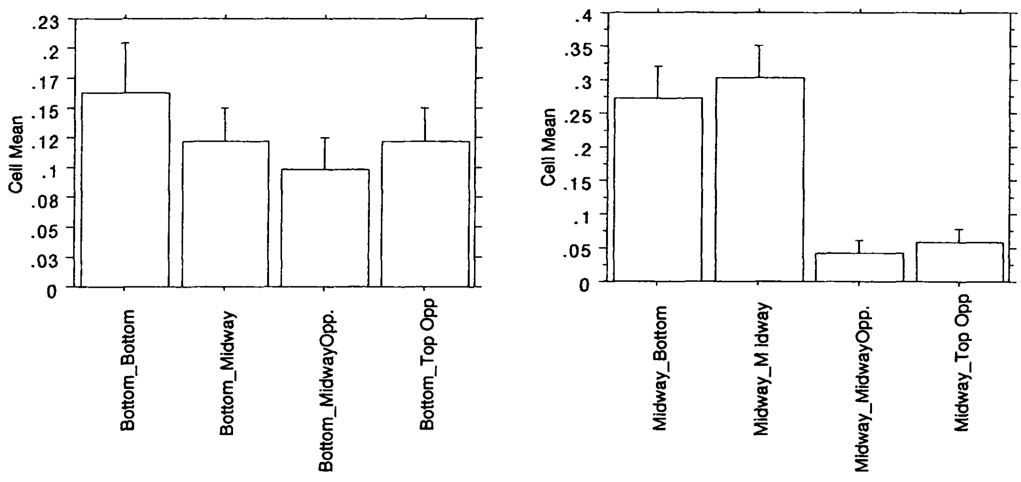
Figure 3.5: Total errors made by clicking on Bottom and Midway of the valley (**Unsuccessful Outcomes**) and on Midway-Opposite and Top-Opposite (**Never-occurred Outcomes**). Mean (max score=2), and standard error.



b) The effect of the type of error relative to the type of outcome was analysed with a nonparametric two-way analysis of variance (Friedman test) conducted on each of the two unsuccessful outcomes, Bottom (Figure 3.6a) and Midway (Figure 3.6b). Results revealed that the errors made when the circle ended at Midway of the valley were of a different type (chi-square=15 p<.002), but not when the circle ended at the Bottom of the valley (chi-square=.4 p= not sig.). Nonparametric paired comparisons (Wilcoxon test) were performed on the Midway outcome errors (circle ends at Midway, subjects click on either Midway, Bottom, Midway-Opposite, or Top-Opposite) testing the hypothesis that the error of clicking on Midway was more frequent than the other errors. The analyses revealed that subjects made more the error by clicking on Midway (mean=.3, sd=.4) than on Midway-Opposite (mean=.04, sd=.2) (z=4, p<.0001), or than Top-Opposite (mean=.06, sd=.2) (z=4, p<.0001). However, contrary to prediction, there was no difference between the errors made by clicking Midway or Bottom (mean=.3, sd=.4).

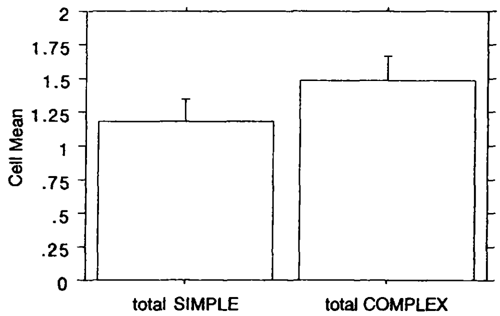
Nonparametric group-comparison tests (Kruskal-Wallis) was conducted to analyze whether children with autism made more errors clicking on Bottom when the circle ended at the Bottom, and on Midway when the circle ended on Midway. Results revealed no significant effect when the circle ended at Bottom, subjects in all groups selected equally Bottom (H=.8, p=not sig.), or Midway (H=.6, p=not sig.), or Midway-Opposite (H=.1, p=not sig.) or Top-Opposite (H=.5, p=not sig.). There was no significant difference relative to the type of errors made by each group when the circle ended at Midway by clicking on Midway (H=.3, p=not sig.), or Bottom (H=1.3, p=not sig.), or Midway-Opposite (H=.2, p=not sig.) or Top-Opposite (H=.1, p=not sig.).

Figure 3.6a,b: Type of errors made when the circle ended at the Bottom of the valley (graph on the left), and Midway between the top and the bottom of the valley (graph on the right). Note that, for example, "Bottom-Midway" means that the circle ended at Bottom and subjects clicked on Midway. Mean (max score=2), and standard error.



c) A non parametric paired comparison (Wilcoxon test) revealed that the errors relative to the Simple motion pattern were less frequent than those of the Complex motion (z=2.2, p=.03), in line with the prediction (see Figure 3.7)

Figure 3.7: Errors made when the circle's motion trajectory includes only forwards rolling (Simple motion pattern), and when the trajectory includes also vertical jumps and backwards rolling (Complex motion pattern). Mean (max score=2), and standard error.



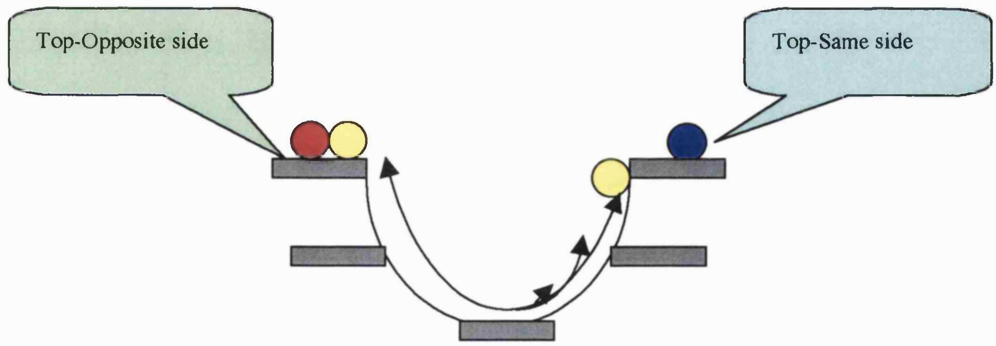
Taken together, these results indicated that, as predicted, subjects made more errors by choosing the agent's unsuccessful outcomes locations rather than the locations, on which it never stopped, that were at the opposite side of the valley where the circle made its attempts to reach the top. As predicted, subjects made more errors when the circle's motion pattern was more complex than simple. However, contrary to prediction, there were no increased number of errors relative to the mistake of associating the actual outcome of the circle with its intended goal (clicking on Bottom when the circle ended at Bottom, and on Midway when the circle ended at Midway) and no difference in the performance of the autism group.

3.4.2 Part II: Intended goal-attribution relative to the direction of the agent's motion

The second part of the study analyses children with autism's performance and the developmental changes in normal controls in the "changing direction condition of the agent's goal-directed motion. The aim of this analysis was to explore the ability of attributing an intended goal to an agent in the presence of an outcome that is on the opposite side of where the goal-directed motion originally "pointed-towards .

The data analyzed are about subjects' responses when the repetitive motion of the agent towards one of the tops of the valley (Top-Same side) suddenly changes direction, rolling all the way down and up towards the top of the opposite side of the valley (Top-Opposite side), landing next to the other stationary target (see Figure 3.8).

Figure 3.8: The agent goal-directed motion of the "Changing Direction condition in a schematic way. The yellow circle rolls on one side of the valley towards the top (Top-Same) three times, and then it suddenly changes direction rolling all the way down and up to the other side of the valley (Top-Opposite side). There are five platforms that subjects can choose to click in response to the question: "Where did the yellow circle want to go?". Only two responses are correct (Top-Same and Top-Opposite). The type of response reflects a different type of intended goal representation (based on final outcome or on persistent motion).



It is an open question whether subjects perceive this end-state of the movement as an intended outcome or an accidental outcome. If subjects choose as intended goal the target towards which the agent repeatedly attempted to reach, then the influence of the motion persistence is greater than the influence of the final outcome. By contrast, if subjects choose as intended goal the target that the agent eventually reached, then the representation of motion persistence is overridden by the contingent representation of the outcome of the movement.

Thus, the questions to be investigated are (1) whether the perception of constant motion towards a target is more salient than the perception of the final outcome (the agent next to a stationary object), (2) whether there is a developmental change in the type of intended goal representation (based on final outcome or on persistent motion), and (3) whether the performance of children with autism differs from that of controls and reflects a developmental delay in their ability to infer intention on the basis of persistent motion or final outcome.

In order to investigate developmental changes across the different age-groups, it was necessary to increase the number of subjects in the Adult group. Table 3.3 shows the chronological age and verbal mental age of all subjects who passed the baseline of part I of the study, divided in five groups. The minimum requirement for a valid performance was to answer either Top-Same or Top-Opposite in the simple motion pattern, in two out of two trials. Consequently, seven subjects were eliminated from the analysis of correct responses: three from the autism group, one from the "under-six" normally developing children, two from the "young" normally developing children, and one from the group of adults. These subjects were considered either not to have understood the test, or not to comply with the instructions.

Table 3.3: Subjects'chronological (CA) and verbal mental age (VMA). The mean and standard deviation (in parentheses) are given in calendar years and months.

Group	CA (yrs) *Mean (sd)*	VMA (yrs) *Mean (sd)*
Autism – children (N = 25)	**12.2** (2.0)	**9.7** (3.0)
Control - "under-six" (N = 16)	**5.2** (0.4)	**6.1** (1.0)
Control - "young" (N = 20)	**9.5** (1.0)	**10.1** (2.6)
Control - "teenager" (N = 10)	**14.2** (4.5)	**15.9** (5.2)
Control – Adult (N = 28)	**30.0** (5.9)	-

Results:

i) Analysis of the type of responses in the Changing Direction condition:

A parametric analysis was performed on subjects' correct responses (Top-Same and Top-Opposite). Table 3.4 shows mean score and standard deviation of each group performance split into Simple and Complex motion pattern, and Figure 3.9 shows in a graphic form the types of response of each group regarless of the motion pattern.

A repeated measure ANOVA 5 (Group) x 2 (Response type) x 2 (Motion pattern) revealed no significant main effects of Group ($F_{(4,87)}$=0.3, p= not sig.), Motion pattern ($F_{(1,87)}$=3.6, p= not sig.), or Response type ($F_{(1,87)}$=3.5, p= not sig.). However, a significant interaction Group x Response type ($F_{(1,91)}$=6.3, p<.001) was revealed. Table 3.5 shows the type of response of each group.

Table 3.4: Subjects' responses in the Changing Direction condition split into Simple and Complex motion pattern. The type of motion pattern had no effect on subjects' correct score. Mean (max score = 2), and standard deviation, in parentheses.

	Subjects' response type			
	Top-Same		**Top-Opposite**	
	Motion pattern:		*Motion pattern:*	
Group	**Simple**	**Complex**	**Simple**	**Complex**
Autism – children (N=22)	**0.7** (0.8)	**0.5** (0.8)	**1.3** (0.8)	**1.5** (0.9)
Under-six – control children (N=15)	**0.5** (0.6)	**0.2** (0.4)	**1.5** (0.6)	**0.7** (0.5)
Young – control children (N=18)	**0.8** (0.9)	**0.7** (0.9)	**1.2** (0.9)	**1.2** (0.9)
Teenager – control children (N=10)	**1.1** (1.0)	**1.2** (0.9)	**0.9** (1.0)	**0.8** (0.9)
Adult – control (N=27)	**1.4** (0.9)	**1.2** (0.8)	**0.6** (0.9)	**0.7** (0.7)

Table 3.5: Subjects' responses in the Changing Direction condition regardless of the motion pattern of the agent's goal-directed motion. Mean (max score = 2) and standard deviation (in parentheses).

	Subjects' response type *(regardless of motion pattern)*	
Group	**Top-Same**	**Top-Opposite**
Autism – children (N=22)	**0.6** (0.8)	**1.4** (0.8)**
Under-six – control children (N=15)	**0.4** (0.6)	**1.6** (0.6)***
Young - control children (N=18)	**0.8** (0.9)	**1.2** (0.9)
Teenager - control children (N=10)	**1.2** (0.9)	**0.9** (0.9)
Adult – control (N=27)	**1.3** (0.8)*	**0.6** (0.8)

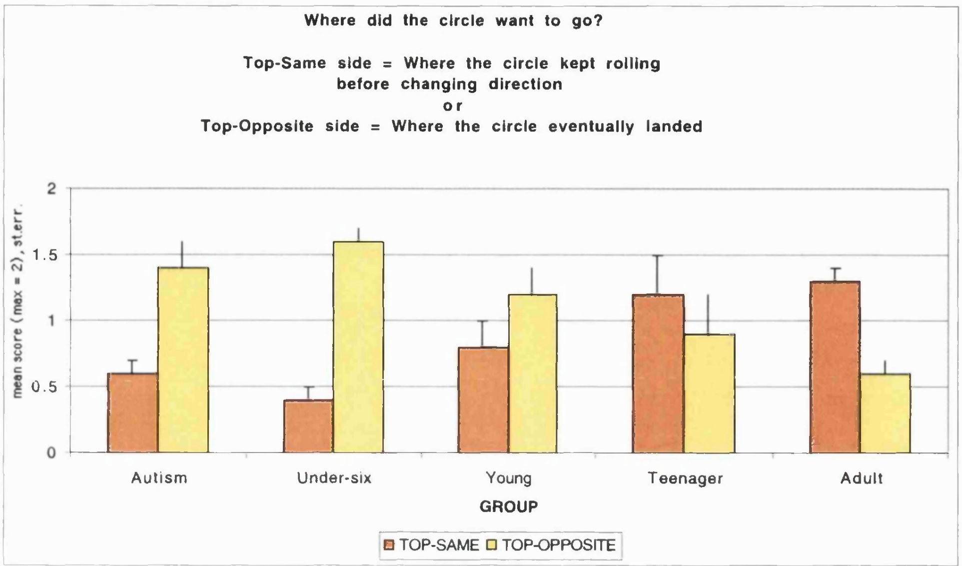
Note: significant level of result of paired t-tests *=p<.01; **=p<.001; ***p=<.0001

A post-hoc analysis (paired t-test) indicated that the group of children with autism (CA=12.2 years; VMA=10.1 years) and the group of normally developing children "under-six" (CA=5.2 years; VMA = 6.1 years) responded significantly more Top-Opposite than Top-Same (autism group: mean diff.=0.8, t=2.6, p= .02; under-six group: mean diff.=1.2, t=6.2, p< .001), and the group of normal adults responded more Top-Same than Top-Opposite (mean diff.=0.7, t=2.8, p< .01)(see Figure 3.7). Unpaired t-tests indicated that the Top-Same responses of the adult group were significantly higher than the autism group (mean diff.=.7, t=3.8, p< .001), the under-six (mean diff.=.97, t=5, p< .0001) and the young group (mean diff.=.58, t=5, p= .01). The responses of the teenager group did not differ either from those of the adult group (mean diff.=.18, t=.6, p= not sig.) nor from the young group (mean diff.=.4, t=1.2, p= not sig.). However, it did differ significantly from responses of the under-six group (mean diff.=.8, t=2.9, p<.01.) and from those of the autism group (mean diff.=.56, t=1.9, p= 0.6 sig. trend).

Unpaired t-tests indicated that the Top-Opposite responses of the autism group did not differ from either the young group (mean diff.=.14, t=5.8, p=not sig.) and the under-six group (mean diff.=.23, t=1.1, p=not sig.). The responses of the teenager group's did not differ from either the young group (mean diff.=.24, t=.9, p=not sig.) and the adult group (mean diff.=.37, t=1.1, p=not sig.). However, the adults' score was significantly lower than that of the young groups (mean diff.=.6, t=2.7, p<.01).

These results support the hypothesis of a developmental change across age groups in the representation of intended-goal, and the hypothesis of a developmental delay in children with autism. The teenager group's responses were intermediate: they did not differ significantly from either the adults or the young group, but the responses of the young and the adult groups were significantly different.

Figure 3.9: Subjects' type of response in the presence of a sudden change in the direction of the agent's goal motion towards a target. The adult group responded that the circle wanted to go to the Top-Same, whereas the autism group and the under-six group responded that it wanted to go to Top-Opposite.



3.4.3 Discussion

The present study sought to investigate the ability in children with autism, normally developing children, and adults to attribute an intended goal to an agent by using simple visual stimuli in motion. The paradigm was based on the perception of an agent's repeated motion (failing and trying again) towards a target, and its final outcome (goal-attained or goal-missed). The first aim of the study was to explore the ability of children with autism to suspend the perception of the agent's failed outcome to reach a target in order to ascribe correctly the intended goal, relying only on the perception of the repetitive motion pattern directed towards a target. The second aim was twofold. First, the relation between the perception of the agent's outcome and of its repeated goal-directed motion towards a target was explored, and consequently the performance of children with autism was investigated to establish whether it would reflect a developmental delay in their ability to infer intended goal on the basis of either persistent motion or final outcome.

With regards to the first aim of the study, the results showed no impairment in children with autism to infer an agent's intended goal on the basis of the perception of its goal-directed motion towards a target. This was shown when they were required to override the visual representation of the agent's unsuccessful outcome after a repetitive motion towards the target. Furthermore, they did not show any difficulty when the motion pattern of the agent towards the target consisted of a complex trajectory including vertical jumps and backward movement rather than a simple linear trajectory aiming at the target. The analysis of the error pattern also revealed no difference across groups. All subjects had more difficulty when the circle's motion pattern was complex rather than simple, and when the circle failed in its attempts by ending at the location between its starting point and the target. Unsurprisingly, subjects did not make the error of choosing as intentional outcome of the agent the two locations it never actually reached (the Midway-opposite and Top-Opposite, placed outside the goal-directed motion trajectory). Surprisingly, children with autism, normally developing children and adults showed an indiscriminate pattern of errors. More specifically, they did not show a particular preference for one of three different types of errors: *a*) choosing as intentional outcome the location where the agent accidentally stopped, *b*) choosing as intentional outcome the location that was further away from the bottom of the valley where the agent started its motion, and *c*) choosing as intentional outcome the location whence the agent started its motion.

There are several points to be made about errors. First, the types of errors made by all participants indicate that the perception of an agent's persistent motion and of the goal-directed outcome (either failed or attained) represent the most salient cues for goal attribution, as suggested by Premack and Premack (1994). In fact, regardless of the wide range of possible answers available (five locations spread along the valley) subjects were not confused by the locations that were neither potential targets nor failed outcomes. Secondly, there is no indication that subjects applied consistently a wrong strategy to infer the agent's intended goal. When subjects made an error, they were generally confused and chose randomly between the locations at the Bottom and Midway, rather than applying a single fallacious reasoning, e.g. "the circle intentionally ends its motion at the bottom of the valley, hence the circle wanted to go there". Children with autism, might have been expected to fall into the mistake of matching the accidental outcome with the intentional outcome. However, on the contrary, they also made, like the controls, two other types of mistakes, e.g. "the circle stops accidentally at the bottom, hence the circle wanted to go midway up the valley", and "the circle accidentally stops midway its trajectory, hence the circle wanted to go to the bottom". The last point concerns the fact that the Midway outcome location induced subjects to make errors more often. This was expected, since this location was indeed designed to be an ambiguous cue. However, children with autism did not perceive it differently from the control children and control adults.

The negative finding of the present study is in contrast with the study by Phillips et al. (1998), which indicated a specific impairment in autism in understanding own intention in the presence of an unsuccessful outcome. Some differences between Phillips et al.'s and the present study have to be taken into account. First, there was a difference in the children's age. In Phillips et al.'s study children with autism had a lower verbal mental age (CA mean: 13.4 yrs, sd.: 3 yrs, VMA: 6.2 yrs, sd.: 2.3 yrs) than those in the present study (CA mean: 12.2 yrs, sd.: 2 yrs, VMA: 9.7 yrs, sd.: 3 yrs). It could be that younger children with autism are less familiar with the possibility that own intentions, while playing a game, might result in failure. However, Montgomery and Montgomery (1999) have shown that normally developing children as young as three-year-old, are able to distinguish an intended outcome from an accidental, failed outcome.

A second difference between the Phillips et al.'s and the present study is in the type of paradigm. The Phillips et al.'s study investigated the ability to self attribute an original intention in conjunction with the experience of both a fulfilled or unfulfilled desire, and of a successful and unsuccessful outcome. Findings indicated that children with autism were able to ascribe to themselves the intention to hit a target regardless whether their desire to win a prize was satisfied or not. However, when they missed the target, they failed to understand that their own intention remained the same despite the outcome being a failure, and attributed to themselves the intention of hitting the accidental target rather than the target they missed. As discussed earlier (paragraph 3.3.1) it is yet to be established whether children with autism failed the task because of impairment in inhibiting a wrong prepotent response, or because they were unable to understand their own intention.

An alternative explanation for the contrasting results of the Phillips et al.'s and the present one is that the former used a much more a demanding task. First of all, it must be noted that the Theory of Mind hypothesis that individuals with autism have an impairment in the ability to represent mental states does not predict any specific difference between the ability to attribute mental states to oneself and to others. Thus, the requirement of self-attribution of intention in the "hitting the target" task as opposed to the third-person attribution of intention of the "agent's goal-directed motion" task cannot account for the higher level of task demands of the Phillips et al.'s study. Thus, what could explain the difference between the two studies?

The purpose of creating the paradigm for the present study was to investigate the ability to represent an agent as acting to bring about a future state of affairs, by using simple visual stimuli in motion. The premise for creating this test was that one of the simplest forms of visual information for judging agent's intentions comes from observing the agent's motion trajectory (Michotte, 1946/1963; Heider and Simmel, 1944). In addition, the perception of a persistent motion towards a target, regardless of the outcome, has been shown to elicit accurate judgements of the intended goal of the agent (Montgomery and Montgomery, 1999). Contrary to Phillips et al.'s study, subjects were shown "on-line", with powerful visual cues, a constant direct motion of the agent towards a target, and the agent's accidental outcome. Subjects were required to disregard the visual cue of the accidental outcome, and to infer the agent's intended goal from the agent's persistent and constantly directed motion. In Phillips et al.'s study, subjects were also required to disregard the accidental outcome, but the only cue of their previously declared intended goal, on which they could base their inference, was distant in time and place. In this sense, the task of Phillips et al.'s study was more demanding than the present study, for it was based on distinguishing two discordant representations with no support of a salient visual cue indicating the intended goal. This increased difficulty might account for the failure of children with autism in understanding that their own intention remained the same regardless the fact that the intention was not attained.

Finally, the discussion turns to the theoretical distinction between different levels in the ability to represent agency suggested by Leslie (1994). According to the tripartite theory of agency, there are two distinct levels for representing an agent beyond its mechanical properties, namely, one for its actional properties, and the other for its mental properties. At a "lower" cognitive level, there might be a system concerned with the pursuit of a future state of affairs (goal-directed action) and, at a "higher" cognitive level, there might be a system concerned with states that are "about something", namely, propositional states (e.g. "believing that", "hoping that"). The present study investigated whether children with autism are able to represent intentionality at a lower level, namely, whether they are able to represent an agent as acting to bring about a state of affairs, as opposed to the representation of an agent as actively holding an attitude to the truth of a proposition. The test of the present paradigm taps the ability to represent possible or future states of affairs in order to understand the agent's goal-directed behaviour without necessarily involving the cognitive process of representing the agent's mental state.

The negative findings of the present study indicate that the "lower" level of the intentional representation system is unimpaired in autism. On the other hand, more than a decade of research on autism has shown that the development of the ability to represent mental states at a "higher" level is compromised in individuals with autism. These findings are of particular interest if their meaning is translated into the context of everyday life of people with autism with high functioning intellectual abilities (e.g. Asperger Syndrome). In fact, their clinical profile indicates that whereas they are able to live independently, have a job and study, they have nevertheless difficulties in understanding other people's thoughts and complex human interactions. The distinction between the ability to represent agency at an "actional" level and at an "attitudinal" level might account for the intact ability to deal with people's actions, and the impairment in dealing with people's thoughts. Supposing, for argument's sake, that individuals with autism lack the ability to understand that agents act in relation to specific goals, then the consequence would be that any single intentional action would appear meaningless to them, to the extent that they would find it difficult to understand why living creatures move about. The value of this hypothetical scenario is only to make clear the consequences of the distinction, suggested by Leslie (1994), between two different levels of metarepresentation abilities, that is, the ability to represent non-contingent events causing the agent's goal-directed behaviour and non-contingent events causing the agent's propositional attitude.

Part II of the experiment, which included a more ambiguous scenario, did show a developmental delay in children with autism in the representation of an agent's intended goal. No previous studies have investigated the nature of goal-directed representation in different age groups. The paradigm, which was based on both the manipulation of the outcome of the agent's goal-directed motion and of the direction of the motion pattern, allowed to test whether the perception of persistent movement and of outcome (either failed or attained) are sufficient and necessary cues to attribute intention to a moving agent across different age-groups. Results showed a developmental change in the type of goal-directed representation between six years of age to adulthood, from an outcome-based representation to a persistent motion-based representation.

For children with a verbal mental age of under-six years, the perception of the agent's persistent motion is not a necessarily relevant cue for goal attribution. Thus, when under-six year olds were presented with the condition where the agent moved repeatedly towards a target and then suddenly changed direction and stopped next to another target, they indicated that the intended goal of the agent was to go where it eventually stopped. Thus, the representation of the persistent motion was overridden by the perception of both the outcome and the proximity of the agent to the target. The same result was shown by the autism group. The relevance of the perception of the proximity between target and agent needs further investigation. In fact, it might be possible that, had the agent not been close to the target, children would attribute the intended goal on the basis of persistent motion.

The situation was reversed in the case of the adults' performance and in the intermediate performance of the teenager group. When adult subjects were presented with the condition where the agent suddenly changes direction and ends at the opposite side from where it was originally directed, they indicated that the intended goal of the agent was to reach the top where it did not stop at. The representation of the outcome and the proximity of the agent to the target were overridden by the perception of the persistent motion. The agent's outcome, contrary to the under-six children and the children with autism, was considered accidental rather than intentional. Results indicated also that the shift between young age and adulthood is rather smooth: indeed, the two in-between age groups did not show any preferential bias towards one way or the other representation. However, the group of children with autism with a verbal mental age of 9.7 years (sd.: 3 yrs) responded in the same fashion as the under-six-year-olds. They valued more the perceptual cue of the end-state of the agent rather than its persistent attempts towards a target.

The picture of the performance of children with autism that comes into focus by combining together the results of the first part and the second part of the study is quite interesting. In fact, when the direction of the agent's repetitive motion towards the target was constant, they showed intact ability in inferring the agent's intended goal regardless of its failed outcome. This result indicated that the representation of the contingent state of the agent was suspended in favour of the representation of a non-contingent state (motion towards a future state of the agent). However, this pattern was reversed when the direction of the agent's motion suddenly changed to the opposite side. In this case, children with autism based their intended goal inference on the representation of the contingent state of the agent rather than the non-contingent state, showing a developmental delay. Thus, the change of direction of the agent's goal-directed motion seems to be the crucial variable in determining the specific difficulties in children with autism. How can this be explained?

The developmental change in normally developing individuals may provide a clue. As already mentioned, it is the sudden change of direction that determines the different types of representation across age. Indeed, all children were able to override the agent's outcome representation when the direction of the agent was tenaciously kept towards one target. Thus, it is plausible that adults considered the change irrelevant compared to a more constant and determined behaviour, and therefore they made a "conservative" decision. On the other hand, young children interpreted the change as a matter of fact, possibly as an intentional act that wiped out the behaviour occurring before the change. This line of speculative reasoning leads towards a distinction between two types of cognitive processes. It is plausible that ambiguous situations trigger executive function in adults, in particular, the ability to take into account discordant but equally valid situations, before making a decision. By comparison, this type of ability is not fully developed in young children, so that the perception of discordant events, which was combined with the absence of a learned rule to apply, would trigger a less demanding process of matching the agent's end state with the agent's intentional state. Finally, this speculative scenario would fit with the findings indicating executive functioning difficulties in individuals with autism, regardless of mentalising abilities.

Chapter 4
Brain mechanisms for the attribution of mental states to animated shapes

The present chapter reports a functional neuro-imaging study with positron emission tomography (PET) in which six healthy adult volunteers were scanned while watching silent computer-presented animations. The characters in the animations were simple geometrical shapes whose movement patterns selectively evoked mental state attribution or simple action description.

4.1 Exploring brain bases of components of social cognition

The ability to recognise, manipulate, and behave with respect to socially relevant information has been termed with the general term of "social cognition". It is generally agreed that social interaction abilities are common to both humans and non-human primates, and furthermore, that they rely on a number of highly functionally organised components that are brain based (Byrne, 1998; Duchaine, Cosmides and Tooby, 2001). In fact, without the existence of neural systems that are geared to process social stimuli from the beginning of human life, it would be difficult to explain the universality and speed of social learning. Furthermore, the interactive behaviour among primates represents an impressive social sophistication, which is shown by their co-operative behaviour in creating and shifting social coalitions, and by their deployment of different but complementary roles within their hierarchical organisation. Thus, both the primates' complex social abilities and the speed of human social learning suggest that new brain regions or systems underpinning these abilities have evolved. Social cognition abilities include, among others, the ability to represent other's mental states, and the issue regarding to the extent to which the ability to mentalise is specific to humans' is highly controversial.

4.1.1 A dedicated brain system for Theory of Mind

Developmental and evolutionary approaches to understanding social cognition are now being combined in some studies based on experiments in both human infants and non-human primates (Tomasello, 1999; Reaux, Theall and Povinelli, 1999). These studies suggested that humans posses different cognitive abilities from those of any other primate, including the ability to adopt another person's point of view. Povinelli and O'Neil (2000) and Povinelli and Giambrone (2000) proposed a view that reconciles the similarities between humans and primates with the striking dissimilarities. Briefly, it is argued that humans and primates have inherited from their common ancestor similar social behaviours, which were originally generated by low-level cognitive mechanism and unrelated to the explicit representation of other mental states of others. In humans, these existing low level mechanisms may have been recruited to support increasingly sophisticated social demands, including the ability of mentalising, to solve adaptive problems. Thus, given the highly sophisticated interaction based abilities in non-human primates, it is plausible that they posses a primitive form of mentalising ability at the very limits of their cognitive skills (Frith and Frith, 2000).

Beyond the specific question whether this ability is or is not specific to humans, the argument for the existence of a brain system specialised for mentalising is supported by evidence from neurodevelopmental disorders. In fact, the ability to mentalise can be either selectively damaged, as in autism (Baron-Cohen, Leslie and Frith, 1985; Frith, Leslie and Morton, 1991; Happé and Frith, 1996), or selectively spared, as in William Syndrome (Karmiloff-Smith, Grant, Bellugi and Baron-Cohen, 1995; Tager-Flusberg, Boshart and Baron-Cohen, 1998). In schizophrenia, a disorder of adult onset, mentalising failure can also be observed (Corcoran, Mercer and Frith, 1995), and in old age, when other cognitive abilities may decline, the ability to mentalise has been shown to increase (Happé, Winner and Brownell, 1998). Finally, the argument for a dedicated, domain-specific, and possibly modular cognitive mechanism for mental state representations is supported on theoretical grounds, as already discussed in chapter 1, by the metarepresentational model, which posits a system that computes mental states independent from other systems (Leslie and Thaiss, 1992; Scholl and Leslie, 1999). In conclusion, if there is a brain system dedicated to mentalising, then it should be possible to localise its components by the techniques of brain imaging.

4.1.2 Functional brain imaging of ToM

There are a growing number of published reports of functional brain imaging studies of Theory of Mind. Most of these studies implicate activation in medial frontal and temporo-parietal regions. Goel, Sadato and Hallet (1995), in a PET study, asked volunteers to judge whether someone living in the 15th century, as Christopher Columbus, would have known the use of a series of objects (contemporary to him or to us). This mentalising task was contrasted with memory retrieval and with simple inferencing. Mentalising was associated with activity in medial prefrontal cortex and left temporo-parietal junction. Fletcher, Happé, Frith et al. (1995), in another PET study, scanned volunteers reading and answering questions about stories involving complex mental states ("ToM" stories) and those involving inferences of physical cause and effect ("Physical" stories). Comparison of activation during ToM versus Physical stories revealed increased activation in the medial frontal gyrus on the left (BA 8/9), as well as in the posterior cingulate cortex and the right inferior parietal cortex (BA 40) at the temporo-parietal junction. More recently, Gallagher, Happé, Brunswick et al. (2000) used the same set of stories adopted by Fletcher et al. (1995) in an fMRI study. In addition to the written stories, subjects were shown figurative drawings (humorous cartoons) which also prompted attribution of mental states. With the greater resolution of fMRI it was possible to distinguish a number of peaks in Brodmann areas 8/9 and the border of 10 and 32, associated with both ToM cartoons and stories. The location of these areas of activity was close to those previously reported by Fletcher et al. (1995) and by Goel et al. (1995), and relates to the paracingulate sulcus. Activity was also observed in the temporo-parietal junction bilaterally. Finally, Baron-Cohen, Ring, Moriarty et al. (1994) in an early study using a single-photon emission computed tomography (SPECT) found increased activation in the orbitofrontal region during a task involving judgement of mental states words as opposed to object words.

4.2 A paradigm based on non-verbal stimuli

Previous brain imaging studies of mental state attribution have tended to use high-level verbal stimuli (Baron-Cohen et al., 1994; Fletcher et al., 1995; Happé et al., 1996; Goel et al., 1998), or visual depictions of humans (Gallagher et al., 2000; Baron-Cohen et al., 1999). Mentalising, however, involves processes at a number of levels, from perceptual to conceptual. The aim of the present study was to examine brain activation during exposure to simple, non-verbal stimuli designed to evoke mental state attribution by their kinetic properties alone.

4.2.1 A classic paradigm

Inspiration for appropriate stimuli came from the classic work of Heider and Simmel (1944), who demonstrated that even simple geometric shapes could elicit by their pattern of contingent movement the attribution of complex internal states, such as intentions and beliefs. Subjects were asked to interpret an movie sequence lasting about two and half minutes three geometrical figures (a small triangle, a large triangle and a circle) were shown moving in various directions and at various speeds. The only other figure in the visual field was a rectangle, a section of which could be opened and closed as a door is. In the first experiment, a groups of subjects (curiously, they were only female) were shown the movie, and instructed to "write down what happened in the picture". In the second experiment, another group of subjects were shown the picture and asked to interpret the movements of the figures as actions of persons. After the viewing they had to answer a detail questionnaire concerning both the nature of the personalities of the three characters, and the reasons behind their actions. Finally, subjects were asked to tell the story of the movie in few sentences. In the third experiment the same movie was shown in reverse, and subjects were again asked what kind of persons were the protagonists, and to tell the story. Results of the first experiment indicated that the majority of subjects perceived the movie in terms of animated beings (most of the times humans, more seldom animals) and reported a connected story. Results of the second experiment indicates that the personalities of the characters are judged with great uniformity, and that the story was about an aggressive big triangle, and a weaker but brave small triangle helping a fearful circle – female - to escape. The results of the third experiment indicated that the interpretation of the human actions varied greatly across subjects, and did not allow for a quantitative analysis. These findings have been replicated a number of times with different populations (Kassin, 1981) and are typically interpreted as evidence of the role of motion in social perception and attribution of animacy. This interpretation was directly tested by Berry and colleagues (Berry, Misovich, Kean and Baron, 1992; Berry and Springer, 1993) using manipulated version of the Heider and Simmel movie with adults and pre-school children. In order to examine the independent contributions of form and motion in the anthropomorphic attribution, Berry et al. (1992) created three altered version of the original movie. In one version, the shapes and size of the characters were disrupted but their movement was preserved. In the second version, the structural aspects of the animation were preserved, but the movement was disrupted. In the third version, both form and motion were disrupted. The original Heider and Simmel movie and the altered version were shown to different groups of adults who were asked to describe what was happening as they watched. Results indicated that the original movie evoked anthropomorphic descriptions in all subjects, whereas the altered versions with the movement disrupted and the structural aspect preserved severely reduced the proportion of adults who provided description of the display in social terms. These findings were replicated by Berry and Springer (1993) with groups of pre-school children, aged three, four and five years, indicating that children as young as three years are sensitive to properties of the movement, rather than of the stimuli/characters. In a subsequent work, Springer and Berry (1996) demonstrated that character attributions to the figures in the original Heider and Simmel movie were more differentiated and more similar to adult attributions in the 5-year-olds than in the younger children. However, regardless the type of descriptions evoked by the animations, these findings revealed that the perception of motion is fundamental to the complex attributions made by adults and children from the late pre-school years onwards. Two studies with children with autism have investigated the ability to mentalise using the original Heider and Simmel animation.

Bowler and Thommen (2000) investigated the production of narrative account of events by showing the Heider and Simmel movie to children with autism/Asperger Syndrome and controls. They found that children with autism were able to distinguish intentional actions from mechanical motions at the same level as the chronological and verbal mental age matched control groups. A closer analysis of the type stimuli displayed by the original animation of Heider and Simmel reveals that they frequently elicited the attribution of goal-directed actions (fighting and chasing), and less often more complex intentional states, e.g. aimed at manipulating someone else mental states (bluffing, surprising). The indication that children with autism have more difficulties with more difficult types of interaction comes from the evidence that they were less likely to comment on interactions between the characters when this did not involve physical contact.

Klin (2000) investigated mentalising ability in adolescents and adults with high-functioning autism and with Asperger Syndrome also using the Heider and Simmel stimuli. Subjects were asked to provide verbal descriptions that were subsequently measured with a specifically developed detailed coding system (Social Attribution Task Index Score). The multidimensional approach allowed the author to avoid the typical dichotomous response approach, (e.g. correct response reflecting "ToM competence", and wrong response reflecting "ToM deficit"), and to quantify the sophistication of social attributions contained in the subjects' narratives. The analysis revealed that the descriptions of the autism group included a small number of mental state terms that were not pertinent to the underlying "social frame" of the animations, compared to those given by the control group.

4.2.2 Developing a new paradigm

The new paradigm that Uta Frith and Francesca Happé created with the aim of investigating Theory of Mind in autism with a nonverbal task, was based on the assumption that motion alone evokes attribution of intentional states, as demonstrated by the above studies. The stimuli of the Heider and Simmel's original movie were suitable to investigate the ability to attribute human properties in general, triggered by perception of motion alone, but were not suitable to distinguish between different type of descriptions of social interactions. In fact, the Heider and Simmel's movie was based on a long and complex script that included different levels of interactive behaviour, from actions reflecting an agent's goal, e.g. "hitting", to actions reflecting an agent's mental state, e.g. "confronting" (these examples are taken from Berry and Springer's (1993) study). By contrast, the aim of the new paradigm was to design different type of animations that would elicit selectively descriptions that did and did not involve mental state descriptions. The scripts of the animations were considerably shorter (about 40 seconds each) than the Heider and Simmel's movie (about 150 seconds), and each of them was aimed at evoking a description of the characters either in terms of goal-directed actions, or in terms of propositional mental states. In addition, a third condition was added aimed at evoking description of the characters in terms of random movement, non-intentional actions. The characters of the animations were two triangles, a big and a small one moving in a self-propelled fashion.

In the "Theory of Mind" condition, interaction between two shapes (big and small triangle) was scripted to imply complex mental states, propositional attitudes, such as the intention to seducing someone. Thus, in these animations one character's actions were readily seen as determined by what the other character *'thought'*. In the second animation type, "Goal-directed", interaction between the two characters was scripted to imply a purposeful type of actions, with no reference to specific mental states, such as dancing together. Thus, in these animations, one character's actions were seen as determined by what the other character *'did'*. In the third animation type, "Random", the two characters did not interact, and their behaviour was not contingent – in effect they were merely floating, or bouncing off the sides, or moving horizontally or vertically within the sides.

4.2.3 Behavioural findings with adults and children

The animation paradigm was first used in a behavioural study which collected descriptions from children with autism, children with developmental delay, normally developing children and adults (Abell, Happé and Frith, 2000). In order to enhance the distinction across the three types of animations, ToM, Goal-directed and random, each of them was presented to subjects by specifying the kind of characters in each animation: people for the ToM sequences, e.g. teacher and boy, animals for the Goal-directed interactions, e.g. duck and duckling, and triangles for the Random animations. After watching each movie, subjects were asked to simply say what was happening. The test was first presented to a group of adults who validated the intended script underling each animated sequence. Consistent with the Klin (2000) study, the findings indicated that children with autism were less accurate in their descriptions of ToM animations than children without autism. The fourteen adults taking part in this study attributed precise mental states, matching the underlying script in 89% of their responses to the ToM animations, with descriptions of purposeful movements for the remaining responses. They attributed precise purposeful interactions in 93% of their responses to the Goal-directed animations with the remaining responses all involving mental state attribution. Descriptions of simple movement without a purposeful component were given in 64% of responses to the Random sequences while purposeful movement was described for the remainder. Even though the vast majority of descriptions of the three types of animations fell into an orderly pattern, the animations were ambiguous enough for interpretations to occur that were either simpler or richer than intended by the designers.

4.3 A neuroimaging study with healthy adult volunteers

Abell et al. (2000) validated the paradigm as a Theory of Mind test based on the perception of geometrical shapes whose movement patterns selectively evoked mental state attribution or simple action description. The task was therefore suitable to explore brain activation during such movement-provoked mental state attribution. As in the Abell et al.' study subjects were presented with silent animations of three types. In the four Theory of Mind (ToM) animations the movement of the two interacting characters evokes description relative to one triangle anticipating or manipulating the mental state of the other. In the four Goal-directed (GD) animations, the movement of the two triangles evokes description in terms of behavioural interaction. In the four Random (Rd) animations, the purposeless movement of the two triangles elicits description with no reference to interaction, goals or intentions. The stimuli could therefore be graded in terms of complexity of predominantly evoked descriptions, from random movements, to goal-directed actions, and finally complex intentional states. Conversely, people's descriptions could be graded in terms of their degree of mentalising regardless of the animation sequences they were describing. Contrary to the Abell et al.'s study, no characters' descriptions (type of people, animals or shapes) were given prior to the view, so that subjects' descriptions were elicited exclusively by the perception of the agents' movement rather than implicitly suggested by the different type of features or roles given to the two triangles. After each scan subjects were asked to describe spontaneously what was happening in the animation. The verbal descriptions were coded along four different dimensions. The aim of the scores was to distinguish in each answer the degree of appreciation of mental states, their appropriateness, that is, how well the underlying script was captured, the certainty of the explanation, and the length of each answer.

The prediction for the present study was that the ToM animations, but not the Random animations, would evoke mental state attributions, and show activation patterns similar to those found in previous functional imaging studies of Theory of Mind (Goel et al., 1995; Fletcher et al., 1995; Baron-Cohen et al., 1999; Gallagher et al., 2000). It was expected the GD animations to have an intermediate status. Going one step further, we predicted activation in ToM related areas for all sequences which provoked mental state interpretations, regardless of the animation condition.

4.3.1 Method

Subjects: Six right-handed male volunteers (aged 20 to 31 years, mean 24.5 years) took part in this study. All subjects were healthy, with no history of significant medical, psychiatric or neurological illness. All gave written informed consent to take part in the study, which was approved by the ethics committee of National Hospital for Neurology and Neurosurgery and the Administration of Radioactive Substances Advisory Committee (ARSAC) UK.

Data Acquisition: All subjects underwent both PET and MRI scanning on the same day. A Siemens VISION (Siemens, Erlangen) operating at 2.0T was used to acquire axial T1 weighted structural MRI images for anatomical coregistration. PET scans were performed with an ECAT EXACT HR+ scanning system (CTI Siemens, Knoxville, TN) in high sensitivity 3-D mode with septa retracted (Townsend et al., 1991). A venous cannula to administer the tracer was inserted in the antecubital fossa vein. Approximately 350 Mbq of $H_2{}^{15}O$ in 3 ml of normal saline were loaded into intravenous tubing and flushed into subjects over 20 seconds at a rate of 10 ml/min by an automatic pump. After a delay of approximately 35 seconds (s), a rise in counts could be detected in the head that peaked 30-40 s later (depending on individual circulation time). The interval between successive administrations was 8 min. The data were acquired in one 90 s frame, beginning 5 s before the rising phase of the head curve. After correcting for background activity, the true counts accumulated during this period were taken as an index of cerebral blood flow (Fox and Minton, 1989). Images were reconstructed by filtered back projection (Hanning filter, cut off frequency 0.5 cycles per pixel) into 63 image planes (separation 2.4 mm) and into a 128x128 pixel image matrix (pixel size 2.1 mm). Twelve scans were acquired per subject.

Statistical analysis: Functional imaging analysis used the technique of Statistical Parametric Mapping implemented in SPM97 (Wellcome Department of Cognitive Neurology, London, UK, (http://www.fil.ion.ucl.ac.uk/spm). For each subject, a set of 12 PET scans was automatically realigned and then stereotactically normalised (Friston et al., 1995b) into the space of Talairach and Tournoux (1988). The scans were then smoothed using a Gaussian kernel of 12mm full-width half maximum.

The analysis of functional imaging data entails the creation of statistical parametric maps that represent a statistical assessment of condition-specific effects hypothesised by the experimenter (Friston et al., 1995a). The effects of global changes in blood flow were modelled as a confound using a subject-specific ANCOVA (Friston et al., 1990). Areas of significant change in brain activity were specified by appropriately weighted linear contrasts of the condition-specific effects and determined using the t-statistic on a voxel by voxel basis. We created the relevant SPM [*t*] for each comparison of conditions, which was then transformed into an SPM [*Z*] and thresholded at a Z-score of 3.09 ($p<0.001$ uncorrected). Clusters of activated voxels were characterised in terms of their peak height and spatial extent conjointly.

Design: A 3x2 repeated measures within subjects design was used. Four different examples of each of three types of animation, ToM, Goal-directed, and Random were displayed over the course of 12 scans, divided into two consecutive counterbalanced blocks: cued animation and uncued animation. In a previous study (Fletcher et al, 1995) subjects were cued before the scan. They were told in advance which kind of stimuli they were going to see (see Appendix 4A). In the present study we counterbalanced cued with uncued animations in order to control for the effect of prior knowledge.

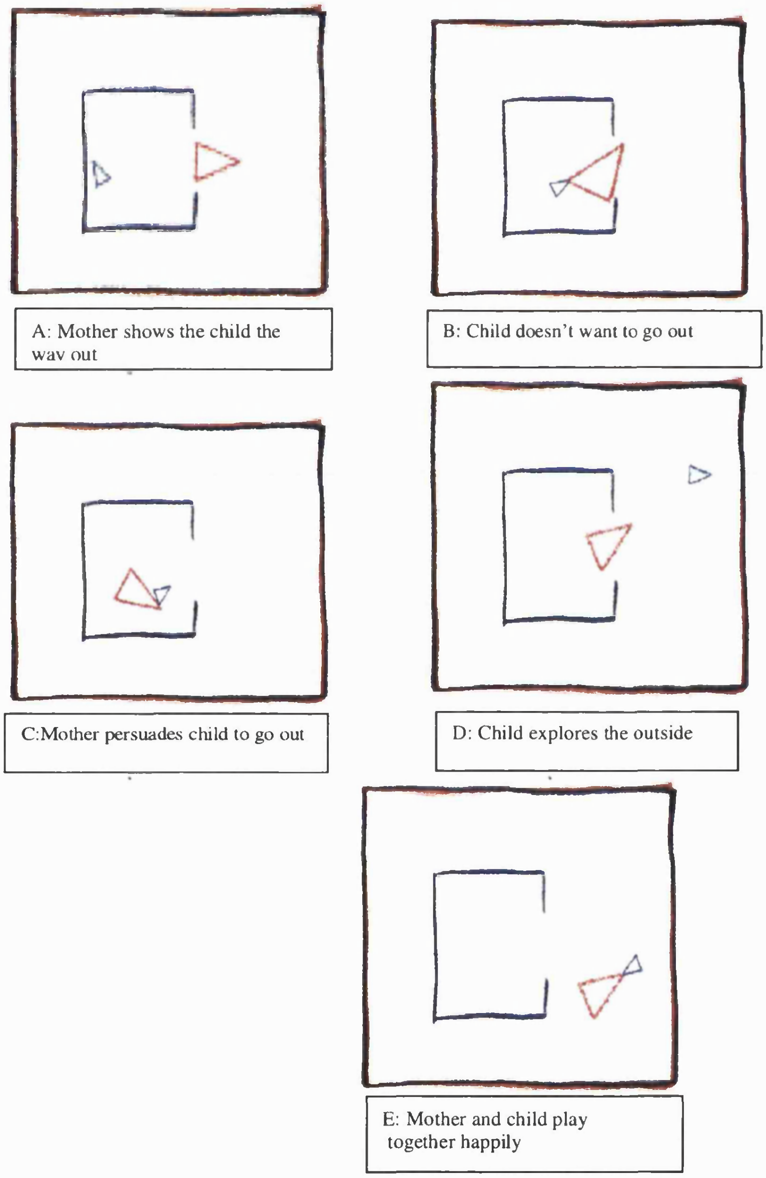
Animation materials: Twelve animations were used during the scanning, and an additional three were shown for practice. All the animations featured two characters, a big red triangle and a small blue triangle, moving about on a framed white background. Each sequence lasted between 34 and 45 seconds, and the three types of animations were matched for length. The 'scripts' for the ToM sequences involved the two triangles persuading, bluffing, mocking and surprising one another (see example in Figure 4.1). The Goal-directed 'scripts' involved the two triangles dancing together, one chasing, one imitating, and one leading the other. The Random movement showed the two triangles bouncing off the walls resembling the movement of billiard balls, or merely drifting about, or moving either horizontally or vertically across the screen. A sample of the animations is available at: www.icn.ucl.ac.uk/groups/UF/research/animations.html. While the type of movement was by definition different between the three conditions, the basic visual characteristics in terms of shape, overall speed and orientation changes were as similar as possible.

Procedure: Subjects were instructed before the scanning (see Appendix 4A) and were given examples of the three types of animations. Neither during the instruction phase nor during the scanning, were subjects given any suggestion of the story or characters roles implicit in the design of the sequences. During scanning, subjects were requested to remain relaxed while watching the animations. The animations were presented on the screen of a Power Macintosh computer suspended on a adjustable cradle at a suitable distance for each subject. Prior to scanning, it was ascertained that the subject could watch the animations comfortably.

Before the cued animations subjects were told either that they were going to see an animation showing "an interaction with feelings and thoughts" (ToM), or "a random movement" (Rd), or "a simple interaction" (GD). Before the uncued animations, subjects were simply told that they were about to see the next animation. Order of cued and uncued blocks was counterbalanced across subjects.

After each scan subjects were asked to tell the experimenter what they thought the triangles were doing. The experimenter always asked the same neutral question: "What was happening in this animation?" Answers were recorded for later scoring. On no occasion was feedback given, but subjects were generally praised for their descriptions.

Figure 4.1: Sequences from a 'Theory of Mind' animation depicting the two triangles interacting. The animation was designed following a script in which Big Triangle is coaxing the reluctant Little Triangle to come out of an enclosure. Subjects were presented with the animations without any suggestion relative to a story, or characters roles. The captions relative to each of the following frames are meant to help the reader to understand the type of interaction represented through movement patterns.



Scoring verbal descriptions: The verbal descriptions given after each animation (in between scans) were transcribed verbatim and coded in terms of four different dimensions (details of all scoring criteria are given in Appendix 4B).

The Intentionality score reflected the use of mental state terms, with scores ranging from 0 to 5. The degree of attribution of mental states to the triangles (agents) of the animations was calculated by analysing the content of each description given by the subjects. In the effort to control as much as possible the use of subjective methods in interpreting someone else's language, terminology, idioms and so forth, the analysis has been conducted exclusively on the type of verb contained in each sentence used to describe the triangles' actions. The degree of intentionality reflected in each action was measured with a numerical scale from zero to five. The scale was created explicitly for this task and the type of stimuli that subjects had to describe somehow influenced it. Thus, it is clear that a measure applied to verbal descriptions concerning two agents interacting along a complex motion pattern, cannot and should not reflect the complexity of mental states appreciation contained in any language describing any social context.

In developing the score, the "intentionality ladder" came into shape, with an agent moving upwards, appreciating step by step both actions, and mental states of another agent. At the bottom of the ladder, where there is no appreciation of another agent, nor actions or mental states (score=0), the agent acts with no intention, and no interaction, randomly, e.g. "moving around", or "floating". A further step up in the ladder (score=1), the agent acts with a purpose, a goal, with no interaction with another agent, e.g. "walking", or "swimming". The following step up (score=2), is when the agent acts with a purpose with another agent, e.g. "fighting" or "following": the actions of the two agents are parallel in time. A further step up (score=3) is when the agent not only interacts with another agent but acts in response to the other's action, e.g. "chasing", or "restraining": the actions of the two agents are sequential in time. Finally, the two steps at the top of the ladder concern the agent's appreciation of mental states. The lower step (score=4) is when the agent acts in response to a mental state, e.g.

"arguing", "wanting" or "encouraging". The upper step (score=5) is when the agent acts with the goal of affecting or manipulating the other agent's mental states, e.g. "pretending", "deceiving" or "coaxing". Details of the Intentionality score with examples are given in Appendix 4C.

The Appropriateness score measured the understanding of the event depicted in the animations, as intended by the designers. The score, ranging from zero to three, was based on the underlying script for each animation. Details of criteria for rating the appropriateness of each animation are given in Appendix 4D. The degree of appropriate description of the animation was calculated by analysing the agents' actions and interactions. For example, an appropriate description (score=3) for the animation where the big triangle persuades the little one to go out, need to convey the idea of little triangle's reluctance to go out *and* big triangle's attempts to get the little one out, e.g. "persuading" or "coaxing". A less appropriate description (score=2) would focus on one aspect of the story or one character only. e.g. little one doesn't want to go out; or, big one is pushing little one to go out. An inappropriate description (score=1) concerns actions that do not relate to the events or relate to a very minor aspect of the sequence only, e.g. "the two triangles didn't like each other". Finally, when the subject did not provide any description, the score was zero.

The Certainty score (0-3) graded the degree of hesitation present in the verbal description. When the subject did not provide a description, the score was zero. Low certainty (score=1) was for high hesitation in describing the animations, e.g. when subject utters few words, does not finish sentences, and need to be prompted. Medium certainty (score=2) was for alternative answers, with some kind of hesitation. The score for a clear, quick answer was the highest (score=3). Finally, the Length score classified the number of clauses in each answer (0 to 4).

The intentionality score was measured by two raters who gave an identical score 65% of the time, and had an average discrepancy of only 1.4 points in the remaining

35% of the cases. On the Appropriateness score the two raters reached full agreement. Only one rater measured both the Length score and the Certainty score.

4.3.2 Results

Behavioural data: The ratings of the four verbal description scores (Intentionality, Appropriateness, Certainty and Length) were analysed with nonparametric tests, since these were in the form of ordinal data. Table 4.1 shows the ratings of the descriptions of each type of animation.

Nonparametric two-way analysis of variance (Friedman test) on the Intentionality score indicated that all ToM animations evoked the same type of mental state attribution (chi-square=3.4, p.=not sig.), as well as the four Goal-directed animations (chi-square=6.6, p.=not sig.) and the four Random animations (chi-square=0.5, p.=not sig.). As expected on the basis of Abell et al.'s results (2000) subjects attributed more intentionality to the characters' behaviour during ToM animations than during GD (Wilcoxon test, z=2.2. p.=0.3) and Rd animations (Wilcoxon test, z=2.2. p.=0.3). Random animations evoked significantly fewer mental state attributions than Goal-directed animations (Wilcoxon test, z=2.2. p.=0.3).

Nonparametric two-way analysis of variance (Friedman test) on the Appropriateness score reavealed no difference in the score for the four ToM animations (chi-square=3.5, p.=not sig.), and the four Goal-directed animations (chi-square=1.1, p.=not sig.) and the four Random animations (chi-square=0.5, p.=not sig.). Wilcoxon tests revealed that subjects described with the same degree of appropriatenss the ToM animations than the GD animations (z=1.2. p.= not sig) and Rd animations (z=1.6. p.= not sig). Random animations and Goal-directed animations were also equally described (z=0 p.=not sig).

Same analyses were performed on the Certainty score, indicating no differences in the ratings of all ToM animations (chi-square=2.4, p.=not sig.), all Goal-directed animations (chi-square=0.7, p.=not sig.) and all Random animations (chi-square=0.4, p.=not sig.). Subjects' certainty scoredid not differ across types of animation, ToM did not differ from GD animations (z=0.5, p.=not sig.), nor from Rd animations (z=1.5, p.=not sig.), and the GD animations did not differ from the Rd animations (z=1.8, p.=not sig.).

Same analyses were performed on the Length score, indicating no differences in the ratings of all ToM animations (chi-square=1.3, p.=not sig.), all Goal-directed animations (chi-square=0.6, p.=not sig.) and all Random animations (chi-square=4.1, p.=not sig.). Differences were revealed by the Wilcoxon tests on the total scores. Subjects' gave longer description for the ToM animations than both the GD animations (z=2.5, p.=0.01), and the Random animations (z=2.7, p.<0.01). The descriptions for the GD animations were longer than the Random animations (z=2, p.=0.05).

There were no difference in the Intentionality score (z=0.1, p=not sig.), nor in the Appropriateness score (z=0, p=not sig) when subjects were "cued" before watching the animations.

Table 4.1: Verbal descriptions given by the six subjects for ToM, Goal-directed and Random animations rated on four dimensions. Mean and standard deviation (in parentheses).

Score type (range)	Animation Type ToM	 Goal-directed	 Random
Intentionality (0-5)	**3.9** (0.5)	**2.4** (0.3)	**0.2** (0.5)
Appropriateness (0-3)	**1.6** (0.2)	**1.9** (0.2)	**1.9** (0.3)
Certainty (0-3)	**2.6** (0.3)	**2.5** (0.4)	**2.8** (0.3)
Length (0-4)	**2.8** (0.8)	**1.9** (0.5)	**1.6** (0.5)

Neuroimaging data:

i) Subtraction analysis:

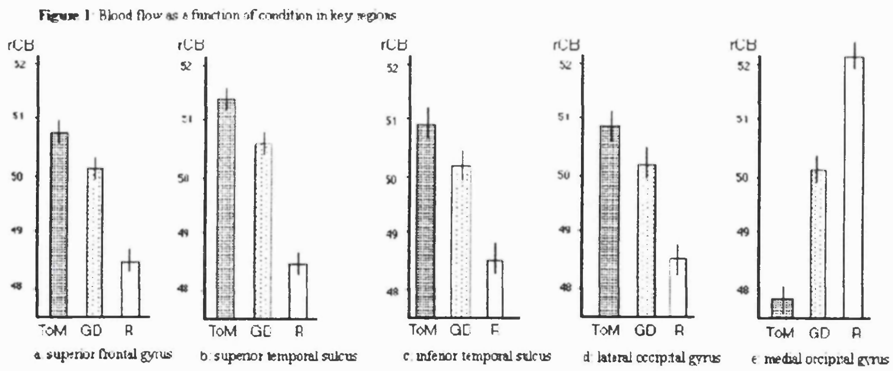
There were no significant differences between cued and uncued presentations, nor were there any order effects, or any significant interactions. Data for cued and uncued sequences were therefore combined. There were significant differences between the three types of animation. ToM animations elicited more activity than Random animations in four regions: medial prefrontal cortex, temporal-parietal junction (at the end of the superior temporal sulcus), basal temporal region (fusiform gyrus and temporal poles, immediately adjacent to the amygdala), and extra-striate cortex (occipital gyrus) (see Table 4.2). All these differences were observed in both hemispheres, but were more significant in the right hemisphere, except for the medial prefrontal cortex. For all these regions differences occurred between the ToM and the Random condition, with the Goal-directed condition showing intermediate activity that was more similar to the Random condition (see Figure 4.2). Direct comparison of ToM with GD confirmed that the differences apparent in Figure 4.2 were significant in the case of temporo-parietal regions and the temporal pole at a level of $p <.0001$ uncorrected and for occipital gyrus and fusiform gyrus at $p < .01$. Random movement when compared to ToM movement, elicited more activity in one region of medial occipital cortex (-2x, -94y, 14z). The locations of the activations are shown superimposed on a standard brain in Figure 4.3.

Table 4.2: Subtraction analysis. Regions where ToM animations elicited more activity than Random animations. The coordinates are given in the stereotactic space of Talairach and Tournoux, 1988. Numbers in bold type indicate regions where differences in activity were significant when corrected for multiple comparisons. Numbers in plain type indicate regions where differences in activity were significant at p<.0001, uncorrected.

Foci of activation		LEFT			RIGHT		
	BA	x, y, z	Z score	p<	x, y, z	Z score	p<
Temporal-parietal junction							
STS	22/39	**-58, -48, 4**	**4.3**	**.06**	**60, -56, 12**	**6.2**	**.001**
Basal temporal							
FuG	37	-38, -44, -22	3.8		**36, -56, -20**	**5.1**	**.01**
TmP/Am	38	-38, -4, -32	3.2		**34, 6, -26**	**4.0**	**.05**
Occipital lobe							
OcG	19/18	**-30, -94, -12**	**4.6**	**.02**	**38, -96, -10**	**5.0**	**.01**
OcG	19/18	-32, -82, -24	4.1				
Medial frontal							
SFG	9	-4, 60, 32	4.1				

Brain regions are identified by name and by putative Brodmann Area (BA) on the basis of the atlas of H.M. Duvernoy (1999) The Human Brain: Surface, Three-Dimensional Sectional Anatomy with MRI, and Blood Supply. Springer Wien New York.
STS - superior temporal sulcus
TmP/Am – temporal pole adjacent to amygdala
FuG – fusiform gyrus
OcG - occipital gyrus
SFG - superior frontal gyrus

Figure 4.2: Blood flow as a function of condition in key regions.



ii) Correlational analysis:

A further analysis was performed in which Intentionality scores, regardless of condition, were correlated with blood flow response. This analysis was carried out within subjects, thus avoiding the assumption that different individuals use the same range of descriptions. An assumption inherent in this analysis is a linear relationship between intentionality scores and blood flow response. However, further analysis which allowed for a non-linear relationship did not produce a significant increase in variance accounted for.

The results of the correlational analysis were clear-cut. The same four areas were identified as more active as in the comparison of the three conditions (see Table 4.3). These results were not affected when the length of the descriptions given by the subjects was entered as a confounding covariate.

Table 4.3: Correlation analysis. Regions where there were significant correlations between blood flow response and Intentionality score.

Foci of activation		**LEFT**			**RIGHT**		
	BA	**x, y, z**	**Z score**	**p<**	**x, y, z**	**Zscore**	**p<**
Temporal-parietal junction							
MTG	21/37	**-60, -48, 4**	**4.6**	**.02**			
STS	39				**62, -58, 12**	**6.6**	**.001**
Basal temporal							
FuG	37	-36, -42, -22	3.7		**38, -54, -22**	**4.9**	**.01**
TmP/Am	34/38				30, 4, -24	2.9	
					22, 0, -16	3.05	.08
Occipital lobe							
OcG	18				**40, -96, -10**	**4.8**	**.01**
OcG	17	**-16, -100, -8**	**4.2**	**.05**			
Medial frontal							
SFG	8/9	-6, 58, 32	3.0				

Note: Z scores (p-value <.0.05 corrected for multiple comparisons in bold, p-value <0.001 uncorrected in plain text).Brain regions are identified by name and by putative Brodmann Area (BA) Brain regions are identified by name and by putative Brodmann Area (BA) on the basis of the atlas of H.M. Duvernoy (1999).

STS - superior temporal sulcus

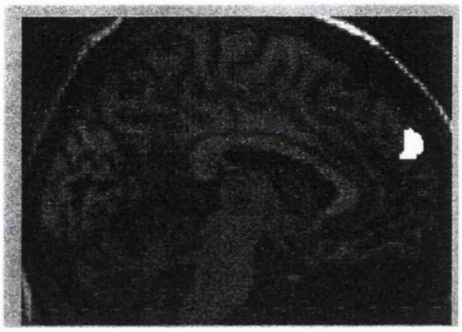
TmP/Am – temporal pole adjacent to amygdala

FuG – fusiform gyrus

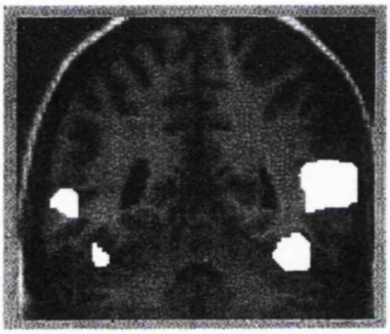
OcG - occipital gyrus

SFG - superior frontal gyrus

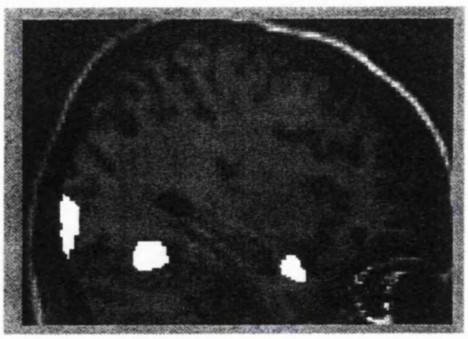
Figure 4.3a,b,c: Regions of significant cerebral blood flow (rCBF) change associated with the perception of ToM animations vs Random animation. **(a)** Saggital view of activation in superior frontal gyrus. **(b)** Coronal view of activation in superior temporal sulcus and fusiform gyrus. **(c)** Saggital view of activation in temporal pole adjacent to the amygdala, fusiform gyrus, and occipital gyrus.



(a)



(b)



(c)

4.3.3 Discussion

The present experiment took as its starting point the pervasive tendency to perceive intentions in complex movement patterns even when no human forms are depicted. We showed that different types of silent animations selectively evoked descriptions of what the characters were thinking, or descriptions of what the characters were doing. The different types of descriptions occurred spontaneously, since alerting subjects in advance to the nature of a particular sequence they were going to see had no effect.

The main aim of this study was to locate a brain system associated specifically with the attribution of mental states evoked by kinetic stimulus properties. At the same time we wished to relate the findings to earlier studies of mentalising with different kind of stimuli. Subtraction analysis (contrasting the ToM sequences with Random or Goal-directed sequences) gave the same picture as correlational analysis (correlating blood flow with degree of mental state description across all animations). The results showed increased activation in four main areas bilaterally. These areas include medial prefrontal cortex, temporo-parietal junction (superior temporal sulcus), basal temporal region (fusiform gyrus and temporal poles adjacent to the amygdala), and occipital cortex. All of these areas have been implicated in prior studies of mentalising. This suggests that a system can be delineated which is to some extent independent of the mode of stimulus input, visual or verbal.

The results of this study do not enable us to identify the functions of these four regions, but clues to their significance can be gained by considering previous studies involving different paradigms.

The medial prefrontal region: An as yet unpublished fMRI study using Heider and Simmel type silent animations has recently been summarised by Klin, Schultz and Cohen (2000). The results appear to be highly consistent with our findings. In particular, these authors mention strong activation in medial prefrontal cortex. The medial prefrontal region activated during the attribution of mental states to animated triangles has also been shown to be specifically activated by other stimuli evoking attribution of beliefs and intentions. These studies are summarised in Table 4.4.

For example, Goel et al. (1995), found left medial prefrontal gyrus activation associated with reasoning about other people's thoughts regarding a novel object. Fletcher et al.'s (1995) story comprehension task, requiring inference about a character's intentions, showed peak activation in a dorsal region of medial frontal cortex. This region was not activated in individuals with Asperger Syndrome, who show delays and deficits in Theory of Mind (Happé et al., 1996). Gallagher et al. (2000) have compared the same story task with a nonverbal comprehension task, using static single frame cartoons. They found a convergence between activations in response to verbal and visual stimuli that prompt mental state attribution (reading a text and viewing a cartoon, respectively), with bilateral activation in a ventral area of the medial prefrontal cortex. In addition, medial prefrontal areas have been shown to be activated during a rather different task that may, nonetheless, require attribution of mental states. In a task of metaphor comprehension which, according to some theorists (Sperber and Wilson, 1986; Happé, 1993), requires recognition of the speaker's intentions, Bottini, Corcoran, Sterzi et al. (1994) found activation in several loci, including left rostral anterior cingulate cortex, very close to the area implicated in the studies of mentalising mentioned above. It appears, therefore, that a number of very different mentalising tasks across several modalities (e.g. verbal, nonverbal) and with differing stimulus qualities (e.g. static, moving), activate regions of medial frontal cortex (see Table 4.4 for coordinates).

Table 4.4: Coordinates for activation of medial frontal regions in present and related studies.

TASK	COGNITIVE PROCESS	STUDY	LEFT	RIGHT
			x, y, z	x, y, z
Judge others' knowledge	mental state attribution	Goel et al. (1995)	-12, 38, 32	
Story comprehension	mental state attribution	Fletcher et al. (1995)	-12, 42, 40	
Story comprehension (Asperger Syndrome)	mental state attribution	Happé et al. (1996)	-12, 36, 36	
Story and cartoon comprehension	mental state attribution	Gallagher et al. (2000)	-10, 48, 12	8, 22, 46
Metaphor comprehension	Attribution of speaker's communicative intention	Bottini et al. (1994)	-2, 42, 8	
Intended speech monitoring	Monitoring own mental states	McGuire et al. (1996a)	-2, 36, 36 -10, 32, 24	2, 52, -4
Self generated thoughts	Monitoring own mental states	McGuire et al. (1996b)	-8, 38, 24 0, 38, 36	
Perceiving pain	Monitoring own mental states	Rainville et al. (1997)		3, 20, 30
Perceiving tickle	Monitoring own mental states	Blakemore et al. (1998)		2, 42, 6
Reporting emotions	Monitoring own mental states	Lane et al. (1997)		0, 50, 16
Intended response monitoring	Monitoring own mental states	Carter et al. (1998)		4, 25, 43
Observing human body movement	Perception of biological motion	Bonda et al. (1996)	-7, 58, 26	

Studies of *self-monitoring* have also shown increased activity in areas including medial prefrontal and cingulate cortex. This suggests that when subjects have to reflect on their *own* mental states, they may use neural pathways similar to those underlying attribution of mental states to others. For example, subjects required to monitor their intended speech, in order to judge whether distorted feedback was their own or another person's voice (McGuire, Silbersweig and Frith, 1996), showed activation of bilateral medial frontal cortex and anterior cingulate gyrus/medial prefrontal cortex as well as temporo-parietal junction bilaterally. The prefrontal region was also activated in a study where subjects reported self-generated thoughts independent from stimuli in the immediate environment (McGuire, Paulesu, Frackowiack and Frith, 1996) A quite different type of self-monitoring task investigated the neural substrates of perceived pain (Rainville, Duncan, Price et al., 1997). The anterior cingulate cortex showed increased activity when subjects perceived (under hypnosis) the increasing unpleasantness of hot water on their hand. Blakemore, Wolpert and Frith, (1998) found anterior cingulate activity associated with reporting a tickling sensation from self-produced tactile stimulation. Activity in anterior cingulate, extending into the medial prefrontal region, was also observed when subjects reported their own emotional responses to pleasant, unpleasant and neutral pictures (Lane, Fink, Chau and Dolan, 1997). A more complex self-monitoring task elicited activity in anterior cingulate cortex when subjects were required to choose between competing responses (Carter, Braver, Barch et al., 1998). Taken together, these results seem to indicate that online monitoring of inner states - own or others' - may engage the anterior cingulate cortex and neighbouring medial frontal regions, regardless of the specific source of information.

Interestingly, the activity shown by the control group in the prefrontal area during the observation of ToM animations overlaps with the area activated by two studies using nonverbal tasks (Bonda, Petrides, Ostry and Evans, 1996; Brunet, Sarfati, Hardy-Bayle and Decety, 2000). Bonda et al (1996) used two biological movement conditions; a dancing figure (human body movement), and a grasping hand simulating the act of reaching out for a glass and bringing it to the mouth (goal-directed action). The comparison of activation during the two conditions showed that perception of a dancing figure versus a grasping hand elicited a network of activation, including left medial prefrontal cortex, close to that activated by our ToM animations. Subjects in the Brunet et al. study were presented with short comic strips and asked to complete the sequence with one of three possible endings. The comic strips were depicting either mentalistic events (e.g. a prisoner wants to escape from the jail and knots together the sheets of his bed) or physical cause-effect event with human characters (e.g. a ball rolls down a slide and hits a girl) or physical cause-effect event with objects (eg. a ball hit a vase and breaks it). The common denominator across these findings seems to be the absence of explicit language processing; however, further investigation is needed to discriminate stimuli properties that elicit mental state attribution.

Grady (1999) provides an exhaustive list of activations observed in prefrontal cortex classified in terms of putative Brodmann areas. The vast majority of these are lateral. However, some medial activations have been observed in the vicinity of the area reported in the present study. The only study observing a relevant activation in Brodmann area 10 was that of Bottini et al. (1994) on metaphor comprehension. Activation in relevant regions of Brodmann area 9 have been observed in motor learning tasks and working memory tasks, but the majority of the activations observed during such tasks are more lateral and more posterior. Activity is also reported in medial Brodmann area 8 for some language tasks and for some object processing tasks, but here again all the activations are more posterior than the one observed in the present study.

In previous studies of mentalising the activity in medial frontal cortex lies at the border of anterior cingulate cortex and medial frontal cortex in the paracingulate sulcus (Gallagher et al., 2000). In an exhaustive examination of studies that have activated anterior cingulate cortex, Paus, Koski, Caramanos and Westbury (1998) conclude that this region has distinct functions. The posterior part of ACC is primarily engaged by motor tasks while the more anterior portions are particularly engaged when emotions are involved. The areas associated with mentalising are clearly anterior to the motor region of anterior cingulate cortex.

Temporo-parietal region: Increased activation in the junction between parietal and temporal lobes has been observed using a story comprehension task and static cartoons (Gallagher et al., 2000). Again this area was highly active in response to stimuli which share properties of biological motion. Bonda et al. (1996), for example, reported activity in the left caudal–most part of the superior temporal sulcus when viewing grasping hand movement compared to random movement. Puce, Allison, Bentin et al. (1998) found increased superior temporal sulcus activation when viewing faces in which eye gaze repeatedly changed direction, and faces in which the mouth opened and closed. Similarly Calvert, Bullmore, Brammer et al. (1997) observed increased activation in a region of the superior temporal gyrus during silent lip-reading of numbers versus still lips, and Grezes, Costes and Decety (1999) reported activation of the superior/middle temporal region during viewing of meaningful hand gestures with tools and objects compared to stationary hands. Taken together these studies implicate the superior temporal sulcus and adjacent cortex in the perception of a variety of human body movements. This region is anterior and superior to the visual motion area MT/V5 (Puce et al., 1998), indicating that these activations are not attributable to movement *per se*. It is notable, too, that all our animations (including Random) displayed self-propelled movement as might be expected of animate agents. Our triangles, when described as moving purposefully and intentionally, activated the key brain regions that have been activated by viewing biological motion. Human-like face or body characteristics thus do not appear to be necessary to trigger the attribution of mental states. Future investigations are needed to clarify what particular properties of biological motion are functionally associated with temporo-parietal activation, and whether distinct regions respond preferentially to specific visual attributes of biological stimuli.

Basal temporal cortex: The ToM animations also elicited bilateral activation in the basal temporal region, with peak components in the caudal part of the fusiform gyrus and in the temporal poles adjacent to the amygdala. Baron-Cohen et al. (1999) reported increased activation in the amygdala region during a mentalising task involving judgement of a person's eyes, as well as activation in medial prefrontal cortex and the temporo-parietal region. Connections between these areas are known to be strong (Amaral, Price, Pitkanen and Carmichael, 1992). Temporal pole activation has previously been associated with narratives (Mazoyer, Tzourio, Frak et al., 1993; Fletcher et al., 1995) and this fits with the idea that subjects inferred the scripts underlying ToM animations. These animations had certainly more narrative content compared to the other sequences.

The studies of biological movement perception discussed above, also reported peaks of activity in left fusiform gyrus and left temporal pole in response to observing meaningful hand gestures compared to stationary hands (Grezes et al., 1999). Left fusiform gyrus activation was found during observation of a dancing human figure compared to random movement (Bonda et al., 1996). The ventral temporal area has also been implicated in visual processing of static stimuli: while reading words and naming pictures (e.g. Vandenberge, Price, Wise et al., 1996) and while reading Braille words, versus letter-strings (Buchel, Price and Friston, 1998). Several imaging studies have reported specific regions of the fusiform gyrus to be more active during face viewing compared to assorted pictures, hands, scrambled faces and houses (e.g. Kanwisher, McDermott and Chun, 1997), and more active during face than letter-string and texture perception (e.g. Puce et al., 1996). Gorno-Tempini, Price, Josephs et al. (1998) reported increased activity in bilateral temporal poles associated with famous and non-famous face and proper name processing. Activity in bilateral fusiform gyri was increased while processing faces relative to names and scrambled faces. Thus different areas of the fusiform gyrus appear to be specialised for recognition of different kinds of objects, including animate agents.

Occipital cortex: In the present study, the ToM animations (relative to Random) elicited increased bilateral occipital activation in a lateral area, as was also found in Gallagher et al.'s (2000) study using a mentalising task involving static cartoons. In contrast, the reverse comparison (Random versus ToM) activated a medial region of occipital cortex. This result indicates a task specific effect not found in other studies of mentalising that deserves further exploration. These regions lie outside the region of the extrastriate cortex identified as retinotopically organised (Mendola et al., 1999). In addition, they are implicated in recent studies of global and local processing of complex visual stimuli (Fink, Halligan, Marshal et al. 1997a; Fink, Halligan, Marshal et al. 1997b). In the Fink et al. study (1997a) subjects were presented with large letters made out of small letters, and required to switch attention between global and local perceptual levels. Attentional modulation between local and global processing was associated with differential activity in prestriate cortex along the medio-lateral axis. Local processing elicited increased left lateral activation, whereas global processing elicited increased right medial activation. This distinction between lateral and medial occipital regions was replicated in a second study using objects rather than letters (Fink et al., 1997b): local processing elicited increased lateral activation, while global processing elicited increased medial activation. It is notable that the comparisons between our ToM and Random animations showed similar differential activity: lateral during ToM stimuli, and medial during Random stimuli. An important difference between the method used in the present study and in Fink et al.'s studies is that the latter reported peaks of activation associated with global and local processing resulting from a 'top-down' (endogenous) process. Subjects were specifically instructed to attend to the stimuli at either the global or the local level, whereas in our study subjects were not instructed how to view the stimuli. It makes some intuitive sense, however, that participants may have attended to global patterns of movement in the, effectively meaningless, Random condition (floating, bouncing), and paid more attention to the specifics of movement, interaction and character details (e.g. which way a triangle is pointing) in the ToM scenarios.

Taken together, these studies suggest that occipital sites may be implicated in the perception of movement patterns that engage attention at different (local-global) levels relevant to the attribution of animacy and intention. Although this speculation is unsupported with regard to the present animations, it is amenable to empirical testing.

Of necessity, the movements in the ToM animations were more complex in terms of greater variation of speed and direction of movement. It may be this greater complexity that results in increased activity in extrastriate regions. Thus, it remains possible that the pattern of activation we attribute to mentalising reflects in part extraneous tasks differences in, for example, psychophysical properties of the stimuli or resulting eye-movement differences. Future tests in which psychophysical properties are systematically varied, are clearly needed.

In conclusion, the present study has shown that abstract movement patterns activate regions previously associated with mentalising in stories and static pictures. Our ToM animations revealed increased activation in a network of brain regions, including the medial prefrontal cortex, the temporal pole adjacent to the amygdala region, and the temporo-parietal junction. All these regions have been repeatedly implicated in previous studies of mental state attribution and might reflect different components of this process. Two particularly important components, paracingulate sulcus and temporo-parietal junction, show overlap with previous mentalising studies as well as studies of self-monitoring and perception of biological motion. We tentatively suggest that the ability to make inferences about other people's mental states evolved from the ability to make inferences about other creatures' actions and movements. This fits with the observation that we commonly infer intentions on the basis of observed action outcomes. The activity of the prefrontal cortex and temporo-parietal junction in our study is combined with activity in a ventral visual pathway, from the extrastriate cortex to the inferior and middle temporal gyri. Thus the regions activated by viewing artfully animated triangles appear to reveal a network for processing visual-kinetic information about intention in action.

Chapter 5

Autism and the brain mechanisms for the attribution of mental states to animated shapes

This chapter presents a neuroimaging study with positron emission tomography (PET) in which ten high-functioning adult volunteers with autistic disorder and ten normal adult volunteers were scanned while watching the animations used in the previous study. The simple, non-verbal stimuli were designed to evoke selective attribution of mental states and purposeful actions in the absence of human-like features and verbal cues. This task, based on the ability to attribute mental states by perceiving a agents' kinetic property alone, aimed to bypass learned strategies and tap real life impairments in high-functioning individuals with autism. Free descriptions of the stimuli sequences were elicited following scanning. The study therefore allowed investigation of ToM abiliy at both the behavioural and the biological levels of the causal model of autism.

5.1 Neuroimaging studies on autism

Overwhelming evidence indicates that autism results from abnormal brain development, which is likely to stem from genetic factors (Happé and Frith, 1996). Much information has been obtained from neuroimaging and neuropathological studies in relation to structural brain abnormalities associated with autistic symptoms, but there is, as yet, no significant convergence of results to suggest what is specific and universal in autism (Gillberg and Coleman, 2000).

Increasingly interest has turned to investigating brain activity associated with social impairments in high-functioning individuals with autism (including Asperger Syndrome, the subgroup without language or cognitive delay). Functional neuroimaging studies of social cognition in autism fall into two broad types: those addressing attribution of mental states, and those addressing processing of faces without an explicit requirement for mentalising. I will briefly summarise the paradigms and the results of four imaging studies on autism, all of which have indicated functional abnormalities compard to control groups. Interestingly, the nature of these abnormalities seems to depend upon the task being performed, suggesting that the abnormal activity is the secondary consequence of primary pathology located elsewhere.

5.1.1 Studies on mentalising

To date there are two published studies investigating the neural substrates of mentalising in high-functioning people with autism (Happé et al., 1996; Baron-Cohen et al., 1999). In a PET study, Happé et al. used a language based paradigm, comparing brain activation in five individuals with autism and six controls. Subjects were scanned while reading stories and answering questions requiring inferences about complex mental states (Theory of Mind, or ToM, stories) or non-mental inferences ("Physical" stories), against a baseline of reading and remembering unconnected sentences. During ToM stories, both groups showed increased bilateral activation in the temporal pole (Brodmann Area 21) and the left temporoparietal junction (BA22/39). The groups' activation patterns differed in the medial prefrontal cortex: the autism group showed no activation than controls in BA 8/9 and increased activation more ventrally (BA 9/10). The authors suggested that high-functioning people with autism solve social problems in a different way from controls, and do not have access to the dedicated neuro-cognitive system for mentalising seen in normal adults. In a fMRI study, Baron-Cohen et al. (1999) compared a group of six adults with autism with a group of twelve controls. Subjects were asked to judge inner states from photographs of the eye region alone, deciding which of two simultaneously presented words best described their mental/emotional state. The baseline condition involved judging gender from the eyes. Compared to the control group, people with autism demonstrated less extensive activation in the left inferior frontal gyrus (BA 44/45) and right insula, and no activation in the amygdala. The autism group showed greater activation in the superior temporal gyrus (BA 22) bilaterally. The authors suggested that the amygdala is an important part of the neural basis of social behaviour, and is abnormal in autism. The rather different findings of these two studies may reflect the very different types of stimuli and tasks adopted; further studies are clearly needed.

5.1.2 Studies on face processing

Distinct from Theory of Mind, but of relevance to social cognition in autism, neuroimaging studies have also begun to explore face processing in autism. Two recent fMRI studies have adopted different tasks based on either the perception of facial expression of emotions, or neutral faces (Critchley, Daly, Bullmore et al. 2000; Schultz, Gauthier, Klin et al., 2000). Critchley et al. scanned nine people with autistic disorder during explicit and implicit processing of facial expressions displaying anger and happiness compared to neutral expressions. Subjects either judged the emotional content of the faces (explicit task) or the gender of the faces (implicit task). Subjects with autism differed significantly from controls in the activity of cerebellar, mesolimbic and temporal lobe cortical regions of the brain when processing facial expressions. In particular, they did not activate the fusiform gyrus during the explicit task, and the left amygdala region and left cerebellum during the implicit task.

Since subjects with autism made significantly more errors than controls in discriminating the two emotional expressions and their performance was correlated with face recognition memory errors, a more generalized deficit in processing faces is suggested. In Schultz et al.'s study, fourteen participants with autism spectrum disorder were scanned while discriminating between pairs of pictures of non-expressive faces, pairs of familiar objects or pairs of patterns. The autism group showed greater activation than controls in the inferior temporal gyrus, and reduced activation relative to controls in the fusiform gyrus, while processing faces. The control group showed greater activation in the inferior temporal gyrus during the object discrimination condition. The authors suggest that people with autism process faces using feature-based strategies that are more a kin to non-face object perception. In sum, in both studies activity was seen in a region of fusiform gyrus widely accepted to be specialized for the perception of faces (Kanwinsher et al., 1997), and this activity was significantly lower in both autistic groups. The autistic groups showed greater activation than control in adjacent regions of temporal cortex, but the precise location of these regions was different in the two studies.

5.2 A neuroimaging study on individuals with autism

At a behavioural level, mentalising ability in individual with autism has been tested with various paradigms. However, as already reported in chapter 1, a high proportion of individuals with autism do pass Theory of Mind tests. Thus, it seems that these high-functioning people, usually with superficially good language skills, acquire a Theory of Mind with time and experience. Nevertheless, they show persistent social difficulties just as other individuals with autism. Thus, is it plausible that individuals with autism do not fully overcome the inability concerning reading people's minds by adopting compensatory strategies. Learned, explicit processes seem to be not sufficiently fine-grained to compensate the impairment at an automatic level. In fact, according to the metarepresentational model (Morton and Frith, 1995) the ability of mentalising is primarily implicit and not a result of learning processes. In particular, it is a property of the system that comes into action when triggered by particular stimuli and allows understanding other's and one's own behaviour fully automatically.

The aim of the present study was to examine brain activation in individuals with autism during on-line processing of moving agents in the absence of either verbal stimuli or visual depictions of humans. By minimizing the factors which are likely to promote ToM tasks' performance but not necessary real life social adaptation, the paradigm allows for exploring the gap between mentalising tests performance and social impairment in high-functioning people with autism.

Unlike the two previous studies on mentalising in autism, inferences concerning mental states in the present study were based solely on the perception of movement patterns. The seminal work of Heider and Simmel (1944), described in details in the previous chapter (section *4.2.1*), revealed the pervasive human tendency to attribute mental states, even to simple shapes in motion. These stimuli, and similar animations, have been used with normally developing children and those with autism (Abell et al, 2000; Bowler and Thommen, 2000; Klin, 2000). For example, the study by Abell et al. (described in detail in section *4.2.3*), found that even children with autism who passed standard false belief tests, used mental state descriptions less appropriately than normally developing and intellectually-impaired children, in response to animations designed to elicit mental state attributions. In contrast, animations designed to display goal-directed and random movements did not discriminate the groups. The same animations were used in a PET study with six normal adults (described in the previous chapter, section 4.3). Consistent with the Abell et al. study, the group of adults used more mental state descriptions for the Theory of Mind animations than for the goal-directed or random animations. Increased regional cerebral blood flow related to mental state attribution (subtraction and correlational analyses) was found in medial prefrontal cortex, temporal parietal junction (superior temporal sulcus), fusiform gyrus and temporal poles adjacent to the amygdala, and extrastriate cortex (occipital gyrus). All these areas, with the exception of the occipital cortex, have been implicated in a prior study of mentalising (Gallagher et al., 2000) regardless of the modality of the stimulus input, visual or verbal.

This study reports an extension of the neuroimaging study described in the previous chapter, to examine the neural correlates of mental state attribution in high-functioning adults with autism. Adults with high-functioning autism and controls were scanned while watching the three different types of silent animations. These depicted a large and a small triangles moving in a self-propelled fashion. The three conditions are briefly summarized as follows:

- In the four Theory of Mind (ToM) animations the movement of the two interacting characters suggests that one triangle anticipates or manipulates the mental state of the other (e.g. the little triangle tries to deceive the big one).
- In the four Goal-directed action animations (GD), the interaction between the two triangles evokes description primarily in behavioural terms (e.g. the two triangles are dancing together).
- In the four Random animations (Rd) the purposeless movement of the two triangles elicits description with no reference to intentions or specific goals (e.g. bouncing around).

Both the quality of the verbal descriptions given to these animations and the related pattern of brain activation in high-functioning adults with autism were the focus for the present study. Unlike the behavioural study with children (Abell et al., 2000), subjects watched the animations without any suggestion as to the story content or the characters' roles. After each scan they were asked to describe what had happened in the animation. Based on the previous studies, it was predicted less accurate use of mental state explanation in the autism group, despite their high IQ and success on standard Theory of Mind tasks. Furthermore, differences in brain activation within the network associated with mentalising tasks (Castelli et al., 2000; Gallagher et al., 2000) were predicted. In particular, it was expected to replicate Happé et al's (1996) finding of differential activity in medial prefrontal cortex using very different stimuli (silent animations versus written stories) and more stringent analytic techniques (random versus fixed effects model).

5.2.1 Method

Subjects: The autism group consisted of 10 adults (mean age = 33 years, sd. 7.6) diagnosed with autistic disorder or Asperger Syndrome according to DSM-IV criteria (American Psychiatric Association, 1994). Their high level of functioning was reflected by their education, social independence and employment: all were living independently or semi-independently, 8 had completed an undergraduate degree or further education courses; 8 had a regular job (Table 5.1).

The present study specifically targeted people with high-functioning autism or Asperger Syndrome, and no additional neurological abnormalities, to ensure that the group was able to understand the experimental procedure, undertake the tasks, and give informed consent for the PET scanning. The control group consisted of 10 subjects recruited from university students and staff (mean age 25 years, sd. 4.8). Table 5.2 shows chronological age, IQ and performance on standard Theory of Mind tests for the two participant groups. Subjects were tested with both verbal test (The Quick Test, Ammons and Ammons, 1962) and non-verbal test (Raven Standard Progressive Matrices) IQ tests. The two groups did not differ in verbal ability (percentile mean=61, sd.24 for autism group, mean 76, sd. 11 for controls) or non-verbal ability (percentile mean 73, sd. 30 for autism group, mean 88, sd. 9.4 for controls). The Sally-Ann test (Baron-Cohen, Leslie & Frith, 1985) and Smarties test (Perner, Frith, Leslie & Leekham, 1989) assessed the ability to understand a first-order false belief ("she thinks *x*"). The Ice-Cream story (Perner and Wimmer, 1985) and Birthday Puppy test (Sullivan, Zaitchik and Tager-Flusberg, 1994) examined the ability to attribute a second-order false belief ("she thinks that he thinks *x*"). The two groups did not differ on standard false belief test: six of the autism group and 8 controls passed all 4 tests, 1 autistic and 2 control subjects passed 3 out of 4 test, and 3 autistic subjects passed only the 2 first order tests. Table 5.3 shows single subjects'performance on IQ tests and ToM tests.

Table 5.1: Social profile, occupation and education of participants with autistic disorder.

Subject No	CA(yrs)	Social independence	Occupation	Education
1	25	Independent	Manual	University
2	39	Semi/independent	None	Secondary
3	46	Semi/independent	None	Secondary and further
4	24	Independent	Assistance	Secondary
5	37.5	Dependent	Technical	Secondary and further
6	28.5	Independent	Technical	Secondary and further
7	38	Dependent	Manual	Secondary
8	31	Independent	Manual	Secondary
9	23	Semi/independent	Art	University
10	35	Independent	Manual	Secondary and further

Table 5.2: Groups chronological age, non-verbal and verbal IQ tests score and Theory of Mind tests performance.

Group	CA	Nonverbal IQ (Raven test)		Verbal IQ (Quick Test)			False belief tests*
	Y:M	Max score=60	Percentile	Max score=50	Percentile	IQ	Max score=2
Autism							
mean	**32.8**	**47.5**	**72.2**	**43.7**	**61.1**	**105**	**1.5**
sd	(7.7)	(9.6)	(30.0)	(3.9)	(23.7)	(10.6)	(0.5)
Control							
mean	**24.4**	**54**	**87.3**	**46.1**	**76.5**	**112**	**1.8**
sd	(4.11)	(4.7)	(15.1)	(1.4)	(11.2)	(5.9)	(0.4)

*Note: Standard false belief test score: 0= fail 1st and 2nd order; 1= pass 1st and fail 2nd order; 2= pass 1st and 2nd order.

Table 5.3: Individual subjects' performance on IQ tests and ToM tests.

	Group	CA Y	non-verbal IQ Percentile	verbal IQ Percentile	ToM 1st order (max=2)	ToM 2nd order (max=2)
1	autism	25	95	55	2	2
2	autism	39	75	--	2	1
3	autism	46	75	60	2	2
4	autism	24	25	20	2	2
5	autism	37.5	95	90	2	2
6	autism	28.5	95	75	2	1
7	autism	38	95	60	2	2
8	autism	31	90	85	2	1
9	autism	23	10	30	2	0
10	autism	35	75	75	2	1
1	control	23	95	85	2	1
2	control	23	95	70	2	2
3	control	28	75	90	2	2
4	control	30	95	90	2	2
5	control	21.5	95	85	2	2
6	control	21.5	90	75	2	2
7	control	20	95	60	2	1
8	control	26	75	75	2	2
9	control	21	75	60	2	2
10	control	35	95	75	2	2

Design: A 3x2 repeated measures within subjects design was used. Four different examples of each of three types of animation, ToM, Goal-directed, and Random were displayed in a semi-random order over the course of 12 scans, divided into two consecutive counterbalanced blocks: cued and uncued animations, to be consistent with the previous study reported in chapter 4. In the cued condition only, subjects were told in advance what kind of animation was going to be displayed (see procedure) .

Materials: Twelve animations were used during the scanning, and an additional three were shown for practice. A full description of the animations is reported in chapter 4.

Procedure: The procedure was identical as in the previous study (reported in chapter 4).

Scoring: The verbal descriptions given after each presentation (in between scans) were transcribed verbatim (transcriptions are reported in Appendix 5A) and coded in terms of three different dimensions. Scoring criteria are the same of the previous study (details are given in Appendix 4B). The aim of the scores was to distinguish in each answer 1) the implied "intentionality", that is, the degree of attribution of mental states, 2) their appropriateness, that is, how well the underlying script was captured, 3) the length of each answer. Since the degree of hesitation present in subjects' tone of voice while describing the animations was found a difficult factor to assess in the first study, the present analysis did not rate the Certainty score.

Two raters independently scored each verbal description, reaching agreement in 94% of the cases for the Intentionality score (Kappa value=92). In particular, they agreed for the ToM animations score in 82% of the cases in the control group (Kappa value=.70) and 95% of the cases in the autism group (Kappa value=.93). The Goal-directed animations received exactly the same score in the case of the control group (Kappa value=1.0) and 95% agreement in the autism group (Kappa value=.93). The Random animations were given the same score in 93% of cases in the control group (Kappa value=.80) and 88% of cases in the autism group (Kappa value=.50). On the Appropriateness score the two raters reached full agreement except in two cases in the autism group (99.2%). The Length score was given by only one rater.

Neuroimaging Data Acquisition: All subjects underwent both PET and MRI scanning on the same day. A Siemens VISION (Siemens, Erlangen) operating at 2.0T was used to acquire axial T1 weighted structural MRI images for anatomical coregistration. A full description of the $H_2{}^{15}O$ PET activation technique and data analysis is reported in chapter 4 of the present thesis.

Neuroimaging Statistical analysis: Data were analysed with statistical parametric mapping (using SPM99 software from the Wellcome Department of Cognitive Neurology, London, UK; http//www.fil.ion.ucl.ac.uk/spm) implemented in Matlab (Mathworks Inc.Sherborn, MA, USA) using standardised procedures (Friston et al., 1995 a,b), including realignment for head movements, spatial normalisation to the Montreal Neurological Institute template brain (Evans et al., 1994) in the space of Tailerach and Tourneaux (1988) and smoothing. The smoothing kernel was a 3D Gaussian filter of 16 mm. Condition and subject effects were estimated according to the general linear model at each voxel. To test hypotheses about regionally specific condition effects, these estimates were compared using linear compounds or contrasts. The resulting set of voxel values for each contrast is an SPM of the t- statistic.

A Random effects analysis was carried out in order to evaluate common and differential areas of response (Friston, 1992) in the autism and control groups during processing ToM animations versus Random animations, Goal-directed animations versus Random animations, Random versus ToM animations, and during the Cue condition versus the No-cue condition. Since in the random effects model the variance estimate is between-subject rather than within-subject, and the degrees of freedom are related to the number of subjects rather than the number of scans, a single mean image of the contrast of interest was first generated for each subjects, and then three main analyses were carried out.

- A main effect analysis allowing for identification of regions that were activated for both groups.
- A conjunction analysis of the common activity with the differential activity of the two groups allowing for identification of differential responses of individuals with autism and controls.

- An analysis of functional connectivity (using the measures of functional connectivity available in SPM99 for fixed effects models) to identify significant differences in connectivity between the two groups.

The activated areas reported below as main effect for processing ToM versus Random animations consisted of voxels that survived a voxel-wise multiple comparison correction of $p < 0.05$. The only exception is represented by activation in prefrontal cortex ($p < .0001$ uncorrected) that was specifically predicted on the basis of previous studies of mentalising tasks.

5.2.2 Results

Behavioural data: The ratings of three verbal description scores (Intentionality, Appropriateness and Length) were analysed with nonparametric tests, since these were in the form of ordinal data. Table 5.4 shows the total scores for each type of animation split into groups and Figures 5.1-5.3 show the results in graphic form.

The effect of the animation type on the subjects' ratings was analysed with a nonparametric two-way analysis of variance (Friedman test) conducted on each of the three scores. The results revealed that for both groups, descriptions differed in degree of intentionality (chi-square=31.3 $p<.0001$), appropriateness (chi-square=10.3 $p <.006$) and length (chi-square=20.1 $p<.0001$), according to the animation type. Nonparametric paired comparisons (Wilcoxon test) revealed that ToM animations evoked attribution of more mental states (i.e. higher Intentionality scores) than either the Goal-directed ($z =3.4$, $p<.001$) or Random animations ($z=3.9$, $p<.0001$), and that fewer mental state terms were attributed to the Random than to the Goal-directed animations ($z=3.9$, $p<.0001$). The analysis of the scores within group (Wilcoxon test) revealed that the Intentionality score of the autism group was the same for the ToM and the GD animations ($z=1.4$ $p= 0.16$), whereas it was higher for ToM than GD animations ($z=2.8$, $p=.005$) in the control group. The Wilcoxon test on the Appropriateness score revealed that the ToM animations were described less appropriately than either the Random ($z=3.2$, $p=.002$) or the Goal-directed animations ($z=2.3$, $p<.02$), but that this effect was entirely attributable to the autism group (ToM$<$GD $z=2.7$, $p<.01$; ToM$<$Rd $z=2.7$, $p<.01$). The same analysis on the Length score revealed that the descriptions of the ToM animations were longer than those for either the Random ($z=3.6$, $p<.001$) or the Goal-directed animations ($z=3.3$, $p<.001$).

Group comparisons, with Mann-Whitney test, revealed that subjects with autism used mentalistic terms significantly less than did controls in response to the ToM animations ($z=3.3$, $p<.001$), did not differ either in the Random and Goal-directed animations. They also gave significantly less appropriate descriptions for the ToM animations ($z=3.8$, $p<.001$) than for the two nonmentalistic animation conditions. There was no significant difference between groups in the length of their descriptions.

The Intentionality and Appropriateness scores of each group were analysed separately to investigate the effect of being alerted with regard to the nature of the animations. Table 5.5 shows the scores split into Cue/No-Cue conditions in both groups. Only the autism group showed any difference in the Intentionality score. However, the results indicated only a trend ($z=1.7$, $p=.08$) of increased use of mental states terms across all animations when cued (mean=2.1, sd=0.6) versus when not cued (mean=1.8, sd=0.8). The control group showed no difference in the use of mental state descriptions for the cued and uncued animations (mean=2.3, sd=0.4). No differences were observed for the Appropriateness score: the autism group mean scores were 1.1 (sd=0.3) and 1.2 (sd=0.4) respectively for the cued and uncued animations; control group mean scores were 1.9 (sd=0.3) and 1.8 (sd=0.4) respectively.

All 10 of the control participants, and 8 of the autism group, gave descriptions of the animations that included mention of mental states, as reflected in high Intentionality ratings (4 or 5). The autism group was distinguished, however, by giving in many cases descriptions that referred to inappropriate mental states (reflected in low appropriateness ratings, 1 or 2). Across the 8 autism participants who gave high Intentionality descriptions, 3 participants gave Intentional descriptions rated as inappropriate in 50% of cases, and 2 gave inappropriate descriptions in every high-intentionality case. Thus, for the majority of the autism group who ever used mental state explanations, these often referred to inappropriate mental states. By contrast, inappropriate mental state explanations were extremely rare among the control group; just one control participant produced 7 high intentionality descriptions, of which 2 (14%) were rated as inappropriate.

Table 5.4: Groups verbal descriptions for Theory of Mind, Goal-directed, and Random animations, rated on three dimensions. Mean and standard deviation (in parentheses)

Score type (range)	**Group**	**Animation Type** ToM	Goal-directed *Mean score (sd)*	Random
Intentionality (0-5)	Autism	**2.8** (1.1)	**2.4** (1.0)	**0.7** (1.2)
	Control	**4.1** (0.7)	**2.4** (0.9)	**0.4** (1.0)
Appropriateness (0-3)	Autism	**0.5** (0.2)	**1.3** (0.2)	**1.5** (0.5)
	Control	**1.7** (0.2)	**1.7** (0.3)	**1.8** (0.4)
Length (0-4)	Autism	**2.5** (1.2)	**2.1** (1.3)	**2.0** (1.0)
	Control	**2.8** (1.1)	**1.9** (0.9)	**1.6** (0.8)

Note: **Intentionality total mean (sd) score:** ToM=3.4 (1.1), GD= 2.4 (0.9), Rd= 0.6 (1.1)
Significant group effect: Control>Autism in ToM (z = 3.3, p< .001)
Appropriateness total mean (sd) score: ToM= 1.1 (0.7)), GD=1.5 (0.5), Rd=1.6 (0.5)
Significant group effect: Control>Autism in ToM (z = 3.8, p<.001)
Length total mean (sd) score:: ToM=2.7(1.1), GD=2 (1.0), Rd=1.8 (0.9)
No significant group effect

Figure 5.1: Groups' verbal descriptions for Theory of Mind, Goal-directed, and Random animations, rated with Intentionality score

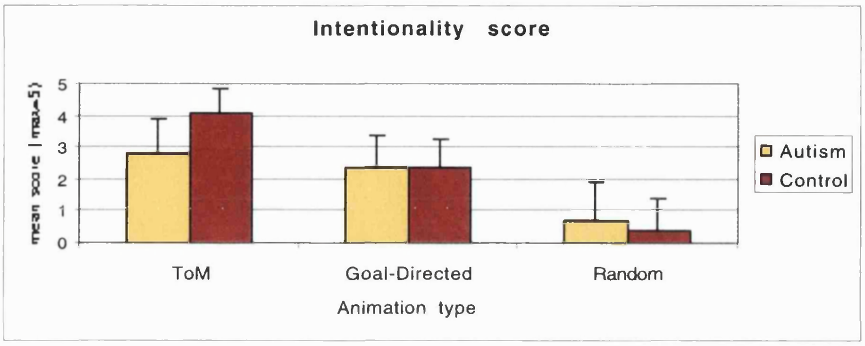


Figure 5.2: Groups' verbal descriptions for Theory of Mind, Goal-directed, and Random animations, rated with Appropriateness score.

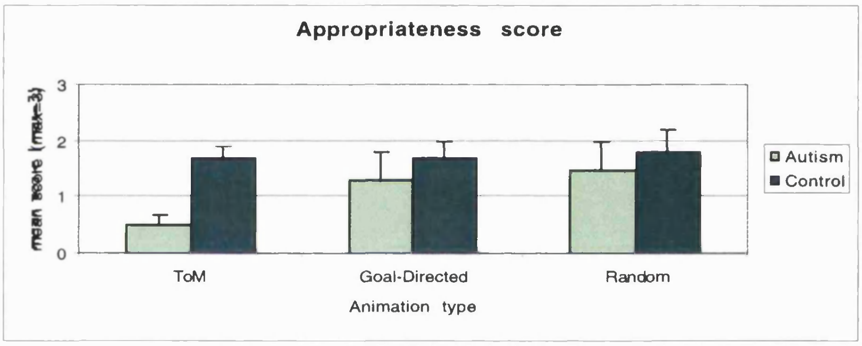
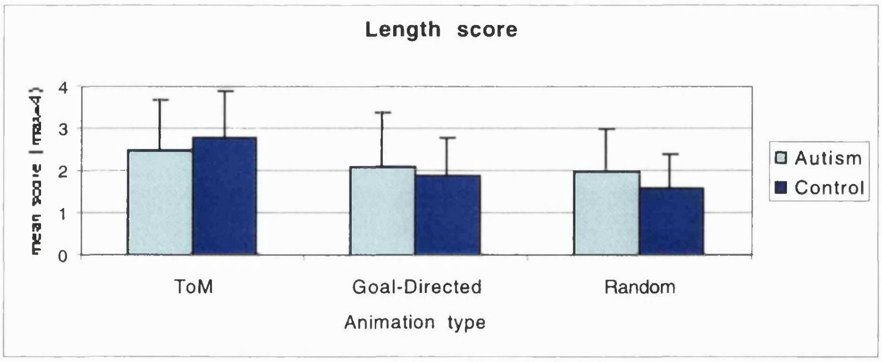


Figure 5.3: Groups' verbal descriptions for Theory of Mind, Goal-directed, and Random animations, rated with Length score.



*Notes: The spontaneous descriptions for ToM animations were rated as reflecting less mental states attribution for the autism group than the control group ($z=3.3$, $p<.001$) and as reflecting less appropriate understanding of the story line for the autism group than the control group ($z=3.8$, $p<.001$). No other group differences were significant. Intentionality score was higher for ToM than GD animations ($z=3.4$, $p<.001$), and higher for GD than Rd animations ($z=3.9$, $p<.0001$). The autism group did not differentiate ToM and GD in terms of attribution of mental states ($z=1.4$, $p=0.16$), and described GD animations more appropriately than ToM animations ($z=2.7$, $p<.01$).

Table 5.5: The effect of the Cue and No-Cued condition. Before the Cue condition subjects were alerted with regard to the nature of the three animations: "interaction with feelings and thoughts" (ToM), "random movement" (Rd), or "simple interaction" (GD). Before the No-Cue condition subjects were told that they were going to see the next animation. Mean and standard deviation (in parantheses)

Group	Score type (range)	Cue Condition	
		CUE	NO-CUE
		Mean (sd)	
Autism	**Intentionality** total score (0-5)*	**2.1** (0.6)	**1.8** (0.8)
	Appropriateness total score (0-3)	**2.1** (0.3)	**2.2** (0.4)
Control	**Intentionality** total score (0-5)	**2.3** (0.4)	**2.3** (0.4)
	Appropriateness total score (0-3)	**2.9** (0.3)	**2.8** (0.4)

*Note: The autism group showed a trend of significant increase of mental states attribution when cued ($z=1.7$, $p= .08$).

Neuroimaging data:

i) Random effects analysis of Theory of Mind animations versus Random animations:

The autism group and the control group activated a network of four regions during ToM contrasted with Random animations (Table 5.6 and Figures 5.1 a,b,c,d): extrastriate cortex (inferior occipital gyrus, see Figures 5.6a,b), basal temporal area (inferior temporal gyrus extending to anterior fusiform gyrus and temporal pole adjacent to amygdala, see Figure 5.6ab), superior temporal sulcus (STS, see fig. 5.6c) at the temporo-parietal junction, and medial prefrontal cortex (SFG, see Fig. 5.6d). Within this network, direct comparison between the groups (Table 5.7) revealed significantly reduced activation in subjects with autism in the following four regions: basal temporal area, STS and the prefrontal area. The extra-striate regions were activated to the same extent in both groups.

ii) Random effects analysis of Random animations versus Theory of Mind animation:

The main effect of this comparison elicited more activity in the medial region of the occipital cortex:

-8 x, -92 y, 16 z; z score =5.6, p corrected for multiple comparisons =.001

-12 x, -76 y, 6 z; z score =5.0, p corrected for multiple comparisons =.015

14 x, -84 y, 16 z; z score =5.3, p corrected for multiple comparisons =.006

This finding confirmed that of the previous study, discussed in chapter 4, and will not be considered in this chapter. Other contrasts were not significant, that is, Goal-Directed animations, which mainly elicited descriptions of simple intentional actions and sometimes but not necessarily mental states descriptions could not be distinguished from the other two conditions in terms of activation. I will therefore consider only the contrast between ToM and Random animations.

iii) Random effects analysis of Cue versus NoCue condition, and NoCue versus Cue condition:

The main effect of these contrasts revealed no significant increased activation in neither contrasts.

Table 5.6: Peaks of activation in the autism group and the control group, during perception of ToM animations versus Random animations. Note: Z scores (p-value <.0.05 corrected for multiple comparisons in bold, p-value <0.001 uncorrected in plain text). Brain regions are identified by name and by putative Brodmann Area (BA) Brain regions are identified by name and by putative Brodmann Area (BA) on the basis of the atlas of H.M. Duvernoy (1999).

Left Right Medial	Foci of common activation Autism and Control	Co-ordinates x	y	z	(Z) score p <
	Basal Temporal area				
L	ITG (BA37)	**-46**	**-60**	**-10**	**(5.5) .002**
L	FuG (BA20)	-38	-14	-30	(4.5) 0001
R	TmP/Am (BA38)	42	6	-28	(4.2) .0001
	Temporo-parietal junction:				
R	STS (BA 22)	**64**	**-48**	**16**	**(5.6) .001**
L	STS (BA 21/22)	**-58**	**-52**	**4**	**(5.4) .003**
	Extra-striate cortex				
R	IOcG (BA 18; V3)	**22**	**-104**	**-8**	**(5.0) .015**
L	IOcG (BA 18; V3)	**-18**	**-106**	**-10**	**(5.0) .02**
R	IOcG (BA18; LO)	**42**	**-82**	**-8**	**(4.8) .04**
L	IOcG (BA18; LO)	**-26**	**-94**	**-12**	**(4.8) .03**
	Prefrontal area				
M	SFG (BA9)	10	54	30	(3.4) .0001

BA = Brodmann Area
ITG = Inferior Temporal Gyrus
TmP/Am = temporal pole adjacent to amygdala
FuG = Fusiform Gyrus
STS = Superior Temporal Sulcus
IOcG = Inferior Occipital Gyrus
SFG = Superior Frontal Gyrus

Figure 5.4a,b,c,d: Regions of significant cerebral blood flow (rCBF) change associated with the perception of ToM animations versus Random animations. (a)**Top figure**: Saggital view of activation in extrastriate cortex (inferior occipital gyrus) and basal temporal area (inferior temporal gyrus extending to anterior fusiform gyrus and temporal pole adjacent to amygdala). (b) **Middle figure**: Horizontal view of activation in extrastriate cortex and inferior temporal gyrus. (c) **Bottom/left figure**: Lateral view showing activation in superior temporal sulcus (STS) at the temporo-parietal junction. (d) **Bottom/right figure**: Saggital view showing activation in medial prefrontal cortex (SFG).

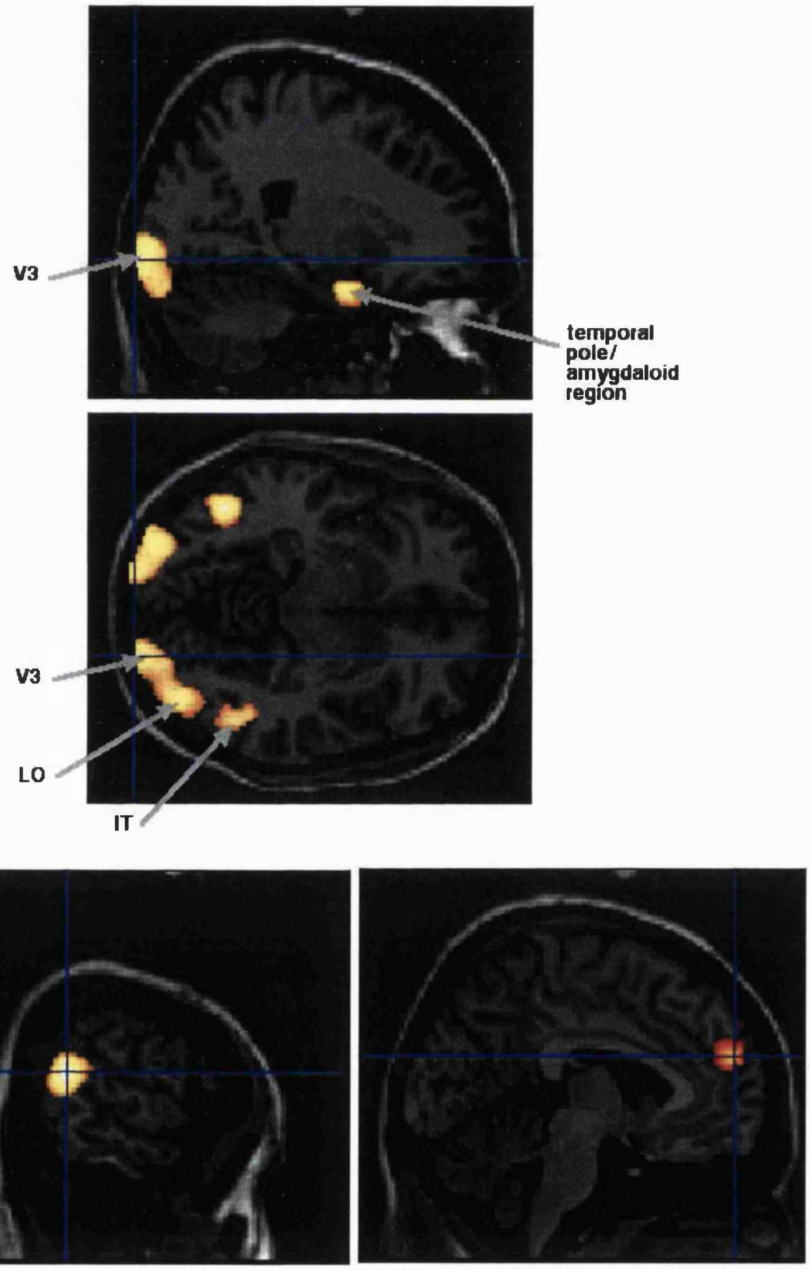


Table 5.7: Peaks of reduced activation in the autism group relative to control group, during perception of ToM animations versus Random animations.

Left Right Medial	Foci of reduced activation Autism<Control	Co-ordinates x	y	z	(Z) score, p <
	Basal Temporal area				
L	FuG (BA20)	**-38**	**-14**	**-26**	**(5.3) .004**
R	TmP/Am (BA38)	**42**	**6**	**-28**	**(6.2) .0001**
	Temporo-parietal junction				
R	STS (BA 22/40)	**52**	**-46**	**24**	**(4.8) .04**
L	STS (BA 21)	**-66**	**-52**	**8**	**(4.9) .02**
	Prefrontal area				
M	SFG (BA 9)	-4	56	22	(4.5) .0001

Note: Analysis was restricted to the areas shown in table 5.4 (p-value <.0.05 corrected for multiple comparisons is shown in bold text, p-value <0.001 uncorrected is shown in plain text).

iv) Connectivity analysis:

The comparison of the groups identified brain regions where ToM animations elicited as much or more activation in the autism group. However, this activation was clearly not sufficient for a full understanding of the scenarios, as shown by the behavioural data. It was hypothesized that these regions were not interacting appropriately with the rest of a larger network, which showed reduced activation. Therefore the connectivity of these areas with the rest of the brain was investigated using the measures of functional connectivity available in SPM for fixed effects models. The results are shown in Table 5.8. The extra-striate region showed significantly less connectivity with the STS in the autism group.

Table 5.8: Connectivity analysis. The right extra-striate cortex (volume of interest of 6mm radius) shows reduced connectivity with superior temporal sulcus in the autism group versus the control group.

Foci of significant connectivity Coordinates: x, y, z		**(Z) score, p <**
IocG	**STS**	
24, -100, -10	66, -46, 4	**(5.05) .004** voxel level
	68, -56, 18	(3.44) .001 uncorrected
	-68, -46, 0	(3.29) .001 uncorrected

5.2.4 Discussion

Behavioural results: The present experiment investigated brain activity in highly functioning individuals with autistic disorder during an on-line mentalising task based on visual perception of kinetic patterns. The three silent animations evoked spontaneous verbal descriptions of the interaction displayed by two geometrical shapes. The behavioural data show evidence for persistent mentalising deficit in able adults on the autism spectrum who pass standard false-belief tests. These findings parallel those reported by Abell et al. (2000) for children with autism, normally developing children and adults. In both studies, the ToM animations triggered in the control groups the tendency to attribute intentional states to the moving triangles. By contrast, people with autism failed to fully understand the interactive movement of the two characters in the ToM animations where it was necessary to appreciate that behind actions there are specific beliefs and intentions. However, they performed as well as the controls where it was necessary to take into account what each triangle was "doing" (Goal-directed animations) rather than "thinking" (ToM animations). They had no difficulties in understanding that some moving patterns (Random animations) were not prompted by any particular intentional stance. Interestingly, the autism group did not differentiate the ToM animations from the Goal-directed ones indicating their tendency to decode even complex social interactions in terms of simple actions (e.g. the "Seducing" animation was described as: "Big red [triangle] was a bully to the little, trying to stop the little one to get out. At the end the little escaped".) Another finding analogous to the Abell et al.'s is that individuals with autism used mental state terms, but often used these inappropriately, that is, they misinterpreted the underlying script of the ToM animations (e.g. the "Surprise" animation was described as: "[...] the one outside (little triangle) breaks in and was happy to see the other"; " [...] the blue triangle [...] came up to the wall and realized it was a door, tapped it and moved to the side. The red one came up to the door opened it and had a look outside".)

The result is consistent with other studies using advanced mentalising tasks based on the interpretation of static visual stimuli or language (Baron-Cohen et al., 1997; Happé et al., 1996; Gallagher et al., 2000; Klin, 2000). It is worth noting that the difficulties of people with autism in describing the ToM animations cannot be ascribed to deficits in face processing or emotion recognition, or mentalistic terminology, since the present paradigm was designed to elicit mental states attribution in absence of any human display of emotions or mental states and mentalistic terminology. Kinetic stimuli have been recently used by Klin (2000). In his study, young adults with high-functioning autism or Asperger Syndrome were shown the original Heider and Simmel (1944) animation and asked to provide narratives that were coded along a 7-indeces coding system. In line with the present findings, the descriptions of the autism group included a significant lower number of pertinent mental states terms than those given by the control group. One possible explanation of the persistent difficulty in on-line mentalising is that individuals with autism use an alternative strategy based on a "coarse" analysis of the displayed action patterns without reaching the more "fine grained" analysis of the kinetic properties that signal uniquely a particular intentional state. This compensatory strategy would thus result in a lack of mental state descriptions or in an inappropriate use of them. Future research is clearly needed to investigate the consequences of the fragility of the mechanism that enables keeping track of mental states and various components of agents' kinetic patterns such as purposeful action, biological movement and cause-effect movement.

In conclusion, it seems that the present real-time mentalising task taps real life difficulties in understanding complex social interactions and it can be added to the repertoire of advanced mentalising test with particular advantage of adopting the "fine-cuts" methodology (Frith and Happé, 1994; Happé and Frith, 1996). Indeed, the three very similar type of animations of which only one requires mentalising ability provide a within subjects control for extraneous factors that also affect task performance besides mentalising requirements.

Neuroimaging results: The aim of the study was to locate pattern of brain activity associated with mentalising in high functional individuals with autism using non-verbal stimuli. The findings, based on the largest sample to date and a more stringent statistical analysis than ever before - random effects model - confirm the previously found network which is active while watching silent animation evoking the attribution of intentional states as opposed to nonintentional movement (Castelli et al., 2000). As in previous studies of mentalising (Goel et al. 1995; Fletcher et al., 1995; Gallagher et al., 2000; Brunet et al., 2000), which are summarized in the discussion of chapter 4 of the present thesis, greater activity was seen for ToM than for Random animations in medial prefrontal cortex, temporal pole and STS. New to this study, and probably task specific, were activations observed during mentalising in extra-striate regions of occipital cortex. The more medial and posterior of these extra-striate regions can probably be identified with V3 while the more lateral may be part of the lateral occipital complex, LO (S. Zeki, personal communication). Area V3 is responsive to form and motion (Felleman, 1987) and has inputs dominated by the magnocellular processing stream (Felleman, 1997), while area LO is involved in the early stages of object recognition (Malach, 1995). The greater activation of these regions suggests that the ToM animations were more visually demanding than the Random animations, despite our attempts to control for physical characteristics of the stimuli. However, the basic movement parameters seem to have been well controlled since ToM animations did not elicit more activity in V5/MT, the visual movement area.

The group of subjects with high-functioning autism showed a different pattern of activation in the mentalising network, with less activity than the control subjects in the following components: bilateral superior temporal sulcus at the temporo-parietal junction, the basal temporal area (left fusiform gyrus and right temporal pole adjacent to amygdala) and the medial prefrontal cortex. This last component also showed reduced activation in autism during ToM story comprehension in an earlier study (Happé et al., 1996) while lack of amygdala activation was found during another mentalising task

(Baron-Cohen et al., 1999). Reduced activation in the region of the amygdala was also observed in subjects with autism while they processed facial expressions implicitly (Critchley et al., 2000).

Given that the subjects with autism were impaired in making correct mental state attributions and in distinguishing between goal directed movement and actions driven by mental states, one might expect to see a general reduction in activity associated with the mental state scenarios. However, although a reduction was seen in areas previously associated with ToM tasks, we did not see a reduction of activity in the extra-striate regions that were specifically associated with our ToM animations. This is evidence that, in early visual processing stages, brain activity in able subjects with autism just as in controls, was greater for these more visually demanding animations. However, despite the detection of this greater complexity, this perceptual information failed to reach the multi-modal brain systems that are associated with mentalising regardless of task. STS (superior temporal sulculs) in particular showed reduced connectivity from extra-striate regions.

This region of STS has been activated in a number of previous studies when subjects observed biological motion (Puce et al., 1998; Bonda et al. 1996; Allison et al., 2000). The region is probably the homologue of area STP (superior temporal polysensory area) in the macaque. STP contains cells with large receptive fields, which also respond to biological motion (Perrett et al., 1989). This region of STS is one of the major targets of extra-striate visual areas and "is in a unique position to integrate motion, spatial and object information" (Boussaoud et al., 1990). In addition, in the macaque, STS also has strong reciprocal connections with the basolateral amygdala and adjacent regions of temporal pole (Amaral et al., 1992).

It seems plausible that the difficulty experienced by our autism group in understanding the ToM animations occurred because important information about the motion of the triangles was failing to be transmitted from V3 to STS. But what is the cause of this transmission failure? The location of the reduced activity associated with autism in previous studies depends upon the nature of the task being performed. When faces must be processed less activity is seen in the fusiform "face area", when biological motion must be processed less activity is seen in STS. A single impairment consistent with both observations would be a failure of feedback signals from the anterior components of the mentalising system (amygdala, temporal pole and medial prefrontal cortex) which normally indicates the social significance of the signals being processed. This would be consistent with the evidence that structural abnormalities in the brains of people with autism are likely to be found in these anterior regions (Abell et al., 1999; Howard et al., 2000; Bauman et al., 1994). It has been proposed that in autism the processing of face information may be compromised due to weak feed-back connections from amygdala to fusiform gyrus, which in turn are due to the developmental effects of lack of signals for the emotional importance of faces (Shultz et al., 2000). In the present experiment, lack of feedback from temporal pole and/or medial prefrontal cortex to STS would mean that the social significance of the moving triangles would not be recognised. This is in line with the suggestion that amygdala's feedback "induces attentional amplification of STS activity evoked by salient social stimuli" (Amaral et al., 1992).

Evidence for such top-down effects is provided by work using single cell recording (Sugase et al., 1999). Activity in cells in the inferior temporal cortex was measured while Macaque monkeys looked at faces or geometrical shapes. The initial activity in the cells simply reflected whether the monkey was seeing a face or a shape, while the later occurring activity also distinguished between facial expressions. The authors suggest that these different processing modes over time reflect intra-area contributions and feedback from higher-level processing areas, necessary for the finer grain analysis of expressions. Likewise, in humans, as shown with intracranial ERPs, context can enhance visual processing by late top-down modulation of temporally earlier activity in visual cortex (Olson, 2001). The weaker connectivity in autism between these areas may reflect a lack of top-down modulation from more anterior regions such as the amygdala and surrounding temporal pole and/or medial prefrontal cortex.

In conclusion, the present study confirmed the hypothesis that able individuals with high-functioning autism read minds differently: they gave less accurate interpretations specifically of those animations that elicited mentalising, even without the need to process facial expressions or other human cues. The controls, on the other hand, were highly accurate in inferring the putative mental states of triangles from movement cues alone. These data parallel those obtained in previous studies and suggest that continuing impairments in individuals with autism are revealed in characteristic inaccuracies in mental state attribution. While viewing ToM animations, in contrast to randomly moving shapes, people with autism showed reduced activation in several regions of a previously identified mentalising network, but showed normal activation in extrastriate cortex. In the autism group this extrastriate region showed reduced functional connectivity with the superior temporal sulcus at the temporo-parietal junction, an area associated with the processing of biological motion as well as mentalising.

Chapter 6
General discussion

The thesis has described a series of experimental studies designed to test the Theory of Mind deficit hypothesis for autism with non-language based paradigms. In this chapter I will attempt to draw together the findings in the context of a model - the metarepresentational model proposed by Leslie (1994) - by summarizing and discussing them in relation to the hypothesis of Theory of Mind deficit in autism, namely, the inability to represent mental states.

6.1 Summary of the background

The general premise of this work is that much cognitive development is based on dedicated domain-specific mechanisms as opposed to a domain-general intelligence and learning capacity. The issues concerning these contrasting positions are discussed in Hirschfeld and Gelman (1994). The domain-specific or "modularity" approach is essentially multidisciplinary and seeks to investigate the nature and scope of specific cognitive abilities, their evolutionary origin, their development and their effects on culture (Sperber and Hirschfeld, 1999; Sperber, 2000). The debate concerning the modularity of the Theory of Mind ability focuses essentially on whether or not this ability has an innate basis, and to what extent environment shapes its development. It is important to underline that these issues, which are discussed in detail in a recent paper by Scholl and Leslie (1999), are not tested in the present studies and do not effect the interpretation of the findings. The focus of the thesis was to investigate the nature of on-line processing cues that were not mediated by language abilities, and to investigate sensitivity to these inputs in individuals with autism. There are several reasons why I found the metarepresentational model particularly interesting for investigating Theory of Mind deficit in autism. First, this model is endowed with the heuristic power of bringing both normal and abnormal patterns of development within a single neuropsychological explanatory framework. Second, it allows to make a clear distinction between competence and performance in autism. Third, the model implies that the ability to attribute mental states is triggered by specific inputs, which needed to be investigated. Finally, the modularity approach for understanding human cognition, which the model stems from, is particularly suitable for research in neuroscience and neuroimaging. I will briefly recapitualte the origin and the structure of the metarepresentational model described in chapter 1, then I will discuss the model in the light of the present findings.

6.1.1 Autism and a multi-level metarepresentational model

The concept of Theory of Mind, which has its roots in philosophy of mind, refers to people's pervasive tendency to explain their own and others' actions in terms of beliefs, desires and goals. The development of this ability is severely compromised in the case of individuals with autism. The Theory of Mind deficit hypothesis of autism originated from a model of normal social cognitive development proposed by Leslie (1987). His model aimed at explaining a fundamental question concerning the mentalising ability, namely: "How does the brain of young children attend to mental states when mental states cannot be seen, heard, or felt?" In more philosophical terms, mental states are "opaque", that is, they may be distinct or even contrasting and yet refer to the same state of affairs, they may be directed upon a nonexistent state of affairs or objects and yet they can be true or false.

In Leslie's *metarepresentational* model (1987), it was suggested that children attend to behaviour and infer mental states from which the behaviour arises. For the child's brain to move attention from behaviour to the mental states from which the behaviour arises, it is necessary to postulate a system of representation capable of representing mental states. According to the metarepresentational model, a specialised cognitive mechanism has the competence of providing agent-centered interpretation of behaviour by constructing metarepresentations. This mechanism operates post-perceptually and spontaneously whenever an agent's behaviour is attended. How is the structure of this system?

According to Leslie (1994), intentions and propositional attitudes concern two different levels of metarepresentation. A lower, "minimal" level of metarepresentation allows for interpreting the behaviour of agents in relation to events that are at distant times and places (future state of affairs). In other words, the lower metarepresentational level is concerned with the agent's intentions. By contrast, the higher, "full-fledged" level of metarepresentation allows for interpreting the behaviour of agents in relation to states that are beyond spatiotemporal circumstances, namely, mental states, or propositional attitudes.

Baron-Cohen (1994, 1995) proposed a more "fine-grained" model, comprising four distinct levels, or systems, developing at different stages. For the purpose of the present work, I suggest that these distinct components can be seen as belonging to two levels of metarepresentational ability, which varies in terms of complexity. In fact, both the systems dedicated to the representation of goals (the Intentionality Detector, or ID) and of eye movements (the Eye-Direction Detector, or EDD) deploy simple, dyadic representation of an agent's volitional state (e.g. agent *wants* 'x'), or of an agent's perceptual state (e.g. agent *sees* 'x'). More complex representations, with a triadic structure, are built by the mechanism (Shared Attention Mechanism, or SAM) dedicated to the representation of the volitional or perceptual state of both the self and another person towards the same object (e.g. I *see* an -agent *wants* 'x'-). A different type of representations with a triadic structure, opaque representations, are built by the higher component of the metarepresentational system, the Theory of Mind mechanism (ToMM) (e.g. agent *believes* "x"). It is therefore plausible, on the basis of the type of representation, with a dyadic or triadic structure, deployed by each system, to refer to the ID and EDD components as "low" level metarepresentational mechanisms, and the SAM and ToMM components as "high" level metarepresentational mechanisms. Furthermore, Baron-Cohen suggested that the system processes inputs within its components in a "cascade" fashion, and in particular, that ToMM is triggered by the triadic representations of SAM. Thus, if the latter component is impaired, then the former will also necessarily be impaired (Baron-Cohen and Swettenham, 1996). The findings in this thesis are entirely compatible with either the Leslie's model (1994) or the Baron-Cohen's model (1994, 1995). Since I made no distinction of metarepresentation ability at the fine-grain level of the Baron-Cohen's model, I will henceforth refer only to a metarepresentational ability with a minimal level and a full-blown level.

The different experiments carried out in this thesis explored the ability of individuals with autism to understand mental states in relation to a "lower" and a "higher" level of the metarepresentation model, using non-language based triggering stimuli. The link between specific types of input and Theory of Mind deficit hypothesis of autism is explained in the next sections.

6.1.2 Autism and ToM tasks

One problem that this thesis confronted concerns the fact that a proportion of individuals with autism, high-functioning people usually with superficially good language skills, are successful in passing false-belief tests – the litmus test for ToM ability - despite evidence of real-life difficulties in understanding what others think and feel. In general, Theory of mind research on autism has mainly been applied to high-level cognitive processing, such as understanding that other people can have different beliefs and desires from one's own. Verbal ability appears to be a key factor for a successful performance on language-based ToM tasks without necessarily promoting real-life social adaptation.

The aim of this thesis was to create experimental mentalising test with non-verbal stimuli. If Theory of Mind failure persists even in able individuals with autism, then it should be possible to investigate the sufficient and necessary cues that may trigger this ability.

6.1.3 Investigating triggering inputs of ToM ability

The starting point of the thesis was to make a clear and simple theoretical distinction of different mental states and then to identify potential inputs for the attribution of different types of mental state. Thus, if different types of mental states exist, then it might be the case that humans attend to different and specific aspects of an agent's behaviour in order to understand and predict its actions.

Mental states hold the property of being "intentional states", namely, they *point towards* certain objects or state of affairs, but these objects or state of affairs need not exist. All mental states are intentional states. These include propositional attitudes (beliefs and desires), emotional states, and intentions. Their difference can be expressed in terms of their intentional contents. Beliefs and desires point towards the truth of a proposition whereas emotional states point towards a state of affairs. Intentions, or goals, point towards a *future* state of affairs. As I argued in chapter 1, the most important distinction is between desire and intention. The simplest form of intention implies an action, whereas desire implies always an attitude towards a proposition.

The definition of emotional states as intentional states is more elusive than the other intentional states. Emotions can be either caused by autonomic responses to states of affairs or by specific propositional attitudes. For the purpose of investigating visual triggering inputs for the metarepresentational mechanism, I have considered emotional states at a basic level. Emotional states are characterized as rapid and fail-safe responses to stimuli that are correlated with basic survival needs. One characteristic of these emotions is that they are clearly visible on the face of a person. They provide therefore salient stimuli to investigate the ability to attribute mental states to others in autism. The first three experiments of this thesis investigated the ability to recognise basic emotions from humans' facial expressions. The questions to be answered were therefore whether facial expression of basic emotions constitutes a sufficient and necessary cue for the metarepresentational system, and furthermore, whether a test based on such stimuli indexed impaired neuropsychological functions.

The other paradigms used in this thesis are based on the perception of kinematic cues. In fact, the display of human or human-like features, e.g. puppets, cartoon characters, does not constitute the minimal requirement for a triggering input of the metarepresentational system. The social environment in which we live provides us with abundant visual stimuli that are in motion. We understand mental states by reading micro-movements of the facial muscles or the changes in body posture, or the motion of the whole body in relation to another body or object. By stripping an agent of face and body features we are left only with one of the simplest forms of visual information for judging an agent's mental states: its motion trajectory. Thus, by watching an agent's kinetic properties alone, we are prompt to attribute intentionality. Again, the question to be answered concerns the sufficient and necessary triggering properties of kinematic cues. What type of motion features the metarepresentational system attends to? Furthermore, to what extent are high-functioning individuals with autism sensitive to different types of motion cues? In the next section I will summarize and discuss the main findings that were presented in this thesis.

6.2 Overview of findings and conclusions

6.2.1 Understanding emotions from facial expression

The aim of the first three experiments was to investigate the perceptual and semantic abilities in children with autism to recognise basic emotional states of others through their facial expressions.

The first experiment investigated the ability to discriminate facial expressions of all six basic emotions (anger, disgust, fear, happiness, sadness and surprise). It was based on the very simple task of sorting and matching pictures depicting different levels of emotional expressions of a man with the expressions of a woman. The aim of the second experiment was to investigate children's semantic ability to discriminate emotions from a wide range of adults' expressions. The aim of the third experiment was the same as that of the second, using more difficult stimuli, namely, facial expressions with different levels of emotional intensity. The study used fine-grained visual stimuli depicting facial expressions of all six basic emotions (anger, disgust, fear, happiness, sadness, surprise) derived from a standard set of pictures of facial effect (Ekman and Friesen, 1976; Calder et al., 1996a; Young et al. 1997).

The three experiments tested the prediction that children with autism have difficulties in recognising only the emotions that are triggered by propositional attitudes (surprise) and have no difficulties with emotions that are triggered by states of affairs (anger, disgust, fear, happiness and sadness). In addition, the wide range of stimuli allowed for monitoring children's performance in relation to the amygdala hypothesis of autism. This hypothesis suggests a correlation between amygdala abnormality and socio-affective impairments. Since two recent studies testing the amygdala hypothesis (Howard et al., 2000; Adolphs et al., 2001) reported contradictory findings on fear recognition in high-functioning adults, the additional purpose of the present investigation was to observe children's performance on fear recognition.

The study revealed that children with autism were as able as controls to recognise all six basic emotions from facial expressions. This was shown not only when they were required to match pictures of emotional expressions with different intensity levels, but also when they were asked to provide a label for expressions with a normal intensity.

Two possible explanations for these negative findings were considered: a) autistic individuals are able to bypass the impairment by compensatory strategies, and b) they have no impairment in recognizing emotion displays that have evolved with adaptive functions, but have difficulties in linking the perceptual level of emotion recognition with the higher level of understanding the social meaning of different expressions. Future research on emotion processing in autism should be oriented towards investigating very young children possibly with measures that bypass explicit strategies.

Interestingly, these negative findings are inconsistent with Howard et al.'s study showing concomitant evidence of amygdala abnormality and selective fear recognition impairment in a group of high-functioning individuals with autism. Since a study with amygdala patients (Adolphs et al., 1999) showed that individuals' deficit in fear recognition ranged from extremely impaired to almost normal, it is plausible that Howard et al.'s study showed impairment in only a particular subgroup of high-functioning autism. Hence, a replication of these findings, along with a description of individual subjects' performance is needed to gain more information on the link between amygdala functioning and fear recognition impairment in autism. In particular, it would be of interest to investigate the incidence of a subgroup of high-functioning individuals with autism with specific fear recognition deficits.

Finally, the present negative findings on recognition of basic emotions are consistent with Adolphs et al.'s (2001) study on individuals with autism and patients with amygdala damage indicating that both groups pass basic emotion recognition tests but fail to understand the mental states of others, and to predict their behaviour on the basis of their faces. It is plausible that compensation strategies, which bypass the amygdala, support basic emotion recognition tasks but do not support tasks based on high-cognitive processing demands, such as mentalising. Future research should investigate the role of the amygdala in the normal development of ToM ability with neuroimaging methods, and the effect of amygdala lesions on mentalising performance by testing large samples of amygdala patients.

A general issue of this study on emotion recognition concerns the link between recognising facial expression of basic emotion and the different levels of metarepresentational ability. In particular, it is unclear whether a test based on such stimuli can index or not impaired neuropsychological functions.

The ability to recognise basic emotions constitute adaptive behaviour: the rapid comprehension and prediction of another agent's behaviour depend on the ability to understand signal of danger or of survival. Compared to complex emotional states, basic emotions constitute a short and fixed list of representations. On the basis of a representation of an agent and a state of affairs, animals are able to detect a source of, for example, danger, by identifying the expression of fear in a conspecific. There is no need to understand the agent's *attitude* towards the mental state's content, or, in other words, to deploy an opaque representation. Thus, it seems that in order to decode the meaning of basic emotional expressions, it is sufficient to possess a rudimentary, minimal, metarepresentational ability. Emotional expressions are sufficient triggers for the lower level of metarepresentation, and individuals with autism are sensitive to such stimuli. Hence, the mentalising deficit in autism may not involve a deficit in detecting basic emotion from facial expressions. What about other types of triggering input?

6.2.2 Understanding intention from an agent's goal-directed motion

The fourth study of the thesis investigated whether children with autism are able to represent intentionality at a lower, "minimal" level of the metarepresentational system. In particular, would they be able to represent an agent's goal-directed behaviour without necessarily involving the higher, "full-blown" cognitive process of representing the agent's beliefs and desires?

A new paradigm was created using computer animated sequences depicting a small circle rolling up and down a valley trying to reach one of two targets resting on top of either sides of the valley. Subjects were presented "on-line" with powerful visual cues: a constant direct motion of the agent towards a target, and the agent's accidental outcome.

The study had two distinct aims: the first one was to investigate the ability to attribute an agent's intended goal in the presence of its unsuccessful outcome. The second aim was to explore the developmental changes, from childhood to adulthood, in the ability to attribute an agent's intended goal in the presence of an ambiguous outcome: the agent changes its motion direction and lands at the opposite side of where the goal-directed motion originally "pointed-towards". It is up to the subject to decide whether the agent reached or failed to reach its goal. The type of answer provided by children and adults reflected a different type of representation for attributing intended goal to an agent: (a) a spatially-based representation when subjects decided the intended goal on the basis of its final outcome next to a target, regardless of its persistent motion towards the target opposite the outcome, and (b) a motion-based representation when subjects decided the intended goal on the basis of its persistent motion toward a target that was accidentally missed.

The findings of the first part of the study revealed that children with autism were as able as controls to attribute an intended goal to an agent in the presence of its unsuccessful outcome. Thus, the minimal level of metarepresentation, which employs representations of an agent's mental states but lacks access to the content of an agents' attitude, is not impaired. In a way, this finding is not entirely surprising, since an unimpaired ability to understand goal-directed actions is essential in making sense of people's behaviour caused by non-contingent behaviour. After all, all human actions are driven by intentions, except for actions based on physiological changes. Interestingly, the findings indicated that the perception of an abstract agent in motion constitutes a sufficient trigger for representing intended goal, and individuals with autism are sensitive to such stimuli. However, since the metarepresentational mechanism operates automatically whenever an agent's behaviour is attended, it was particularly interesting to investigate the two distinct triggering visual inputs: the agent's persistent motion and its outcome. More specifically, the aim of the second part of the study was to determine whether or not the agent's persistent motion was a necessary cue to represent an agent's intended goal.

Results showed a developmental change in the type of goal-directed representation between the age of six years and adulthood, from an outcome-based representation to a persistent motion representation. The shift between young age and adulthood was rather smooth, since children between the age of seven and fifteen years did not show any preferential bias towards one way or the other representation. Children with autism, older than six-years, represented the indented goal on the basis of the final outcome of the agent rather than its persistent attempts towards a target. In addition, they valued the agent's sudden change of direction more than the repeated motion in one particular direction. Future investigation is needed to clarify the role of two other triggering inputs of the metarepresentational mechanism, namely, the proximity of the agent to the target, and the sudden change of the agent's direction.

In conclusion, the paradigm based on the perception of an agent's goal-directed motion was sensitive enough to capture a developmental delay, which indicates that children with autism may have a difficulty in executive function processing, which impacts on their metarepresentational ability. However, they showed no difficulty to represent an agent's goal-directed intention, indicating no impairment at the minimal level of the metarepresentational system. Finally, the present study indicated that the perception of an agent in motion constitutes a sufficient triggering input for the metarepresentational system, but an agent's persistent motion is not a necessary cue for the attribution of intended goal.

6.2.3 Understanding mental states from agents' complex patterns of motion

The last two experiments of the thesis were focused on the full-blown ability to metarepresent agents' behaviour. This ability allows the representation of a richer and more flexible repertoire of mental states, namely, it allows understanding mental states in an "opaque" way.

The last paradigm adopted in the two neuroimaging studies was based on the perception of the motion patterns of two interacting agents. The different animations evoked descriptions of the agents in mentalistic terms or in behavioural terms. If we assume that different types of input trigger different levels of the metarepresentational mechanism, then it can be said that the complex patterns of the Goal-Directed animations were triggering inputs for the lower metarepresentational level, whereas the ToM animations were triggering inputs for the higher-level. However, of necessity, the motion trajectories of the ToM animations were more complex than the Goal-Directed animations in terms of greater variation of speed and direction of movement. Further investigations are required to make a distinction between kinetic properties of agents' movement that are necessary to evoke different types of metarepresentations.

The aim of neuroimaging techniques is to link cognitive processes to their implementation in terms of brain activity. The framework for investigating how the brain works is based on the principles of functional segregation and functional integration: different areas of the brain have different functions, and brain function depends on the interaction between areas. Hence, the model of brain function adopted to interpret neuroimaging results is akin to that of "neural networks", which assumes distribution of function over several brain regions. Indeed, even if mentalising is a domain-specific cognitive ability, it should be possible to localize a distributed brain system with the technique of neuroimaging.

The first neuroimaging study, reported in chapter 4, with a group of healthy individuals was aimed at identifying the normative neurocorrelates of mentalising ability. A brain system dedicated to mentalising was localized in a network of brain regions, including the medial prefrontal cortex, the temporal pole adjacent to the amygdala region, and the temporo-parietal junction. All these regions have been repeatedly implicated in previous studies of mental state attribution, self-monitoring and perception of biological motion. Simple localization of different components, however, is not enough to have a clear picture of the brain mechanism underlying ToM (Frith and Frith, 2000). Further research is needed to investigate the functions of the different areas comprising the mentalising system by gaining and integrating information from different neuroscience disciplines, e.g. single-cell studies in non-human primates.

The second PET study, reported in chapter 5, investigated brain activity in a group of high-functioning individuals with autism while watching the silent animations. Furthermore, the spontaneous descriptions of the animations provided an interesting behavioural result on their ability to appreciate that actions may also be driven by complex mental states, and not only by goals. The analysis of verbal descriptions confirmed previous findings that able individuals with high-functioning autism read minds differently: they gave less accurate descriptions than controls specifically for those animations that elicited mentalising, even without the need to process facial expressions or other human cues. In addition, the autistic participants indicated that they were not able to discriminate between goal-directed movements and movements driven by mental states. At a very speculative level, it can be said that a spontaneous strategy for autistic people who are impaired in mindreading would be to "minimize" their interpretation of people's behaviour at the level of goal-directed actions. Thus, following the terminology of Baron-Cohen's model, the level that deploys dyadic representations would be used inappropriately, rather than the level that deploys triadic representations. That is the equivalent of saying that one "weaker" mechanism compensates for the lack of functioning of a "stronger" mechanism. However, this suggestion begs the question of how the two mechanisms have developed - in parallel or in sequence - and how they are connected (*if* they are connected).

According to Leslie (1994), each level of the metarepresentational system constitutes a learning device with a specific way of organizing the inputs it receives. This suggestion implies that compensation strategies may occur within the metarepresentational system, and that the "shift" from impaired higher level and unimpaired lower level is likely to be graded across individuals. Hence, it should be possible to investigate the existence of sub-groups that would be sensitive to different properties of perceptual triggering inputs of the metarepresentational system.

In summary, the picture emerging from the analysis of the spontaneous descriptions of the high-functioning individuals with autism is compatible with the suggestion that their mentalising impairment is due to an inaccurate *modus operandi* of the metarepresentational system. Within the neuroimaging framework, this suggestion can be, to some extent, tested. In fact, according to the model of neural networks, if a system is impaired, then the *function* of the network is expected to be abnormal, rather than the structure of a single component of the network being "lesioned", or a single component "missing".

In line with this suggestion, the neuroimaging findings revealed that the autism group showed reduced activation in several areas of the identified mentalising network, but showed normal activation in extrastriate cortex. Furthermore, the analysis of the functional connectivity of these regions with the rest of the network, which showed reduced activation, revealed that in the autism group the extrastriate region was functionally less connected with the superior temporal sulcus at the temporo-parietal junction, an area associated with the processing of biological motion as well as mentalising. Future neuroimaging studies should investigate the neurocorrelates of systems involving the perception of biological motion with clear-cut paradigms, for example, contrasting the ability to understand an agent's bodily changes with an agent's motion, or different properties of motion itself, e.g. repetitive versus complex or exaggerated trajectories.

6.2.4 Final remarks

The work presented in this thesis showed that visual inputs that trigger the metarepresentational system are perceived differently in individuals with autism. The ability to represent mental states of others is not compromised at a minimal processing level, which involves attending to an agent's mental states by constructing dyadic representations of the relation between an agent and a proposition. This claim is supported by the findings that children with autism are not impaired in recognizing facial expression of basic emotions, and are not impaired in understanding an agent's intended goal by perceiving its movement pattern towards a target. However, the functioning of the full-fledged level of metarepresentation, which allows representing the agent's attitude towards the truth of a proposition, is compromised. This claim is supported by the findings that adult high-functioning individuals with autism were not able to interpret complex movement patterns of two abstract agents in terms of belief, desire, deception or pretence.

Little is known about the cognitive architecture of the system underpinning metarepresentational ability. It is plausible that distinct levels of metarepresentational ability exist. For the sake of theoretical simplicity, I've discussed the findings of the thesis in the light of two distinct levels. However, the heuristic advantage of postulating a system with a complex and flexible architecture is that it would be possible to test different hypotheses about the normal and abnormal functioning of different components of the system, their input and output organization. The ultimate goal of a research program would uncover the different evolutionary stages of the transition from a minimal to a full-bown metarepresentational system. That is, from the ability to represent a limited repertoire of others' behaviour to the ability to represent their mental states in a rich and flexible way, by employing representations of indefinitely varied contents. In the meantime, from a more limited perspective, it seems that a more fine-grained picture of the autistic mind has emerged from the present work. Since the impairments in real-life social interaction vary greatly in individuals with autism, it is important to work towards a full, detailed cognitive picture of autism by identifying impaired and intact abilities.

Appendix 2A

Computer-manipulated images on the Happiness-Anger continuum .

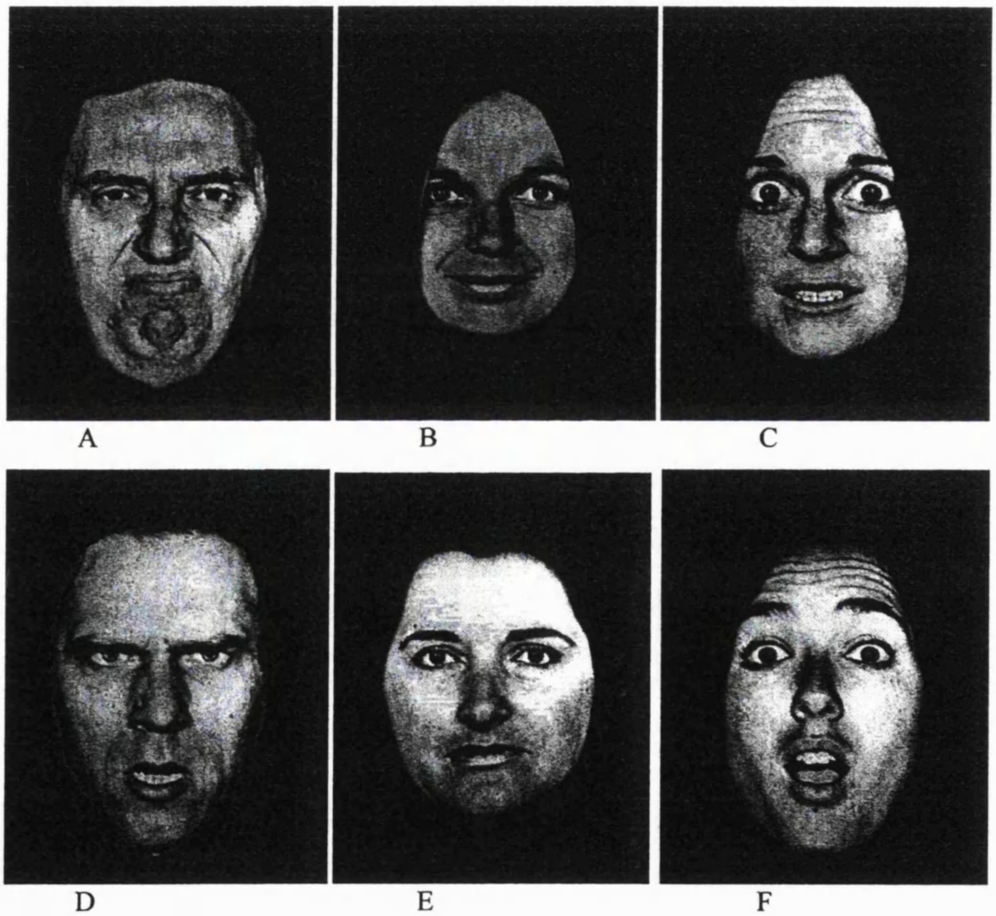
H=happiness, A= anger

90% H	**70% H**	50% H	30%H	10% H
10% A	30% A	50% A	**70%A**	**90% A**

Appendix 2B

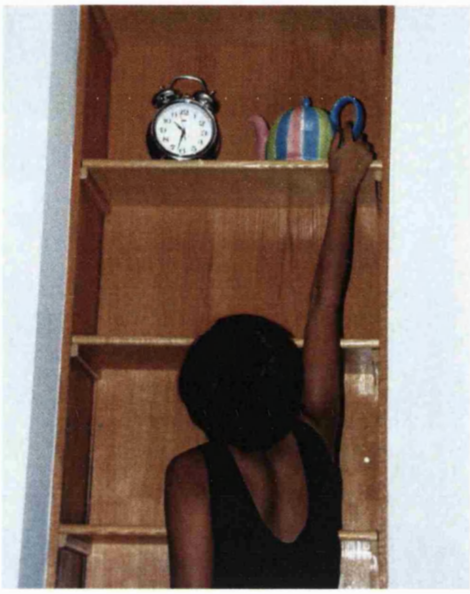
Sample of photographic-quality images of facial expressions with natural intensity

A = disgust
B = happiness
C = fear
D = anger
E = sadness
F = surprise

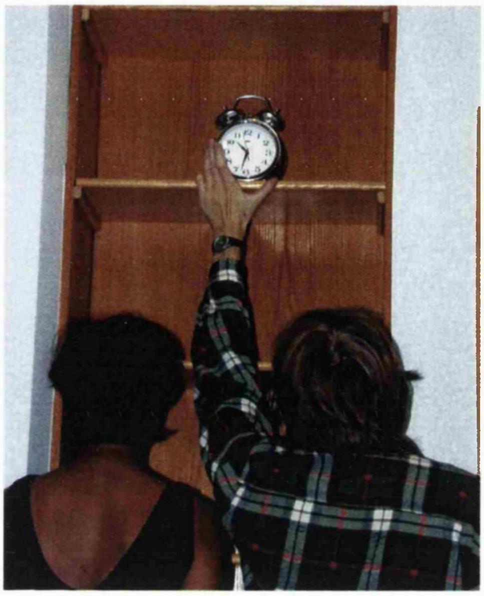


Appendix 3A

Examples of stimuli used in the “wanting test



what does she want?



who wants the clock?

Appendix 4A

Instructions given to participants

The aim of this experiment is to understand which parts of your brain are active while watching a short animated film sequence.

All you have to do is relax, and watch the animations shown on the monitor in front of you. Each animation lasts approximately 40 seconds. The sequences are similar to one another (two triangular shapes moving about) but different in their content. The triangles act as characters performing different movements, for example, dancing, drifting or courting each other.

There are different types of content: In some animations the behaviour of both triangles will appear disconnected from each other. They just move about, with random movement. By contrast, other animations will show the two triangles moving about doing something together, interacting. Their actions are somehow connected to each other, for example, they are imitating each other, or one is feeding the other. Still other animations show the two triangles doing something more complex together, as if they are taking into account their reciprocal feelings and thoughts. By just watching them you will probably imagine they are interacting, for example, courting each other.

In this experiment there is no 'right' or 'wrong' answer. Sometimes I will tell you in advance what kind of animation you are going to see, for example, a random movement, a simple interaction or an interaction involving thoughts and feelings. While you are watching the animations, be relaxed, and ... enjoy them! After each cartoon is over, I will ask you what you think the triangles were doing, whether they were randomly moving about, or whether they were doing something more specific.

Now I will show you some examples of animations you are going to see. If you have any questions, feel free to ask.

Appendix 4B

Scoring verbal descriptions

Score (0-5) for Intentionality:

0= action, non-deliberate.

1= deliberate action with no other.

2= deliberate action with another.

3= deliberate action in response to other's action.

4= deliberate action in response to other's mental state.

5= deliberate action with goal of affecting other's mental state.

Score (0-3) for Appropriateness:

0= no answer, "I don't know".

1= inappropriate answer: reference to the wrong type of interaction between triangles.

2= partially correct answer: reference to correct type of interaction but confused overall description.

3= appropriate, clear answer.

Score (0-3) for Certainty (based on voice tone):

0= long hesitation or silence.

1= hesitation, few words, sentences unfinished, need to be prompted to say more.

2= hesitation between words, alternative answers.

3= no hesitation at all, quick answer, description correctly reflects the script underlying the animation.

Score (0-4) for Length:

0= no response

1= one clause

2= two clauses

3= three clauses

4= more than three clauses

Appendix 4C

Examples of intentionality score for verbal descriptions

0= action, non deliberate.
e.g. :"Bouncing"; "Moving around"

1= deliberate action with no others.
e.g. : "Ice-skating"; "Playing"

2= deliberate action with somebody else.
e.g. : "Blue and red are fighting"; "Parent is being followed by a child"; 'They are playing together"

3= deliberate action in response to other's action.
e.g. : "Big is chasing the little one"; "Red is allowing Blu to get close to him"; "Big guarding the little who was trying to escape"

4= deliberate action with reference to mental states.
e.g. : "The little is mocking the big one"; "The little is mimicking the big one"; "Two people were arguing"; "One tells off the other"; "Parent encourages the child to go out"; "Teasing"; "They are happy together";"They are friendly"; "Big one wasn't interested…at the end he scared off the little one"

5= deliberate action with explicit goal of effecting other's mental state.
e.g. : "Blue triangle wanted to surprise the red one"; "Child pretended not to be doing anything"; "Little triangle tricks his way out".

Appendix 4D

Appropriateness score criteria

General rules:

Each description was scored 3, 2, 1,or 0 according to how accurately it reflected the sequence.

3= spot-on description of the story or the actions represented. It may be concise just capturing gist, as well as discursive.

2= partial description of the sequence: description is related to the sequence, but imprecise or incomplete..

1 = inappropriate answer: reference to the wrong type of interaction between triangles.

0 = no answer, "I don't know".

Random movement sequences:

Billiard, Drifting , Star, Tennis

3 = descriptions implying random or purposeless movement including moving about, bouncing off the walls or dancing as in dancing lights.

2 = purposeful movement without interaction, including turning round and getting dizzy, or dancing in a circle

1 = purposeful movement implying interaction between the triangles including copying each other

Goal directed movement sequences:

Chasing

3 = description that conveys the idea of a chase

2 = description that is related to but somewhat remote from chasing (e.g. following)

1 = action that does not relate to chasing

Fighting:

3 = action implying physical fight, e.g. bashing each other

2 = action that conveys the idea of a conflict, but is either too specific or too vague, e.g. biting; pushing

1 = action that does not relate to conflict, (e.g. following each other), or focus solely on a minor aspect of the sequence

Leading:

3 = description which conveys the idea of one leading the other or one following the other

2 = description that is related to but somewhat remote from following (e.g. copying; chasing)

1 = action that does not relate to following/leading, or focus solely on a minor aspect of the sequence

Dancing:

3 = description that conveys the idea of moving in formation, (e.g dancing; making a pattern)

2 = description that is partially correct or related to dancing, e.g. doing different things- one went one way – the other went the other way

1 = action that is not related to dancing, e.g. galloping along, or focus solely on a minor aspect of the sequence

Theory of Mind movement sequences:

Surprising:

3 = any mention of tricking, surprising, hiding, hide and seek

2 = description which gives part of the story but misses the critical point (see above)

1 = description not related to any of the events in the sequence, or focus solely on a minor part of action (e.g. knocking on the door)

Coaxing:

3 = description that conveys idea of little triangle's reluctance to go out *and* big triangle's attempts to get the little one out (e.g. persuading, coaxing).

2 = partially correct description focusing on one aspect of the story or one character only,
(e.g. little doesn't want to go out; or, big is pushing little to go out)

1 = actions that do not relate to the events or relate to a very minor aspect of the sequence only (e.g. the two triangles didn't like each other)

Mocking:

3 = description that conveys idea little triangle is copying big one with the intention of not being noticed (e.g. pretending, hiding, being naughty)

2 = partially correct description, (e.g. following, pursuing, copying)

1 = description that does not relate to the events (e.g. big triangle not interested) or relate to a very minor aspect of the sequence only (e.g. little triangle ran away)

Seducing:

3 = description that conveys the little triangle is trapped in and escapes by persuading, tricking the big one (e.g. Little convince in a seductive way to let him out)

2 = partial story with minimal action for each character, e.g. Little trying to escape

1 = description which is too minimal, e.g. she got out, or unrelated to the sequence.

Mental state attribution: use of mental state verbs to describe reciprocal interactions, e.g. wanting; hiding; tricking; pretending; being naughty;

NOT: complex goal-directed interaction, e.g. chasing each other round the house; A pushing B out of the way; NOT: solely direct speech, NOT: solely “trying to”.

Appendix 5A

VERBATIM VERBAL DESCRIPTIONS OF ANIMATIONS

Autism Group

- **Coaxing**
- Mummy bird was trying to encourage baby bird to fly on his own. Obviously the blue bird was scared, petrified and apprehensive...whatever you want to call it...So she wanted to force the blue bird out and she did. Eventually the blue bird went out but was prevented going back to the box. Eventually he went out on its own. He is still very apprehensive, and there is all sort of communication going on.
- They are rubbing noses and caressing each other, and they ended up holding hands.
- I don't know, the two triangles supposedly representing people, were kissing and dancing, and at the end they were making a knot glued… magnets, stuck together… magnets, stuck together.
- I really don't know…my best guess is that they were having a row
- The two ...uh... triangles were fighting each other. They obviously ...didn't like each other...they were…uh…one was following another to suggest...fight each other…and occasionally they ...later they clashed...it was quite...the other ones...they were not going on very well...they were obviously angry with each other.
- The red triangle appeared to want the blue triangle to leave the house in the middle. Blue had to be pushed by the red and Red blocked the entrance. The two were rotating one another...and then stopped touching head to head, _. A bit bizarre."
- Tough one! I don't know: Kid steps outside
- Big triangle trying to make little triangle go out, but he doesn't want to.
- It seems to be a little more interacting than before, touching each other (previous animation was dancing)
- Cats running around and playing with each other, and they just frozen when they sniffed each other

- **Surprise**
- Red triangle trying to get out of the box but couldn't do. So the blue triangle flies in and opens the door, which is on the top of the box. Eventually they have some interaction when they negotiate to go together as we now see them [the subject was looking at last frame of the animation when it stops with the two triangles closed together]
- The large went to open the door and kept going in and out. Then the little got in and they seemed to be friendly.
- Twice the one between the confined space tried to get out of his enclosure, the first time retrieved back into and again, the one on the outside couldn't find the way into the enclosure. The one outside breaks in and was happy to see the other person…maybe an autistic and a non-autistic person communicating…they are happy they found each other, and stay in that enclosure.
- Blue triangle knocked at the door, and hide behind the door, then he came back, knocked at the door again, and it seemed forcing his way in.
- The...uh...first of all it looks like the red T was obviously stuck into the box. The Blue T...he obviously wanted to go out but obviously couldn't. Then I realized there is a door and obviously it was almost as it was a house...and obviously the other one came out from the left [see: right] which is the blue one and she was trying to get in and she peeked behind the door and looked as she went out again and eventually he managed to come through the door and they just shuffled together and that's when the animation ended.
- The red triangle was inside the box, possibly trapped in, and was pacing up and down or waiting until the blue triangle appeared, and came up to the wall and realised it was a door, tapped it and moved to the side. The red one came up to the door opened it and had a look outside. (?) The blue came from behind the door shut

the door.... The blue came around, went inside and stopped in the middle, touching one another.

- Two options: Red triangle seems to be pacing around… pacing around, looking for room home or floor plan…I don't know…it wasn't clear at the beginning…it seems to be a mom frantically waiting for her child
- They are looking for each other
- At first…the…hum…. Hold on a minute…I am trying to think something simple…*[experimernter:"just tell me what it comes to your mind*] Well … embracing each other.
- Two animals who are playing with each other, and then decide to take a sleep in the box

- **Chasing**
- Following mother at home
- Red triangle was chasing the blue triangle, as opposite of the previous chase [it was the "leading" animation].
- The big triangle was chasing the little triangle and then he stopped in the cage with the big one blocking the entrance
- It appears as they were following each other in circles. They appear to be following each other like pigeons like cars and then they park in that garage.
- It looks like the red was chasing the blue one.
- ...Uh...the red one was chasing after the blue one and...They didn't ...the blue wasn't really pursuing the red...they were just...uh...it was basically the red, the blue one was going its own way really and the red was...at first...was following continuously and then...it was having to deciding [subject giggles] the blue one was like...going all over the place really, and then they followed each other into the box and so...I suppose they were,,,it seemed really to be friendly...it wasn't... it wasn't nasty...uh...uh..[the experimenter: that's okay, thank you!]

- Initially Blue t was moving around, then Red appeared into the scene, and started following the blue on the exact path, even when the blue triangle went through the edge of the box, the red one squeezed to follow him, the blue went a little further and went into the box and the Red followed it in and stopped there.
- The blue evades the Red. Chasing.
- Following each other around
- Two animals playing with each other, possibly courting

- **Dancing**
- Kind of dancing
- Red triangle was mummy bird, blue triangle baby bird. They fly together, and ever so often they go apart, make battle, fight, and go apart. They were obviously interacting non verbally with each other about the routine, which was the same at the opposite ends...if that make any sense!
- It's a sort of sequence of dancing.
- Ice-skating!
- Don't know my best guess is that they were doing some form of dance, that the only thing I could come up with!
- Two people going along together...obviously ...very much two friends following each other on the screen [street?] ...they looked pretty...they weren't harming each other...they were being pretty social...and they decided to go their own way towards the end...then came back together again and so...that's what I think.
- It was a sort of dance. They came in together and started mirroring each other.
- Swimming
- They were dancing
- Two fishes swimming in a tank

- **Fighting**
- They were kind of holding and kissing each other
- Basically they were fighting each other! Pushing, swung each other around...(long sentences describing the movement of the triangles)
- Two stags beating each other.
- The two of them are very friendly...they were...well...when I say friendly ...they were obviously two people ...that were...like [subject giggles] kissing each other...that were...fighting each other and.in very much sort of conflicting way...well a nice way wasn't aggressive like some of the others with...in a way they were friendly with each other in a way they.uh...I don't think it was...like some of the others where they were actually fighting...it looks like ...if it was like a man and a woman they were kissing and then...[??]...but then it wasn't like that. It wasn't as aggressive as the other one where the big one was blocking the entrance [see: mocking previously shown.
- Don't know...seemed to be friendly. There is some kind of interaction, they seem to be alive, people...then they die, they become non-biological, non-living...rigid...dead.
- It seems they were dancing around together.
- One of the two triangles appeared to knock one another and then they started a sort of fight, then kept facing each other, and went round and round until they stopped against the wall, I suppose.
- School playground...pushing, wrestling, thrown around, then retreating
- They were having a fight
- Two kids squabbling over a toy

- **Leading**
- Red triangle was mummy bird, blue triangle baby bird. The baby is chasing mom, mom hesitates as she could see the baby bird but other than that she couldn't.

Eventually they go home- square in the middle of the screen. Obviously they are having sort of an interaction. From what I could see they are obviously touching.

- It was a bit like a mother and a baby, the baby following the mother.
- Very disciplined… mechanical movements…representing people's cars.
- Playing the game of following the leader.
- It seems that the little one follow the other at home, the little box
- Blue t would follow the exact path of the red triangle although the (?) tantalising to vary slightly. the blue was going round and round and then stopped one another into the box.
- Playing the game following the leader
- One following the other all the way around
- Pretty similar to the one before [see: mocking]. The red one was a bit more responsive and...he stops turning around and she frightened the other one...and I don't know...they ended up together in this box...uh...so this...they rejoined...I don't know, they...uh...I am not sure.... are they going to be very similar, or are they going to change? [*this animation was the third, presented after mocking*]
- Animals. A dog and a puppy walking around a room and then going back into a chair or something and resting.

- **Billiard**
- They are just playing...that's it!
- They are crashing around. It seems they had too much alcohol, they are drunk.
- Bouncing off each other…kept inside the perimeter, and then stopped having contact each other…bizarre…they were going around…then stopped bouncing into each other….
- It looks completely random, but this time they were bumping into each other quite hard, really.

- I think they were just going basically all over [subject giggles] the place, all over the screen. There was no definitive pattern really or doing anything particularly in terms of behaviours or emotions. They were not particularly aggressive...that's all I can say really.
- As you rightly said (cued animation) the triangles appeared to be randomly bouncing off the walls and bouncing off each other just randomly moving about, that's it!
- Random circle. Bouncing off from each other
- Just random
- Kind of...just go dancing
- They were gypsies on the train who were slipped off with the movement

- **Drifting**
- Clockwise, anticlockwise...then they separate.... If that make any sense!
- Lots of swings and roundabouts.
- Very strange...Big one in the bottom line... the blue one ... strange...clockwise, anticlockwise movements...seem to be following each other...shut up in the right end...directions... around the bay or whatever that is.
- They were moving around and seemed little intimate to each other.
- I am not sure .... about that ...but I think they were playing around the square quite independently
- The triangles were moving around the central square, clockwise, and the spanned off
- Movement....they were following their own pattern
- Moving around randomly
- I think it was two triangles...they were both...again...they weren't...they were trying to interact with each other and the purple one was making more an effort. I think the red one just kept bouncing doing what really...they seemed...the purple one was to see the red one but he wasn't really getting anywhere really....he wasn't interested. He was going on the screen on its own regardless of the other. I can't answer really.

- ..hum....birds flying around a bird table

- **Tennis**
- Both were coping each other, imitating each other. They flew to one wall to the other; eventually they went off together.
- They are bouncing off the walls.
- Don't' know, they seem to be going up and down, east and west walls...there is no movements of the walls, there is no interaction in any occasion between the two, if they were supposed to be people. It was a very strange shape; it didn't pass the central gate!
- Two triangles moving back and forward at the same period.
- Triangles were randomly going towards the walls, rotating slightly...I think
- Spinning around
- Just random
- I am not sure, really. It seems to be...let me think...seems to be bouncing independently.
- I think it was just the...the two triangle were obviously ...weren't...they were going through side to side across the screen. Thy weren't doing really anything particularly, they weren't interacting...just going in one direction each time...horizontally.
- Something on ice in a tap water that was moving side to side

- **Star**
- Red triangle was mummy bird, blue triangle little baby bird. There is no interaction at all. They are obviously trying to chase each other, but they go in opposite directions.
- They are bouncing off the walls; they seem to be a little dizzy.

- It's like two models, it's clockwise and anticlockwise...I couldn't quite follow, it was quite complicated, I concentrated more on the living one, blue one. It was trying to do ... seems to do patterns....
- Moving about seemed to be trying to avoid bumping each other.
- Two triangles seemed to bounce off the inner walls, and occasionally one would bounce off to its side, but it seems to be randomly, not connected, but never collided one another
- Implies moving, bouncing
- Moving around bouncing randomly
- The... uh...uh...the two objects weren't actually interacting they didn't touch each other...and so...I suppose it was almost as if they were two people going about minding their own business...basically.
- Very light objects possibly travelling on a back of something
- Two triangles touching the edges of the square. Playing around.

Control Group

- **Coaxing**
- Initially they look content, then emotional. Red decides to leave, try to persuade Blue to leave, and go outside where they are happy.
- Triangles cuddling inside house, wanted to persuade L to get out, he didn't want, cuddling again.
- Red and Blue are quite happy inside house. Red went outside, blue didn't want to go, red drags him out, then the blue doesn't mind to go outside, it's okay.
- Big tries to make little go outside, but he doesn't want to.
- They are playing inside. The little didn't want to go out, the big tucks him out.
- .... Hum...don't know...they were quite intimate initially. Then red one left and blue didn't follow. Red pushed him out.
- Parent wanted child to go out the house, encouraging him to go out.
- Big wants little to go out the nest, eventually they go out.
- Inside the hose, excited to play, anxiety, big tries to pull little out then dance together.
- It was a sort of fight between Red and Blue. Blue didn't want to get out, and the Red forced him out

•

- **Mocking**
- Red triangle looks almost frustrated or angry changing sides quite rapidly. The Blue began to follow him, changing sides. It seems to me as he was trying to make an impression of him. Red id quite sensitive and turns round but he couldn't catch him making this kind of movement (last words indecipherable from tape)

- Little follows Big as it was his brother, he was mimicking L, B got angry and told him off
- Red walking around, blue come along, following in, stops when he turns around, then starts again until R tell him off, and blue runs away.
- Little follows big, pretended not to be following, then big realize and shouted at him.
- Blue coping red one, following him behind his back, which annoyed him.
- Child following the adult. Child mimics the adult's action. When adult turned, child pretended not to be doing anything, eventually run off
- Parent followed by child. Child plays on his own, then went off
- old guy walk, little follows him, molesting him, old tells him off, and he runs away
- Red walking around, limps, blue follows, imitates, mimicking him, red turns around push him away, shouts, little laughs and runs away.
- Blue was following the Red. The Red was annoyed, and told him off

- **Seducing**
- Small tormenting Red. Red become aggressive, blocks exit so that blue can't go out. Small rubbing against big so to say she was not upset anymore giving a false sense of security, but then the Little moves away trying to escape.
- Blue maybe a female tries to convince Red in a very seductive way to get in.
- Red and Blue outside together. Blue signals Red to go inside the box. R doesn't seem particularly keen to get in, try to force him. Blue persuades him and blue runs out.
- Big tries to trap in little one, but little trick his way out.
- Big guarding little one, trying to escape or something [long silence]...trying to be naughty.
- Blue imprisoned, trying to escape, finds the way out. Red left open the door, blue sees the opportunity to get out, leaves.
- Little forced into the house, persuaded the big to get in, then left

- Mother keeps child inside, tells him to stay in. child tries to force his way out, then he talks to the mother and runs away.
- Red pushes blue in unfriendly way. . Blocks the entrance. Blue tries to get out stroking red, checking there is holes to get out. Then he goes in the corner, sight, implies it's a game to get out...
- Red forced Blue inside the box. Then the Blue tried to get out, he made an attempt but with no success, so it started caressing and tickling him. The Red got inside and the blue tricks his way out.

- **Surprise**
- Big quite repressed inside the box, the other comes along, wants to get in, and the Big opens the door but couldn't see the little. Eventually little finds his way in, and they are happy to be together.
- Blue wants to surprise red, double hide and seek, then gets in, happy
- Red at home anxious to see blue. Blue come along, knocks, red opens, amused not to see him, then knock again and surprise him.
- Little surprising the big one
- Blue was surprising red one.
- Blue teasing red, hide and seek, eventually meet and greet.
- Little knocks at the door, plays hide and seek, finally gets in, embrace.
- Red waits at home. Blue plays a trick. They get in, love each other. It is a nice trick.
- Red waits at home. Other comes along. Opens the door, hiding, and plays a practical joke. Eventually they are reunited, very happy.
- The box was the Red's home. Blue knocked to the door, and hid behind the door. Red got out, didn't see him, closed the door. Red knocked again, Red reopened , Blue was being silly hiding again, then showed up. They got inside, hug each other, and that was the end.

- **Chasing**
- Blue trying to get away from Red, he didn't want, very persistent. Eventually got what he wanted and trapped him into the room
- Red follows Blue in every step. Blue doesn't care. They are nervous, going around and around, then they get in a cage and rest.
- Red following Blue, coping his actions as best as he could.
- Little chasing Big one
- Chasing
- Red chasing Blue. Eventually refuge into the building
- Game of chasing
- Red chasing Blue
- Chasing rather than following
- Red following Blue. At a certain point Blue passed close to the box so that the Red had to shrink to pass through.

- **Dancing**
- Ice skating
- Look like dolphins
- Mirroring each other's actions, they were symmetrical
- Copying each other, mirroring
- Dancing
- Mirroring each other or mimicking in dancing
- Strange dance in a way they were imitating each other
- Dancing or ice skating
- Symmetrical behaviour, spinning with each other, travelling in diagonals
- Sort of dance. A mirroring movement.

- **Fighting**
- Knocking each other, impact increases then come to a stationary position which is quite tense
- Male and female playing. They kissed...then it's not clear what happened because they squashed against the wall.
- Pushing each other, wrestling.
- Playing, pushing each other around
- Fighting
- They were fighting each other
- Playing fighting
- Two arguing begun to fight
- Fighting each other
- At the beginning, they were pushing each other, then fighting, and then pushing again. It seemed they were biting each other.

- **Leading**
- Red dominant, blue following the leader Red. Red going ahead assessing the territory before moving into a smaller area and then resting
- Mother and daughter little does everything big does little follows behind big, little is late but mother waits for her.
- Blue following red, exactly each step, then stop, move around, stops, moves, eventually went into the box.
- [*subject said he didn't follow the animation*]
- Little following big one
- Blue following red. Red allowing blue to get close then went into the house.
- Triangles following, stop, start following again
- Mother walks child follows. Alternatively....they are cars going in the same direction and then parking in the garage.

- Blue following Red, Reds turn at 90 degrees, Blue stops and then follows his journey.
- Blue was following the Red, was copying him. At the end they parked inside

- **Billiard**
- Initially B was composed, then R came in, they loose control, and their bouncing was restricted by the perimeter.
- There was no interaction, it was like they were manipulated by an external force, like wind, or like someone shaking them.
- Triangles were bouncing splitting into different directions
- Random
- Random
- Bouncing around
- Bouncing off the walls. Knocking each other
- Random movement: triangles bashing each other and bouncing off
- Bouncing off each other, off the walls, at the same speed
- It starts with the blue in the centre, and the Red hitting it, then both started bouncing around.

- **Drifting**
- ...They were moving around the room, resting...or.... Sort of thinking, maybe they were assessing the room, eventually they left.
- It was like there was no gravitation in the room
- Red and Blue moving around quite independently
- Random
- Random
- Both going in circular motion, in opposite direction
- Moving around randomly, without touching

- Rotating around the square
- Triangles randomly moving around
- Moving randomly around the box

- **Star**
- Moving playfully affected by the barriers, never crashing together.
- It's random. I don't know...it reminds me a certain toy box in which the pieces have to be pushed inside the holes, never succeeding.
- Triangles moving around randomly, bouncing off the walls
- Bouncing about
- Random
- Both trying to find their way home, independently
- Moving around randomly
- Random movement
- Simple random movement , bouncing off the walls
- Bouncing. Random movement

- **Tennis**
- The two triangles wanted to be together, but they didn't have control over their movements. They are manipulated by the boundaries.
- There was no intentional interaction. Random, with the only exception that they were moving horizontally.
- Bouncing independently from each other
- Not sure…there was no interaction
- No interaction, random
- Bouncing off opposite walls, it seems three were mirroring each other
- Triangles moving horizontally without touching and then went off the screen
- Horizontal random movement

- Triangles moving horizontally, one at the top, the other at the bottom, bouncing
- Random movement. They were constrained along two pathways divided by the box in the middle

References

Abell, F., Krams, M., Ashburner, J., Passingham, R., Friston, K., Frackowiak, R., Happé, F., Frith, C., and Frith, U. 1999. The neuroanatomy of autism: a voxel-based whole brain analysis of structural scans. *Neuroreport* **10**, 1647-51.

Abell, F., Happé, F., and Frith, U. 2000. Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Journal of Cognitive Development* **15**, 1-20.

Abravanel, L. and Gingold, F. 1985. Learning via observation during the second year of life. *Developmental Psychology* **218**, 614-23.

Adolphs, R. 2001. The neurobiology of social cognition. *Current Opinion in Neurobiology* **11**, 231-239.

Adolphs, R., Tranel, D., Hammann, S., Young, A. W., Calder, A. J., Phelps, E. A., Anderson, A., Lee, G. P., and Damasio, A. R. 1999. Recognition of facial emotion in nine individuals with bilateral amygdala damage. *Neuropsychologia* **37**, 1111-17.

Adolphs, R., Sears, L., and Piven, J. 2001. Abnormal Processing of social information from faces in autism. *Journal of Cognitive Neuroscience* **13**, 232-240.

Allison, T., Puce, A., and McCarthy, G. 2000. Social perception from visual cues: role of the STS region. *Trends in Cognitive Science* **4**, 267-278.

Amaral, D., Price, J. L., Pitkanen, A., and Carmichael, S. T. 1992. Anatamical organization of primate amygdaloid cortex. In *The amygdala: neurobiological aspects of emotion, memory and mental dysfunction* (J. Aggleton, Ed.), pp. 1-66. Wiley, New York.

American Psychiatric Association 1994. *Diagnostic and statistical manual of mental disorder*, Washington, DC.

Ammons, R. B. and Ammons, C. H. 1962. *The Quick Test.* Psychological Test Specialists.

Bacon, A. L., Fein, D., Morris, R., Waterhouse, L., and Allen, D. 1998. The responses of autistic children to the distress of others. *Journal of Autism and Developmental Disorders* **28**, 129-42.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., and Rutter, M. 1995. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological Medecine* **25**, 63-77.

Bailey, A., Phillips, W., and Rutter, M. 1996. Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *Journal of Child Psychology and Psychiatry* **37**(1), 89-126.

Bailey, A., Palferman, S., Heavey, L., and Le Couteur, A. 1998. Autism: the phenotype in relatives. *Journal of Autism and Developmental Disorders* **28**, 369-92.

Baron-Cohen, S. 1987. Autism and symbolic play. *British Journal of Developmental Psychology* **5**, 139-48.

Baron-Cohen, S. 1988. Social and pragmatic deficits in autism: cognitive or affective? *Journal of Autism and Developmental Disorders* **18**, 379-402.

Baron-Cohen, S. 1989. Are autistic children behaviourist? An examination of their mental-physical and appeareance-reality distinctions. *Journal of Autism and Developmental Disorders* **30**, 285-98.

Baron-Cohen, S. 1989. The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry* **30**, 285-97.

Baron-Cohen, S. 1991. Do people with autism understand what causes emotion? *Child Developement* **62**, 385-95.

Baron-Cohen, S. 1992. Out of sight or out of mind? Another look at deception in autism. *Journal of Child Psychology and Psychiatry* **33**, 1141-55.

Baron-Cohen, S. 1994. How to build a baby that can read minds: cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition* **13**, 513-52.

Baron-Cohen, S. 1995. *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA, MIT Press.

Baron-Cohen, S., Leslie, A. M., and Frith, U. 1985. Does the autistic have a "Theory of Mind"? *Cognition* **21**, 37-46.

Baron-Cohen, S., Leslie, A. M., and Frith, U. 1986. Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology* **4**, 113-125.

Baron-Cohen, S., Spitz, A., and Cross, P. 1993. Do children with autism recognize surprise? A research note. *Cognition and Emotion* **7**, 507-516.

Baron-Cohen, S., Tager-Flusberg, H., and Cohen, D. 1993. *Understanding other minds: Perspectives from autism*, Oxford University Press, Oxford.

Baron-Cohen, S., Ring, H., Moriarty, J., Schmits, B., Costa, D., and 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**, 640-649.

Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., and and Walker, J. 1995. Are children with autism lind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology* **13**, 379-98.

Baron-Cohen, S. and Swettenham, J. 1996. The relationship between SAM and ToMM: two hypotheses. In *Theories of theories of mind* (P. Carruthers and P. K. Smith, Eds.), Cambridge University Press, Cambridge.

Baron-Cohen, S., Baldwin, D. A., and Crowson, M. 1997. Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Developement* **68**, 48-57.

Baron-Cohen, S. and Hammer, J. 1997. Parents of children with Asperger Syndrome: what is the cognitive phenotype? *Journal of Cognitive Neuroscience* **9**, 548-54.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., and 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**, 813-22.

Baron-Cohen, S., Wheelwright, S., and Jolliffe, T. 1997. Is there a "language of the eyes"? Evidence from normal adults and adults with autism or Asperger Syndrome. *Visual Cognition* **4**, 311-31.

Baron-Cohen, S., Whellwright, S., Stott, C., Bolton, P., and Goodyer, I. 1997. Is there a link between engineering and autism? *Autism* **1**, 101-9.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., and Plaisted, K. 1999. Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders* **29**, 407-18.

Baron-Cohen, S., Ring, H. A., Wheelwright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., and Williams, S. C. 1999. Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience* **11**, 1891-8.

Baron-Cohen, S., Wheelwright, S., Stone, V., and Rutherford, M. 1999. A mathematician, a physicist and a computer scientist with Asperger syndrome: Performance on folk psychology and folk physics tests. *Neurocase* **5**, 475-483.

Baron-Cohen, S., Ring, H. A., Bullmore, E. T., Weelwright, S., Ashwin, C., and Williams, S. C. R. 2000. The amygdala theory of autism. *Neuroscience and Behavioual Reviews* 355-64.

Baron-Cohen, S., Tager-Flusberg, H., and Cohen, D. 2000. *Understanding other minds: Perspectives from developmental neuroscience*, Oxford University Press, Oxford.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and 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**, 241-51.

Bartsch, K. and Wellman, H. 1995. *Children talk about the mind*, Oxford University Press.

Bauman, M. L. 1996. Brief report: neuroanatomic observations of the brain in pervasive developmental disorders. *Journal of Autism and Developmental Disord*ers **26**, 199-203.

Bauman, M. L. and Kemper, T. L. 1994. *Neuroanatomic observation of the brain in autism*, John Hopkins University Press, Baltimore.

Berry, D. S., Misovich, S. J., Kean, K. J., and Baron, R. M. 1992. Effects of disruption of structure and motion perceptions of social causality. *Personality and Social Psychology Bulletin* **18**, 237-44.

Berry, D. S. and Springer, K. 1993. Structure, motion, and preschoolers' perception of social causality. *Ecological Psychology* **5**, 273-83.

Blair, R. J. R. 1999. Psychophysiological responsiveness to the distress of others in children with autism. *Personality and Individual Differences* **26**, 477-485.

Blair R.J. and Coles, M. 2000. Expression recognition and behavioural problems in early adolescence. *Cognitive Development* **15**, 421-34

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. 1998. Central cancellation of self-produced tickle sensation. *Nature Neuroscience* **1**, 635-640.

Bleuler E.T. 1952. *Dementia precox or the group of schizophrenias*, International University Press, NewYork.

Bloom, P. and German, T. P. 2000. Two reasons to abandon the false belief task as a test of theory of mind. *Cognition* **77** (1), B25-B31

Blythe, P. W., Todd, P. M., and Miller, G. F. 1999. How motion reveals intention: categorizing social interactions. In *Simple heuristics that makes us smart* (G. Gigerenzer, P. M. Todd, and ABC research group, Eds.), Oxford University Press, NewYork.

Bonda, E., Petrides, M., Ostry, D., and Evans, A. 1996. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience* **16**, 3737-44.

Bottini, G., Corcoran, R., Sterzi, R., Paulesu, E., Schenone, P., Scarpa, P., Frackowiak, R. S., and Frith, C. D. 1994. The role of the right hemisphere in the interpretation of figurative aspects of language. A positron emission tomography activation study. *Brain* **117**, 1241-53.

Boussaoud, D., Ungerleider, L. G., and Desimone, R. 1990. Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *Journal of Comparative Neurology* **296**, 462-95.

Bowler, D. M. 1992. Theory of mind in Asperger's syndrome. *Journal of Child Psychology and Psychiatry* **33**, 877-93.

Bowler, D. M. and Thommen, E. 2000. Attribution of mechanical and social causality to animated displays by children with autism. *Autism* **4**, 147-171.

Brentano F. von 1970. *Psychology from an empirical standpoint*, Rutledge, New York.

Brothers, L. 1997. *Friday's footprint: How society shapes the human mind*, Oxford, Oxford University Press.

Bruner, J. 1976. From communication to language: a psychological perspective. *Cognition* **3**, 255-87.

Brunet, E., Sarfati, Y., Hardy-Bayle, M. C., and Decety, J. 2000. A PET investigation of the attribution of intentions with a nonverbal task. *Neuroimage* **11**, 157-66.

Buchel, C., Price, C., and Friston, K. 1998. A multimodal language region in the central visual pathway. *Nature* **394**,274-77

Buitelaar, J. K., Van der Wees, M., Swabb-Barneveld, H., and Van der Gaag, R. J. 1999. Theory of Mind and emotion-recognition functioning in autistic spectrum disorders and in psychiatric control and normal children. *Development and psychopathology* **11**, 39-58.

Butterworth, G. 1991. The ontogeny and phylogeny of joint visual attention. In *Natural Theories of Mind: evolution, development, and simulation in everyday mind reading* (A. Whitten, Ed.), Basil Blackwell, Oxford.

Calder, A. J., Young, A. W., Rowland, D., and Perrett, D. I. 1996. Facial emotion recognition after bilateral amygdala damage: Differentially severe impairment of fear. *Cognitive Neuropsychology* **13**, 699-745.

Calvert, G. A., Bullmore, E. T., Brammer, M. J., Campbell, R., Williams, S. C., McGuire, P. K., Woodruff, P. W., Iversen, S. D., and David, A. S. 1997. Activation of auditory cortex during silent lipreading. *Science* **276**, 593-6.

Carruthers, P. and Smith, P. K. 1996. *Theories of Theory of Mind*, Cambridge University Press, Cambridge.

Carter, A. S., Volkmar, F. R., Sparrow, S. S., Wang, J. J., Lord, C., Dawson, G., Fombonne, E., Loveland, K., Mesibov, G., and Schopler, E. 1998. The Vineland Adaptive Behavior Scales: supplementary norms for individuals with autism. *Journal of Autism and Developmental Disorders* **28**, 287-302.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., and Cohen, J. D. 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* **280**, 747-9.

Castelli, F., Happé, F., Frith, U., and Frith, C. 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *Neuroimage* **12**, 314-25.

Charman, T. and Baron-Cohen, S. 1992. Understanding drawings and beliefs: a further test of the metarepresentation theory of autism: a research note. *Journal of Child Psychology and Psychiatry* **33**, 1105-12.

Charman, T. and Baron-Cohen, S. 1995. Understanding models, photos, and beliefs: a test of the modularity thesis of metarepresentation. *Cognitive Development* **10**, 287-98.

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., and Drew, A. 2000. Testing joint attention, imitation, and play as infancy precursor to language and theory of mind. *Cognitive Development* **15**.

Corcoran, R., Mercer, G., and Frith, C. D. 1995. Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophrenia Research* **17**, 5-13.

Corona, R., Dissanayake, C., Arbelle, S., Wellington, P., and Sigman, M. 1998. Is affect aversive to young children with autism? Behavioral and cardiac responses to experimenter's distress. *Child Developement* **69**, 1494-502.

Critchley, H. D., Daly, E. M., Bullmore, E. T., Williams, S. C., Van Amelsvoort, T., Robertson, D. M., Rowe, A., Phillips, M., McAlonan, G., Howlin, P., and Murphy, D. G. 2000. The functional neuroanatomy of social behaviour: changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain* **123** , 2203-12.

Csibra, G., Gergely, G., Biro, S., Koos, O., and Brockbank, M. 1999. Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition* **72**, 237-267.

Dahlgren, S. O. and Trillingsgaard, A. 1996. Theory of mind in non-retarded children with autism and Asperger's syndrome. A research note. *Journal of Child Psychology and Psychiatry* **37**, 759-63.

Darwin, C. 1998. *The expression of the emotions of man and animals*, Harper Collins, London.

Davidson, R. J. 1992. Prolegomenon to the structure of emotion. *Cognition and Emotion* **6**, 245-68.

Dennett, D. C. 1978. Beliefs about beliefs. *Behavioural and Brain Sciences* **1**, 568-70.

Dennett, D. C. 1990. True believers: the intentional strategy and why it works. In *Mind and Cognition* (W. J. Lycan, Ed.), pp. 150-167. Blackwell, Oxford, UK.

Dissanayake, C., Sigman, M., and Kasari, C. 1996. Long-term stability of individual differences in the emotional responsiveness of children with autism. *Journal of Child Psychology and Psychiatry* **37**, 461-7.

Duchaine, B., Cosmides, L., and Tooby, J. 2001. Evolutionary psychology and the brain. *Current Opinion in Neurobiology* **11**, 225-30.

Duvernoy, H. M. 1999. *The human brain: Surface, three-dimensional sectional anatomy with MRI, and blood supply*, Springer Wien, New York.

Ekman, P. 1984. Expression and the nature of emotion. In *Approaches to Emotion* (Ekman and Sherer , Ed.), pp. 319-343. Erlbaum, Hillsdale, NJ.

Ekman, P. 1989. The argument and evidence about universals in facial expressions of emotions. In *Handbook of social psychophysiology* (H. M. Wagner, Ed.), pp. 143-64. Wiley, Chicester.

Ekman, P. 1992. Are there basic emotions? *Psychological Review* **99**, 550-553.

Ekamn, P. and Friesen, W. V. 1969. Non verbal leekage and clues to deception. *Psychiatry* **32**, 88-105.

Ekman, P. and Friesen, W. V. 1971. Constraints across cultures in the face and emotion. *Journal of Personality and Social Psychology* **17**, 124-129.

Ekman, P., Friesen, W. V., and Ellsworth, P. 1972. What emotion categories or dimensions can observers judge from facial behaviour. In *Emotion in the human face* (P. Ekman, Ed.), Cambridge University Press, Cambridge.

Ekman, P., Friesen, W. V. and Ellsworth, P. 1972. *Emotion in the human face*, Pergamon Press, Elmsford, NY.

Ekman, P. and Friesen, W. V. 1975. *Unmasking the face*, Prentice Hall, Englewood Cliffs, NJ.

Ekman, P. and Friesen, W. V. 1976. *Pictures of Facial Affect*, Consulting Psychologists Press, Palo Alto.

Ekman, P. and Oster, H. 1979. Facial Expression of emotion. *Annual Review of Psychology* **30**, 527-54.

Etcoff, N. L. and Magee, J. J. 1992. Categorical perception of facial expressions. *Cognition* **44**, 227-240.

Evans, A. C., Kamber, M., Collins, D. L., and MacDonald, D. 1994. A MRI based probabilistic atlas of neuroanatomy. In *Magnetic Resonance Scanning and Epilepsy* (S. Shorvon, D. Fish, F. Andermann, G. M. Bydder, and H. Steffan, Eds.), pp. 263-74. Plenum, New York.

Felleman, D. J. and Van Essen, D. C. 1987. Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *Journal of Neurophysiology* **57** (4), 889-920.

Felleman, D. J., Burkhalter, A., and Van Essen, D. C. 1997. Cortical connections of areas V3 and VP of macaque monkey extratriate visual cortex. *Journal of Comparative Neurology* **379**(1), 21-47.

Fine, C. and Blair, R. J. R. 2000. Mini review: The cognitive and emotional effects of amygdala damage. *Neurocase* **6**, 435-50.

Fine, C., Lumsden, J., and Blair, R. J. R. 2001. Dissociation between mentalizing ability and executive functions in a patient with early left amygdala damage. *Brain* **124**, 287-298.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S., and Dolan, R. J. 1997. Neural mechanisms involved in the processing of global and local aspects of hierarchically organized visual stimuli. *Brain* **120**, 1779-91.

Fink, G. R., Marshall, J. C., Halligan, P. W., Frith, C. D., Frackowiak, R. S., and Dolan, R. J. 1997. Hemispheric specialization for global and local processing: the effect of stimulus category. *Proceedings Royal Society London B Biology Science* **264**, 487-94.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S., and Frith, C. D. 1995. Other minds in the brain: a functional imaging study of Theory of Mind in story comprehension. *Cognition* **57**, 109-28.

Fodor J. 1992. A theory of the child's theory of mind. *Cognition* **44**, 283-96.

Folstein, S. and Rutter, M. 1977. Genetic influences and infantile autism. *Nature* **265**, 726-8.

Folstein, S. and Rutter, M. 1977. Infantile autism: a genetic study of 21 twin pairs. *Journal of Child Psychology and Psychiatry* **18**, 297-321.

Fombonne, E. 1999. The epideomology of autism: a review. *Psychological Medicine* **29**, 769-86.

Fombonne, E., Siddons, F., Achard, S., Frith, U., and Happe, F. 1994. Adaptive behaviour and theory of mind in autism. *European Child and Adolescent*

Psychiatry **3**, 176-186.

Fox, P. T. and Mintun, M. A. 1989. Noninvasive functional brain mapping by change-distribution analysis of averaged PET images of $H_2{}^{15}O$ tissue activity. *Journal of Nuclear Medecine* **30**, 141-9.

Frackowiak, R., Friston, K., Frith, C., Dolan, R., and Mazziotta, J. 1997. *Human Brain Function*, Academic Press, San Diego.

Friston, K. J., Frith, C. D., Liddle, P. F., Dolan, R. J., Lammertsma, A. A., and Frackowiack, R. S. 1990. The relationship between global and local changes in PET scans. *Journal of Cerebral Blood Flow Metabolism* **10**, 458-466.

Friston, K. J., Worsley, K., Frackowiak, R. S., Mazziotta, J. C., and Evans, A. C. 1994. Assessing the significance of focal activations using their spatial extent. *Human Brain Mapping* **1**, 214-220.

Friston, K. J., Ashburner, J., Poline, J.-B., Heather, J. D. and Frackowiak, R.S.J. 1995. Spatial realignement and normalisation of images. *Human Brain Mapping* **2**, 165-189.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. 1995. Statistical parametric maps in functional imaging: a general linear approach. *Human Brain Mapping* **2**, 189-210.

Frith, C. D. and Frith, U. 2000. The physiological basis of theory of mind. In *Understanding other minds: Perspective from developmental neuroscience. Second edition* (S. Baron-Cohen, H. Tager-Flusberg, and D. Cohen, Eds.), 335-356, Oxford University Press, Oxford.

Frith, U. 1989. *Autism: Explaining the Enigma.* Oxford, Blackwell.

Frith, U. 1991 *Autism and Asperger syndrome.* Cambridge, Cambridge University Press.

Frith, U. 1997. The neurocognitive basis of autism. *Trends in Cognitive Sciences* **1**, 73-77.

Frith, U., Morton, J., and Leslie, A. M. 1991. The cognitive basis of a biological disorder: autism. *Trends in Neuroscience* **14**, 433-8.

Frith, U., Happé, F., and Siddons, F. 1994. Autism and Theory of Mind in everyday life. *Social Development* **3**, 108-24.

Frith, U. and Frith, C. 1999. Interacting minds - A biological basis. *Science* **286**, 1692-1695.

Gallagher, H. L., Happé, F., Brunswick, N., Fletcher, P. C., Frith, U., and Frith, C. D. 2000. Reading the mind in cartoons and stories: an fMRI study of Theory of Mind in verbal and nonverbal tasks. *Neuropsychologia* **38**, 11-21.

Gergely, G., Nadasdy, Z., Csibra, G., and Biro, S. 1995. Taking the intentional stance at twelve months of age. *Cognition* **56**, 165-193.

Gillberg, C. and Coleman, M. 2000. *The biology of the Autistic Syndrome*, Mac Keith Press, London.

Goel, V., Sadato, N., and Hallett, M. 1995. Modelling other minds. *Neuroreport* **6**, 1741-46.

Goldman, A. 1993. The psychology of folk psychology. *Behavioral and Brain Sciences* **16**, 15-28.

Gopnick, A., Capps, L., and Meltzoff, N. M. 2000. Early Theories of Mind: what the theory theory can tell us about autism. In *Understanding other minds: Perspective from developmental neuroscience. Second edition* (S. Baron-Cohen, H. Tager-Flusberg, and D. Cohen, Eds.), Oxford University Press, Oxford.

Gorno Tempini, M. L., Price, C., Josephs, O., Vandenberghe, R., Cappa, S. F., Kapur, N., and Frackowiack, R. S. 1998. The neural systems sustaining face and proper name processing. *Brain* **121**, 2103-18.

Grady, C. L. 1999. Neuroimaging and activation of the frontal lobes. In *The human frontal lobes: functions and disorders* (B. L. Miller and J. L. Cummings, Eds.), pp. 196-230. Guildford Press, New York.

Grezes, J., Costes, N., and Decety, J. 1998. Top-down effect of strategy on the perception of human biological motion: a PET investigation. *Cognitive Neuropsychology* **15**, 553-82.

Griffiths, P. E. 1997. *What emotions really are: The problem of psychological categories.*, The University of Chicago Press, Chicago, London.

Grossman, J. B., Klin, A., Carter, A. S., and Volkmar, V. F. 2000. Verbal Bias in recognition of facial emotions in children with Asperger Syndrome. *Journal of Child Psychology and Psychiatry* **41**, 369-79.

Happé, F. 1993. Communicative competence and Theory of Mind in autism: a test of Relevance Theory. *Cognition* **48**, 101-19.

Happé, F. 1994. An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders* **24**, 129-54.

Happé, F. 1994. *Autism: An introduction to psychological theory.* London, UCL Press/Psychology Press.

Happé, F. 1995. The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Developement* **66**, 843-55.

Happé, F., Ehlers, S., Fletcher, P., Frith, U., Johansson, M., Gillberg, C., Dolan, R., Frackowiak, R., and Frith, C. D. 1996. 'Theory of mind' in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport* **8**, 197-201.

Happé, F. and Frith, U. 1996. The neuropsychology of autism. *Brain* **119**, 1377-1400.

Happé, F., Winner, E., and Brownell, H. 1998. The getting widsom: Theory of Mind in old age. *Developmental Psychology* **34**, 358-362.

Happé, F., Briskman, J., and Frith, U. 2001. Exploring the cognitive phenotype of autism: Weak "central coherence" in parents and siblings of children with autism. I. Experimental tests. *Journal of Child Psychology and Psychiatry* **42**, 299-307.

Heider, F. and Simmel, M. 1944. An experimental study of apparent behavior. *American Journal of Psychology* **57**, 1377-259.

Heyes, C. 2001. Causes and consequences of imitation. *Trends in Cognitive Science* **5**, 253-61.

Hirschfeld, L. A. and Gelman, S. A. 1994. *Mapping the mind: Domain specificity in recognition and culture*, Cambridge University Press, New York.

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. *Journal of Child Psychology and Psychiatry* **27**, 671-80.

Hobson, R. P. 1993. Understanding persons: the role of affect. In *Understanding other minds: Perspectives from autism* (S. Baron-Cohen, H. Tager-Flusberg, and D.

Cohen, Eds.), pp. 204-27. Oxford University Press, Oxford.

Hobson, R. P., Ouston, J., and Lee, A. 1988. Emotion recognition in autism: coordinating faces and voices. *Psychological Medicine* **18**, 911-23.

Hobson, R. P., Ouston, J., and Lee, A. 1988. What's in a face? The case of autism. *British Journal of Psychology* **79**, 441-53.

Howard, M. A., Cowell, P. E., Boucher, J., Broks, P., Mayes, A., Farrant, A., and Roberts, N. 2000. Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. *Neuroreport* **11**, 2931-5.

Izard, C. E. 1971. *The face of emotion*, Appleton Century Crofts, New York.

Izard, C. E. 1977. *Human emotion*, Prenum Press , New York.

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

Kanwisher, N., McDermott, J., and Chun, M. M. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience* **17**, 4302-11.

Karmiloff-Smith, A., Grant, J., Bellugi, U., and Baron-Cohen, S. 1995. Is there a social module? Language, face-processing and theory of mind in Williams syndrome and autism. *Journal of Cognitive Neuroscience* **7**, 196-208.

Kasari, C., Sigman, M., and Yiriya, N. 1993. Focused and social attention of autistic children in interactions with familiar and unfamiliar adults. *Development and Psychopatology* **5**, 401-412.

Kasari, C., Sigman, M. D., Baumgartner, P., and Stipek, D. J. 1993. Pride and mastery in children with autism. *Journal of Child Psychology and Psychiatry* **34**, 353-362.

Kassin, S. M. 1981. Heider and Simmel (1944) revisited: Causal attribution and the animated filmtechnique. *Review of Personality and Social Psychology* **3**, 145-70.

Kestenbaum, R. 1992. Feeling happy versus feeling good: the processing of discrete and global categories of emotional expressions by children and adults. *Developmental Psychology* **28**, 1132-42.

Klin, A. 2000. Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The Social Attribution Task. *Journal of Child Psychology and Psychiatry* **41**, 831-46.

Klin, A., Schultz, R., and Cohen, D. J. 2000. Theory of Mind in action, Developmental perspective on social neuroscience. In *Understanding other minds: Perspective from developmental neuroscience* (S. Baron-Cohen, H. Tager-Flusberg, and D. Cohen, Eds.), pp. 357-388. Oxford university Press, Oxford.

Klin, A., Volkmar, F., and Sparrow, S. 2000. *Asperger Syndrome*, Guildford Press.

Lane, R. D., Fink, G. R., Chau, P. M., and Dolan, R. J. 1997. Neural activation during selective attention to subjective emotional responses. *Neuroreport* **8**, 3969-3972.

Lazarus, R. S. 1991. *Emotion and adaptation*, Oxford University Press, Oxford.

Le Couteur, A., Bailey, A., Goode, S., Pickles, A., Robertson, S., Gottesman, I., and Rutter, M. 1996. A broader phenotype of autism: the clinical spectrum in twins. . *Journal of Child Psychology and Psychiatry* **37**, 785-801.

Leekam, S. R. and Perner, J. 1991. Does the autistic child have a metarepresentational deficit? *Cognition* **40**, 203-18.

Leslie, A. M. 1987. Pretense and representation: The origins of Theory of Mind. *Psychological Review* **94**, 412-426.

Leslie, A. M. 1994. ToMM, ToBy, and Agency: Core architecture and domain specificity. In *Mapping the mind: Domain specificity in cognition and culture* (L.A. Hirschfeld and S.A. Gelman, Eds.) pp. 119-148. New York, Cambridge University Press.

Leslie, A. M. and Frith, U. 1988. Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology* **6**, 315-324.

Leslie, A. M. and Happé, F. 1989. Autism and ostensive communication: The relevance of metarepresentation. *Development and Psychopatology* **1**, 205-12.

Leslie, A. M. and Frith, U. 1990. Prospects for a cognitive neuropsychology of autism: Hobson's choice. *Psychology Review* **97**, 122-31.

Leslie, A. M. and Thaiss, L. 1992. Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition* **43**, 225-251.

Leslie, A. and Roth, D. 1993. What representation teaches us about metarepresentation. In *Understanding other minds: Perspectives from Autism* (S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen, Eds.), pp. 83-111. Oxford University Press, Oxford.

Leslie, A. M. and Polizzi, P. 1998. Inhibitory processing in the false belief task: Two conjectures. *Developmental Science* **1**, 247-253.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A. M. N., Levrier, O., Salomon, G., Dehaene, S., Cohen, L., and Mehler, J. 1993. The cortical representation of speech. *Journal of Cognitive Neuroscience* **5**, 467-79.

McGuire, P. K., Paulesu, E., Frackowiak, R. S., and Frith, C. D. 1996. Brain activity during stimulus independent thought. *Neuroreport* **7**, 2095-9.

McGuire, P. K., Silbersweig, D. A., and Frith, C. D. 1996. Functional neuroanatomy of verbal self-monitoring. *Brain* **119**, 907-17.

Meltzoff, A. N. 1995. Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology* **31**, 838-850.

Meltzoff, A. N. and Moore, M. K. 1999. Infant intersubjectivity: broadening the dialogue to include imitation, identity and intention. In *Intersubjective communication and emotion in early ontogeny* (S. Braten, Ed.), Cambridge University Press.

Michotte, A. 1963. *The perception of causality*, Metheun, London.

Montgomery, D. E. and Montgomery, D. A. 1999. The influence of movement and outcome on young children's attributions of intention. *British Journal of Developmental Psychology* **17**, 245-261

Morris, M. V. and Peng, K. 1994. Culture and cause: American and Chinese attribution for social and physical events. *Journal of Personality and Social Psychology* **67**, 949-971.

Morton, J. and Frith, U. 1995. Causal modeling: Structural approaches to developmental psychopathology. In Developmental Psychopathology (D. Cicchetti and D. Cohen, Eds.) pp. 357-390. New York, Wiley.

Mundy, P., Sigman, M., and Kasari, C. 1993. The theory of Mind and joint-attention deficits in autism. In *Understanding other minds: Perspective from autism* (S. Baron-Cohen, H. Tager-Flusberg, and D. Cohen, Eds.), Oxford University Press, Oxford.

Nobre, A. C., Coull, J. T., Frith, C. D., and Mesulam, M. M. 1999. Orbitofrontal cortex is activated during breaches of expectation in tasks of visual attention. *Nature Neuroscience* **2**, 11-12.

Olson, I. R., Chun, M. M., and Allison, T. 2001. Contextual guidance of attention: human intracranial event-related potential evidence for feedback modulation in anatomically early temporally late stages of visual processing. *Brain* **124**, 1417-25.

Ozonoff, S., Pennington, B., and Rogers, S. 1990. Are there emotion perception deficits in young autistic children? *Journal of Child Psychology and Psychiatry* **31**, 343-363.

Paus, T., Koski, L., Caramanos, Z., and Westbury, C. 1998. Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. *Neuroreport* **9**, 37-47.

Perner, J. and Wimmer, H. 1985. "Joseph thinks that Mary thinks that...": Attribution of second order beiefs by 5- to 10-year-old children. *Journal of Experimental Psychology* **39**, 437-71.

Perner, J., Frith, U., Leslie, A. M., and Leekam, S. R. 1989. Exploration of the autistic child's theory of mind: knowledge, belief, and communication. *Child Development* **60**, 688-700.

Phillips, W., Gomez, J. C., Baron-Cohen, S., Laa, V., and Riviere, A. 1995. Treating people as objects, agents, or "subjects": how young children with and without autism make requests. *Journal of Child Psychology and Psychiatry* **36**, 1383-98 .

Phillips, W., Baron-Cohen, S., and Rutter, M. 1998. Understanding intention in normal development and in autism. *British Journal of Developmental Psychology* **16**, 337-348.

Povinelli , D. J. and Giambrone, S. Inferring other minds: failure of the argument by analogy. *Philosophical Topics* (in press)

Povinelli, D. J. and O'Neill, D. K. 2000. Do Chimpanzees use their gestures to instruct each other? In *Understanding other minds: perspectives from developmental cognitive neuroscience* (S. Baron-Cohen, H. Tager-Flusberg and D. Cohen, eds.), pp. 459-87. Oxford University Press, Oxford.

Premack, D. 1990. The infant's theory of self-propelled objects. *Cognition* **36**, 1-16.

Premack, D. and Woodruff, G. 1978. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* **1** (4), 515-526.

Premack, D. and Premack, A. J. 1994. Moral belief: form versus content. In *Mapping the mind: Domain specificity in cognition and culture.* (L. A. Hirschfeld and S. A. Gelman , eds.), Cambridge University Press, New York.

Price, C. J., Moore, J., Humphreys, G. W., and Wise, R. J. 1997. Segregrating semantic from phonological processes during reading. *Journal of Cognitive Neuroscience* **9**, 727-33.

Puce, A., Allison, T., Asgari, M., Gore, J. C., and McCarthy, G. 1996. Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *Journal Neuroscience* **16**, 5205-15.

Puce, A., Allison, T., Bentin, S., Gore, J. C., and McCarthy, G. 1998. Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience* **18**, 2188-99.

Quine W.V. 1961. *From a logical point of view*, Harvard University Press, Cambridge.

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., and Bushnell, M. C. 1997. Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* **277**, 968-971.

Raven, J. C. 1958. *Standard Progressive Matrices*. Cambridge, Cambridge University Press.

Reaux, J. E., Theall, L. A., and Povinelli, D. J. 1999. A longitudinal investigations of chimpanzees' understanding of visual perception. *Child Development* **70**, 275-290.

Ritvo, E. R., Freeman, B. J., Mason-Brothers, A., Mo, A., and Ritvo, A. M. 1985. Concordance for the syndrome of autism in 40 pairs of afflicted twins. *American Journal of Psychiatry* **142**, 74-7.

Rochat, P., Morgan, R., and Carpenter, M. 1997. Young infants' sensitivity to movement information specifying social causality. *Cognitive Development* **12**, 537-561.

Roeyers, H., Buysse, A., Ponnet, K., and Pichal, B. 2001. Advancing advanced mind-reading tests: empathic accuracy in adults with a pervasive developmental disorder. *Journal of Child Psychology and Psychiatry* **42**, 271-8.

Roth, D. and Leslie, A. M. 1998. Solving belief problems: towards a task analysis. *Cognition* **66**, 1-31.

Schlottmann, A. and Surian, L. 1999. Do 9-month-olds perceive causation at-a-distance? *Perception* **28**, 1105-13.

Scholl, B. J. and Leslie, A. M. 1999. Modularity, development and Theory of Mind. *Mind and Language* **14**, 131-153.

Scholl, B. J. and Tremoulet, P. D. 2000. Perceptual causality and animacy. *Trends in Cognitive Science* **4**, 299-309.

Schopler, E. a. M. G. 1988. *Diagnosis and assessment in autism*, Plenum, New York.

Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., Skudlarski, P., Lacadie, C., Cohen, D. J., and Gore, J. C. 2000. Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Archive of General Psychiatry* **57**, 331-40.

Siegal, M. and Beattie, K. 1991. Where to look first for children's knowledge of false belief. *Cognition* **38**, 1-12.

Sigman, M. D., Kasari, C., Kwon, J., and Yirmiya, N. 1992. Responses to the negative emotions of others by autistic, mentally retarded, and normal children. *Child Development* **63**, 796-807.

Slater, A., Morison, V., and Rose, D. 1984. Habituation in the newborn. *Infant Behaviour and Development* **7,** 183-200.

Sperber, D. 2000. Metarepresentations in an evolutionary perspectives. In *Metarepresentations: a multidisciplinary perspectives* (D. Sperber, Ed.), pp. 117-37. Oxford University Press, Oxford.

Sperber, D. and Wilson, D. 1986. *Relevance: Communication and cognition*, Blackwell, Oxford.

Steffenburg, S., Gillberg, C., Hellgren, L., Andersson, L., Gillberg, I. C., Jakobsson, G., and Bohman, M. 1989. A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *Journal of Child Psychology and Psychiatry* **30**, 405-16.

Stevens, D., Charman, T., and Blair, R. J. 2001. Recognition of emotion in facial expressions and vocal tones in children with psychopatic tendencies. *Journal of Genetic Psychology* **162** 201-11

Sugase, Y., Yamane, S., Ueno, S., and Kawano, K. 1999. Global and fine information coded by single neurons in the temporal visual cortex. *Nature* **400**, 869-73.

Sullivan, K., Zaitchik, D., and Tager-Flusberg, H. 1994. Preschooloers can attribute second-order beliefs. *Developmental Psychology* **30**, 395-402.

Surian, L. and Leslie, A. M. 1999. Competence and performance in false belief understanding: a comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology* **17**, 141-55.

Tager-Flusberg, H. 1993. What language reveals about the understanding of minds in children with autism. In *Understanding other minds:Perspectives from autism* (S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen, Eds.), Oxford University Press.

Tailarach, J. and Touroux, P. A. 1988. *Co-planar stereotaxic atlas of a human brain*, Thieme Verlag, Stuttgard.

Tantam, D., Monaghan, L., Nicholson, H., and Stirling, J. 1989. Autistic children's ability to interpret faces: a research note. *Journal of Child Psychology and Psychiatry* **30**, 623-630.

Tomasello, M. 1999. *The cultural origins of human cognition*, Harvard University Press, Cambridge, MA.

Townsend, D. W., Defrise, M., Geissbuhler, A., Spinks, T. J., Bailey, D. L., Gilardi, M. C., and Jones, T. 1991. Normalisation and reconstruction of PET data acquired by a multi-ring camera with septa retracted. *Medical Progress Through Technology* **17**, 223-8.

Vandenberghe, R., Price, C., Wise, R., Josephs, O., and Frackowiack, R. S. 1996. Functional anatomy of common semantic system for words and pictures. *Nature* **383**, 254-6.

Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happé, F., Falkai, P., Maier, W., Shah, N. J., Fink, G. R., and Zilles, K. 2001. Mind reading: neural mechanisms of Theory of Mind and self-perspective. *Neuroimage* **14**, 170-81.

Weeks, S. J. and Hobson, R. P. 1987. The salience of facial expression for autistic children. *Journal of Child Psychology and Psychiatry* **28**, 137-51.

Wellman, H. M., Cross, D., and Watson, J. 2001. Meta-analysis of Theory of Mind development: the truth about false belief. *Child Development* **72**, 655-84.

Whiten, A. and Brown, J. D. 1999. Imitation and the reading of other minds: Perspectives from the study of autism, normal children and non-human primates.

In *Intersubjective communication and emotion in early ontogeny* (S. Braten, Ed.), Cambridge University Press.

Wimmer, H. and Perner, J. 1983. Beliefs about beliefs: Representation and the constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103-128.

Wimmer, H. and Hartl, M. 1991. Against the Cartesian view of mind: young children's difficulty with own false beliefs. *British Journal of Developmental Psychology* **9**, 125-136.

Wing, L. 1981. Language, social and cognitive impairments in autism and severe mental retardation. *Journal of Autism and Developmental Disorders* **11**, 31-44.

Wing, L. 1981. Asperger's syndrome: a clinical account. *Psychological Medicine* **11**, 115-29.

Wing, L. 1988. The continuum of autistic charecteristics. In *Diagnosis and assessment in autism* (E. Schopler and G. B. Mesibov, Eds.), Plenum, New York.

Wing, L. and Gould, J. 1979. Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders* **9**, 11-29.

World Health Organization 1993. *International classification of diseases*, Geneva.

Young, A. W., Rowland, D., Calder, A. J., Etcoff, N. L., Seth, A., and Perrett, D. I. 1997. Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition* **63**, 271-313.

Zaitchik, D. 1990. When representations conflict with reality: the pre-schooler's problem with false beliefs and "false" photographs. *Cognition* **35**, 41-68.